An Empirical Calibration of the Helium Abundance in HII Regions based in Literature and CALIFA Survey data

M. Valerdi1*, J. K. Barrera-Ballesteros,1 S. F. Sánchez,1 C. Espinosa-Ponce,1 L. Carigi1 and A. Mejía-Narváez,1
1Instituto de Astronomía, Universidad Nacional Autónoma de México, AP 70264, 04510 Mexico City, Mexico

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT
He is the second most common chemical species in the Universe. The study of helium abundance has the potential to unravel the chemical evolution of and within galaxies. In this study, we provide an empirical calibration for the singly ionized helium abundance: \[ \log_{10}(\frac{\text{He}}{\text{H}}) \], based on the emission line flux ratio \( \text{He}\alpha/\text{H}\alpha \) from Galactic and extragalactic H II regions compiled from the literature. Based on this calibrator, we explore for the first time the helium abundance in a large sample of H II regions located in galaxies representative of the nearby Universe from the CALIFA survey. Furthermore, this calibrator allows us to explore the variations of the helium abundance with respect to the oxygen abundance. The observed trends are in agreement with a change in the chemical enrichment with mass/oxygen abundance similar to the one observed due to the inside-out model in a MW-galaxy (highlighting the connection between resolved and global trends in galaxies). Our calibrator provides an empirical proxy to estimate the helium abundance at kpc scales as well as to constrain chemical evolutionary models.

Key words: galaxies: statistics – galaxies: ISM – galaxies: abundances

1 INTRODUCTION
The formation of the elements is a key to understand the evolution of the Universe. In particular, the formation of helium has been fundamental to the study of cosmology and the chemical evolution of galaxies. The majority of the cosmic helium was produced just after the Big Bang, during the primordial nucleosynthesis phase (Kobayashi et al. 2020b). Afterward, the first stars that formed contained only hydrogen and helium. The heavier elements have been formed from nuclear reactions in the interior of successive generations of stars (e.g. Bromm & Larson 2004; Schneider et al. 2006).

In the present helium constitutes about 24–25% of the baryonic mass in the Universe and it is the second most abundant element. Helium is produced by hydrostatic nucleosynthesis in the interior of stars of all initial masses. Massive stars produce this element through the CNO cycle and low-mass stars via the proton-proton chain (e.g. Weaver & Woosley 1980; Arnett & Thielemann 1985). Finally, there is an amount of helium that enriches the ISM through stars that ejected it. This production depends on the progenitor star’s initial mass, metallicity, stellar rotation, and stellar winds (e.g. Peimbert & Serrano 1980; Peimbert 1986).

In principle, due to the helium nature, one might expect that their abundance to be well determined. However, it is not possible to observe all ionization stages, particularly when it is in the neutral form. Due to the high efficiency of the absorption of ultraviolet radiation by hydrogen, the boundary between ionized and neutral gas is very well defined. For this reason, helium may exist in its neutral form in the outermost parts of these regions. The neutral helium cannot be measured, for this reason, the total helium abundance can only be estimated in H II regions where all helium is ionized (high excitation H II regions, e.g. Orion nebula Esteban et al. 2004). To obtain the total abundance of an element one would have to sum over all their ionic stages (and correct for the neutral component). However, for most elements, we cannot observe all the ionic stages, and the helium is not the exception. To solve this problem it is usually used Ionization correction factors (ICFs) (e.g. Peimbert & Torres-Peimbert 1977a; Stasińska 1978; Vilchez 1989; Pérez-Montero et al. 2007; Delgado-Inglada et al. 2014a; Amayo et al. 2020).

Observationally, the spectra of H II regions only show recombination lines of He I and sometimes of He II. However, as we mentioned before, it is possible that it is found in a neutral form in the external areas of these regions. Vilchez (1989); Deharveng et al. (2000) suggested that within the H II regions with low and medium ionization degrees the presence of neutral helium is important and the ICF(He) > 1. However, for objects with a high ionization degree, we can consider the presence of neutral helium within the nebula to be negligible. To obtain the total helium abundance, there are several ICF schemes for helium available in the literature and based on different ionic ratios (e.g., Peimbert & Torres-Peimbert

* E-mail: mvalerdi@astro.unam.mx

© 2020 The Authors
The main goal of this study is to obtain a calibrator to estimate the single ionized helium abundance using the best compilation of H II regions available in the literature, and compare with helium abundances obtained from the catalog of H II regions of CALIFA. The helium abundance is usually measured using various emission lines in the optical range and using spectra with high spectral resolution, thus having reliable observations on weak lines. The helium abundance has been studied in several works (e.g. Izotov et al. 2007; Peimbert et al. 2007; Aver et al. 2015; Méndez-Delgado et al. 2020), the largest sample is reported by Izotov et al. (2007), with ~350 H II regions. Most of these works aim to study the primordial helium abundance and therefore seek a highly accurate determination. In this study, we derive an empirical calibration for the singly ionized helium abundance, and we derive it in one of the largest sample of H II (5386 H II regions). This calibrator is important because helium has not frequently used to trace chemical evolution. However, this study could help to constrain the chemical evolution models, and consequently, we could understand better the chemical enrichment process in galaxies. In addition our sample was extracted from an statistical significant sample of galaxies representative of the population in the nearby Universe.

