Boletim da Sociedade Astronômica Brasileira, **34**, no. 1, 15-19

Extreme emission line galaxies at z<0.05 with S-PLUS

A.R. Lopes[1,2] & E. Telles[2]

[1] Instituto de Astrofísica de La Plata, CONICET–UNLP, Argentina e-mail: amandalopes1920@gmail.com
[2] Observatório Nacional, Brasil

Abstract. Extreme emission line galaxies (EELGs) are low-metallicity star-forming galaxies that can be regarded as local analogs of galaxies in the early Universe. Due to their intense emission lines and low continuum, these objects are ideal targets to be studied using narrow-band photometry. Here, we will discuss the approach and results in our search for EELG candidates at $z < 0.05$ using S-PLUS, a wide field 12 (5 broad and 7 narrow) bands survey that observes the Southern Sky. The methodology relies in an initial sample obtained from the detection of an excess in a given narrow band due to the presence of an emission line. EELGs at $z < 0.05$ are [OIII] emitters, and consequently should present an excess in S-PLUS $J0515$ filter. Based on $\sim$ 440 known EELGs from SDSS cross-matched with S-PLUS DR3 and unWISE data in the Stripe82 area, we remove contaminants (QSOs, stars) from our sample. We find a total of 248 EELGs, and after additional analysis, including SED-fitting, we confirm that this sample contains galaxies with low mass, low metallicity, and high equivalent widths (for [OIII]: $\gtrsim$ 200Å, for Hα: $\gtrsim$ 100Å). Spectroscopic follow-ups are needed to confirm our findings.

Resumo. Galáxias com linha de emissão extrema são galáxias com formação estelar e baixa metalicidade que podem ser consideradas análogas locais de galáxias do Universo Primoridal. Devido às suas linhas de emissão intensas e baixo contínuo, esses objetos são alvos ideais para serem estudados usando fotometria de bandas estreitas. Aqui nós discutiremos a abordagem e resultados da nossa busca por candidatas à EELGs em $z < 0.05$ usando S-PLUS, um levantamento de dados em 12 bandas (5 largas e 7 estreitas) que observa o Céu Austral. A metodologia depende de uma amostra inicial obtida através da detecção do excesso em uma dada banda estreita devido a presença um linha de emissão. EELGs em $z < 0.05$ são emissores de [OIII], e consequentemente devem apresentar excesso no filtro S-PLUS $J0515$. Baseado em $\sim$ 440 EELGs conhecidos do SDSS com correspondência cruzada com S-PLUS DR3 e UnWISE na área do Stripe82, removemos contaminantes (QSOs, estrelas) da nossa amostra. Encontramos um total de 248 EELGs, e após análises adicionais, incluindo ajuste de SED, nós confirmaamos que essa amostra contém galáxias com baixa massa, baixa metalicidade e grandes larguras equivalentes (para [OIII]: $\gtrsim$ 200Å, para Hα: $\gtrsim$ 100Å). Acompanhamentos espectroscópicos são necessários para confirmação dos nossos resultados.

Keywords. galaxies: photometry – galaxies: dwarf – galaxies: starburst

1. Introduction

Extreme emission-line galaxies (EELGs) are low-mass, low metallicity galaxies dominated by intense bursts of star formation. Their spectra present high-excitation emission lines with large equivalent widths (EW), a consequence of the photoionization of gas by hot massive stars in a young burst of star formation (Sargent & Searle 1970), and low blue continuum. These properties make them local analogs to high redshift galaxies (e.g. Lyman-break galaxies and Lyman-α), and therefore an unique tool for connecting physical properties (e.g. mass, SFR, metallicity, gas, dust content) to different processes (e.g. star formation, feedback, chemical enrichment).

In the local Universe, these systems receive different nomenclatures, depending on the observational technique or selection criteria (see Kunth & Östlin 2000, for a review), such as HII galaxies (Terlevich et al. 1991), blue compact dwarfs (BCDs, Thuan & Martin 1981), green peas (Cardamone et al. 2009; Amorín et al 2010) and blueberries (Yang et al. 2017). In addition, they are not defined by an unique threshold for their emission lines in the literature. Often, the rest-frame equivalent width (EW) of the [OIII]5007 line is used but with different inferior limits, for example Amorín et al (2015) assumes 100Å while Jiang et al. (2019) uses 300Å. For our purposes, we will assume as EELGs all objects with EW[OIII]>100Å.

