

# The O3N2 and N2 abundance indicators revisited: improved calibrations based on CALIFA and $T_e$ -based literature data<sup>★,★★</sup>

R. A. Marino<sup>1</sup>, F. F. Rosales-Ortega<sup>2,3</sup>, S. F. Sánchez<sup>4,5</sup>, A. Gil de Paz<sup>1</sup>, J. Vílchez<sup>4</sup>, D. Miralles-Caballero<sup>2</sup>, C. Kehrig<sup>4</sup>, E. Pérez-Montero<sup>4</sup>, V. Stanishev<sup>9</sup>, J. Iglesias-Páramo<sup>4,5</sup>, A. I. Díaz<sup>2</sup>, A. Castillo-Morales<sup>1</sup>, R. Kennicutt<sup>6</sup>, A. R. López-Sánchez<sup>7,8</sup>, L. Galbany<sup>9</sup>, R. García-Benito<sup>4</sup>, D. Mast<sup>4,5</sup>, J. Mendez-Abreu<sup>10,11</sup>, A. Monreal-Ibero<sup>4</sup>, B. Husemann<sup>12</sup>, C. J. Walcher<sup>12</sup>, B. García-Lorenzo<sup>10,11</sup>, J. Masegosa<sup>4</sup>, A. del Olmo Orozco<sup>4</sup>, A. M. Mourão<sup>9</sup>, B. Ziegler<sup>13</sup>, M. Mollá<sup>14</sup>, P. Papaderos<sup>15</sup>, P. Sánchez-Blázquez<sup>2</sup>, R. M. González Delgado<sup>4</sup>, J. Falcón-Barroso<sup>11,10</sup>, M. M. Roth<sup>12</sup>, G. van de Ven<sup>16</sup>, and the CALIFA team<sup>4</sup>

<sup>1</sup> CEI Campus Moncloa, UCM-UPM, Departamento de Astrofísica y CC. de la Atmósfera, Facultad de CC. Físicas, Universidad Complutense de Madrid, Avda. Complutense s/n, 28040 Madrid, Spain  
e-mail: ramarino@ucm.es

<sup>2</sup> Departamento de Física Teórica, Universidad Autónoma de Madrid, 28049 Madrid, Spain

<sup>3</sup> Instituto Nacional de Astrofísica, Óptica y Electrónica, Luis E. Erro 1, 72840 Tonantzintla, Puebla, Mexico

<sup>4</sup> Instituto de Astrofísica de Andalucía (CSIC), Camino Bajo de Huétor s/n, Aptdo. 3004, 18080 Granada, Spain

<sup>5</sup> Centro Astronómico Hispano-Alemán, Calar Alto, (CSIC-MPG), C/Jesús Durbán Remón 2-2, 04004 Almería, Spain

<sup>6</sup> University of Cambridge, Institute of Astronomy, Madingley Road, Cambridge, CB3 0HA, UK

<sup>7</sup> Australian Astronomical Observatory, PO Box 915, NSW 1670 North Ryde, Australia

<sup>8</sup> Department of Physics and Astronomy, Macquarie University, NSW 2109, Australia

<sup>9</sup> CENTRA – Instituto Superior Técnico, Av. Rovisco Pais 1, 10 49-001 Lisbon, Portugal

<sup>10</sup> Depto. Astrofísica, Universidad de La Laguna (ULL), 38206 La Laguna, Tenerife, Spain

<sup>11</sup> Instituto de Astrofísica de Canarias (IAC), 38205 La Laguna, Tenerife, Spain

<sup>12</sup> Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, 14482 Potsdam, Germany

<sup>13</sup> University of Vienna, Department of Astrophysics, Türkenschanzstrasse 17, 1180 Vienna, Austria

<sup>14</sup> CIEMAT, Departamento de Investigación Básica, Avda. Complutense 40, 28040 Madrid, Spain

<sup>15</sup> Centro de Astrofísica and Faculdade de Ciências, Universidade do Porto, rua das Estrelas, 4150-762 Porto, Portugal

<sup>16</sup> Max Planck Institute for Astronomy, Königstuhl 17, 69117 Heidelberg, Germany