The content of the article is distributed as follows: In Section 2, we describe two different datasets on which our study is based: I) the compilation of Galactic and extragalactic H II regions available in the Literature with helium abundance determinations, II) the largest catalog of H II regions with spectroscopic information extracted from the CALIFA dataset (Espinosa-Ponce et al. 2020). In Section 3, we present the details on the analysis procedures and the determination of helium abundance. Finally, Section 4, summarizes the discussion and conclusions of this work.

2 SAMPLE

The current study is based on two different datasets. The first one is a compilation of all the available H II regions with helium abundance in the literature. The second one comprises the largest catalog of H II regions with spectroscopic information extracted from the CALIFA dataset (Espinosa-Ponce et al. 2020). In Sec 2.1 we describe the compilation from the literature while in Sec 2.2 we describe the CALIFA survey, and in Sec 2.2 we describe the H II regions catalog derived from the CALIFA survey.

2.1 Compilation of helium abundances in H II regions

We perform a search in the literature of H II regions in order to compile a sample of them with a good estimation of helium abundances. We search those targets in previous works with estimations of helium abundances using the He I lines in the optical range that include \( \lambda 3889, \lambda 4026, \lambda 4471, \lambda 4713, \lambda 4922, \lambda 5876, \lambda 6677 \). The final compilation comprises 174 H II regions from 8 different works in the literature. Table 1 shows the references from these studies. We summarize here the main properties of those samples.

Izotov et al. (1997) observed high-quality spectra of 45 supergiant H II regions in low-metallicity blue compact galaxies (BCGs) with the aim to compute the primordial helium abundance. The observations were made with the Ritchey-Chrétien RC2 spectrograph at the Kitt Peak 4m Mayall Telescope. The spectral range is 3500 – 7500Å and a spectral resolution of ~ 5Å. They computed ionized helium abundance in each H II region through the helium emission-lines: \( \lambda 4471, \lambda 5876, \lambda 6677 \), as obtained by a self-consistent procedure. They used two different sets of theoretical recombination He I line emissivities: one by Brocklehurst (1971) and another one by Smits (1996). To obtain the total helium abundance, they estimated the radiation softness parameter \((\eta)\), to correct for neutral helium (Vilchez & Pagel 1988).

Deharveng et al. (2000) analyzed spectra of 34 Galactic H II regions. The observations were made with ESOP, Fabry-Perot spectrophotometer at the 1.5m telescope of the Observatorio Astronómico Nacional at San Pedro Mártir. The spectral range is 3400 – 7000Å with a spectral resolution of ~ 1Å. To calculate ionized the helium abundance they used the He I \( \lambda 5876/\lambda 6677 \) ratio, corrected for extinction and they assumed a two temperature H II region model. They used the \( \lambda 5876 \) emissivity from Storey & Hummer (1995) and the He I \( \lambda 5876 \) emissivity from Benjamin et al. (1999) (which include the effect of collisional excitation from the 2s levels, and thus eliminate the need to correct for collisional enhancement of the He I lines). Through the \( O^+/O \) ratio, they observed that most of the sample has values \( O^+/O < 0.4 \), and therefore contain neutral helium. When the ratio value is \( 0.70 < O^+/O < 0.85 \), they assumed that He\( ^+/H^+ = He/H \). Finally, they concluded that an 06.5 star may ionize an H II region that contains a significant fraction of neutral helium.

Peimbert (2003); Peimbert et al. (2005, 2012); Valardi et al. (2019, 2021) studied small samples (from 1 to 5) of H II regions in comparison with the previous ones. Their main goal was to estimate the primordial helium abundance. The spectra analyzed in these works have been observed with the Ultraviolet Visual Echelle Spectrograph (UVES) at the 8m VLT Kueyen Telescope. The spectral range was 3100 – 10360Å and the FWHM resolution was \( \Delta \lambda ~ 1/8800 \). Data were observed at the VLT Telescope using Spectrograph FORS1. Three grism settings were used: GRIS-600B+12, GRIS-600R+14 with filter GG435, and GRIS-300V in addition to the filter GG375, the spectral range is 3450 – 5900Å with a spectral resolution of \( \Delta \lambda/\Delta \lambda ~ 1300, 5350 – 7450Å \) with a spectral resolution of \( \Delta \lambda/\Delta \lambda ~ 1700, \) and 3850 – 8800Å with a spectral resolution of \( \Delta \lambda/\Delta \lambda ~ 700, \) respectively. To derive the He\( ^+/H^+ \) value of a region, they used the following He I lines in the optical range: \( \lambda 3889, \lambda 4026, \lambda 4387, \lambda 4471, \lambda 4713, \lambda 4922, \lambda 5876, \lambda 6677, \) and \( \lambda 7065 \). In order to optimize the data, they used the HeI\( ^+ \) code (Peimbert et al. 2012), which is an extension of the maximum likelihood method to search for the physical and chemical conditions (\( T_e, t^*, \lambda 3889, \) and He\( ^+/H^+ \)) that would give them the best simultaneous fit to all the measured lines. They used recombination coefficients were those by Storey & Hummer (1995) for H and those by Smits
Table 1. Bibliographic references to the original works for the compiled helium abundance sample.