Based on their characteristic spectra, it is natural to expect that EELGs are mostly detected using spectroscopy. Indeed, SDSS has a large spectra database of EELGs. However, the samples were pre-selected and limited to the brighter objects of this class. More recently, photometry has been proving as an advantageous observational alternative to find these objects at different redshift ranges (Yang et al. 2017; Lumbreras-Calle et al. 2022; Iglesias-Páramo et al. 2022).

In the present work, we aim at identifying EELGs at $z < 0.05$ using the Southern Photometric Local Universe Survey (S-PLUS, Mendes de Oliveira et al. 2019), which is a twelve optical band imaging survey. The multiband data allows the detection of objects with different emission lines at different redshift regimes. In addition, as an imaging survey, S-PLUS has the opportunity to obtain the emission in the whole galaxy, differently from fiber spectroscopy surveys (e.g. SDSS). For the third data release, S-PLUS has an observing area $\sim$ 1772 square degrees. The S-PLUS filter configuration (5 SDSS-like broad bands and 7 narrow bands) allow us to target [OIII] emitters at $z < 0.05$, which will present an excess in magnitude (or flux) in J0515 filter. Our goal is to complement or extend previous photometric and spectroscopic samples of low redshift EELGs by searching for these objects in the Southern Hemisphere. We will also assist in defining new samples for follow-up observations.

Throughout this text, [OIII] will always means the doublet [OIII], unless otherwise specified.

2. Control sample of confirmed EELGs

We create a subsample of spectroscopically confirmed EELGs from SDSS Data Release 16 (Ahumada et al. 2020) using

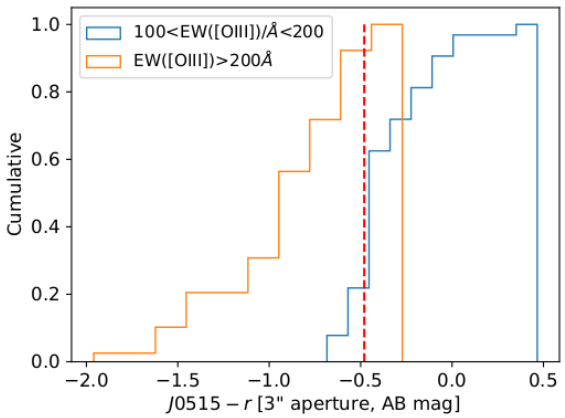

Figure 1. Cumulative distribution of color ($J0515 - r$) for EELGs_SDSS. The red dashed line represents a color cut that selects ~ 90% of galaxies with EW([OIII]λ5007)>200Å. Figure from Lopes et al. (in prep.).

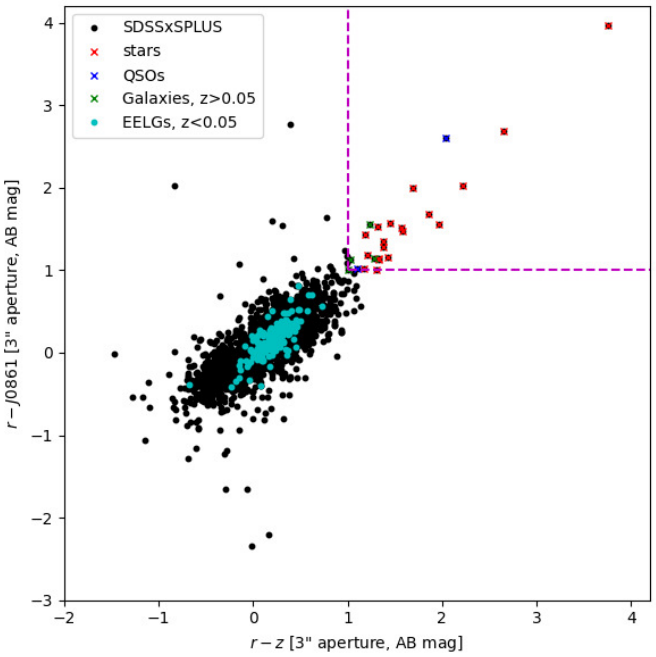

Figure 2. Color-color diagram to identify stars with low r-band flux in SDSSxSPLUS sample. The region where we expect to find most low-r band stars is delimited by purple dashed line, where we find mostly stars (red cross), and a few QSOs (blue cross) and galaxies at z>0.05 (green cross). EELGs_SDSS are represented in cyan circles. Figure from Lopes et al. (in prep.).

