On the discovery of stars, quasars, and galaxies in the Southern Hemisphere with S-PLUS DR2

L. Nakazono1*, C. Mendes de Oliveira1, N. S. T. Hirata2, S. Jeram3
C. Queiroz4, Stephen S. Eikenberry3, A. H. Gonzalez3, R. Abramo4, R. Overzier5,
M. Espadoto2, A. Martinazzo2, L. Sampedro1, F. R. Herpich6, F. Almeida-Fernandes1,
A. Werle1,6, C. E. Barbosa1, L. Sodré Jr.1, E. V. Lima1, M. L. Buzzo1, A. Cortesi7,
K. Menéndez-Delmestre7, S. Akras8*, Alvaro Alvarez-Candal5,9,10,
A. R. Lopes5, E. Telles5, W. Schoenell11, A. Kanaan12, T. Ribeiro13

1 Instituto de Astronomia, Geofísica e Ciências Atmosféricas da U. de São Paulo, Cidade Universitária, 05508-900, São Paulo, SP, Brazil
2 Departamento de Ciência da Computação, Instituto de Matemática e Estatística da USP, Cidade Universitária, 05508-090, São Paulo, SP, Brazil
3 Department of Astronomy, University of Florida, 211 Bryant Space Center, Gainesville, FL 32611, USA
4 Departamento de Física Matemática, Instituto de Física, Universidade de São Paulo, SP, Rua do Matão 1371, São Paulo, Brazil
5 Observatório Nacional / MCTIC, Rua General José Cristino 77, Rio de Janeiro, RJ, 20921-400, Brazil
6 The remaining institutions are at the end of the paper.

ABSTRACT
This paper provides a catalogue of stars, quasars, and galaxies for the Southern Photometric Local Universe Survey Data Release 2 (S-PLUS DR2) in the Stripe 82 region. We show that a 12-band filter system (5 Sloan-like and 7 narrow bands) allows better performance for object classification than the usual analysis based solely on broad bands (regardless of infrared information). Moreover, we show that our classification is robust against missing values. Using spectroscopically confirmed sources retrieved from the Sloan Digital Sky Survey DR16 and DR14Q, we train a random forest classifier with the 12 S-PLUS magnitudes + 4 morphological features. A second random forest classifier is trained with the addition of the W1 (3.4 μm) and W2 (4.6 μm) magnitudes from the Wide-field Infrared Survey Explorer (WISE). Forty-four percent of our catalogue have WISE counterparts and are provided with classification from both models. We achieve 95.76% (52.47%) of quasar purity, 95.88% (92.24%) of quasar completeness, 99.44% (98.17%) of star purity, 98.22% (78.56%) of star completeness, 98.04% (81.39%) of galaxy purity, and 98.8% (85.37%) of galaxy completeness for the first (second) classifier, for which the metrics were calculated on objects with (without) WISE counterpart. A total of 2,926,787 objects that are not in our spectroscopic sample were labelled, obtaining 335,956 quasars, 1,347,340 stars, and 1,243,391 galaxies. From those, 7.4%, 76.0%, and 58.4% were classified with probabilities above 80%. The catalogue with classification and probabilities for Stripe 82 S-PLUS DR2 is available for download.

Key words: methods: data analysis - surveys - catalogues - quasars: general - stars: general - galaxies: general

1 INTRODUCTION
The next generation of astronomical surveys will map large volumes of the sky, collecting data for millions, and even billions of objects. For instance, GAIA DR2 (Gaia Collaboration et al. 2018)
contains measurements of 1.7 billion sources; the Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST; Ivezic et al. 2019) aims to provide 32 trillion observations of 40 billions objects; the Javalambre-Physics of the Accelerated Universe Astrophysical Survey (J-PAS; Benitez et al. 2014, Bonoli et al. 2020) will cover at least 8,000 deg$^2$ with 54 overlapping narrow bands, 2 medium-bands extending to the UV and the near-infrared, and complemented by the u, g, r, i Sloan Digital Sky Survey (SDSS) broad-band filters; the Javalambre Photometric Local Universe Survey (J-PLUS; Cenarro et al. 2019) will cover the same area of the Northern Sky as J-PAS with the Javalambre optical filter system consisting of 7 narrow bands and 5 SDSS-like bands. Regarding the Southern Sky, the Southern Photometric Local Universe Survey (S-PLUS; Mendes de Oliveira et al. 2019) will cover ~9,300 deg$^2$ using the same 12 optical bands as J-PLUS. With such amount of data, object classification is essential for all these surveys, requiring an efficient and automated methodology.

Generally, galaxies and stars can be easily separated due to their different morphologies (e.g. Pimbblet et al. 2001, Moore et al. 2006), while quasars can be confused with stars at any magnitude range, especially when quasar targets are selected in broad-band colour-colour diagrams (see e.g. Fig. 14 from Richards et al. 2002). In the low signal-to-noise regime, however, galaxies and point-like sources can also be difficult to separate. The performance of the object classification can be improved when spectroscopy is available, since it brings more information than the imaging in just a few optical bands. However, given that complete spectroscopic surveys are more limited in area and magnitude than photometric surveys, many techniques have been developed to perform star/quasar separation with photometric data, such as: colour-colour cuts (e.g. Wu & Jia 2010, Wu et al. 2012, Schindler et al. 2017, Päriv et al. 2018); proper motion criteria (e.g. Guo et al. 2018, Heintz et al. 2018); Bayesian Statistics (e.g. Kirkpatrick et al. 2011, Peters et al. 2015, Yang et al. 2017a), and Machine Learning algorithms (e.g. Peng et al. 2012, Carrasco et al. 2015, Jin et al. 2019). For star/galaxy separation, methods that rely on morphological features of the photometric images are widely used to distinguish the classes (e.g. Costa-Duarte et al. 2019, López-Sanjuan et al. 2019 and Baqui et al. 2021, for S-PLUS, J-PLUS and miniJ-PAS, respectively). Few works have tackled both problems at once in a 3-class (star/quasar/galaxy) separation (e.g. Clarke et al. 2020, Yang et al. 2017b, Kurcz et al. 2016, Ball et al. 2006). Brescia et al. (2015) have shown, using SDSS magnitudes, that a 3-class separation improved accuracy in comparison with a 2-step approach that consisted of a galaxy/star and a star/quasar classification.

Object classification in photometric surveys is much improved if there are images available in a number of different filters, covering a low-resolution spectrum. In this sense, combining optical with infrared information – usually from the Wide-field Infrared Survey Explorer (WISE; Wright et al. 2010) – greatly improves object classification, especially for finding quasars (e.g. Bowy et al. 2012, Wu et al. 2012, Schindler et al. 2017). This is due to the dust surrounding the nucleus of quasars that absorbs light and re-radiates in the near-infrared (NIR) and mid-infrared, together with stellar NIR emission of the host galaxies (see Hernán-Caballero et al. 2016 and references therein). Narrow-band surveys – such as J-PAS, J-PLUS and S-PLUS – can also play an important role on finding quasars, given that the broad-emission features of the quasars can be easily identified when they coincide with the narrow bands, as illustrated in Fig. 1. In the top panel, we show the spectra of a quasar, a galaxy, an M6 star, and a white dwarf with their respective observed fluxes in each of the 12 S-PLUS filters and WISE bands. In the bottom panel, we show the redshift ranges where the main quasar emission lines (Lyα, CIV, CIII, MgII, Hγ, Hβ, Hα) can be observed for each of the seven S-PLUS narrow-band filters. With the J-PAS unique set of 54 narrow-band filters, one will not only detect the quasar photometric redshift but will immediately obtain a high-accuracy photometric redshift measurement.

In this era of Big Data, we face the need of migrating from more traditional techniques (e.g. colour-colour cuts), to computational, statistical and mathematical methods, which is corroborated
by the growth of works in many astronomical fields with artificial intelligence and machine learning applications (see review from Fluke & Jacobs 2020). In this work we adopted a supervised learning approach, given that a sufficient number of spectroscopically confirmed stars, quasars and galaxies are available in the observed S-PLUS fields of Stripe 82 region (see Section 2). We have chosen to test Random Forest and Support Vector Machines, as both are non-parametric models and, therefore, require fewer assumptions about the data (such as probability distribution). A good understanding of the S-PLUS data (as described in Section 3) is required to best construct our evaluation and model fitting strategy (see Section 4).

In this process, we also want to statistically: (1) assess the potential of the 7 narrow-band filters in comparison with analyses based only on the broad-band magnitudes, which is novelty and (2) reaffirm that the infrared bands and morphological parameters improve the classification performances. We discuss the comparison of different models through cross-validation in Section 5, from which we select the best classifiers (one for objects with WISE counterpart, and another for objects without it). We perform a statistical hypothesis test (Wilcoxon signed-rank) in order to compare the performances with confidence level of 90%, which are shown in Appendix B.

The expected performances of the chosen models are calculated on a separate test dataset and are shown in Section 6. We finally use our validated random forest models to classify the ~3 million photometric sources from the Stripe 82 region in S-PLUS DR2. We describe the details of this catalogue containing object classification and probabilities in Section 7. Unless specified otherwise, all magnitudes here are presented in the AB system. In the Data Availability section we give the links to download the datasets used in this work and the access to our ADQL queries. All appendices are only available as supplementary online material.

2 DATABASE

We use photometric catalogues from S-PLUS and WISE surveys, which are further introduced in §2.1. In order to perform a supervised-learning approach, we retrieved spectroscopically confirmed stars, quasars and galaxies from SDSS. We established a few selection criteria on these samples, as further described in §2.2, aiming to avoid spurious sources for training and validating the algorithms.

2.1 Surveys

2.1.1 The Southern Photometric Local Universe Survey (S-PLUS)

S-PLUS is a photometric survey performed by the T80-South robotic telescope, located at Cerro Tololo, Chile. By the end of 2023, it will have covered ~9300 deg² of the Southern Sky in 12 optical filters.

S-PLUS uses the Javalambre 12-band system, which includes the narrow-band filters J0378, J0395, J0410, J0430, J0515, J0660 and J0861, centered on [OII], Ca H+K, Hγ, G-band, Mg b triplet, Hα and Ca triplet features, respectively. The broad-band filters are the SDSS-like u, g, r, i, z. These filters are similar to the SDSS filter system, except for the u band, which differs significantly (see Cenarro et al. 2019). The information of their effective wavelength, width and photometric depth are available in Table 1. The photometric depth for narrow bands is typically 20.0 mag, except for Hα that is 21.1 mag. For broad bands, the depth is typically 21 mag, except for z band that is 20.1 mag.

Table 1. Characteristics of the S-PLUS 12-bands filter system. The effective wavelength and width information were taken from Mendes de Oliveira et al. 2019. The depths were taken from Almeida-Fernandes et al. 2021.

| Filter | Effective Wavelength (Å) | Width ** (Å) | Depth for S/N > 3 (mag) | Comments |
|--------|--------------------------|--------------|------------------------|----------|
| u      | 3536                     | 352          | 21.0                   | Javalambre u |
| J0378  | 3770                     | 151          | 20.4                   | [OII]    |
| J0395  | 3940                     | 103          | 19.9                   | Ca H+K   |
| J0410  | 4094                     | 201          | 20.0                   | Hγ       |
| J0430  | 4292                     | 201          | 20.0                   | G-band   |
| J0515  | 5133                     | 207          | 20.2                   | Mg b Triplet |
| r      | 6258                     | 1465         | 21.3                   | SDSS-like r |
| J0660  | 6614                     | 147          | 21.1                   | Hα       |
| i      | 7690                     | 1506         | 20.9                   | SDSS-like i |
| J0861  | 8611                     | 408          | 19.9                   | Ca Triplet |
| z      | 8831                     | 1182         | 20.1                   | SDSS-like z |

* Wavelength corresponding to the middle point at half maximum.
** Width at half maximum.

For this work, we used data from S-PLUS DR2 in the Stripe 82 area (0° < RA < 60°; 300° < RA < 360° and -1.4° < Dec < +1.4°), covering ~336 deg² close to the Celestial Equator. To construct our catalogue with class labels and probabilities, we consider a subset of 3076191 sources from this area that are within 13 < r ≤ 22 and PhotoFlag_r == 0, which will be further discussed in Section 7. To build our training sample for our classifiers, we establish more specific criteria for each spectroscopic class (see §2.2). The photometry was obtained with SExtractor dual mode using a detection image built from the sum of g, r, i, z bands.