Received 24 May 2013 / Accepted 19 July 2013

## ABSTRACT

The use of integral field spectroscopy is since recently allowing to measure the emission line fluxes of an increasingly large number of star-forming galaxies, both locally and at high redshift. Many studies have used these fluxes to derive the gas-phase metallicity of the galaxies by applying the so-called strong-line methods. However, the metallicity indicators that these datasets use were empirically calibrated using few direct abundance data points ( $T_e$ -based measurements). Furthermore, a precise determination of the prediction intervals of these indicators is commonly lacking in these calibrations. Such limitations might lead to systematic errors in determining the gas-phase metallicity, especially at high redshift, which might have a strong impact on our understanding of the chemical evolution of the Universe. The main goal of this study is to review the most widely used empirical oxygen calibrations, O3N2 and N2, by using new direct abundance measurements. We pay special attention to (1) the expected uncertainty of these calibrations as a function of the index value or abundance derived and (2) the presence of possible systematic offsets. This is possible thanks to the analysis of the most ambitious compilation of  $T_e$ -based H II regions to date. This new dataset compiles the  $T_e$ -based abundances of 603 H II regions extracted from the literature but also includes new measurements from the CALIFA survey. Besides providing new and improved empirical calibrations for the gas abundance, we also present a comparison between our revisited calibrations with a total of 3423 additional CALIFA H II complexes with abundances derived using the ONS calibration from the literature. The combined analysis of  $T_e$ -based and ONS abundances allows us to derive their most accurate calibration to date for both the O3N2 and N2 single-ratio indicators, in terms of all statistical significance, quality, and coverage of the parameters space. In particular, we infer that these indicators show shallower abundance dependencies and statistically significant offsets compared to others'. The O3N2 and N2 indicators can be empirically applied to derive oxygen abundances calibrations from either direct abundance determinations with random errors of 0.18 and 0.16, respectively, or from indirect ones (but based on a large amount of data), reaching an average precision of 0.08 and 0.09 dex (random) and 0.02 and 0.08 dex (systematic; compared to the direct estimations), respectively.

**Key words.** Galaxy: abundances – Galaxy: evolution – ISM: abundances – H II regions – techniques: spectroscopic

\* The compiled catalog is only available at the CDS via anonymous ftp to [cdsarc.u-strasbg.fr](ftp://cdsarc.u-strasbg.fr) (130.79.128.5) or via

<http://cdsarc.u-strasbg.fr/viz-bin/qcat?J/A+A/559/A114>

\*\* Based on observations collected at the German-Spanish Astronomical Center, Calar Alto, jointly operated by the Max-Planck-Institut für Astronomie Heidelberg and the Instituto de Astrofísica de Andalucía (CSIC).

## 1. Introduction

Studying H II regions offers a powerful tool for understanding the physical properties and chemical evolution of the interstellar medium (ISM) in galaxies. They also represent perfect laboratories for deriving chemical abundances of gaseous nebulae and stars clusters across the surface of nearby galaxies

(Osterbrock & Ferland 2006, hereafter OF06). From the study of their characteristic emission-line spectra we can get key insights not only into the amount of massive stars being formed and the rate star formation occurs but also into nucleosynthesis processes taking place in these stars and into the physical conditions of the gas surrounding them (see Shields 1990, for a review).

With regard to the gas conditions, the depth and spatial resolution of the spectroscopic observations carried out in the Local Universe have allowed study not only of the global (stars and gas) metallicity of galaxies but also of the variation across galaxies and the abundance discrepancies among different phases of the ISM and between these and the abundance of the stars. However, these measurements require determining of the electron temperature ( $T_e$ ), which is obtained from ratios of faint auroral to nebular emission line intensities, such as  $[\text{O III}]\lambda 4363/\lambda 5007$ . This is often referred to as the *direct* (or  $T_e$ -based) method (OF06). A well-known difficulty, however, arises from the inherent weakness of the auroral lines, and they become fainter as metallicity increases (as the temperature decreases, owing to the more efficient cooling via metal lines). This has led to the need to use *indirect* abundances based on relations between metallicity and the (either predicted or empirically calibrated) intensity of strong emission lines. This is particularly relevant when the issue of the chemical evolution of the Universe as a whole is addressed because only the brightest emission lines (from the brightest H II regions) are usually accessible at high redshift (Pettini et al. 2001; Kobulnicky & Phillips 2003; Lilly et al. 2003; Steidel et al. 2004).