emission-LinePort table (Portsmouth stellar kinematics and emission line flux measurements, Thomas et al. 2013), according to the following criteria,

- subclass = STARBURST;
- equivalent widths of [OIII]λ5007 > 100Å and Hβ > 30Å;
- $0 < \log([OIII]/H\beta) < 1.2$
- $-2.5 < \log([NII]/H\alpha) < -0.8$
- redshifts in the range $0.005 < z < 0.4$ to avoid local giant HII regions in nearby galaxies.

These conditions ensure a selection of extreme star-forming galaxies with high excitation and low abundances. Cross-matching this list of EELGs with the S-PLUS DR3, we have a final number of 665 galaxies, where 173 are at $z < 0.05$. This sample will be referred as "EELGs_SDSS" throughout this text, and will be used for comparisons with our candidates.

3. Objects with J0515 excess

For EELGs at $0.008 < z < 0.05$, the [OIII] doublet lines fall inside narrow-band J0515. Considering $(J0515 - r) < 0$ as an indication of the presence of [OIII] emission lines, we tested the sensibility of this color in respect to the equivalent width (EW) of [OIII]λ5007 using EELGs_SDSS sample. As shown in Figure 1, in order to optimize the detection to more extreme cases, we limit our color selection to $(J0515 - r) < -0.2$, which will include all objects with EW([OIII])>200Å .

First, in order to perform a query using the Astronomical Data Query Language (ADQL) interface in the S-PLUS database[1], we assumed that the continuum in the $J0515$ filter is roughly equivalent to the flux of the r band. This is a good enough assumption to derive an initial sample of emitters.

With the purpose of ensuring the quality of our selection, aside from $(J0515 - r) < -0.2$, we add other restrictions, such as:

- signal-to-noise (S/N) greater than 3 in g, r and $J0515$ bands;
- PhotoFlag<4, includes objects with good photometry, or with close neighbors, or originally blended with another one, or a combination of those;
- BrightStarFlag=0, in order to exclude objects affected by bright stars;
- a significant excess in $J0515$ in comparison to g, considering the uncertainties in each band (e_$J0515$, e_g), i.e. $J0515$+e_$J0515$ < g+e_g;
- errors in magnitude lower than 0.2 in the g, r and $J0515$;
- $14 < r_{\rm AUTO} < 22$, to avoid saturated objects and fainter sources with unreliable photometry.

The result sample have 41634 objects with J0515 excess. This initial selection recovers all EELGs_SDSS at $0.008 < z < 0.05$ with EW([OIII]λ5007)>200Å, and a few others with 100Å<EW([OIII]λ5007)<200Å.

4. Removing contaminants

4.1. SDSSxSPLUS dataset

In an effort to effectively remove different types of interlopers (stars, QSOs), we cross-match the list of $J0515$ excess objects with SDSS DR16 spectroscopic information. Assuming a maximum separation of 3 arcsec, we find 2777 objects. This list of sources will be referred as "SDSSxSPLUS".

4.2. Stars with low r-band flux

We identify objects with $J0515$ excess in our SDSSxSPLUS sample that are, in fact, stars with low r-band flux. In these cases, the color $(J0515-r) < 0$ is just a consequence of the deficit in the broad-band, not an emission in the narrow one. Such contamination was also found by Lumbreras-Calle et al. (2022), where they search for EELGs in the Northern Hemisphere using J-PLUS. According to these authors, we can eliminate such sources by comparing the r-band flux with redder filters, for instance $J0861$ and z. We test this approach by applying the following color cuts to SDSSxSPLUS data:

[1] https://splus.cloud/

– $(r - z) > 1$,
– $(r - J0861) > 1$.