| Reference                          | Number of H II regions |
|------------------------------------|------------------------|
| Izotov et al. (1997)               | 58 \(^a\)              |
| Deharveng et al. (2000)            | 64 \(^b\)              |
| Peimbert (2003); Peimbert et al. (2005, 2012) | 13 \(^a\)          |
| Valerdi et al. (2019, 2021)        | 8 \(^a\)               |
| Méndez-Delgado et al. (2020)      | 15 \(^b\)              |
| Aver et al. (2020)                 | 16 \(^a\)              |
| Total                              | 174                    |

\(^a\) Extragalactic H II regions  
\(^b\) Galactic H II regions

(1996) and Benjamin et al. (1999) for He. Using the results from Sawrey & Berrington (1993) and Kingdon & Ferland (1995), they estimated the collisional contribution, and using the computations by Benjamin et al. (2002) they estimated the optical depth effects. To estimate the helium total abundance, they used different ICF(He\(^{+}\)) expressions to correct for neutral helium.

Méndez-Delgado et al. (2020) determined the radial abundance gradient of helium in the Milky Way disc from spectra of 19 H II regions and ring nebulae surrounding massive O-type stars. The data were compiled from previous works of the same group. The spectra were obtained with different telescopes and instruments: (i) The UVES spectrograph at the VLT; the spectral range was 3100 – 10400 Å and the resolution at a given wavelength is given by \(\Delta \lambda \sim \lambda/8800\), that corresponds to an average FWHM ~ 0.265 Å; (ii) The Optical System for Imaging and low-Intermediate-Resolution Integrated Spectroscopy (OSIRIS) spectrograph at the 10.4m Gran Telescopio Canarias (GTC); two grism were used: R1000B and R2500V, the spectral range is 3640 – 7870 Å with a spectral resolution of 6.52 Å, and 4430 – 6070 Å with a spectral resolution of 2.46 Å, respectively; and finally (iii) the Magellan Echellette (MagE) spectrograph at the 6.5m Clay Telescope; the spectral range is 3100 – 10000 Å and a spectral resolution of \(\lambda/\Delta \lambda \sim 4000\). They determined the He\(^+\)/H\(^+\) abundance with the code PHOEN (Luridiana et al. 2015) using all or several of the following list of He lines: \(\lambda 3364, \lambda 3889, \lambda 3965, \lambda 4026, \lambda 4421, \lambda 4438, \lambda 4471, \lambda 4713, \lambda 4922, \lambda 5016, \lambda 5048, \lambda 5087, \lambda 6678, \lambda 7065, \lambda 7281, \lambda 9464\). The recombination coefficients used were those by Porter et al. (2012, 2013) for He lines and Storey & Hummer (1995) for H lines. To obtain the total helium abundance, they used four different ionization correction factor (ICF; He) schemes.

Like in the previous explorations Aver et al. (2020) were interested in the study of the primordial helium. They studied 16 low-metallicity extragalactic H II regions to lower limit the systematic uncertainties in helium abundance determinations. The sample contain 15 objects from Aver et al. (2015) and they added the brightest H II region in the extremely metal-poor (\(\sim 3\% Z_{\odot}\)) galaxy Leo P observations. Last one were taken with the optical & near-IR spectrum, from the UV atmospheric cutoff to 1μm. They used the LBT’s Multi-Object Double Spectrograph LBT/MODS, and the LBT Utility Camera in the Infrared LBT/LUCI to obtain a near-IR spectrum, from 0.95 to 1.35 μm (to measurement He I 110830 emission line).

2.2 Catalog of H II regions on the CALIFA survey

The CALIFA Survey (Calar Alto Legacy Integral Field Area Survey) is an astronomical project to map galaxies with integral field spectroscopy (IFS), to allow detailed studies of the spatially resolved spectroscopic properties of these objects. The data was taken using the PPAK Integral Field Unit (IFU) (Kelz et al. 2006), of the PMAS spectrograph (Roth et al. 2005) mounted at the Calar Alto 3.5m telescope, with a hexagonal field-of-view of 74arcsecx64arcsec, and a 100% covering factor by adopting a three pointing dithering scheme. The optical wavelength range is covered from 3700 to 7000Å(Sánchez et al. 2012). CALIFA is the survey that samples galaxies with the largest independent number of spatial elements for the largest field of view, and allows us to study a statistically significant sample of galaxies, representative of the population at the nearby Universe (e.g. Espinosa-Ponce et al. 2020; Sánchez et al. 2016c; Sánchez 2020).

CALIFA provides IFS data for a sample of galaxies in the local Universe (0.005 < z < 0.03). These galaxies were originally selected from the Sloan Digital Sky Survey imaging survey to have a similar projected size, covering any morphological type. The survey has been designed to allow to build two-dimensional maps of the following quantities: I) stellar populations: ages, metallicities and star formation histories; II) ionized gas: two-dimensional distribution of the flux and EW for each emission line, and chemical abundances; and III) kinematic properties: both from stellar and ionized gas component. The analyzed data fulfill the expectations of the original observing proposal. The process of data reduction and analysis is provided by (Sánchez et al. 2012, 2016c).