Calibrators based on strong emission lines are therefore required in such cases. Most of these calibrators rely on comparing of the ratios between the strongest emission lines with direct estimations of the oxygen abundances (e.g., Pilyugin et al. 2012, and references therein, hereafter P12). A few of them rely alternatively on the comparison of these ratios with photoionization models (e.g., McGaugh 1991; Kewley & Dopita 2002), which had the advantage of covering any possible range of parameters but carry our limitations in knowing the actual physical conditions of the nebulae. These most widely used strong-line indicators are mainly focused on the relative intensity of the following lines:  $[\text{O II}]\lambda 3727$ ,  $H\beta$ ,  $[\text{O III}]\lambda 5007$ ,  $H\alpha$ ,  $[\text{N II}]\lambda 6583$ ,  $[\text{S II}]\lambda 6717$ , 6731. López-Sánchez et al. (2012, hereafter LS12) have recently presented a revision of the different methods, showing their main strengths and caveats and illustrating in which range/conditions they can be safely applied.

The goal of our work is to provide updated calibrations for two ( $\text{O3N2}$  and  $\text{N2}$ ) of the most widely used indicators of the oxygen abundance. These indicators are very popular at low and high redshifts for different reasons: (1)  $\text{O3N2}$  is weakly affected (compared to the  $R_{23}$  or  $[\text{N II}]\lambda 6583/[\text{O II}]\lambda 3727$  indices) by differential extinction and makes use of the strongest and most easily accessible emission lines in the rest-frame optical spectroscopy, (2)  $\text{N2}$  uses two emission lines that are very close in wavelength and are also accessible in the near-infrared at moderate-to-high redshifts (see works by Cresci et al. 2012; Queyrel et al. 2012). However,  $\text{N2}$  does not consider an ionization parameter of the gas, which may be very important in some cases, specially when dealing with integral field spectroscopy (IFS) data (López-Sánchez et al. 2011, hereafter LS11; Kehrig et al. 2013). The most popular calibration of the  $\text{O3N2}$  and  $\text{N2}$  indicators was introduced by Pettini & Pagel (2004, hereafter PP04). Their calibration combine direct derivations of the oxygen abundance for the most part with a few estimations based on photoionization models, mainly at the

high-metallicity range, where direct measurements of the metal abundance were scarce at that time. This is prone to large uncertainties owing to the well-known difference in the absolute scale of abundances between both methods (Kewley & Ellison 2008). For the sake of completeness, we also compare our calibrations with the ones proposed by Nagao et al. (2006, hereafter NAG06) and Pérez-Montero & Contini (2009, hereafter PMC09), which are based, respectively, on Sloan Digital Sky Survey data (SDSS, York et al. 2000; Strauss et al. 2002) and literature measurements.

In this work we take advantage of two complementary datasets for the recalibration of both indicators by anchoring them to direct estimations of the oxygen abundance. First, we use the recent compilations of direct abundance measurements by P12 and PMC09, together with other recently published values. We also present here, for the first time, new  $T_e$ -based measurements belonging to the CALIFA survey (16 H II regions with  $[\text{O III}]\lambda 4363$ ). This compilation led to a notorious increase in the number of regions with direct abundance measurements even at the elusive high-metallicity range. For the whole  $T_e$ -sample, and due to the heterogeneous nature of our compilation, we recalculated the electron temperatures and the oxygen abundances of the 603 H II regions using the recipe proposed by Pilyugin et al. (2010, hereafter P10) and P12 to homogenize our sample.

We then use the large catalog of extragalactic H II regions created by the CALIFA survey (see Sánchez et al. 2012b, 2013), which comprises emission-line flux measurements for thousands of H II complexes. The limits set to the CALIFA sample in terms of diameter and distance (see Sect. 2.2) prevent auroral lines from being detected except for the case of  $[\text{O III}]\lambda 4363$  and only in the nearest and most metal-poor galaxies in the sample. However, this dataset provides abundance measurements with unprecedented statistics from measuring of multiple strong emission-line fluxes and their corresponding line ratios.