For this work, we only considered the I80 magnitudes, in AB system. We use a subscript (e.g. rSDSS) whenever we refer to the modelmag magnitudes from SDSS in order to distinguish from the S-PLUS broad-band filters.

2.1.2 Wide-field Infrared Survey Explorer (WISE)

The WISE mapped the entire sky with a 40-cm telescope in orbit of the Earth. It has observed the sky in infrared using the W1, W2, W3 and W4 bands centered at 3.4, 4.6, 12 and 22 μm, respectively. The Vega magnitudes can be converted to AB magnitudes using the following relation $m_v= m_v\text{_{vega}} + \Delta m$, where $\Delta m = 2.699$ and 3.339 for W1 and W2 bands, respectively. Using different magnitude systems do not interfere in our analyses. Thus, we use the Vega magnitudes to train our models. For this work, we are only using W1 and W2 magnitudes, whose limiting Vega magnitudes are 16.5 and 15.5 respectively. We disregard W3 and W4, whose limiting Vega magnitudes are 11.2 and 7.9, respectively, due to their shallowness.

For this work we are using the AllWISE catalog (Cutri et al. 2013).

2.2 Spectroscopic sample selection

Among many wide-field spectroscopic surveys, the SDSS has provided the largest number of discovered quasars (QSOs). Recently, the quasar catalog from DR14 (DR14Q; Pâris et al. 2018) was...
The following selection criteria were applied in order to avoid any contamination from spurious sources in our samples: we decided to keep these duplicates as this number will not affect our models. There are 62 S-PLUS observations that were cross-matched twice as both quasar and galaxy. We decided to allow multiple matches; 47,448 WISE sources would be attributed multiple times, within a total of 1,459,524 S-PLUS DR2 objects with $\text{WISE}$ counterpart (47.44% of the total photometric sample). Those would be mostly cross-matched with two different close sources of S-PLUS, which in principle could represent a source of duplicate WISE assignments (one S-PLUS object could be assigned to more than one WISE source). We try to minimize this problem by not allowing the same WISE source to be cross-matched more than once. Here we are assuming that most of the emission from the WISE match is coming from the closest counterpart, in angular separation, to the S-PLUS source. If we were to allow multiple matches, 47,448 WISE sources would be attributed multiple times within a total of 1,459,524 S-PLUS DR2 objects with WISE counterpart (47.44% of the total photometric sample). Those would be mostly cross-matched with two different close sources of S-PLUS, had we allowed this to happen. This exercise was done in order to have an idea of the magnitude of the problem, although it is not solved here (the only solution would be to have higher resolution images in the infrared).

Table 2 shows the total number of quasars, stars and galaxies retrieved from SDSS. There are 62 S-PLUS observations that were cross-matched twice as both quasar and galaxy. We decided to keep these duplicates as this number will not affect our models. The following selection criteria were applied in order to avoid any contamination from spurious sources in our samples:

- PhotoFlag $r = 0$ and $r \leq 22$ for all classes: this output from SExtractor allows us to select only the objects with reliable photometry from the S-PLUS images;
- $z$Warning $= 0$ for galaxies: this flag is retrieved from SDSS catalogue. Any non-zero values indicate problems with the pipeline fit to the SDSS spectrum;
- $r > 13$ for stars: this condition was set from a visual inspection on the S-PLUS images, to avoid saturated images.

In the S-PLUS catalog, non-detected bands have magnitudes set to 99. These missing values were not excluded from our analysis or treated in any special way, as one of our goals is to build a model that is robust against missing data. Some sources have a discrepant magnitude in S-PLUS when compared to SDSS, even considering the filtering conditions listed above. This can be explained by either intrinsic variable sources or problems in the image reduction in released with 526,356 quasars over 9,376 deg$^2$, including new discoveries from SDSS-IV extended Baryon Oscillation Spectroscopic Survey (eBOSS; Dawson et al. 2016) and identified quasars from SDSS-III (DR7Q; Schneider et al. 2010) and SDSS-III programs (DR12Q; Pâris et al. 2017). The DR14Q is provided with the cross-matched information from the AllWISE catalogue for sources within 2″, with 2% rate of false positives (Pâris et al. 2018). We cross-matched Stripe 82 S-PLUS DR2 with DR14Q with a matching radius of 1″ to retrieve the quasar sample. We selected all spectroscopically confirmed stars and galaxies in Stripe 82 from the SDSS DR16. We cross-matched the catalogues from SDSS DR16 and Stripe 82 S-PLUS DR2 within 1″, that was also cross-matched with ALLWISE catalogue within 2″.

Note that there is a difference in resolution between WISE and S-PLUS, which in principle could represent a source of duplicate WISE assignments (one S-PLUS object could be assigned to more than one WISE source). We try to minimize this problem by not allowing the same WISE source to be cross-matched more than once. Here we are assuming that most of the emission from the WISE match is coming from the closest counterpart, in angular separation, to the S-PLUS source. If we were to allow multiple matches, 47,448 WISE sources would be attributed multiple times within a total of 1,459,524 S-PLUS DR2 objects with WISE counterpart (47.44% of the total photometric sample). Those would be mostly cross-matched with two different close sources of S-PLUS, had we allowed this to happen. This exercise was done in order to have an idea of the magnitude of the problem, although it is not solved here (the only solution would be to have higher resolution images in the infrared). Table 2 shows the total number of quasars, stars and galaxies retrieved from SDSS. There are 62 S-PLUS observations that were cross-matched twice as both quasar and galaxy. We decided to keep these duplicates as this number will not affect our models. The following selection criteria were applied in order to avoid any contamination from spurious sources in our samples:

- PhotoFlag $r = 0$ and $r \leq 22$ for all classes: this output from SExtractor allows us to select only the objects with reliable photometry from the S-PLUS images;
- $z$Warning $= 0$ for galaxies: this flag is retrieved from SDSS catalogue. Any non-zero values indicate problems with the pipeline fit to the SDSS spectrum;
- $r > 13$ for stars: this condition was set from a visual inspection on the S-PLUS images, to avoid saturated images.

In the S-PLUS catalog, non-detected bands have magnitudes set to 99. These missing values were not excluded from our analysis or treated in any special way, as one of our goals is to build a model that is robust against missing data. Some sources have a discrepant magnitude in S-PLUS when compared to SDSS, even considering the filtering conditions listed above. This can be explained by either intrinsic variable sources or problems in the image reduction in either or both surveys. To be conservative and select the purest samples possible, an additional 4-sigma clipping was applied on the difference between SDSS and S-PLUS magnitudes ($\Delta r = r_{\text{SDSS}} - r_{\text{S-PLUS}}$) per $r$ bin of size 0.5. The number of lost objects in these processes and the size of the final samples are described in Table 2.

### Table 2. Description of the spectroscopic samples from the cross-match between S-PLUS DR2, SDSS DR14Q (quasars), SDSS DR16 (stars and galaxies) and AllWISE catalogues. Selection criteria are $z$Warning $= 0$ for galaxies, and $r > 13$ for stars. For all classes, the criteria are PhotoFlag $r = 0$ and an upper limit of 22 magnitudes in $r$. After the selection criteria, we applied a 4-sigma clipping of the difference between $r$ magnitudes from SDSS and S-PLUS.

|                | Total matches | Number of lost objects from criteria | Number of lost objects from 4-sigma clipping | with_WISE sample | no_WISE sample | Final sample |
|----------------|---------------|-------------------------------------|-----------------------------------------------|------------------|----------------|--------------|
| QSO            | 28,275        | 9,217                               | 69                                            | 15,538           | 3,451          | 18,989       |
| STAR           | 78,217        | 22,464                              | 1,702                                         | 39,595           | 14,456         | 54,051       |
| GALAXY         | 154,529       | 87,667                              | 951                                           | 60,059           | 5,553          | 65,612       |

3 DESCRIPTION OF THE SPECTROSCOPIC SAMPLES

In order to evaluate the inclusion of WISE data for classifying objects, it is important to distinguish between two complementary subsets of our spectroscopic sample, which will be called as with_WISE and no_WISE. These two subsets contain objects with and without WISE magnitudes ($W1$ and $W2$), respectively. The no_WISE sample includes sources with null information on the error or signal-to-noise for $W1$ or $W2$, containing a total of 3,451 QSOs (18.17% out of 18,989), 14,456 stars (26.74% out of 54,051) and 5,553 galaxies (8.46% out of 65,612). The total number of sources in each of these samples is shown in Table 2. Figure 2 shows the relation between redshift and $r$ magnitude of the 15,358 known quasars and 60,059 known galaxies in the with_WISE sam-

![Figure 2](image-url)
For the 39,595 stars in the with_WISE sample, we only show the magnitude distribution. In this figure, we also show the marginal distributions for both with_WISE and no_WISE samples, shown in filled and dashed curves, respectively. We note that the with_WISE sample is distributed in brighter magnitudes compared to no_WISE.

For a classifier that does not include the WISE magnitudes as features, this model would be better generalized if trained with more total number of sources in each class in the stars detected in a certain number of the S-PLUS filters. From the typically lower.

The majority of the sources in our samples were detected in all twelve bands of S-PLUS. All sources were detected in at least 3 filters. However, we note that there are some missing bands and these are mostly in the bluer region of the spectra, where the S/N is typically lower. In Fig. 3, we show the fraction of QSOs, galaxies and stars detected in a certain number of the S-PLUS filters. From the total number of sources in each class in the with_WISE sample, 86.2% (68.6%) QSOs, 52.9% (53.6%) galaxies and 70.7% (79.4%) stars were detected in all 12 bands. When we consider objects that have been detected in at least 9 bands, the fractions are: 99.5% (99.1%), 94.9% (95.4%) and 98.4% (98.6%), respectively.

The target selection of SSQS quasar surveys has relatively low completeness at \( z = 2.2 - 3.5 \), due to overlapping of quasars and the stellar locus in the five-band optical colour space (Richards et al. 2006; Ross et al. 2013). Moreover, Schindler et al. (2017) showed that the survey has systematically missed many bright \( m_r < 18.5 \) quasars at \( z > 3 \). We also have shown in Fig. 1 that some quasars at specific redshifts will not have any emission lines detected in a narrow band filter (we further discuss this in §6.2). Therefore, our sample is not representative of all quasars.

We show the distribution of early- and late-type galaxies in Fig. 4, using the empirical relation from Strateva et al. (2001). The with_WISE (no_WISE) sample is composed of 60.7% (33.7%) early-type and 39.3% (66.3%) late-type galaxies. Both with_WISE and no_WISE samples have a majority of A, F, G, K, M stars, according to SDSS classification. The with_WISE sample has: 37.6% M-type, 27.6% K-type, 24.7% F-type, 7.3% G-type and 2.2% A-type stars. The remaining 0.6% correspond to cataclysmic variables, sub-dwarfs, white-dwarfs, carbon-, O-, B-, L-, and T-type stars. The no_WISE sample has: 39.4% F-type, 17.2% K-type, 16.5% A-type, 8.5% M-type, 7.6% G-type, 4.8% white-dwarfs, 1.6% B-type stars. Cataclysmic variables, sub-dwarfs, carbon-, L-, T-, and O-type stars represent the remaining 4.4% out of the no_WISE sample.