Along this article we use the catalog of spectral properties of H II regions obtained by Espinosa-Ponce et al. (2020). This catalog is based on the integral field spectroscopy (IFS) data of the extended CALIFA sample comprising 941 galaxies. It includes galaxies observed with the same setup extracted either from the original sample and a subset of extended sub-samples comprising galaxies under-represented in the original sample (e.g., large elliptical galaxies, dwarf galaxies, mergers, etc. Barrera-Ballesteros et al. 2015) that fulfill most of the selection criteria (in particular the diameter selected), as described in Sánchez et al. (2016c). The largest of these sub-samples (~ 150 galaxies) corresponds to the PISCO survey (Galbany et al. 2018), that include all SN-hosts within the footprint of the CALIFA selection criteria. The Espinosa-Ponce et al. (2020) catalog of spectroscopic properties is the largest catalog of H II regions. The detection of H II regions was based on: (i) their shape (clumpy/peaked) structures in the H\(\alpha\) emission maps extracted from the datacubes; (ii) the EW of H\(\alpha\) (> 6Å); (iii) the presence of a fraction of ionizing young stellar populations larger than a 4% (in the V-band).

To derive these parameters (H\(\alpha\) fluxes and EW, and stellar population properties), the CALIFA data were analyzed using the Pipe3D pipeline. This pipeline has been extensively used and described in previous articles (Sánchez et al. 2016a,b). As a brief summary, it performs an automatic decomposition of the stellar population (using a multi Single Stellar Population modelling), and the emission lines (using both a Gaussian fitting and a momentum analysis), for all the spectra corresponding to each position within the IFS datacubes. For details on the fitting procedure, the currently adopted SSP library, the full set of emission lines analyzed, and the low and high order data products provided we refer the reader to Pipe3D articles Sánchez et al. (2016a,b). The final data products of Pipe3D are a set of maps comprising the spatial distribution of the different derived properties. From these maps, Espinosa-Ponce et al. (2020) extracted the H II region catalog that is the basis of this study. This catalog provides information of 26,408 individual regions including (i) the flux intensities and equivalent widths of 51 emission lines covering the wavelength range between 3745 and 7200Å corrected by dust with the extinction law by Cardelli et al.
and the line with the best S/N is the \(\lambda 5876\) line of the spectra corresponding to three \(\text{H} II\) regions with \(\text{O/H}\) between 7.16 – 8.93 in units of \(12 + \log_{10}(\text{O}/\text{H})\).

3.2 Helium abundance \(\text{He}^+/\text{H}^+\) in CALIFA

The set of emission lines fluxes included in the \(\text{H} II\) regions catalog by Espinosa-Ponce et al. (2020) comprises the following \(\text{He} I\) emission lines \(\lambda 3819, \lambda 4026, \lambda 4471, \lambda 4713, \lambda 4922, \lambda 5876,\) and \(\lambda 6678\). Prior to any analysis, we explore the signal-to-noise ratio of these seven \(\text{He} I\) emission lines in each cataloged \(\text{H} II\) region, and the line with the best S/N is the \(\text{He} I\) line \(\lambda 5876\) (we only use \(\lambda 5876\) line to develop this work). In Figure 1, we show an example of the spectra corresponding to three \(\text{H} II\) regions extracted from the NGC6090, NGC3395, and NGC3553 galaxies. These correspond to those spectra with the best S/N in the \(\text{He} I\) line \(\lambda 5876\) (red vertical line) from the Espinosa-Ponce et al. (2020) catalog. We highlight with vertical lines the seven \(\text{He} I\) recombination lines contain in the CALIFA catalog. Our final sample contain those \(\text{H} II\) regions where \(\text{He} I\) line \(\lambda 5876\) has S/N> 5, corresponding to 5386 \(\text{H} II\) regions extracted from 465 galaxies.

To calculate the helium abundance, several codes have been designed, including PYNEB (Luridiana et al. 2015) and Helio14 (Peimbert et al. 2012). In a recent work, Méndez-Delgado et al. (2020) present the comparison between the helium abundances obtained in both codes for the same set of \(\text{H} I\) lines. The general deviation between two methods is less than 0.008 dex. In this work, we use PYNEB, an update and expansion code of the IRAF package NEBULAR, which is designed to be more user-friendly and powerful than its predecessors.

We use the version 1.1.13 of PYNEB. The code analyses emission lines from gaseous nebulae and solves the equilibrium equations for a n-level atom. The results depend crucially on the input for the atomic data for each element (Luridiana et al. 2015). For recombination lines, PYNEB computes the emissivity of a given line by either interpolating in tables or using a fitting function. We use the effective recombination coefficient computations by Storey & Hummer (1995) for \(\text{H} I\) lines and Porter et al. (2012, 2013) for \(\text{He} I\) lines, which include collisional effects.