Thus, our objectives are to provide i) an accurate derivation of the oxygen abundance from the single-ratio indicators  $\text{O3N2}$  and  $\text{N2}$  anchored to *direct* abundance measurements. However, due to the still somewhat limited number of regions available and poor coverage in metallicity, it will have a limited capacity to estimate the uncertainties at specific values of the aforementioned  $\text{O3N2}$  and  $\text{N2}$  indices. To partly mitigate this problem, we then carry out ii) a comparison between the new  $T_e$ -calibrations found with the best fit obtained using  $\sim 3400$  H II complexes from the CALIFA survey (Sánchez et al. 2012b) where oxygen abundances could be indirectly computed using the multiple line-ratio ONS calibration (P10) and combine the results with those from the previous  $T_e$ -based analysis. Thus, we use the large collection of CALIFA H II regions to analyze not only the most probable value of the oxygen abundance for a given value of these single-parameter indicators but also the corresponding uncertainty and in the widest range of physical parameters possible. This is particularly critical when the evolution of the metallicity of galaxies with redshift ought to be analyzed because any potential evolutionary effect should be compared with the individual uncertainties achieved at each specific redshift. In this regard, it should be emphasized that these are largely affected by the uncertainties associated with the calibration of the indices and the variation in the prediction interval with the index itself. The *indirect* derivations of the oxygen abundances analyzed in this work are  $T_e$ -anchored calibrations that make use of multiple strong-line ratios. We will make use of these indirect metallicities of the CALIFA sample in the knowledge that some offsets between them could still be present because the CALIFA sample

is mapping a larger universe of physical conditions with respect to our  $T_e$  sample.

The content of the article is distributed as follows. Sect. 2 describes the sample of H II regions, including those with  $T_e$ -based abundances in the literature and those in the CALIFA survey. In Sect. 3 we give details on the analysis procedures and our main results, while Sect. 4 summarizes the conclusions of this work.

## 2. The sample

This work is based on the largest accessible database of H II regions ever accomplished, including a compilation of 603 H II regions with accurate measurements of the electron temperature, together with 3423 H II regions provided by the CALIFA survey (Sánchez et al. 2013). As we describe below, most of the regions in the CALIFA catalog (all except 16) could only be used to understand the behavior pattern of the different single-line ratio estimators compared to *indirect* abundance measurements. On the other hand, the emission-line data from the literature (plus 16 CALIFA H II complexes) are used to determine *direct* oxygen abundances for the empirical calibrations of O3N2 and N2 indices.

### 2.1. Compilation of $T_e$ -based H II regions

We performed a comprehensive search in the literature for H II regions within spiral and irregular galaxies in order to compile our  $T_e$ -based sample. We looked for those targets in previous works that measured bright emission lines (typically [O III]  $\lambda$ 5007 and [N II]  $\lambda$ 6584) and at least one of the auroral emission lines [O III]  $\lambda$ 4363, [N II]  $\lambda$ 5755, and [S III]  $\lambda$ 6312, because we want to recalculate all the temperatures and the indices in a homogeneous way. The compilation consists of a set of 603 calibrating H II regions from 17 different works in the literature whose references are given in Table 1. In our compilation we eliminated H II regions that were found to be duplicated among different works that actually came from the same original observational dataset. For this reason we have selected only the 84 H II regions from the work of PMC09 not in common with the P12 sample.

### 2.2. The CALIFA catalog of H II regions

The Calar Alto Legacy Integral Field Area (CALIFA) project is one of the most ambitious 2D-spectroscopic surveys to date. It will provide the scientific community, with a very large number of spectra of individual H II regions in nearby galaxies. This ongoing large project will use 250 observing nights already awarded with the Centro Astronómico Hispano-Alemán (CAHA) 3.5 m telescope. CALIFA comprises a diameter ( $45'' < D_{25} < 80''$ ) selected sample of  $\sim 600$  galaxies in the Local Universe ( $0.005 < z < 0.03$ ) and addresses several fundamental issues in galactic structure and evolution of galaxies.

The CALIFA observations started in July 2010 and are performed using the Potsdam Multi Aperture Spectrograph (PMAS, Roth et al. 2005) at CAHA 3.5 m in the PPAK mode (Verheijen et al. 2004; Kelz et al. 2006). The PPAK integral field unit (IFU) has a hexagonal field-of-view (FoV) of  $74'' \times 62''$ , sufficient to cover the full optical extent of the galaxies up to two to three effective radii ( $R_{\text{eff}}$ ), on average. This is possible due to the diameter selection of the sample (Walcher et al., in prep.). PPAK is made up of 331 science fibers, 36 sky background fibers, and 15 calibration fibers with diameters of  $2.7''$ , and