In Fig. 5 we show a 2-dimensional projection of the (a) 12 S-PLUS bands; (b) 12 S-PLUS bands + 4 morphological features; and (c) 12 S-PLUS bands + 4 morphological features + 2 WISE bands from our spectroscopic sample of quasars, stars and galaxies. These morphological features are: full width at half maximum, major semi-axis, minor semi-axis and Kron radius. The projection was performed by the t-Distributed Stochastic Neighbour Embedding (t-SNE) algorithm (van der Maaten & Hinton 2008), which is a nonlinear dimensionality reduction technique. The chosen hyper-parameters for the t-SNE run are shown in Appendix C. The stars and galaxies seem to be better discriminated when we include the morphological parameters as seen in Fig. 5(b), compared to (a). When including WISE features, there is less confusion between stars and quasars, as seen in Fig. 5(c). For this last plot, only objects with WISE counterparts are being shown. The objects detected in all 12 bands are concentrated in the centre of all figures, whereas the objects with at least one missing band are in the small islands surrounding the centre, outside the contour. From the projection, we can see that the observed stars, quasars and galaxies are generally concentrated in their own loci, even for the small islands outside of the contours. This indicates that, in spite of missing data, objects of the same class tend to cluster, thus suggesting that classification of those objects is possible. We therefore decided to keep sources with missing-band values (mag = 99) in our classification strategy based on machine learning. In Section 6.2 we discuss in more detail if the missing band-values are affecting our performances.
Figure 5. Two-dimensional projection of the (a) 12 S-PLUS bands; (b) 12 S-PLUS bands + 4 morphological features; and (c) 12 S-PLUS bands + 4 morphological features + 2 WISE bands, where these four morphological features are: FWHM, A, B and KRONE_RADIUS. The projection of the features was done using t-SNE with perplexity = 30. A total of 18,989 quasars (pink), 54,051 stars (blue), and 65,612 galaxies (yellow) are shown in (a) and (b). In (c), 15,538 quasars, 39,595 stars and 60,059 galaxies are shown. The contour in the three plots was hand-drawn to roughly separate the cases with and without missing bands. The small islands outside the contours are related to objects with at least one missing value in the S-PLUS bands. These figures show that, despite the missing-band values, the three classes are well clustered in their corresponding loci. Panel (b) clearly shows less confusion between stars and galaxies due to the inclusion of the four morphological features. The confusion between stars and quasars is diminished when the WISE features are included, as shown in panel (c).
4 METHODOLOGY BASED ON SUPERVISED-LEARNING ALGORITHMS

Initially, in order to facilitate the data visualization, we created colour-colour diagrams following the star/quasar separation methodology from Jeram et al. (in preparation). We use simulated AB magnitudes of a theoretical stellar library which covers 3,000 K < $T_{\text{eff}}$ < 25,000 K from 130 nm to 100 $\mu$m (Coelho 2014) and a single QSO template in the range 0.1 - 11 $\mu$m (Hernán-Caballero et al. 2016). The quasar template was, then, shifted from $z = 0 - 5$ in steps of 0.25. From the simulated data, we were able to select the combination of colours that provide the best separation of stars and quasars. We then included in the plot the known stars, galaxies and QSOs from the with_WISE sample. These are shown in Fig. 6, where we selected the following best colour combinations: J0395 – W1 versus $z$ – W2 and J0515 – i versus J0861 – W2. These diagrams show a good separation between stars and QSOs, for both observed and simulated data. In both colour spaces, the observed galaxies lie between stars and QSOs. We can see from the panels for $r > 19$ that the overlap of the three regions is greater at fainter magnitudes. For $r < 19$, however, the overlap is minimum. The figure also displays a smooth gradient for the spectroscopic redshift (retrieved from SDSS) of the known QSOs, following the trend from low to high redshift, from right to left, along the tracks. In section 5.4 of Mendes de Oliveira et al. (2019), preliminary results using the methodology above are given. While this method works well, we have the caveat of classifying only objects described in §4.3, considering different feature spaces described in $K$ from 130nm to 100 nm.

4.1 Introduction to Support Vector Machine (SVM)

SVM is a supervised machine learning algorithm, first introduced for binary classification (for more details, see Cortes & Vapnik 1995). For linearly separable cases, it consists in finding a hyperplane with maximum margin that separates the two classes of objects. Positive objects should stay on one side and negative ones on the opposite side with respect to the hyperplane and no object should lie within the margin. The objects on the margin are the support vectors. As the two classes may not be linearly separable, the soft-margin formulation allows violation of the margin, and the optimal solution is one that maximizes margin at the same time it minimizes classification error. There is a parameter $C$ that controls the trade-off between these two factors. Small $C$ corresponds to larger margins, but possibly with more classification errors while large $C$ tends to generate smaller margins and fewer classification errors (which might lead to over-fitting, however). To account for non-linearity, the kernel trick is used to fit the hyper-plane in a higher dimensional space using non-linear transformations (Scholkopf & Smola 2001).

The SVM can be extended to a multi-class classification, decomposing it into binary sub-classifications. We adopted the one-against-one strategy (decision_function_shape = "ovo"), that consists in building one SVM for each combination of pairs. For this work we chose the radial basis function kernel (kernel = "rbf") and other parameters that we did not mention were set using default values (see Appendix C).

4.2 Introduction to Random Forest (RF)

Random Forest is a method based on multiple decision trees and commonly used for classification and regression problems (for more details, see Breiman 2001). In a decision tree, the feature space is split successively into smaller sub-spaces. At each splitting step a node is added to the tree, and branches for the corresponding sub-spaces are created. This splitting process is successively applied on each branch subspace until the training objects contained in the resulting subspace is "pure" according to some purity criteria (e.g., contains only or predominantly objects of same class or value). Nodes with no splitting are the leaf nodes. The final prediction is done by traversing the tree from the root to a leaf node, following thus a single set of decisions. In a random forest, multiple trees are constructed. Each tree is grown independently of one another, from a sample with replacement of the training set. This procedure is known as bootstrap aggregating (or bagging). Each node is determined from a randomly chosen subset of the input features, instead of all of them. The class prediction is then given by the majority vote from all trees, resulting in predictions with reduced variance. Such strategy is robust against over-fitting, due to the Strong Law of Large Numbers. Probabilities for each class are estimated as the number of votes per total trees. In order to calculate the features importances, we use the implementation of scikit-learn that takes the Gini Importance approach (Breiman et al. 1984), also known as Mean Decrease in Impurity. It is defined as the total decrease in node impurity weighted by the fraction of the sample that reaches that node and then averaged over all trees of the ensemble. Oshiro et al. (2012) suggest using a range between 64 and 128 trees in a forest, as larger numbers of trees increase the computational cost and do not bring significant performance gain. In a first attempt, we fixed a total of 100 trees (n_estimators = 100), with the square root of the total number of features being sampled per node (max_features = "auto"). Parameters not described here were set using default values (see Appendix C).

4.3 Cross-validation strategy

The goal of doing cross-validation (Burman 1989) is to find the best configuration for classifying objects in terms of features, algorithms and their hyper-parameters. For that, we only use the with_WISE dataset, allowing us to make direct comparisons between different feature space configurations. The no_WISE dataset will be considered in further analyses.

We sampled 75% out of with_WISE (no_WISE) as our training set and 25% as our testing set in a stratified way, i.e. maintaining the
Figure 6. Colour-colour diagrams containing observed stars (blue contours), galaxies (orange contours) and quasars (points coloured by their spectroscopic redshift, according to the bottom right colour bar) with good photometric flag (PhotoFlag_R = 0) and signal-to-noise greater than 3 for each considered filter. These diagrams include simulated colours for a stellar template library (Coelho 2014), indicated by star-shaped points coloured by their temperature, according to the upper right colour bar. Colour simulations of a quasar template (Hernán-Caballero et al. 2016) at different redshifts are shown as a yellow track. We also indicate with yellow symbols the colours at specific redshifts $z = \{0, 1, 2, 3, 4, 5\}$. Upper panels: $J0395 - W1$ versus $z - W2$. Bottom panels: $J0515 - i$ versus $J0861 - W2$. From left to right: without limiting magnitude; including only $r < 19$ observed objects; including only $r > 19$ observed objects.

4.4 Feature spaces

Several works have used the five SDSS broad-bands for either star/quasar or star/quasar/galaxy separation (e.g. Brescia et al. 2015, Peters et al. 2015). We want to evaluate the advantage of the 7 S-PLUS narrow-bands against using only the 5 broad-bands. Thus, we tested the following feature spaces:

(i) 12 S-PLUS bands
(ii) 12 S-PLUS bands + 2 WISE bands
(iii) 5 S-PLUS broad-bands
(iv) 5 S-PLUS broad-bands + 2 WISE bands

As fixed hyper-parameters are considered for SVM and RF during the feature and algorithm evaluation, we ended up with 8 unique experiments (4 feature spaces $\times$ 2 algorithms). After selecting the best feature spaces among the ones listed above and the best algorithm, we also want to evaluate the improvements in our classification when adding morphological features, such as Full Width at Half Maximum (FWHM), major semi-axis ($A$), minor semi-axis ($B$) and Kron Radius (KRON_RADIUS in DR2 catalogue; Kron 1980). As the FWHM is variable field-to-field, we use the value normalized to the detection image seeing (FWHM_n in DR2 catalogue, but hereafter we will mention it simply as FWHM).

For further references, we define the following naming system to refer to each experiment:

Algorithm_\alpha.S + 2W + 4M,
Discovering stars, quasars, and galaxies in S-PLUS

Figure 7. Distribution of r magnitudes for 15,538 quasars (top left), 39,595 stars (top middle), and 60,059 galaxies (top right) that have WISE counterpart. In the bottom panels we show the distribution for objects without WISE counterpart (3,451 quasars, 14,456 stars, and 5,553 galaxies). All panels show the distributions for the train and test sets that correspond to 75% and 25% of the data, respectively.

where “Algorithm” are RF or SVM, “S” refers to the number of features from S-PLUS bands (i.e. 12 or 5), “2W” refers to the inclusion of W1 and W2 magnitudes from WISE, and “4M” refers to the four morphological features mentioned above. Any term is omitted if the corresponding features were not included in the experiment. For instance, RF_5S+2Wis the model trained with Random Forest, 5 S-PLUS broad bands and 2 WISE bands, without including the morphological features.

4.5 Performance Metrics

Our dataset is imbalanced, as the number of stars and galaxies in our spectroscopic sample are much larger than the number of quasars. In that case, accuracy is not a good performance measurement, due to bias towards the classes with higher frequencies. We will consider other metrics based on the confusion matrix (Fig. 8) as further described.

Considering $i, j \in \mathcal{Y}$, we can define precision ($P$) and recall ($R$) as follows:

$$P_i = \frac{(TP)_i}{(TP)_i + \sum_{j \neq i} (FP)_{i,j}},$$

$$R_i = \frac{(TP)_i}{(TP)_i + \sum_{j \neq i} (FN)_{i,j}},$$

where $TP$, $FP$, and $FN$ denotes the true positives, false positives and false negatives, respectively. In other words, $P_i$ is the percentage of objects assigned as class $i$ that are indeed class $i$, while $R_i$ is the percentage of class $i$ objects that were classified as class $i$. In astronomy, precision is also known as purity, and recall is also known as completeness or efficiency. We can also define another metric that evaluates $R$ and $P$ in a single value, which is the harmonic mean of $P_i$ and $R_i$ called F-measure ($F_i$):

$$F_i = 2 \left( P_i^{-1} + R_i^{-1} \right)^{-1} = 2 \times \frac{P_i R_i}{P_i + R_i}.$$  \hspace{1cm} (3)

Note that $P_i$, $R_i$ and $F_i \in [0,1]$. A perfect classification would yield $P_i$, $R_i$ and, consequently, $F_i$ to be 1. For an overall evaluation of the model performance (i.e. considering the classification of quasars, stars and galaxies altogether), we define the macro-averaged F-measure ($\bar{F}$) as follows:

$$\bar{F} = \frac{1}{N \sum_{y} F_i}.$$  \hspace{1cm} (4)
where \( N = 3 \) in our case. This definition differs from the micro-averaged F-measure that first computes an overall \( P \) and \( R \) and then calculates the harmonic mean. The micro-average is very sensitive to the imbalance of the classes towards the one with higher frequency, while macro-average will give more importance to the class with smallest frequency.

5 MODEL VALIDATION

In this section we show the performance metrics from the cross-validation for each of the experiments (Table 3). We will refer to each experiment by the naming system previously described in §4.4. From the cross-validation results, we first compare the overall performances from SVM and RF (see §5.1). Then we discuss the feature selection by comparing the performances from different feature spaces within the same algorithm (see §5.2). All comparisons were statistically performed by a Wilcoxon signed-rank test that are fully described in Appendix B. Throughout this section, we only consider an improvement on the performance of a given classifier if there is a statistically significant increase at 90% confidence level (i.e. p-values < 0.1), when comparing two experiments. All measured p-values are shown in Appendix B. In §5.3 we show the cross-validation results for the experiment without WISE features when we consider an increased dataset. Finally, in §5.4 we perform a randomized grid search to look for better hyper-parameter values. In this section we select two models that achieve the best performance: one that includes WISE features and another without WISE features. The expected performance of the chosen models will be shown in Section 6. We then provide classifications for unlabelled sources in Stripe 82 S-PLUS DR2 using both classifiers, which will be further discussed in Section 7.