To determine helium abundances, PYNEB uses as an input the electron temperature, electron density, and the intensities of recombination helium lines. Due to the inherent weakness of the emission lines required to estimate the electron temperature, it is not possible to derive this parameter. Therefore, we create a grid in the space defined by these parameters at the specific values \(T_e = 10,000, 12,000, 13,000, 15,000,\) and \(20,000\ K\) correspond to the typical \(T_e\) for \(\text{H} II\) regions (e.g. Peimbert et al. 2012; Izotov et al. 2014). On the other hand, the electron density \(n_e\) \((10 – 250\ \text{cm}^{-3})\) is calculated using the ratio \([\text{S} II] \lambda 6717/\lambda 6731\). To compute the helium abundance for each \(\text{H} II\) regions, we use a combination of temperatures range and adopted as the final value the average of all the estimations. Helium

---

1 http://research.iac.es/proyecto/PyNeb
abundance is generally determined by weighting the results from different He I lines with the brightest line measured has the greatest weight. At visible wavelengths, the brightest He I recombination line is one of the fine structure He I, called "Balmer-α of helium". It corresponds to $^3P \rightarrow ^3P (n = 3 \rightarrow 2)$ transition at 5876Å (Benjamin et al. 1999). For this reason, we decided to use only this line to estimate the final singly ionized helium abundance. In addition, this line has a high probability of being measured in other surveys and therefore can be used to estimate the helium abundance broadly.

We use a MC simulation to estimate the uncertainty of the He$^+$/H$^+$ ratios. We generate 5000 random values for each diagnostic line assuming a Gaussian distribution with a standard deviation equal to the associated uncertainty of the line intensity involved in the diagnostic. In addition, due to the nature of our sample, we incorporate the error due to our inability to constrain $T_e$. As we mentioned before, it is not possible to calculate the electron temperature with the current data. This introduces a systematic uncertainty in our helium abundance estimates due our lack of knowledge of the temperature of each H II regions. Considering a range between 10,000 and 20,000 K in our MC simulation we estimate that this systematic uncertainty is of the order ~ 0.06 dex. We considered this uncertainty in any further calculation.

### 3.3 Helium calibrator

We use the sample describe in the Sec. § 3.1 to provide an empirical calibration of the singly ionized helium abundance. In Figure 2 we show helium abundance $12 + \log_{10}(\text{He}^+/\text{H}^+)$ versus the He I emission line flux $\lambda 5876$ for the collection of 95 extragalactic H II regions (blue circles), and 79 Galactic H II regions (blue triangles) with singly ionized helium abundance. It is known that the helium abundance and the He I emission line flux should present a linear dependence when the rest of the physical parameters of the H II are similar (electron temperature, electron density, the shape of the nebulae). This is indeed the relation predicted by codes like PTNEB (Luridiana et al. 2015). Therefore, to obtain our calibrator, we perform a least-squares linear regression to characterize the relation between the two parameters (dark blue line), corresponding to the interval from $-2.14$ to $-1.25$ dex for $\log_{10}(\text{HeI5876}/\text{Hz})$ (vertical blue dashed line). For the current study we perform the largest compilation of public data taking into account the errors introduced by the different observational effects (slit apertures, H II selection in the different studies, differences in the ionization structures, etc.). All these effects introduce possible systematic offsets between the different datasets. However we do not expect them to strongly bias or affect the final derived relation. On the other hand, we show the singly ionized helium abundance obtained from our sample (5386 H II regions, gray circles) from CALIFA data, using only the procedure described in the text (Sec. 3.2). The helium calibration is plotted with a dark blue solid line with its applicability interval from $-2.47$ to $-0.23$ dex. In blue shadow we show the standard deviation for the range between $-2.14$ to $-1.25$ dex (vertical blue dashed line). In gray shadow we show the standard deviation for the range between $-2.47$ to $-2.14$ dex and $-1.25$ to $-0.23$ dex (vertical gray dashed-dotted line). Finally, we show on right lower the average random error associated with the computation of the singly helium abundance for different abundance bins. The vertical histogram represents the distribution of helium abundance for both samples, blue for literature data, and gray for CALIFA data (the number of H II regions is presented in log scale).

The calibrator error is obtained from the difference between the best fit in the literature (calibrator) versus the best fit of the abundances obtained in H II regions of CALIFA. This way, the final error covers both samples, and this generates a systematic effect as the literature data do not cover the same dynamical range of parameters. The following equation describes the calibrator:

$$12 + \log_{10}(\text{He}^+/\text{H}^+) = x \cdot (0.891 \pm 0.039) + (12.161 \pm 0.055),$$  

where $x$ is the flux ratio $\log_{10}(\text{HeI5876}/\text{Hz})$. We report the nominal errors of the coefficients are independent of the range of $x$. However, we have to add a systematic error that is indeed different for different ranges of the $x$ value: i) for the range covered by literature data, (i.e. when $x$ is between $-2.14$ and $-1.25$ dex), the statistical (total) error is $\pm 0.047$ ($\pm 0.083$) dex (blue shadow); ii) for values outside this range (from $-2.47$ to $-2.14$ dex or $-1.25$ to $-0.23$ dex), the statistical (total) error that must be adopted is $\pm 0.053$ ($\pm 0.086$) dex (gray shadow). Moreover, we have made a test where we clear outliers from literature (the blue triangles described in the previous paragraph), and we do not observe a significant difference in the fit.