Indeed we find that such constraint selects 19 stars, 4 galaxies at z>0.05 and 2 QSOs, but no galaxy at z<0.05. In a $(r - z)$ vs $(r - J0861)$ diagram, EELGs_SDSS sample is located in a different region, as exemplified by Figure 2.

Applying these color cuts to our whole initial selection, we exclude 1184 objects, leaving 40450 candidates.

4.3. C III] emitters at $z \sim 1.7$

Another possible contaminants in our sample are QSOs at $z \sim 1.7$. Such objects present a prominent CIII]1909 emission line that creates an excess of flux in the $J0515$ filter, mimicking [OIII] emitters at low redshift. Studies of AGN detection using narrowband surveys (e.g. Stroe et al. 2017a,b) show that QSOs at $z \sim 1.7$ have also very intense CIV 1549 line, which at this redshift corresponds to an excess in the $J0430$ filter. Therefore, CIII] emitters have excess in both $J0430$ and $J0515$ bands, and a criteria of $(J0515 - J0430) > 0$ should select most of such contaminants. We find 4553 objects with an excess in $J0430$, which results in 35897 candidates. From SDSSxSPLUS sample, we check that none of the sources with $(J0515 - J0430) > 0$ is a low redshift galaxy.

4.4. S-PLUS classification and WISE photometry

In order to exclude remaining stars and QSOs in our sample, we use the 3 type object (GALAXY, STAR, QSO) classification available in the S-PLUS database. This classifier is based on two random forest algorithms denoted as model_flag (== 0 or 1). In both cases, the models trained with the same features: 12 S-PLUS magnitudes and 4 morphological features, Full Width at Half Maximum normalized to the detection image seeing (FWHM_n), major semi-axis, minor semi-axis and Kron Radius, except for the WISE photometry. The model_flag=0 includes WISE magnitude, while model_flag=1 do not.The algorithms were trained on spectroscopically confirmed data. For more details see (Nakazono et al. 2021).

From EELGs_SDSS, we find that this methodology is efficient to single out these objects, with ~ 88% being classified as galaxies, 3% as stars and 9% as QSOs. However, for the redshift range of our analysis, the classification is even better, identifying correctly 163 out of 173 EELGs by setting CLASS=2 and PROB_GAL>0.5. We apply such cuts in our data, selecting 1161 objects.

Nevertheless, we still have a contamination of a few galaxies at $z > 0.05$. In order to check if the alternative approach presented by Lumbreras-Calle et al. (2022) improves our selection, we cross-match our list of objects, before the classification, with the unWISE catalog (Schlafly, Meisner, & Green 2019). We obtain 10912 objects in unWISE, with a maximum separation of 3 arcsec. In a $(g - W2)$ vs. $(r - W1)$ plot, we find that most low redshift galaxies are located in the same region as Lumbreras-Calle et al. (2022), see Figure 3.

Considering objects that follow both criteria, S-PLUS classification and unWISE, and then excluding HII regions from large local galaxies, we have a final sample of 248 EELGs.

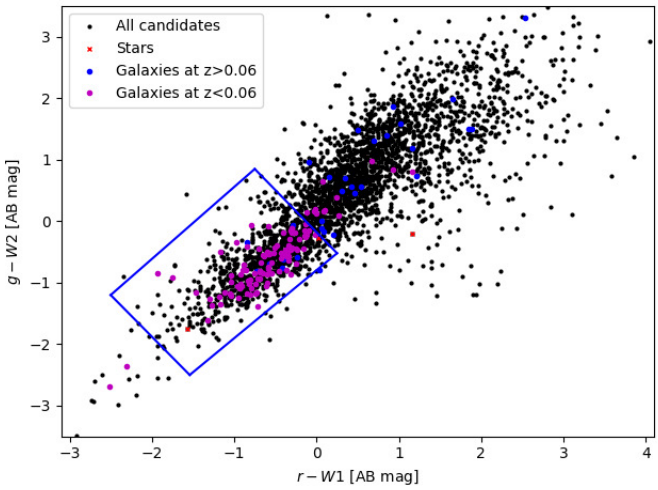

Figure 3. Color-color diagram to identify EELGs. The region defined by Lumbreras-Calle et al. (2022), where we expect to find local EELGs, is delimited by a blue solid line. Figure from Lopes et al. (in prep.).