5.1 Algorithm evaluation

First, we evaluated which algorithm gave us the best performance in overall terms. To answer this question, we only compared the macro-averaged F-score within each feature space. Table 3 shows that, in this setup, RF provides a better performance in all feature spaces compared to SVM. These differences are statistically significant (see Appendix B). Moreover, we can see that RF generally requires less than half of the SVM computational time for the same feature space. Thus, we will only consider results from RF for the further comparisons.

5.2 Feature evaluation

One aim of this work is to assess if the narrow-bands — which is the major S-PLUS advantage — improve the classification. Then we compared RF_12S versus RF_5S, i.e. when WISE magnitudes are not being considered, and RF_12S+2W versus RF_5S+2W, when considering WISE magnitudes. With addition of WISE photometry, the narrow-bands do not significantly increase the performance of quasar recall and galaxy precision with 90% confidence. This is possibly due to the fact that WISE plays a more important role on distinguishing these objects. For stars, all metrics have been improved. The macro-averaged F-score does significantly improve for RF_12S+2W. On the other hand, the S-PLUS narrow-bands significantly increase the precision and recall of all classes when training RF without WISE magnitudes. This means that, for sources without a WISE counterpart, the 12-band magnitude system is able to provide a more reliable star/quasar/galaxy separation in comparison to classifications based solely on broad-band filters. This is important to note, given that 43.81% of our photometric sample from S-PLUS DR2 sources do not have WISE detection. Nevertheless, we evaluated the inclusion of WISE magnitudes by comparing RF_12S versus RF_12S+2W, and RF_5S versus RF_5S+2W, where all metrics were statistically improved. This confirms the importance of WISE magnitudes for the classification.

Comparing our findings with results from other works is not straightforward as we would be comparing metrics that are calculated on completely different validation datasets. Nevertheless, it can guide us through possible interpretations. For instance, Brescia et al. (2015) achieves \( \text{F}_{\text{QSO}} = 88.65\% \), \( \text{F}_{\text{STAR}} = 89.95\% \), and \( \text{F}_{\text{GAL}} = 96.32\% \) with their best three-class experiment using only broad-bands (\text{psfMag} and \text{magNode1}) from SDSS DR10. These results would be more related to our RF_5S experiment. We can possibly assume that the morphological aspect may be covered by using both \text{psfMag} and \text{magNode1} in their models, which could explain a \( \text{F}_{\text{GAL}} \) better than ours. On the other hand, they find a similar confusion between quasars and stars, which could be explained by the fact that their model also does not consider WISE features. This similarity is reassuring, since different samples and photometry are used and SDSS is about 1 magnitude deeper than S-PLUS. For instance, the SDSS photometric depth in \( r \) is 22.7 mag, whereas the S-PLUS depth is 21.3 mag. The exception is the \( z \) band for which the depths are comparable (see SDSS DR16 website\(^1\) and table 4 from Almeida-Fernandes et al. (2021). Clarke et al. (2020) use a method that is more similar to ours by also applying a random forest algorithm on SDSS and WISE features, albeit they include W3 and W4 from WISE and they use all-sky data. For their analyses based only on SDSS magnitudes, they achieve \( \text{F}_{\text{QSO}} = 88.1\% \), \( \text{F}_{\text{STAR}} = 92.7\% \), and \( \text{F}_{\text{GAL}} = 97.0\% \). Our RF_5S experiment shows similar behaviour for quasars and worse performance for stars and galaxies. It is not clear what could be the reason for this difference in performance, although percentages can be misleading performance indicators for smaller samples, which is our case. Including WISE features, they achieve \( \text{F}_{\text{QSO}} = 94.2\% \), \( \text{F}_{\text{STAR}} = 97.4\% \), and \( \text{F}_{\text{GAL}} = 98.8\% \). Our RF_5S+2M performances seem to be in agreement with the results from Clarke et al. (2020), although our galaxy F-score is slightly worse. For further evaluation we only consider the best feature spaces with and without WISE using a RF algorithm: 12 S-PLUS bands (RF_12S) and 12 S-PLUS bands + 2 WISE bands (RF_12S+2W). For those, we made new experiments including morphological information as features, which can distinguish resolved galaxies from unresolved stars (and quasars) as shown in Costa-Duarte et al. (2019) for the S-PLUS data. The results including four morphological parameters (FWHM, A, B, and \( \text{KRON\_RADIUS} \)) statistically improved the star/quasar/galaxy classification for both experiments, as expected (see two top blocks from Table 4). The only exception is the quasar precision when WISE features are included. For that, the difference on quasar classification precision is very small (from 95.48% to 95.54%) for the classifier trained with WISE features. Morphological features would likely help decreasing the number of galaxies being misclassified as quasars or stars. Therefore, the small improvement in quasar precision could be due to the WISE features providing enough information to separate galaxies and quasars.

Comparing RF_12S+2W+4M versus RF_12S+4M, the results when considering WISE bands remain statistically superior over RF_12S+4M. Moreover, the metrics had greatly improved from

---

\(^1\) https://www.sdss.org/dr16/imaging/other_info/
Table 3. Cross-validation results. Fit time and macro-averaged F-score are shown for each model. Besides that, the metrics precision, recall and F-score are shown for each class. We highlight the highest observed metrics among all experiments in bold. Experiments in bold were chosen for further improvements (see Table 4). Note that these results were based only on the with_WISE dataset.

| EXPERIMENT | Fit time (s) | Prediction time (s) | $\bar{F}$ (%) | CLASS | Precision (%) | Recall (%) | F (%) |
|------------|-------------|--------------------|---------------|-------|---------------|-----------|-------|
| RF_12S     | 37.61 ± 1.26| 0.426 ± 0.042      | 92.17 ± 0.3   | QSO   | 91.99 ± 0.41  | 89.31 ± 0.97 | 90.63 ± 0.52 |
|            |             |                    |               | STAR  | 95.55 ± 0.46  | 89.59 ± 0.44 | 92.47 ± 0.31 |
|            |             |                    |               | GAL   | 91.26 ± 0.41  | 95.7 ± 0.29  | 93.42 ± 0.2  |
| RF_12S+2W  | 36.56 ± 4.39| 0.354 ± 0.010      | 96.85 ± 0.05  | QSO   | 95.48 ± 0.12  | 94.83 ± 0.26 | 95.15 ± 0.12 |
|            |             |                    |               | STAR  | 98.61 ± 0.08  | 97.06 ± 0.25 | 97.83 ± 0.15 |
|            |             |                    |               | GAL   | 96.97 ± 0.15  | 98.14 ± 0.08 | 97.55 ± 0.08 |
| RF_5S      | 23.99 ± 0.2 | 0.431 ± 0.012      | 90.19 ± 0.25  | QSO   | 90.58 ± 0.46  | 88.4 ± 0.7  | 89.48 ± 0.53 |
|            |             |                    |               | STAR  | 91.93 ± 0.73  | 87.68 ± 0.53 | 89.75 ± 0.26 |
|            |             |                    |               | GAL   | 89.74 ± 0.35  | 93.03 ± 0.45 | 91.36 ± 0.14 |
| RF_5S+2W   | 21.77 ± 0.2 | 0.335 ± 0.010      | 96.62 ± 0.13  | QSO   | 94.98 ± 0.2   | 95.08 ± 0.32 | 95.03 ± 0.18 |
|            |             |                    |               | STAR  | 98.14 ± 0.12  | 96.93 ± 0.34 | 97.53 ± 0.18 |
|            |             |                    |               | GAL   | 96.92 ± 0.17  | 97.68 ± 0.16 | 97.3 ± 0.12  |
| SVM_12S    | 142.49 ± 7.64| 81.29 ± 0.537     | 76.96 ± 0.24  | QSO   | 90.89 ± 0.13  | 77.53 ± 1.17 | 83.67 ± 0.7  |
|            |             |                    |               | STAR  | 76.17 ± 0.39  | 61.36 ± 0.4  | 67.96 ± 0.36 |
|            |             |                    |               | GAL   | 73.59 ± 0.25  | 85.82 ± 0.18 | 79.24 ± 0.13 |
| SVM_12S+2W | 127.48 ± 13.89| 68.81 ± 2.50     | 89.7 ± 0.18   | QSO   | 95.46 ± 0.31  | 83.56 ± 1.0  | 89.11 ± 0.63 |
|            |             |                    |               | STAR  | 94.39 ± 0.34  | 84.06 ± 0.48 | 88.93 ± 0.37 |
|            |             |                    |               | GAL   | 86.77 ± 0.2   | 95.83 ± 0.24 | 91.07 ± 0.16 |
| SVM_5S     | 100.13 ± 1.38| 74.91 ± 0.11      | 78.47 ± 0.3   | QSO   | 89.46 ± 0.54  | 82.87 ± 0.84 | 86.04 ± 0.6 |
|            |             |                    |               | STAR  | 74.6 ± 0.4    | 65.98 ± 0.48 | 70.02 ± 0.24 |
|            |             |                    |               | GAL   | 75.91 ± 0.23  | 83.14 ± 0.38 | 79.36 ± 0.19 |
| SVM_5S+2W  | 54.71 ± 0.62| 39.62 ± 0.24      | 95.11 ± 0.15  | QSO   | 95.37 ± 0.26  | 91.43 ± 0.45 | 93.36 ± 0.32 |
|            |             |                    |               | STAR  | 95.8 ± 0.15   | 96.4 ± 0.21  | 96.1 ± 0.08  |
|            |             |                    |               | GAL   | 95.57 ± 0.16  | 96.2 ± 0.09  | 95.88 ± 0.09 |

RF_12S to RF_12S+4M, with a percentage increase of: 2.28% for quasars, 5.7% for stars, and 4.56% for galaxies, in terms of F-score. From RF_12S+2W to RF_12S+2W+4M the percentage increase for F-score is: 0.36% for quasars, 0.84% for stars, and 0.74% for galaxies. Thus, morphological information plays an important role mainly for the classification when WISE features are not available. In Fig. 9, we show the estimated feature importance rankings, where we can see that the FWHM, followed by $A_0$, is the most useful information for both classifiers. This reassures the importance of morphological features in our models. However, it is not clear from our analyses if these features are able to distinguish complex structures (such as interacting galaxies). This could be further investigated in future work, as it could potentially help to avoid confusion in the bright end of saturated stars being classified as galaxies, allowing us to build a classification catalogue with more sources. This confusion is not present within our spectroscopic sample results due to our selection criteria (PhotoFlag $r = 0$). Along with morphology, the reddest S-PLUS filters ($i$, J0861 and $z$) seem to be more relevant for classification than the bluest ones, being also more relevant than $W1$ and $W2$ features.

5.3 Increasing the dataset for the classifier without WISE features

The analyses presented in §5.1 and §5.2 were done on the same training dataset from with_WISE sample in order to obtain comparable results. However, the classifier that does not use WISE features as input (i.e. RF_12S+4M) is therefore biased towards objects with WISE counterpart. One can see the differences in distribution between with_WISE and no_WISE samples in Fig. 2. Thus, we should evaluate how RF_12S+4M performs with an increased training dataset consisting of with_WISE and no_WISE datasets. To distinguish between the two experiments, we will call this one with_WISE and RF_12S+4M performs in both no_WISE and with_WISE testing sets for a better understanding of this model.

5.4 Tuning models with randomized grid search

We attempted to tune RF_12S+4M and RF_12S+2W+4M by employing a 3-fold randomized grid search using RandomizedSearchCV from the Python library scikit-learn. We detail in Table 5 the evaluated hyper-parameters, the list of values to be validated for each hyper-parameter, and the best values for RF_12S+4M and RF_12S+2W+4M. Besides n_estimators, min_samples_split, and max_depth, all other hyper-parameters found by the grid search are the same default values that we used in our previous analyses. The cross-validation results are shown in the two bottom blocks from Table 4, denoted by $\star$ symbols. From these results, we can see that our metrics changed very slightly and they are not worth the increase in computational time. For
We can roughly estimate an order of ~85 million sources in the total area of S-PLUS (~9300 deg²) within 13 < r < 22 and PhotoFlag_r = 0. For N ~ 85 million objects and n_estimators = 100, the prediction time would be approximately 21.25 minutes. For the RF_12S+4M@ with n_estimators = 400 and fixing N ~ 85 million, we estimate ~3.53 days to complete the predictions. Therefore, we maintain the same hyper-parameters used to train RF_12S+2W+4M and RF_12S+4M@ that we described previously in §4.2.