---

2 Calibrator without outliers $12 + \log_{10}(\text{He}^+/\text{H}^+) = x \cdot (0.876 \pm 0.056) + (12.139 \pm 0.079)$. 

---
It is observed that some of the CALIFA H II regions present a helium abundance near or even larger than 12.0 dex. These values are physically not feasible, since it would indicate that helium is more abundant than hydrogen. Due to the systematic and statistical errors, it is possible that the measured values reach that nonphysical regime if their number is low. In Figure 2, we include a histogram comparing the distributions of the helium abundances for both literature and CALIFA subsamples. As expected, literature data are well peaked at \( \sim 10.75 \) dex, with a tail towards lower values. The CALIFA data present a peak at similar He abundances, but with a wider range of covered values and larger tails. This is a combination of the wider range of H II region types covered by CALIFA data and the larger systematic and statistical errors. The fraction of regions with high He abundance is very low. Indeed, less than a 0.5% of them have a value larger than 11.75, and only a 0.03% present unphysical values. We consider that given the described uncertainties this fraction is not abnormal.

4 DISCUSSION

In this study, we explore two different datasets, one conformed by a compilation of literature H II regions with measured He abundances, and another one extracted from the catalog of H II regions derived from the IFS data provided by the CALIFA data. This is so far the largest extragalactic H II regions spectroscopic database. Using these data we provide for the first time an empirical calibration for the singly ionized helium abundance describe in Sec. § 3.3.

To further verify the calibrator, we explore the variations of the helium abundance obtained from our calibrator with respect to the oxygen abundance in both compilations. The oxygen abundances are obtained using different methods; the direct method for the literature data and the strong lines method using the Ho (2019) calibrator for CALIFA data. This strong-line calibrator was derived using an state-of-the-art neural network modelling using several line ratios and anchoring the oxygen abundance to a catalog of H II regions with values estimated using the direct method.

In Figure 3, we show the oxygen versus helium abundance distribution. For literature data, we show the abundances for each individual Galactic (blue triangles) and extragalactic (blue circles) H II regions. For the CALIFA data, we present the abundances for each individual H II regions (gray circles), the average abundances for each galaxy (pink circles), and the average helium abundance from the integrated data within 10 regular bins in oxygen abundance of 0.13 dex width (white circles). We observe that the H II regions of CALIFA (gray data-points) cover a broader range in oxygen and helium abundances than literature data. CALIFA data covers a metallicity range of 7.51–9.69 for 12+\( \log_{10}(O/H) \), while literature H II regions are located between 7.16 – 8.93. As mentioned above (see § 2.1), most H II regions in literature data have been used to determine the primordial helium abundance, and therefore are selected to be metal-poor regions. This may be the reason why there is a bias in helium abundances relative to the literature data in Figure 2 too.

We analyze the median helium abundance in the CALIFA galaxies within 10 bins in oxygen abundance (white circles) in order to obtain the general behavior of helium abundance. Both data from the literature compilation and the H II regions from CALIFA show a similar trend where at low metallicity (12+\( \log_{10}(O/H) \) < 8.5), the variation of helium abundance is negligible and the scatter around the average in the distribution of helium abundance in galaxies decreases appreciably. And at high metallicity (12+\( \log_{10}(O/H) \) > 8.75), we see that there is an apparent increasing trend for both the median He abundance and the scatter around these values. However, due to the high dispersion in the data, this result, although consistent with what is expected in the models, is not totally conclusive.

To understand the observed trends, we compare them with the values for the oxygen and helium abundances predicted by the inside-out chemical evolution model derived for the Milky Way (MW) described in (hereafter CP11, Carigi & Peimbert 2011; Carigi et al. 2019). This model was built to reproduce the radial distributions of the total baryonic mass, and the oxygen abundances of the ISM for different galactocentric radii of our Galaxy. We note that the CP11 model yields a total He/H abundance instead of a He\(^+\)/H\(^+\) one. Nevertheless, it is expected that the difference between the singly ionized helium and total abundances is of the order of 0.008 dex, see Izotov et al. (1997); Peimbert et al. (2000, 2012); Valerdi et al. (2019, 2021); Aver et al. (2020). It is known that MW-like galaxies grow from the inside-out following a sequence of SFHs (Star Formation Histories) and ChEhE (Chemical Evolution Histories) at each radii which shape is similar to that galaxies of different masses (Sánchez 2020; Sánchez et al. 2020). Thus, a chemical evolution model that reproduces the observed radial gradients in the MW would also recover the differential chemical evolution of galaxies of different masses.
In Figure 3, we show the present-time oxygen and helium abundances for different galactocentric distances (red stars, from 3 to 19 kpc), derived from the above model and are in increasing order from right to left as indicated by the size of the stars. The CP11 model presents an increasing behavior for helium abundance, mainly for innermost radii corresponding to high metallicity. This is due to the high efficiency in the He production by young, massive, metal-rich stars and old metal-poor low-mass stars. At more recent times, quasi simultaneously, both types of stars enrich the ISM with great amount of helium due to the strong difference in the stellar lifetime of the young-massive stars and the old low mass stars (for more details, see Carigi & Peimbert 2008; Romano et al. 2010; Carigi et al. 2019; Kobayashi et al. 2020a). The observed trend depicts a radial negative gradient for the helium abundance. When compared to the known radial gradient of oxygen abundance (between ~0.04 and ~0.06 dex/kpc 2020) it is clear that the helium gradient is weaker, reaching a mild ~0.01 dex/kpc. This trend agrees with the main behavior observed in the average distribution traced by the binned data (white circles). This means, as indicated before, that the MW at different radii behaves like a set of different spiral galaxies of different masses. This is a direct consequence of the connection between the global and resolved trends recently reviewed in Sánchez (2020).