5. Results

5.1. Three Filter Method

We estimate the fluxes and equivalent widths of the emission lines (EL) employing the Three Filter Method (3FM). This approach calculates the contribution of the EL fluxes inside a given narrow band filter (NB) by its contrast to the continuum. Such methodology presupposes that the object continuum is well traced by a linear function over the three filter. A complete discussion can be found in Pascual et al. (2007).

Considering the S-PLUS filter system, we include in our analysis the NB that contains the EL, the broad-band that partially overlaps with the NB (hereafter with the index "BB+EL"), and a third band with only continuum (hereafter adopting the index "BB+C"). Based on these observational fluxes $F^{\rm obs}$, the flux of the EL can be derived by (Vilella-Rojo et al. 2015)

$$F_{\rm EL} = \frac{\left(F^{\rm obs}_{\rm BB+EL} - F^{\rm obs}_{\rm BB+C}\right) - \left(\frac{\alpha_{\rm BB+EL} - \alpha_{\rm BB+C}}{\alpha_{\rm NB} - \alpha_{\rm BB+C}}\right)\left(F^{\rm obs}_{\rm NB} - F^{\rm obs}_{\rm BB+C}\right)}{\beta_{\rm NB}\left(\frac{\alpha_{\rm BB,C} - \alpha_{\rm BB+EL}}{\alpha_{\rm NB} - \alpha_{\rm BB+C}}\right) + \beta_{\rm BB+EL}} \quad (1)$$

with

$$\alpha_{\rm band} = \frac{\int \lambda^2 T_{\rm band}(\lambda) {\rm d}\lambda}{\int T_{\rm band}(\lambda)\lambda {\rm d}\lambda}, \quad (2)$$

$$\beta_{\rm band} = \frac{\lambda_{\rm EL} T_{\rm band}(\lambda = \lambda_{\rm EL})}{\int T_{\rm band}(\lambda)\lambda {\rm d}\lambda} \quad (3)$$

where $T_{\rm band}$ is the transmission curve for band = NB, BB+EL or BB+C.

By definition the equivalent width is the ratio between the fluxes from the EL and the continuum,

$$EW_{\rm SPLUS} = \frac{F_{\rm EL}}{F_{\rm cont}(\lambda = \lambda_{\rm EL})}, \quad (4)$$

where $F_{\rm EL}$ is given by Eq.(1) and $F_{\rm cont}$ is assumed to be a linear relation,

$$F_{\rm cont}(\lambda) = A\,\lambda + B. \quad (5)$$

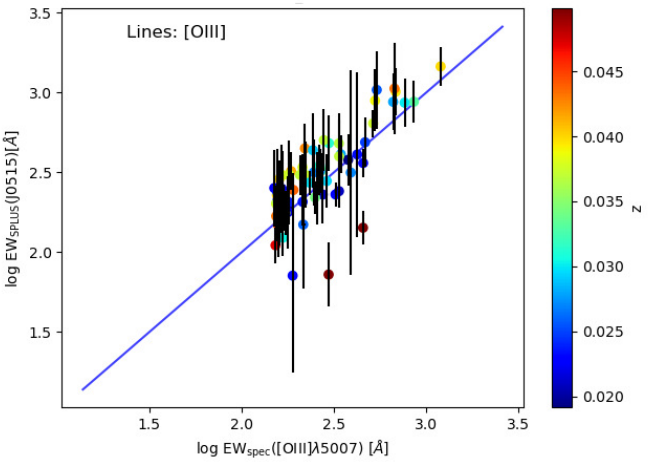

Figure 4. The equivalent width of [OIII] estimated using S-PLUS photometry ($EW_{\rm SPLUS}$) versus the spectroscopic measurements from SDSS ($EW_{\rm spec}$) for EELG_SDSS sample. The blue solid line represents the 1:1 relation. Figure from Lopes et al. (in prep.).