6 DESCRIPTION OF THE FINAL MODELS

We have so far been selecting the best algorithm (see §5.1) and the best feature spaces (see §5.2) for our classification problem through k-fold cross-validation. The statistical comparisons were based only on the with_WISE sample, in order to minimize possible bias effects due to differences in the features distributions. The experiment that achieved the best performance is a random forest algorithm trained with 12 S-PLUS bands + 2 WISE bands and 4 morphological features (RF_12S+2W+4M). However, the S-PLUS catalogue has a significant fraction of objects with no WISE counterparts. Therefore it is desirable to have another classifier that can be used on those objects. Thus, we have also selected the experiment with best performance among the ones without WISE features, which is a random forest algorithm trained with 12 S-PLUS bands + 4 morphological features (RF_12S+4M@). For RF_12S+4M@ we use an increased sample that includes no_WISE, in order to better generalize our model.

For each of the two selected models we show the overall expected performance, and the performance in terms of magnitude and probability threshold (see §6.1). In §6.2 we discuss and investigate some sources of bias that are possibly affecting our models.

6.1 Performance on test dataset

Considering classifier RF_12S+2W+4M, we achieved a macro-averaged F-score of 97.69% on the with_WISE testing set. In terms of quasar classification, we have 138 galaxies (from a total of 15015)
testing set has 3884 quasars, 9899 stars, and 15015 galaxies. The testing set classified either as quasar (top left), star (top middle) and galaxy (top right). The contamination rates of quasars, stars and galaxies, we obtained \(\mathcal{R}_{\text{QSO}} = 95.68\%\), similarly to RF_12S+2W+4M. We have 298 and 242 quasars being classified as stars and galaxies, respectively (\(\mathcal{R}_{\text{QSO}} = 95.68\%\)), which is 10.2% decrease compared to the same metric from RF_12S+2W+4M. For stars, we obtained \(\mathcal{P}_{\text{QSO}} = 95.59\%\) and \(\mathcal{R}_{\text{QSO}} = 95.59\%\). For galaxies, we obtained \(\mathcal{P}_{\text{QSO}} = 95.26\%\) and \(\mathcal{R}_{\text{QSO}} = 95.26\%\). Specifically, 149 stars were classified as galaxies and 42 galaxies were classified as stars. For classifier RF_12S+4M@ considering the same with_WISE testing set, we achieved a macro-averaged F-score of 95.04%. We have 123 galaxies and 28 stars being misclassified as quasars (\(\mathcal{P}_{\text{QSO}} = 95.68\%\)), similarly to RF_12S+2W+4M. We have 298 and 242 quasars being classified as stars and galaxies, respectively (\(\mathcal{R}_{\text{QSO}} = 95.68\%\)), which is 10.2% decrease compared to the same metric from RF_12S+2W+4M. For stars, we obtained \(\mathcal{P}_{\text{QSO}} = 95.59\%\) and \(\mathcal{R}_{\text{QSO}} = 95.59\%\). For galaxies, we obtained \(\mathcal{P}_{\text{QSO}} = 95.26\%\) and \(\mathcal{R}_{\text{QSO}} = 95.26\%\). Specifically, 173 stars were classified as galaxies and 149 galaxies were classified as stars. Note that the scores for RF_12S+2W+4M in the with_WISE testing set are mostly better or similar to its corresponding results in the cross-validation shown in Table 4. This is also true for the testing results on with_WISE+no_WISE from RF_12S+4M@ and its corresponding cross-validation results shown in Table 4. The fact that we do not see any of the two models performing too differently from their corresponding results in cross-validation indicates no evidence of over- or under-fitting.

The testing (Table 6) and validation (Table 4) results for RF_12S+2W+4M@ and RF_12S+4M@ considering with_WISE sample show that using the classification from the former is preferable over the latter whenever the WISE information is available. Otherwise, one can use the classification from RF_12S+4M@. Therefore, it is desirable to know how RF_12S+4M@ performs when no WISE counterpart is found. The results shown in Table 6 for the no_WISE testing set shows a great performance for star precision (\(\mathcal{P}_{\text{QSO}} = 95.68\%\)) and quasar recall (\(\mathcal{R}_{\text{QSO}} = 95.68\%\)); an acceptable performance for star recall (\(\mathcal{P}_{\text{QSO}} = 95.26\%\)), galaxy precision (\(\mathcal{P}_{\text{QSO}} = 95.26\%\)), and galaxy recall (\(\mathcal{R}_{\text{QSO}} = 95.26\%\); and a very poor quasar precision (\(\mathcal{P}_{\text{QSO}} = 95.26\%\)). The results do not improve in overall if we consider other training dataset (e.g. only no_WISE or only with_WISE) for the classifier RF_12S+4M@. The quasar precision is due to 546 stars being mis-classified as quasars (9.0% are white dwarfs). Checking the r distribution of these stars (Fig. 10, top panel), we see that most of them are in the faint end. From that we can understand that RF_12S+4M@ is typically classifying faint point-like sources as quasars and it can indicate a possible bias. Depending on one’s science purposes, one should consider an lower-upper-limit magnitude than 22. On the other hand, RF_12S+2W+4M does not present the same issue (Fig. 10, bottom panel), thus one can consider going down to 22 magnitudes with this classifier.

One can also use the information present in Fig. 11 to choose an upper-limit magnitude. We show in the top panels of Fig. 11 the cumulative percentage of contamination (i.e. \(1 - \mathcal{P}_r\), also known as false discovery rate) in the classification per model, per class and per magnitude in \(r\). In each of these panels, we also show the corresponding distribution of \(r\) magnitudes for all the sources in the testing sets classified either as quasar (top left), star (top middle) or galaxy (top right). The contamination rates of quasars, stars and galaxies are 0.05% (0.46%), 0.21% (0.10%) and 0.18% (0.07%) for RF_12S+2W+4M@ (RF_12S+2W+4M) at magnitude \(r = 19\), respectively. Down to magnitude \(r = 21\), these rates are: 1.93%
with WISE the colour-colour diagram analyses from Mendes de Oliveira et al. (2019). In the bottom panels of Fig. 11 we show the cumulative miss rates (0.21% (0.47%) and 0.13% (0.22%) for quasars, stars and galaxies, respectively. Down to magnitude $r = 21$, these rates are: 1.93% (3.13%) for quasars, 0.54% (9.24%) for stars, and 1.25% (4.39%) for galaxies. Note that the contamination and miss rates are less than 10% for all classes and for both models (except quasar contamination for RF_12S+4M@) down to magnitude $r = 21$, which is within the photometric depth of the S-PLUS r band. Moreover, the classification of quasars, stars, and galaxies from RF_12S+2W+4M are still reasonably good down to magnitude $r = 22$.

In the star/galaxy separation using RF on 12 S-PLUS bands + Morphology from Costa-Duarte et al. (2019), the obtained accuracy (i.e. fraction of correctly classified objects per total objects) is 95%. We reached an star/galaxy accuracy (i.e. number of correctly classified stars and galaxies per total number of stars and galaxies in our test sample) of 95.15% when also applying RF on 12 S-PLUS bands + Morphology on the with_WISE-no_WISE testing set. For the with_WISE and no_WISE testing sets separately, the accuracy is 98.1% and 79.09%, respectively. Moreover, our star/galaxy accuracy reaches 98.59% when WISE is considered.

Along with the classification, probabilities of being a quasar, star or galaxy were also assigned for each source. We expect that known quasars will present high probabilities of being a quasar, and low probabilities of being a star or a galaxy, as the three probabilities must sum up to 1. The same idea can be applied to the known stars and galaxies. Therefore, we can check the probabilities for each class of our spectroscopic test set. In Fig. 12 we show the probabilities distributions for each class in the with_WISE test set for RF_12S+2W+4M (RF_12S+4M@). We show the distribution in cumulative frequency for the probability of each quasar (left panel), star (middle panel) and galaxy (right panel) being quasar (pink curve), star (blue curve) or galaxy (yellow curve). Most of the known sources from with_WISE are being correctly classified with high probabilities by the classifier RF_12S+2W+4M, as shown by the high peaks near probability 1 in each panel. The classifier RF_12S+4M@ returns less certain probabilities for no_WISE sources, as seen by their broader distribution. The plot shows that the classifier RF_12S+4M@ is mostly confusing the known quasars as stars (vice-versa) and known galaxies as quasars.

In Fig. 13 we show the relation between precision and recall when a threshold $p$ in probability is set for both classifiers. These metrics were calculated using Eq. 1 and 2 on a subset of objects from the test sample that follows the condition $P_i(X = i) > p$ for each corresponding class, where $X$ denotes the event of a source being classified as class $i$, $i \in \mathcal{Y} = \{\text{QSO, STAR, GAL}\}$. Here again we consider the with_WISE test set for RF_12S+2W+4M and the no_WISE test set for RF_12S+4M@. As illustrated in the figure, using a higher probability thresholds results in higher precision and lower recall.

Note that the precision, recall and F-score metrics shown in Table 6 can be overestimated for the photometric sample, which is likely fainter. It is important to stress that a supervised learning approach can only recognize the patterns learned from the training sample, meaning that any bias in our spectroscopic sample will be reproduced in our classification. In other words, the metrics from Table 6 are only reliable for the part of the photometric data within the same feature space distribution of the spectroscopic sample. On the other hand, Clarke et al. (2020) have shown that the random forest probabilities are correlated to F-score, thus one can consider them as a reliability measurement for the classification. Therefore, we recommend that one selects the quasars, stars or galaxies from our catalogue using a probability threshold that best suits one's science by analysing Fig. 12 and Fig. 13.
Figure 11. Cumulative contamination (top panels) and miss rates (bottom panels) of quasars (left panels), stars (middle panels), and galaxies (right panels). The models referring to RF_12S+4M (12 S-PLUS bands + Morphology) and RF_12S+2W+4M (12 S-PLUS bands + 2 WISE bands + Morphology) in a Random Forest algorithm are represented by a solid and dashed curves, respectively. In the top panels, we show the $r$ magnitude distribution of all objects classified as quasars (left panel), stars (middle panel) or galaxies (right panel) in the corresponding test set for RF_12S+4M (dashed blue) and for RF_12S+2W+4M (grey). For RF_12S+4M we consider only the no_WISE testing set. In the bottom panels, we show the $r$ magnitude distribution of the known sources in these test sets. These distributions are not described by the y-axis, however the scale of the histograms is the same for all panels. The numbers on top of each histogram bar are the absolute frequency for each magnitude bin. The cumulative contamination and miss rates are relative to the total counts (i.e. sum of all frequencies of the corresponding distribution). One should avoid using the results here presented to compare the two classifiers as they are calculated on different datasets.

Figure 12. From the RF_12S+2W+4M (RF_12S+4M) test sample, we show the cumulative distribution of the estimated probabilities for the spectroscopically confirmed: 3,884 (863) QSOs, in the left panels; 9,899 (1,388) stars, in the middle panels; 15,015 (3,614) galaxies, in right panels. The cumulative distributions of the probability of an object being a quasar, star or galaxy are shown in pink, blue and yellow curves, respectively. Results from classifier RF_12S+2W+4M, referring to 12 S-PLUS bands + 2 WISE bands + Morphology, are shown in solid lines, whereas RF_12S+4M, referring to 12 S-PLUS bands + Morphology, are shown in dotted lines. One should avoid using the results here presented to compare the two classifiers as they are calculated on different datasets.
Figure 13. Relation between recall and precision that were calculated in a subset of objects that follows the condition $P_i(X = i) > p$, where $X$ is the event of an source be classified as class $i$, $i \in \mathcal{Y} = \{\text{QSO, STAR, GAL}\}$, and $p$ is the threshold shown below the points. Results for the classifiers RF_12S+4M (no_WISE test set) and RF_12S+2W+4M (with_WISE test set) are shown bottom and top panels, respectively. The results for known stars, quasars, and galaxies are shown in blue, pink, and yellow, respectively. One should avoid using the results here presented to compare the two classifiers as they are calculated on different datasets.

Figure 14. Redshift distribution of the missing and recovered quasars from the corresponding test set for RF_12S+4M (top panel) and RF_12S+2W+4M (bottom panel).

not ideal for finding underrepresented sources (e.g. high-redshift quasars) from our spectroscopic samples. For objects that are naturally rare, supplementary techniques would be necessary to properly find them. For non-rare objects that were still underrepresented in our spectroscopic samples, the solution is to increase the spectroscopic observations of these objects. Albeit there are available data from other spectroscopic surveys that could be used in our models, the construction of a better training dataset demands a detailed study, which is not in the scope of this work.