The increase of He/H with O/H observed for our CALIFA subsample suggests a coupled evolution of both elements. Since the O/H presents a well-known global and local relation with the stellar mass and mass surface density, it is expected that the helium follows a similar relation too. However, as the range of values covered by the Helium abundance is relatively narrower than that of oxygen, it is expected that these relations are shallower than the ones described for O/H.

Helium is not frequently used to trace chemical evolution because the range of evolution is wider in oxygen. However, the joint analysis of the two elements could help to understand the chemical enrichment process in galaxies. CALIFA sample allows us to expand the dynamical range on the exploration of the helium abundance in star-forming regions in galaxies with a larger redshift and wider metallicity ranges. With our sample covering a wide range of galaxy properties, it is possible to constrain chemical evolutionary models at kpc scales.

5 CONCLUSIONS

We propose an empirical calibration for the singly ionized helium abundance based on a compilation of 174 literature H II regions. For this purpose we use the most intense He emission line at λ5876. The reported calibrator for this sample is:

\[ 12 + \log_{10}(\text{He}^+/\text{H}^+) = x + (0.891 \pm 0.039) + (12.161 \pm 0.055), \]

where \( x \) is the flux ratio \( \log_{10}(\text{He}^+/\text{H}) \) (see Sec 3.3 for more details). On the other hand, we estimate the singly ionized helium abundance using PYNEB in the largest sample of H II regions collected from CALIFA survey (5386 H II regions in 465 galaxies), observed with the technique integral field spectroscopy. Although our calibrator is derived from the literature data, we find that it represents well the two explored samples (literature and CALIFA H II regions). We use helium abundances estimated from CALIFA data to define the calibrator validity, and estimate its associated error.

We further explore the validity of the calibrator using the distribution of oxygen versus helium abundance (obtained from our calibrator), and we note an apparent increasing behavior for helium abundance for high oxygen abundance. Our result suggests that, at least for the studied sample, the average values for each galaxy are consistent with our understanding of the chemical evolution of the Milky Way.

This is a first attempt of characterization in helium abundance for a large sample of H II regions located in galaxies representative of the nearby Universe. It would be possible to improve our study with the help of higher resolution instruments, higher quality data, and higher redshift ranges. In this way, we could be obtained a better determination of helium abundance using more than one recombination line.

ACKNOWLEDGEMENTS

We acknowledge support from the grant IA-100420 (DGAPA-PAPIIT, UNAM), PAPIIT-DGAPA-IN100519 project, and funding from the CONACYT grants CF19-39578, CB-285080 and FC-2016-01-1916. M.V. acknowledges to Andrés Sixtos for several constructive suggestions in PYNEB. Finally, we would like to acknowledge an anonymous referee for a carefully reading and many helpful suggestions.

Data Availability. We use along this article literature data that are accessible in the quoted articles (Sánchez et al. 2012, 2016a,b,c), and data from the H II regions catalog by Espinosa-Ponce et al. (2020), publicly available in the following webpage http://ifs.astrosun.unam.mx/CALIFA/HII_regions/