Following Spinoso et al. (2020)[2], the parameters A and B are defined as

$$A = \frac{\left(F^{\rm obs}_{\rm NB} - F^{\rm obs}_{\rm BB+C}\right) - \frac{\beta_{\rm NB}}{\beta_{\rm BB+EL}}\left(F^{\rm obs}_{\rm BB+EL} - F^{\rm obs}_{\rm BB+C}\right)}{(\alpha_{\rm NB} - \alpha_{\rm BB+C}) - \frac{\beta_{\rm NB}}{\beta_{\rm BB+EL}}(\alpha_{\rm BB+EL} - \alpha_{\rm BB+C})}, \quad (6)$$

$$B = F^{\rm obs}_{\rm BB+C} - \alpha_{\rm BB+C}\ A. \quad (7)$$

By solving Eqs. (1), (6) and (7) for $\lambda = \lambda_{\rm EL}$, we find the EW expressed in Eq. (4). In our case where EL= [OIII], NB= $J0515$, BB+EL= r and BB+C =i.

Based on our sample of known EELGs, we perform a test to verify our estimates of EW for [OIII] at $z_{\rm EL} \sim 0.035$. Figure 4 shows the comparison between our computed values and the spectroscopic measurements ($EW_{\rm spec}$) derived from SDSS. We find a spread around the 1:1 relation, with many objects being overestimated, mostly due to our EW including contribution from multiple lines while the $EW_{\rm spec}$ has only one line being measured. Moreover, the $z_{\rm EL}$ is not the same as the spectroscopic redshift, which may lead to an over/underestimation of the resulting EW. The [OIII] emitters are particularly complex, because depending on the object redshift, the line can be located at the wing of the transmission filter, leading to a smaller contribution to the photometric flux, or closer to maximum of the response curve, adding more flux to the NB filter. In the Figure 4, we see that the 3 offset points are sources in the upper ($z \sim 0.05$) and lower ($z \sim 0.019$) limits of our redshift interval.

5.2. SED-fitting

In order to obtain the physical properties of our EELGs, we perform a SED-fitting using the code CIGALE[3] (e.g. Boquien et al. 2019). This is a package consisting of a series of modules that are combined to create a set of theoretical SEDs, then compared to the observed data.

Following Lopes et al. (2021), we use a star formation history (SFH) composed of three main episodes of star formation (3BF). In this procedure we use AUTO magnitudes in the 12 bands. Table 1 presents the main input parameters in our SED-fitting analysis.

Table 1. List of input parameters applied to CIGALE.

Parameter	Value
Module: Simple Stellar Population (SSP)	
Model	Bruzual & Charlot (2003)
Initial Mass Function (IMF)	Chabrier (2003)
Metallicity [Z]	[0.0001, 0.0004, 0.004, 0.008, 0.02]
Module: Nebular Emission	
Ionization parameter	$\log U = -2.0$
Module: Dust Attenuation	
Name	Calzetti et al. (2000)
E(B−V)$_{\rm line}$	0. – 0.45
E(B−V)$_{\rm factor}$	0.44, 1
Slope delta of the power law	-0.5
Module: Dust Emission	
Dust template	Updated models of Draine & Li (2007)
Mass fraction of PAH	0.47, 1.12
Module: SFH	
Name	3BF
$t_{\rm old}$	10 Gyr
$\tau_{\rm old}$	[100,300,500] Myr
$t_{\rm int}$	[100,500,1000] Myr
$\tau_{\rm int}$	[10,50,100,200,300] Myr
$t_{\rm y}$	[5,10] Myr
$\tau_{\rm y}$	5 Myr

5.3. Description of results

The top panels in Figure 5 show histograms with output results for mass and metallicity derived with CIGALE for all 248 galaxies. They belong to the low mass regime, mostly below 10^9 M$_\odot$ and mostly low metallicity, with a mean value of $Z \sim 0.0084$. The left bottom panel presents the distribution of EW([OIII]) as obtained by the 3FM. As expected, the objects have high EW[OIII], mostly > 200Å. The right bottom panel shows the results of EW(Hα) as derived by CIGALE, where most galaxies have EW >100Å.

We must stress that the SED-fitting approach underestimates most of the EW([OIII]) in contrast to the ones derived from 3FM. This is not surprising, as the SED gives the contribution of [OIII]λ5007 along, while 3FM can eventually include Hβ.