As shown in Fig. 1, quasars at some specific redshifts will not have emission lines detected in any of the S-PLUS narrow bands, which could possibly cause a bias effect. We can investigate that by checking the redshift distribution of the quasars that are not being recovered by our classifications. Figure 14 shows that we may have problems classifying quasars at redshift higher than four for both RF_12S+4M (top panel) and RF_12S+2W+4M (bottom panel), although this is due to under-representation in our spectroscopic sample. At lower redshifts, it does not seem that the non-detection of quasar emission lines in any of S-PLUS narrow bands are particularly affecting the classification.

The model RF_12S+4M is probably biased towards faint point-like sources being quasars, as previously discussed in §6.1. This means that faint stars will be often classified as quasars. On the other hand, we retrieve most of the known quasars. This problem can perhaps be solved after gathering more spectroscopic confirmed sources to train the classifier. This will be possible extending this work for areas beyond Stripe 82 and also using data from future releases. Meanwhile, one can avoid this problem by restricting the source selection in an upper magnitude below 22 for objects without WISE counterpart.

One could argue that the absence of magnitude measurements (missing bands) in any of the 12 S-PLUS filters could add bias to our models. We have shown in Fig. 5 that the missing bands are likely not affecting our classification, as our data seems to be fully separable within the considered features. In order to better investigate that, we can compare the performances between a model...
trained with a subset of our spectroscopic data that have information in all bands, and another model trained with the same subset but with simulated missing values. To simulate the missing values, we are computing the empirical proportions of each combination of missing bands from our entire dataset. We then use those proportions to sample the sources that will have the corresponding observed magnitudes replaced by a missing value (i.e. 99). Therefore, we can run a 5-fold cross-validation similarly to what was done in Section 5. Moreover, we can see how each model performs on the data that originally have missing-band values, which were not included in those models. The results for these analyses are shown in Appendix D. We see that there is no loss in performance due to the presence of missing-band values in any of our S-PLUS features. Therefore, there is no evidence to conclude that our models are being biased by the inclusion of features with missing-band values, which is in agreement with the t-SNE visualisation. We furthermore argue that missing-band values are usually a physically important information (e.g. the Lyman break) and excluding these objects from our models is not desirable.

7 NEW CATALOGUE OF STARS, QUASARS AND GALAXIES FOR S-PLUS DR1

In order to build our catalogue, we re-trained the classifiers using all spectroscopically confirmed objects from SDSS. The final model performance shown in Table 6 can change for these new models trained with the entire dataset and we cannot quantitatively account for this. Nevertheless, we would expect very similar performances as the training and testing sets were randomly drawn from the same distribution and their class were ratios maintained. Therefore, the classifier RF_12S+2W+4M to be used on sets of unknown objects was trained with the entire withWISE set, whereas RF_12S+4M was trained with all of our spectroscopic sample. Their expected performances were described in the last section.

The catalogue with the classification for sources in Stripe 82 from S-PLUS DR2 can be found on https://splus.cloud/files/SPLUS_SQGTool_DR2_B22_v1.0.0.csv. In this catalogue, there are 3076191 objects that are within the selection criteria we used to build our models, i.e. Photo_Flag_r = 0, 13 < r < 22, from a total of ~1M in the S-PLUS DR2 for the Stripe 82 region. A total of 1 347 637 objects (43.81%) have WISE counterpart and were provided with predictions from RF_12S+2W+4M. On the other hand, we provide predictions for all sources using the classifier RF_12S+4M. These numbers include sources that were in our spectroscopic training sample. The columns in our catalogue that are being classified as galaxies, stars, or quasars with very confident probabilities (> 80%). This plot does not include any source with spectroscopic information and we only show objects without WISE counterpart. We can conclude from this plot that we are able to identify patterns with RF, otherwise, our classes would not appear clustered. We cannot conclude, however, that we are labelling these patterns correctly by only looking at the t-SNE plots.

Figure 17 shows how the labels from the RF classifier are related to an unsupervised clustering using t-SNE algorithm, which is an independent process from RF. In this figure we show sources from our catalogue that are being classified as galaxies, stars, or quasars with very confident probabilities (> 80%). This plot does not include any source with spectroscopic information and we only show objects without WISE counterpart. The t-SNE plot itself does not prove that we are labelling these objects correctly, nevertheless, it would be an indication of problems in the classification if the colours from RF labels were randomly dispersed throughout the t-SNE clusters. Although the t-SNE algorithm is totally different from the RF process, the figure clearly shows that the spatial clusters are in agreement with the RF labels. In addition, in the great majority of the cases, the three classes do not significantly overlap and most
Figure 15. Histograms show the corresponding probabilities distribution of classified quasars (left panels), stars (middle panels), and galaxies (right panels) that were not part of our spectroscopic training sample. Top panels shows the probabilities distribution of sources without WISE counterpart. Bottom panels are for the remaining sources with WISE counterpart. For those, we show the probabilities distribution from both classifiers, but we only consider the classified objects from RF_12S+2W+4M. We can see from bottom-left panel that many objects classified as quasars by RF_12S+2W+4M (black) are not being classified as such by RF_12S+4M. For each panel we also show the probability distribution from their corresponding test set (gray).

Figure 16. Distribution of $\mathcal{W}1$ for the sources we classify as quasars (pink), stars (blue), or galaxies (yellow). We also show in gray the distribution for the spectroscopically confirmed objects.

of the groups show a clear separation between the labels from RF and quasars are clearly in distinct islands, in most cases.

The S-PLUS DR2 fields in the interval 20h < RA < 21h (300° < RA < 320°) are extremely populated by stars, given they are in the Galactic Plane, i.e. $|b| < 30$. Thus, we might have greater number counts of stars being confused as quasars or galaxies compared to less populated regions. This can also happens in regions around resolved stellar clusters. Figure 18 shows the total number of objects from Stripe 82 S-PLUS DR2 being classified as quasar, star or galaxy with both of our classifiers as a function of Galactic coordinates. As the number counts for galaxies and quasars are expected to be homogeneous in the spatial distribution, it is clear from Fig. 18 that there is a high absolute number of contaminated sources in the quasar sample at lower galactic latitudes for classifier RF_12S+4M. In contrast, this behaviour is alleviated when WISE magnitudes are taken into consideration, as seen from the right panel of Fig. 18.

The SDSS has a photometric pipeline that performs a star/galaxy separation based on the difference of the composite model magnitudes ($cmodelMag$) and PSF magnitudes ($psfMag$) with 95% confidence level to r_{SDSS} = 21 (Stoughton et al. 2002). The sources are classified as extended source (type == 3) if they satisfies the relation: $psfMag - cmodelMag > 0.145$. Otherwise, they will be classified as point-like source (type == 6). We compared this information with our obtained results, assuming that extended sources are likely to be galaxies, whereas point-like sources are likely to be stars or quasars. In Fig. 19 we show the distribution of all extended (top panels) and point-like sources (bottom panels) for our two classifiers down to magnitude $r = 22$. For the classifier RF_12S+2W+4M (RF_12S+4M), 0.78% (1.97%) out of the extended sources are being classified as quasars (most at the faint end) and 1.35% (0.83%) are being classified as stars. On the other hand, we are classifying 1.52% (4.91%) out of the point-like sources as galaxies.

There are 195 galaxies in Stripe 82 S-PLUS DR2 that were not in our training sample, but were classified by Galaxy Zoo DECaLS (Walmsley et al. 2021). From those, only three galaxies were clas-
magnitudes. We also compared RF_12S versus RF_12S+2W and when narrow-band information is available along with broad-band concluded, with 90% confidence, that the classification is improved RF_12S versus RF_5S and RF_12S+2W versus RF_5S+2W. We sess the importance of the S-PLUS narrow bands, we compared PLUS broad-bands + 2 WISE bands (5S+2W). In order to as-

8 SUMMARY AND CONCLUSIONS

In the context where many large photometric surveys are emerging, we need an efficient, accurate, optimized and automated technique for classifying objects. We compared two supervised learning methods based on Support Vector Machines (SVM) and Random Forest (RF) algorithms for the star/quasar/galaxy classification of S-PLUS DR2 for the Stripe 82 region. From our analysis we conclude that S-PLUS has a great advantage in distinguishing these three classes due to its Javalambre filter system, consisting of 7 narrow and 5 broad bands. Our classification is robust against missing bands, as first indicated by the t-SNE results and further evidenced from an analysis of simulated missing data. We were also capable of improving the star/galaxy classification with respect to Costa-Duarte et al. (2019), extending the analysis to r = 22. Here we summarize our main results from the cross-validation:

- Algorithm selection: we compare the performance of RF and SVM through macro-averaged F-score and we concluded that the best results came from RF. It also generally required less computational time to train the classifier.
- Feature space evaluation: we tested the performance of four feature spaces: 12 S-PLUS bands (12S); 12 S-PLUS bands + 2 WISE bands (12S+2W); 5 S-PLUS broad-bands (5S); and 5 S-PLUS broad-bands + 2 WISE bands (5S+2W). In order to assess the importance of the S-PLUS narrow bands, we compared RF_12S versus RF_SS and RF_12S+2W versus RF_SS+2W. We concluded, with 90% confidence, that the classification is improved when narrow-band information is available along with broad-band magnitudes. We also compared RF_12S versus RF_12S+2W and RF_5S versus RF_5S+2W, confirming the importance of WISE bands in the classification. From these results, we selected RF_12S and RF_12S+2W for new tests including morphological parameters (FWHM, A, B and KRON_RADIUS), which we further referred to as RF_12S+4M and RF_12S+2W+4M, respectively. All metrics were statistically improved with the addition of morphological information in our models. In order to provide a more generalized model, we increase the dataset for RF_12S+4M by including sources without WISE counterpart, which is now denoted by RF_12S+4M.

- Final model performances: the expected performances were calculated from two unseen test samples: one that only contains objects with WISE counterpart, and another that only contains objects with no WISE information. Considering the former, we calculated the overall F-score, which is the harmonic mean of precision and recall, macro-averaged over all three classes as 97.69% (95.04%) for RF_12S+2W+4M (RF_12S+4M). In terms of quasar classification, we achieved 95.76% (95.68%) for precision/purity and 95.88% (86.1%) for recall/completeness. In terms of star classification, we achieved a precision of 99.44% (95.59%) and recall of 98.22% (97.97%). Finally for galaxies, a precision of 98.04% (97.26%) and a recall of 98.8% (98.19%) were obtained. These metrics demonstrate the efficacy of both our classifiers. For RF_12S+4M, we also calculated the performances on objects without WISE counterpart, for which we obtained an overall F-score of 95.04%; 52.47% precision and 92.24% recall for quasars; 98.17% precision and 78.56% recall for stars; 81.39% precision and 85.37% recall for galaxies. A total of 546 stars in the faint-end of r distribution are being misclassified as quasars, which reflects in the bad quasar precision and star recall.

We trained two classifiers using the complete spectroscopic sample from SDSS:

- Classifier RF_12S+4M: 12 S-PLUS bands + Morphology parameters in a Random Forest algorithm
- Classifier RF_12S+2W+4M: 12 S-PLUS bands + 2 WISE bands + Morphology parameters in a Random Forest algorithm (trained only with objects that have WISE counterparts)

In our catalogue, we provided two classifications (using RF_12S+4M and RF_12S+2W+4M) and their corresponding estimated probabilities of a source being a star, quasar or galaxy. The classifier RF_12S+2W+4M achieved a 2.79% increase in macro-averaged F-score over classifier RF_12S+4M within objects with WISE counterpart, dominated by a 11.36% increase in recall for quasars. Thus we recommend that one uses the classification from RF_12S+2W+4M whenever the WISE information is available. For sources with no WISE counterpart, we can only rely on the classifier RF_12S+4M that has a reasonable performance down to magnitude r ≥ 21, for which all metrics are higher than 90% (except for quasar precision that is 83.59%). Moreover, we have shown that using the 12 S-PLUS bands provides better performances than using only 5 broad bands, improving macro-averaged F-score in 2.19% (see §5.2). For any of the classifiers, we recommend that one uses Fig. 12 and 13 shown in §6.1 in order to select a probability threshold that gives the best star, quasar or galaxy sample that suits one’s science. One should also consider Fig. 11 for selecting an upper limit magnitude.