REFERENCES

Amayo A., Delgado-Inglada G., García-Rojas J., 2020, MNRAS, 492, 950
Arellano-Córdova K. Z., Esteban C., García-Rojas J., Méndez-Delgado J. E., 2020, MNRAS, 496, 1051
Arnett W. D., Thielemann F. K., 1985, ApJ, 295, 589
Aver E., Olive K. A., Skillman E. D., 2015, J. Cosmology Astropart. Phys., 2015, 011
Aver E., Berg D. A., Olive K. A., Pogge R. W., Salzer J. J., Skillman E. D., 2020, arXiv e-prints, p. arXiv:2010.04180
Barrera-Ballesteros J. K., et al., 2015, A&A, 582, A21
Benjamin R. A., Skillman E. D., Smiths D. P., 1999, ApJ, 514, 307
Benjamin R. A., Skillman E. D., Smiths D. P., 2002, ApJ, 569, 288
Brocklehurst M., 1971, MNRAS, 153, 471
Bromm V., Larson R. B., 2004, ARA&A, 42, 79
Cardelli J. A., Clayton G. C., Mathis J. S., 1989, ApJ, 345, 245
Carigi L., Peimbert M., 2008, Rev. Mex. Astron. Astrofis., 44, 341
Carigi L., Peimbert M., Peimbert A., 2011, Rev. Mex. Astron. Astrofis., 47, 139
Carigi L., Peimbert M., Peimbert A., 2019, ApJ, 873, 107
Delavergne L., Peña M., Caplan J., Costero R., 2000, MNRAS, 311, 329
Delgado-Inglada G., Morisset C., Statiska G., 2014a, MNRAS, 440, 536
Delgado-Inglada G., Morisset C., Statiska G., 2014b, MNRAS, 440, 536
Espinosa-Ponce C., Sánchez S. F., Morisset C., Barrera-Ballesteros J. K., Galfan L., García-Benito R., Lacerda E. A. D., Mast D., 2020, MNRAS, 494, 1622
Esteban C., Peimbert M., García-Rojas J., Ruiz M. T., Peimbert A., Rodríguez M., 2004, MNRAS, 355, 229
Galfan L., et al., 2018, ApJ, 855, 107
Ho I. T., 2019, MNRAS, 485, 3569
Izotov Y. I., Thuan T. X., Lipovetsky V. A., 1997, ApJS, 108, 1
Izotov Y. I., Thuan T. X., Statiska G., 2007, ApJ, 662, 15
Izotov Y. I., Thuan T. X., Guseva N. G., 2014, MNRAS, 445, 778
Kelz A., et al., 2006, PASP, 118, 129
Kingdon J., Ferland G. J., 1995, ApJ, 442, 714
Kobayashi C., Leung S.-C., Nomoto K., 2020a, ApJ, 895, 138
Valerdi, M. et al.

Kobayashi C., Karakas A. I., Lugano M., 2020b, ApJ, 900, 179
Kunth D., Sargent W. L. W., 1983, ApJ, 273, 81
Luridiana V., Morisset C., Shaw R. A., 2015, A&A, 573, A42
Méndez-Delgado J. E., Esteban C., García-Rojas J., Arellano-Córdova K. Z., Valerdi M., 2020, MNRAS,
Peimbert M., 1986, PASP, 98, 1057
Peimbert A., 2003, ApJ, 584, 735
Peimbert M., Serrano A., 1980, Rev. Mex. Astron. Astrofis., 5, 9
Peimbert M., Torres-Peimbert S., 1977a, MNRAS, 179, 217
Peimbert M., Torres-Peimbert S., 1977b, MNRAS, 179, 217
Peimbert M., Torres-Peimbert S., Ruiz M. T., 1992, Rev. Mex. Astron. Astrofis., 24, 155
Peimbert M., Peimbert A., Ruiz M. T., 2000, ApJ, 541, 688
Peimbert A., Peimbert M., Ruiz M. T., 2005, ApJ, 634, 1056
Peimbert M., Luridiana V., Peimbert A., 2007, ApJ, 666, 636
Peimbert A., Peña-Guerrero M. A., Peimbert M., 2012, ApJ, 753, 39
Pérez-Montero E., Hägele G. F., Contini T., Díaz Á. I., 2007, MNRAS, 381, 125
Porter R. L., Ferland G. J., Storey P. J., Detisch M. J., 2012, MNRAS, 425, L28
Porter R. L., Ferland G. J., Storey P. J., Detisch M. J., 2013, MNRAS, 433, L89
Romano D., Karakas A. I., Tosi M., Matteucci F., 2010, A&A, 522, A32
Roth M. M., et al., 2005, PASP, 117, 620
Sánchez S. F., 2020, ARA&A, 58, 99
Sánchez S. F., et al., 2012, A&A, 538, A8
Sánchez S. F., et al., 2016a, Rev. Mex. Astron. Astrofis., 52, 21
Sánchez S. F., et al., 2016b, Rev. Mex. Astron. Astrofis., 52, 171
Sánchez S. F., et al., 2016c, A&A, 594, A36
Sánchez S. F., Walcher C. J., Lopez-Coba C., Barreto-Ballesteros J. K., Mejía-Narvaez A., Espinosa-Ponce C., Camps-Fariña A., 2020, arXiv e-prints, p. arXiv:2009.00424
Sawey P. M. J., Berrington K. A., 1993, Atomic Data and Nuclear Data Tables, 55, 81
Schneider R., Salvaterra R., Ferrara A., Ciardi B., 2006, MNRAS, 369, 825
Smits D. P., 1996, MNRAS, 278, 685
Stasińska G., 1978, A&A, 66, 257
Storey P. J., Hummer D. G., 1995, MNRAS, 272, 41
Valerdi M., Peimbert A., Peimbert M., Sixtos A., 2019, ApJ, 876, 98
Valerdi M., Peimbert A., Peimbert M., 2021, arXiv e-prints, p. arXiv:2105.12260
Vilchez J. M., 1989, Ap&SS, 157, 9
Vilchez J. M., Pagel B. E. J., 1988, MNRAS, 231, 257
Weaver T. A., Woosley S. E., 1980, in Ninth Texas Symposium on Relativistic Astrophysics, pp 335–357, doi:10.1111/j.1749-6632.1980.tb15942.x
Zhang Y., Liu X. W., 2003, A&A, 404, 545

This paper has been typeset from a LaTeX file prepared by the author.