6. Conclusion remarks

Based on a sample of known EELGs from SDSS DR16, we determine the best color relations to find EELGs $z < 0.05$. After removing contaminants, using the S-PLUS classification and unWISE data, we obtained a final sample of 248 EELGs.

We characterized our EELG sample with SED-fitting and using the Three Filter Method to estimate equivalent width (EW) for [OIII] and Hα. Our findings can be summarize as:

- low mass: $M_{\rm stellar} < 10^9 M_\odot$;
- low metallicity: mean Z$\sim$ 0.0084;
- EW([OIII])>100Å, with the majority presenting EW larger than 200Å;

[2] The full derivation of these relations can be found in Spinoso et al. (2020).

[3] Code Investigating GALaxy Emission is available at `https://cigale.lam.fr/`.

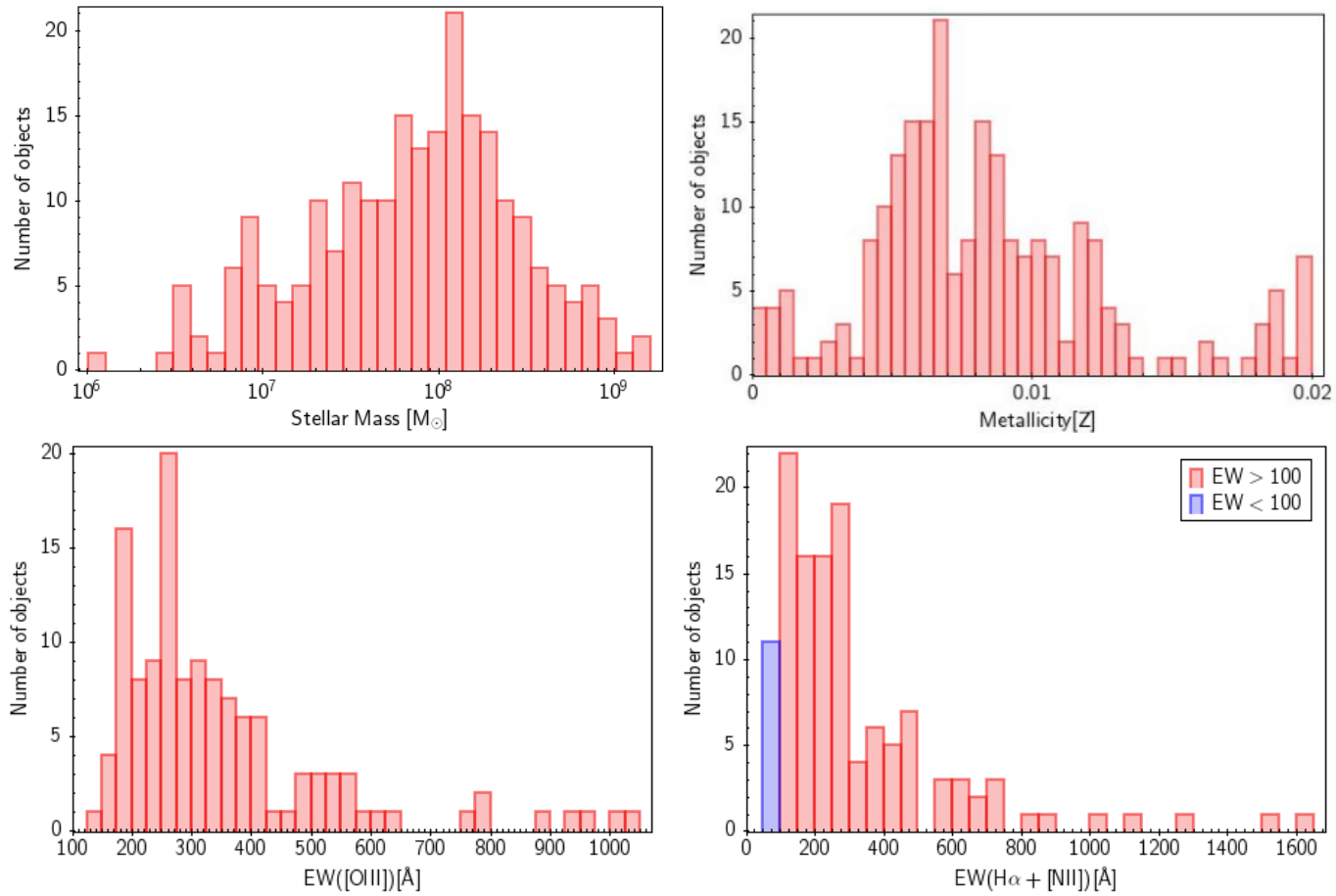

Figure 5. Basic characteristics of EELGs candidates. Upper panels: mass (left) and metallicity (right) results derived from SED-fitting analysis. Bottom panels: equivalent widths of [OIII] based on 3FM (left) and Hα derived by CIGALE (right). As expect, most sources have low mass ($\lesssim 10^9$ M$_\odot$), low metallicity (peak at $Z \sim 0.008$) and high equivalent widths (EW$\gtrsim$ 100 – 200Å).

– EW(Hα) are mostly greater than 100Å.

We intend to perform spectroscopic follow-ups for our best candidates to confirm our selection.

This work shows the great potential narrow band surveys, such as S-PLUS, have in the search of EELGs. Together with other narrow band surveys, such as J-PLUS and J-PAS, they are a good window of opportunity to populate the faint regime of this class of objects.

Acknowledgements. A.R.L acknowledges the financial support from Conselho Nacional de Pesquisa (CNPq) through PDJ (Pós-doutorado Júnior) fellowship, and from CONICET.

References