Another useful approach for the star/quasar/galaxy separation are methods that directly process the broad- and narrow-band images, instead of using catalogued datasets created from source-finding algorithms. Our group is currently investigating the use of deep learning techniques for such image-based classifications. Pre-
Figure 18. Number counts of total objects from Stripe 82 S-PLUS DR2 being classified as quasars (top panels), stars (middle panels) and galaxies (bottom panels) per galactic coordinates. Each pixel corresponds to an area of 1x1 deg\(^2\). Results for the classifiers RF\(_{12S+4M}\) and RF\(_{12S+2W+4M}\) are shown in left and right panels, respectively. We set a top limit of 1500 counts in order to provide a better visualization of the higher longitudes. Thus, the dark red regions at lower longitudes should not be interpreted as homogeneous. We recommend caution when using our classification for crowded regions (e.g. regions close to the Galactic Plane), due to the high contamination of stars.

Figure 19. The stacked histograms show the frequencies of extended and point-like sources (in the top and bottom panels, respectively) that were classified as stars, quasars or galaxies with classifier RF\(_{12S+4M}\) (left panels) and RF\(_{12S+2W+4M}\) (right panels). The photometric classification of extended and point-like sources were retrieved from SDSS DR16. The numbers on the top of each histogram bar refer to the percentage of spectroscopically confirmed galaxies (orange), quasars (pink) and stars (blue) in each bin. In total, 0.78\% (1.97\%) out of the extended sources are being classified as quasars (most at the faint end) and 1.35\% (0.83\%) are being classified as stars with RF\(_{12S+2W+4M}\) (RF\(_{12S+4M}\)). On the other hand, we are classifying 1.52\% (4.91\%) out of the point-like sources as galaxies. Assuming that the extended sources are likely to be galaxies and the point-like sources are likely to be quasars or stars, these results are in fairly good agreement with the estimated contamination and miss rates on the galaxy sample that were shown in Fig. 11.
DATA AVAILABILITY

The S-PLUS DR2 can be obtained from https://splus.cloud/query. The ALLWISE catalogue used in this work can be obtained at https://datalab.noao.edu/. The SDSS DR16 catalogue can be obtained at http://skyserver.sdss.org/CasJobs. The SDSS DR14Q catalogue can be obtained at https://www.sdss.org/dr14/algorithms/qso_catalog/.

The Stripe 82 S-PLUS DR2 catalogue containing classification and probabilities for all sources with PhotoFlag_r = 0 and 13 < r < 22 is available for download at https://splus.cloud/files/SPPLUS_SQGTool_DR2_822_v1.0.0.csv. Classification for all sources in DR2 can also be downloaded from https://splus.cloud/query under dr2_vacs schema as star_galaxy_quasar table. We made available the spectroscopic sample used in this paper through https://splus.cloud/files/Spectroscopic_sample.csv, where we use the flag 1 to identify sources that were used for training, and 0, otherwise (for both cross-validation and final model performance contexts).

The codes for the analyses presented in this paper are available at https://github.com/ledik/MNRAS_nakazono_2021, where the ADQL queries for downloading the above mentioned datasets can also be found.

ACKNOWLEDGEMENTS

At the time this work was accepted for publication, Brazil reached 496,172 deaths by COVID-19. I dedicate my first-first-author paper to all Brazilians who lost your beloved ones. I would like to thank my family, my therapist and my close friends who have been of great support for my mental health throughout this political chaos.

This work has been supported by a PhD fellowship to the lead author from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, 2019/01312-2). LN also acknowledges the support of Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 and FAPESP through process number 2014/10566-4. LN also thanks the staff of the Astronomy department from the University of Florida, where part of this work was done; Marco Antonio dos Santos and Ulisses Manzo Castello for the technical support; Luis Manrique for the technical support, feedback and discussions about Machine Learning; Christian Massao Tsujiguchi Takagi and Vinícius Amaral Haga for the feedback on the accessibility of the figures in this paper; Gustavo Schwarz for the help with the database access, NSTH acknowledges FAPESP (grants 2017/25835-9 and 2015/22308-2), CQ acknowledges support from FAPESP (grants 2015/11442-0 and 2019/06766-1), AM acknowledges FAPESP scholarship grant 2018/25671-9, CEB acknowledges FAPESP, grant 2016/12331-0. FA-F acknowledges funding for this work from FAPESP grant 2018/20977-2. LSJ acknowledges support from Brazilian agencies FAPESP (2019/10923-5) and CNPq (304819/201794). AAC acknowledges support from FAPERJ (grant E26/203.186/2016 and CNPq (304971/2016-2 and 401669/2016-5). EVL acknowledges funding for this work from CNPq grant 169181/2017-0 and CAPES grant 88887.470064/2019-00. MLB acknowledges support of FAPESP grant 2018/09165-6 and 2019/23388-0. KMD acknowledges support from FAPERJ (grant E26/203.184/2017), CNPq (grant 312702/2017-5) and the Serrapilheira Institute (grant Serra-1709-17357). AAC acknowledges support from FAPERJ (grant E26/203.186/2016), CNPq (304971/2016-2 and 401669/2016-5), from the Universidade de Alicante under contract UATALENTO18-02, and from the State Agency for Research of the Spanish MICIU through the “Center of Excellence Severo Ochoa” award to the Instituto de Astrofísica de Andalucía (SEV-2017-0709). ARL acknowledges the financial support from CNPq through the PCI fellowship. The authors would also like to thank all fellow colleagues from the S-PLUS collaboration who have been adding value to this work throughout those years with great suggestions and discussions. In special: Dr João E. Steiner (in memoriam), Dr Marcus V. Costa Duarte, Dr Alberto Molino, Dr Rodrigo C. Thom de Souza, Dr Raimundo Lopes, Dr Maria L. Linhares Dantas, Carlos A. Galarza, Dr Jose L. Nilo Castellón, Leonildo J. de Melo de Azevedo, and Geferson Lucatelli. We would also like to deeply thank Dr Alex Clarke who has carefully read our paper and gave very valuable feedback.

This research made use of Scikit-learn (Pedregosa et al. 2011); Numpy (Harris et al. 2020); seaborn (Waskom & the seaborn development team 2020); matplotlib, a Python library for publication quality graphics (Hunter 2007); Astropy, a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013); TOPCAT, an interactive graphical viewer and editor for tabular data (Taylor 2005).

The S-PLUS project, including the T80-South robotic telescope and the S-PLUS scientific survey, was founded as a partnership between the Fundação de Amaparo à Pesquisa do Estado de São Paulo (FAPESP), the Observatório Nacional (ON), the Federal University of Sergipe (UFS), and the Federal University of Santa Catarina (UFSC), with important financial and practical contributions from other collaborating institutes in Brazil, Chile (Universidad de La Serena), and Spain (Centro de Estudios de Física del Cosmos de Aragón, CEyCA). We further acknowledge financial support from the São Paulo Research Foundation (FAPESP), the Brazilian National Research Council (CNPq), the Coordination for the Improvement of Higher Education Personnel (CAPES), the Carlos Chagas Filho Rio de Janeiro State Research Foundation (FAPERJ), and the Brazilian Innovation Agency (FINEP).

The authors are grateful for the contributions from CTIO staff in helping in the construction, commissioning and maintenance of the T80-South telescope and camera. We are also indebted to Rene Laporte and INPE, as well as Keith Taylor, for their important contributions to the project. We also thank CEIPEA staff for their help with T80-South, specifically we thank Antonio Marín-Franch for his invaluable contributions in the early phases of the project, David Cristóbal-Hornillos and his team for their help with the installation of the data reduction package nrvn version 0.9.9, César Íñiguez for providing 2D measurements of the filter transmissions, and all other staff members for their support.

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is www.sdss.org. SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), National Central University, National Taiwan University, New Mexico State University, New York University, University of Notre Dame, Ohio State University, University of Pennsylvania, University of Portsmouth, Princeton University, Spanish MCIU through the “Center of Excellence Severo Ochoa” award to the Instituto de Astrofísica de Andalucía (SEV-2017-0709). ARL acknowledges the financial support from CNPq through the PCI fellowship. The authors would also like to thank all fellow colleagues from the S-PLUS collaboration who have been adding value to this work throughout those years with great suggestions and discussions. In special: Dr João E. Steiner (in memoriam), Dr Marcus V. Costa Duarte, Dr Alberto Molino, Dr Rodrigo C. Thom de Souza, Dr Raimundo Lopes, Dr Maria L. Linhares Dantas, Carlos A. Galarza, Dr Jose L. Nilo Castellón, Leonildo J. de Melo de Azevedo, and Geferson Lucatelli. We would also like to deeply thank Dr Alex Clarke who has carefully read our paper and gave very valuable feedback.

This research made use of Scikit-learn (Pedregosa et al. 2011); Numpy (Harris et al. 2020); seaborn (Waskom & the seaborn development team 2020); matplotlib, a Python library for publication quality graphics (Hunter 2007); Astropy, a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013); TOPCAT, an interactive graphical viewer and editor for tabular data (Taylor 2005).

The S-PLUS project, including the T80-South robotic telescope and the S-PLUS scientific survey, was founded as a partnership between the Fundação de Amaparo à Pesquisa do Estado de São Paulo (FAPESP), the Observatório Nacional (ON), the Federal University of Sergipe (UFS), and the Federal University of Santa Catarina (UFSC), with important financial and practical contributions from other collaborating institutes in Brazil, Chile (Universidad de La Serena), and Spain (Centro de Estudos de Física del Cosmos de Aragón, CEyCA). We further acknowledge financial support from the São Paulo Research Foundation (FAPESP), the Brazilian National Research Council (CNPq), the Coordination for the Improvement of Higher Education Personnel (CAPES), the Carlos Chagas Filho Rio de Janeiro State Research Foundation (FAPERJ), and the Brazilian Innovation Agency (FINEP).

The authors are grateful for the contributions from CTIO staff in helping in the construction, commissioning and maintenance of the T80-South telescope and camera. We are also indebted to Rene Laporte and INPE, as well as Keith Taylor, for their important contributions to the project. We also thank CEIPEA staff for their help with T80-South, specifically we thank Antonio Marín-Franch for his invaluable contributions in the early phases of the project, David Cristóbal-Hornillos and his team for their help with the installation of the data reduction package nrvn version 0.9.9, César Íñiguez for providing 2D measurements of the filter transmissions, and all other staff members for their support.

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is www.sdss.org. SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astro-
APPENDIX A: AFFILIATIONS

6 INAF - Osservatorio Astronomico di Padova, Vicolo Osservatorio 5, 35122 Padova, Italy
7 Observatório do Valongo, Universidade Federal do Rio de Janeiro, Ladeira Pedro Antônio 43, Saude, Rio de Janeiro, RJ, Brasil, CEP 20080-090
8 Institute for Astronomy, Astrophysics, Space Applications and Remote Sensing, National Observatory of Athens, Penteli GR 15236, Greece
9 Instituto de Astrofísica de Andalucía, CSIC, Apt 3004, E18080 Granada, Spain
10 IUFACyT, Universidad de Alicante, San Vicent del Raspeig, E03080, Alicante, Spain
11 GMTO Corporation 465 N. Halstead Street, Suite 250 Pasadena, CA 91107
12 Departamento de Física, Universidade Federal de Santa Catarina, Florianópolis, SC, 88040-900, Brazil
13 NOAO, P.O. Box 26732, Tucson, AZ 85726

APPENDIX B: STATISTICAL TESTS

We performed some statistical tests to be able to conclude which model had the best performance compared to another with a given significance level, as discussed in Section 5. We used the one-tailed Wilcoxon signed-rank test (Wilcoxon 1945), which is a non-parametric method for testing two hypothesis:

- \( H_0 \): the median of the differences in any considered metric is negative,
- \( H_1 \): the median of the differences in any considered metric is positive

As we are using five folds for the cross validation, we have five independent measurements for each metric. For each classifier/metric, these measurements are ordered from highest to lowest and paired. The test statistics \( W \) is given by:

\[
W = \sum_{i=1}^{N_r} [\text{sgn}(x_{2,i} - x_{1,i})] \times R_i
\]  

where \( \text{sgn} \) refers to the sign function, \( N_r \) is the total number of pairs (excluding pairs that \( |x_{2,i} - x_{1,i}| = 0 \)), \( R_i \) is the rank for each pair, and \( i = \{1, \ldots, 5\} \). To calculate the statistics \( W \), we used the wilcoxon module from scipy.stats in Python, with \( \text{mode} = \text{"exact"} \) and \( \text{alternative} = \text{"greater"} \). We always put in the parameter \( x \) of wilcoxon the set of measurements with highest average between the two. Thus, if the null hypothesis \( H_0 \) is rejected, there is statistical evidence to accept that the corresponding classifier with highest average in a specific metric had performed better than the other. The null hypothesis \( H_0 \) is rejected if the p-value is below 10%, i.e., considering a confidence level of 90%.

We show in Table E2 the statistics \( W \) and the corresponding p-value for each of the comparisons we have done in Sec. 5

APPENDIX C: MODEL SETTINGS

In this section we show the hyper-parameters we used for t-SNE, Random Forest, and Support Vector Machines with sklearn (version 0.24.1) package in Python.

APPENDIX D: MISSING-BAND VALUES ANALYSES

In this section we describe the analyses we made to investigate any bias effect due to missing-band values. For that, we separate our entire spectroscopic sample into two complementary datasets containing objects that:

---

**Table C1. Hyperparameters of the t-SNE implementation.**

| Hyperparameter   | Value |
|------------------|-------|
| n_components     | 2     |
| perplexity       | 30    |
| early_exaggeration| 12    |
| learning_rate    | 200   |
| n_iter           | 1000  |
| n_iter_without_progress| 300  |
| min_grad_norm    | 10^{-7} |
| metric           | euclidian |
| init             | random |
| verbose          | 0     |
| random_state     | 2     |
| method           | barnes_hut |
| angle            | 0.5   |
| n_jobs           | None  |
| square_distances | legacy |

**Table C2. Hyperparameters of the Random Forest implementation.**

| Hyperparameter   | Value |
|------------------|-------|
| n_estimators     | 100   |
| criterion        | gini  |
| max_depth        | None  |
| min_samples_split| 2     |
| min_samples_leaf | 1     |
| min_weight_fraction_leaf | 0     |
| max_features     | auto  |
| max_leaf_nodes   | None  |
| min_impurity_decrease | 0      |
| min_impurity_split | None  |
| bootstrap        | False |
| oob_score        | False |
| n_jobs           | None  |
| random_state     | 2     |
| warm_start       | False |
| class_weight     | None  |
| ccp_alpha        | 0     |
| max_samples      | None  |

**Table C3. Hyperparameters of the Support-Vector Machine implementation.**

| Hyperparameter       | Value |
|----------------------|-------|
| C                    | 1     |
| kernel               | rbf   |
| degree               | 3     |
| gamma                | scale |
| shrinking            | True  |
| tol                  | \(10^{-3}\) |
| cache_size           | 200   |
| class_weight         | None  |
| max_iter             | -1    |
| decision_function_shape | ovo |
| break_ties           | False |
| random_state         | 2     |
Table D1. Cross-validation results for a random forest algorithm. Two different datasets were used within each considered model features. The "original" dataset refers to the subset of our spectroscopic sample for which all sources have information in the twelve S-PLUS bands. The "original with simulations" refers to the same dataset but with imputed missing values in some bands. The performance are described in terms of precision, recall and F-score (F) for each class. The overall performance is shown in terms of macro-averaged F-score ($\bar{F}$).

| Model features | Train/validation data | $\bar{F}$ (%) | CLASS | Precision (%) | Recall (%) | F (%) |
|----------------|-----------------------|---------------|-------|---------------|------------|-------|
| 12 S-PLUS bands + 4 morphological features | original | 93.09 ± 0.2 | QSO | 85.57 ± 0.78 | 89.68 ± 0.53 | 87.57 ± 0.46 |
| | STAR | 96.21 ± 0.3 | 94.31 ± 0.32 | 95.25 ± 0.21 |
| | GAL | 96.42 ± 0.19 | 96.48 ± 0.08 | 96.45 ± 0.11 |
| | original with simulations | 92.92 ± 0.19 | QSO | 85.53 ± 0.62 | 89.03 ± 0.63 | 87.24 ± 0.38 |
| | STAR | 95.87 ± 0.38 | 94.12 ± 0.34 | 94.99 ± 0.21 |
| | GAL | 96.42 ± 0.24 | 96.62 ± 0.07 | 96.52 ± 0.15 |
| 12 S-PLUS bands + 2 WISE bands + 4 morphological features | original | 97.95 ± 0.15 | QSO | 96.12 ± 0.58 | 97.38 ± 0.15 | 96.75 ± 0.3 |
| | STAR | 99.41 ± 0.1 | 98.53 ± 0.11 | 98.97 ± 0.07 |
| | GAL | 98.01 ± 0.12 | 98.23 ± 0.22 | 98.12 ± 0.14 |
| | original with simulations | 98.01 ± 0.2 | QSO | 96.2 ± 0.61 | 97.54 ± 0.34 | 96.87 ± 0.4 |
| | STAR | 99.39 ± 0.06 | 98.55 ± 0.11 | 98.97 ± 0.05 |
| | GAL | 98.11 ± 0.18 | 98.26 ± 0.2 | 98.18 ± 0.17 |

Table D2. Performance on the subset of our spectroscopic sample that have missing-band value in at least one S-PLUS band for the model trained with the subset with no missing-band values. In parenthesis we show the performances for the model trained with the same subset but with simulated missing-band values.

| Model Features | $\bar{F}$ (%) | CLASS | Precision (%) | Recall (%) | F (%) |
|----------------|---------------|-------|---------------|------------|-------|
| 12 S-PLUS bands + 4 morphological features | 85.31 (85.19) | QSO | 64.70 (65.76) | 72.71 (71.00) | 68.47 (68.28) |
| | STAR | 95.13 (95.82) | 87.88 (87.14) | 91.36 (91.27) |
| | GAL | 95.00 (94.41) | 97.19 (97.66) | 96.08 (96.01) |
| 12 S-PLUS bands + 2 WISE bands + 4 morphological features | 93.71 (93.39) | QSO | 84.58 (83.64) | 86.44 (86.11) | 85.50 (84.86) |
| | STAR | 99.44 (99.53) | 95.74 (95.30) | 97.55 (97.37) |
| | GAL | 97.43 (97.22) | 98.76 (98.69) | 98.09 (97.95) |

(i) have information in all 12 S-PLUS bands  
(ii) have missing value in at least one S-PLUS band

The magnitude distribution from each of these subsets are shown in Fig. D1. By using subset (i), we are able to investigate how missing-band values can possibly affect the performance of a random forest modelling. We use subset (ii) to empirically calculate the proportion of bands with missing values for each class. This proportion is calculated considering each combination of bands, totaled 12282 cases. We randomly sample objects from subset (i) following these empirical proportions. Then replace their observed magnitudes in the corresponding bands to missing values (i.e. 99). We run 4-fold cross validation for which we want to compare the performances of a random forest algorithm trained either with subset (i) or with the same subset but with imputed missing values.

If there is any bias effect due to missing-band values, we should see a significant decrease in performance for the latter compared to the former. The cross-validation results are shown in Table D1, for which we do not see any great difference for the performances between the two datasets.

APPENDIX E: CATALOGUE DESCRIPTION

This paper has been typeset from a LATEX file prepared by the author.
Table E1. Description of the columns in our catalogue.

| Column   | Name          | Format  | Description                                                                 |
|----------|---------------|---------|------------------------------------------------------------------------------|
| 1        | RA            | float64 | Right ascension in degrees (J2000)                                           |
| 2        | DEC           | float64 | Declination in degrees (J2000)                                               |
| 3        | FWHM_n        | float64 | Full width at half maximum normalized per field                               |
| 4        | A             | float64 | Major semi-axis                                                              |
| 5        | B             | float64 | Minor semi-axis                                                              |
| 6        | KRON_RADIUS   | float64 | Kron Radius                                                                  |
| 7        | u_iso         | float64 | Magnitude in AB system for u band                                            |
| 8        | J0378_iso     | float64 | Magnitude in AB system for J0378 band                                         |
| 9        | J0395_iso     | float64 | Magnitude in AB system for J0395 band                                         |
| 10       | J0410_iso     | float64 | Magnitude in AB system for J0410 band                                         |
| 11       | J0430_iso     | float64 | Magnitude in AB system for J0430 band                                         |
| 12       | g_iso         | float64 | Magnitude in AB system for g band                                            |
| 13       | J0515_iso     | float64 | Magnitude in AB system for J0515 band                                         |
| 14       | r_iso         | float64 | Magnitude in AB system for r band                                            |
| 15       | J0660_iso     | float64 | Magnitude in AB system for J0660 band                                         |
| 16       | i_iso         | float64 | Magnitude in AB system for i band                                            |
| 17       | J0861_iso     | float64 | Magnitude in AB system for J0861 band                                         |
| 18       | z_iso         | float64 | Magnitude in AB system for z band                                            |
| 19       | w1mpro        | float64 | Magnitude in Vega system for W1 band taken from ALLWISE catalogue            |
| 20       | w2mpro        | float64 | Magnitude in Vega system for W2 band taken from ALLWISE catalogue            |
| 21       | w1snr         | float64 | W1 signal-to-noise ratio taken from ALLWISE catalogue                         |
| 22       | w2snr         | float64 | W2 signal-to-noise ratio taken from ALLWISE catalogue                         |
| 23       | w1sigmpro     | float64 | W1 uncertainty in mag units taken from ALLWISE catalogue                      |
| 24       | w2sigmpro     | float64 | W2 uncertainty in mag units taken from ALLWISE catalogue                      |
| 25       | CLASS_SPEC    | object  | Spectroscopic classification taken from SDSS DR14Q/DR16                      |
| 26       | spec_flag     | int64   | 0, if the object was in our training sample and 1, otherwise.                 |
| 27       | CLASS         | float64 | Object classification from RF_12S+4M@ for sources with no WISE counterpart     |
| 28       | PROB_QSO      | float64 | Probability of the source being a quasar from RF_12S+4M@                     |
| 29       | PROB_STAR     | float64 | Probability of the source being a star from RF_12S+4M@                       |
| 30       | PROB_GAL      | float64 | Probability of the source being a galaxy from RF_12S+4M@                     |
| 31       | CLASS_WISE    | float64 | Object classification from RF_12S+2W+4M for sources with WISE counterpart      |
| 32       | PROB_QSO_WISE | float64 | Probability of the source being a quasar from RF_12S+2W+4M                   |
| 33       | PROB_STAR_WISE| float64 | Probability of the source being a star from RF_12S+2W+4M                     |
| 34       | PROB_GAL_WISE | float64 | Probability of the source being a galaxy from RF_12S+2W+4M                   |

*Not all training objects are in this catalogue due to our selection criteria.
Figure E1. Distribution of (a) FWHM and (b) magnitude in $r$ band for the sources we classify as quasars (pink), stars (blue), or galaxies (yellow). Top (bottom) panels show objects without (with) WISE counterpart that were classified by RF$_{12S+4M}$ (RF$_{12S+2W+4M}$). We also show in gray the distribution for the spectroscopically confirmed objects.
Table E2. P-values for the comparisons from Section 5. P-values below 0.1 indicate evidence that Classifier 1 performs statistically better than Classifier 2.

| Classifier 1     | Classifier 2     | Metric | W   | p-value |
|------------------|------------------|--------|-----|---------|
| RF_5S            | SVM_5S           | All    | 15  | 0.03    |
| RF_12S           | SVM_12S          | All    | 15  | 0.03    |
| RF_5S+2W         | SVM_5S+2W        | All    | 15  | 0.03    |
| RF_12S+2W        | SVM_12S+2W       | P_QSO  | 8   | 0.5     |
|                  |                  | Others | 15  | 0.03    |
| RF_12S           | RF_5S            | All    | 15  | 0.03    |
| RF_12S+2W        | RF_5S+2W         | P_QSO  | 14  | 0.06    |
|                  |                  | R_QSO  | 1   | 0.97    |
|                  |                  | F_QSO  | 12  | 0.16    |
|                  |                  | R_STAR | 13  | 0.09    |
|                  |                  | P_GAL  | 11  | 0.22    |
|                  |                  | Others | 15  | 0.03    |
| RF_5S+2W         | RF_5S            | All    | 15  | 0.03    |
| RF_12S+2W+4M     | RF_12S+2W        | P_QSO  | 10  | 0.31    |
|                  |                  | R_QSO  | 14  | 0.06    |
|                  |                  | F_QSO  | 14  | 0.06    |
|                  |                  | Others | 15  | 0.03    |
| RF_12S+4M        | RF_12S           | All    | 15  | 0.03    |
| RF_12S+2W+4M     | RF_12S+4M        | All    | 15  | 0.03    |