Ahumada R., Prieto C. A., Almeida A., et al., 2020, ApJS, 249, 3
Amorín R. O., Pérez-Montero E., Vílchez J. M., 2010, ApJL, 715, L128
Amorín, R. O, Pérez-Montero, E., Contini, T., et al. 2015, A&A, 578, A105
Benitez N., Dupke R., Moles M., et al., 2014, arXiv e-print, arXiv:1403.5237
Boquien M., Burgarella D., Roehlly Y., et al., 2019, A&A, 622, A103
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Calzetti D., Armus L., Bohlin R.C., et al., 2000, ApJ, 533, 682
Cardamone C., Schawinski K., Sarzi M., et al., 2009, MNRAS, 399, 1191
Chabrier G., 2003, PASP, 115, 763
Draine B. T., Li A., 2007, ApJ, 657, 810
Iglesias-Páramo J., Arroyo A., Kehrig C., et al., 2022, A&A, 665, A95
Jiang, T., Malhotra, S., Rhoads, J. E., & Yang, H. 2019, ApJ, 872, 145
Kunth D., Östlin G., 2000, A&ARv, 10, 1
Lopes A. R., Telles E., Melnick J., 2021, MNRAS, 500, 3240
Lumbreras- Calle A., López-Sanjuan C., Sobral D., et al., 2022, A&A, 668, A60
Mendes de Oliveira C., Ribeiro T., Schoenell W., et al., 2019, MNRAS, 489, 241
Nakazono L., de Oliveira C. M., Hirata N. S. T., et al., 2021, MNRAS, 507, 5847
Pascual S., Gallego J., Zamorano J., 2007, PASP, 119, 30
Thomas D., Steele O., Maraston C., et al., 2013, MNRAS, 431, 1383
Sargent W. L. W., Searle L., 1970, ApJL, 162, L155
Schlafly E. F., Meisner A. M., Green G. M., 2019, ApJS, 240, 30
Spinoso D., Orsi A., López-Sanjuan C., et al., 2020, A&A, 643, A149
Stroe A., Sobral D., Matthee J., Calhau J., Oteo I., 2017a, MNRAS, 471, 2558
Stroe A., Sobral D., Matthee J., Calhau J., Oteo I., 2017b, MNRAS, 471, 2575
Terlevich R., Melnick J., Masegosa J.,et al., 1991, A&AS, 91, 285
Thuan T. X., Martin G. E., 1981, ApJ, 247, 823
Vilella-Rojo G., Viironen K., López-Sanjuan C., et al., 2015, A&A, 580, A47
Yang H., Malhotra S., Rhoads J. E., Wang J., 2017, ApJ, 847, 38