**Table 1.** Bibliographic references to the original works for the compiled  $T_e$ -sample.

Reference	Number of H II regions	Auroral lines <sup>a</sup>
Berg et al. (2012)	2	[O III]
Bresolin et al. (2012)	16	[O III]
Crowther & Bibby (2009)	4	[O III]
Croxall et al. (2009)	2	[O III]
Esteban et al. (2013)	1	[O III]
García-Benito et al. (2010)	3	[O III], [S III]
Guseva et al. (2012)	3	[O III]
Hadfield & Crowther (2007)	6	[O III]
Kehrig et al. (2011)	3	[O III]
Monreal-Ibero et al. (2012)	3	[O III]
Pérez-Montero & Contini (2009)	84	[O III], [N II]
Pilyugin et al. (2012)	414	[O III], [N II], [S III]
Sanders et al. (2012)	5	[O III]
Stasińska et al. (2013)	16	[O III]
Westmoquette et al. (2013)	7	[N II]
Zahid & Bresolin (2011)	9	[O III]
Zurita & Bresolin (2012)	9	[O III], [N II], [S III]
<b>This work</b>	<b>16</b>	<b>[O III]</b>
<b>Total</b>	<b>603</b>	

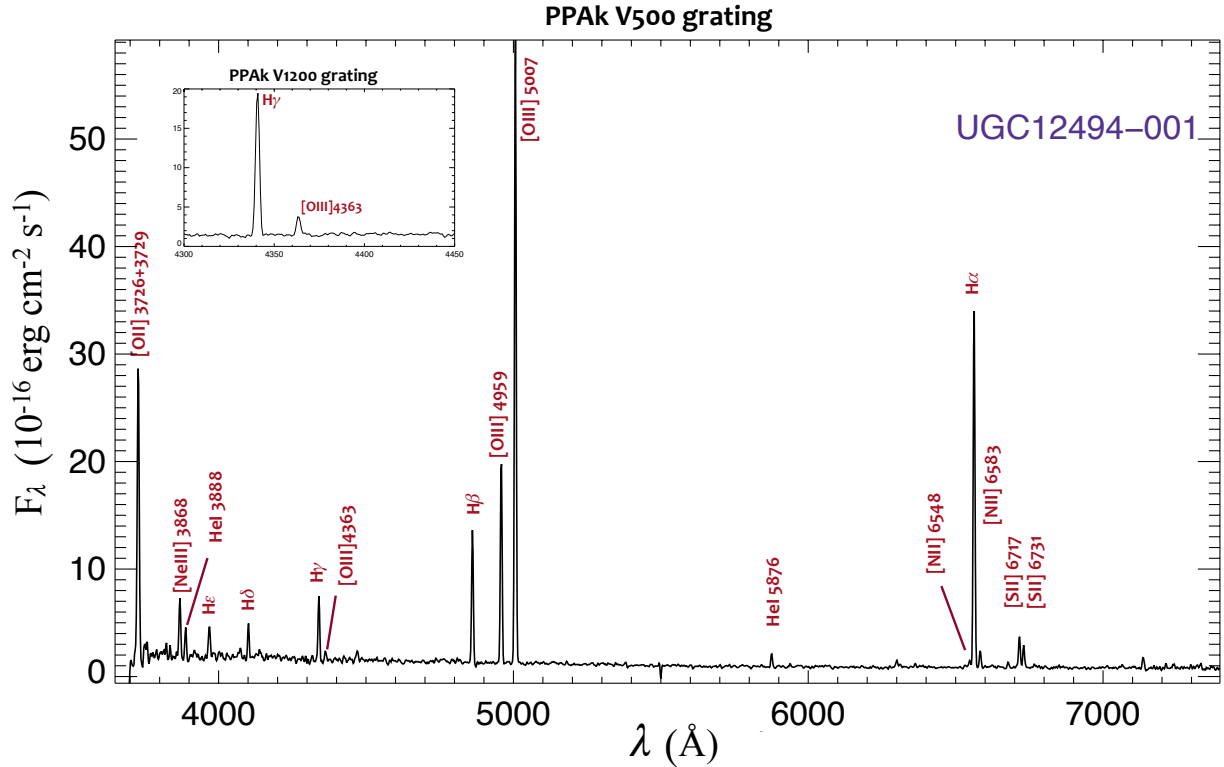
**Notes.** <sup>(a)</sup> The corresponding wavelengths of the auroral lines are [O III]  $\lambda$ 4363, [N II]  $\lambda$ 5755, [S III]  $\lambda$ 6312.

it has a 100% covering factor when a three-pointing dithering scheme is adopted. The CALIFA spectra allow us to investigate the most prominent emission lines from our H II regions covering the wavelength range of 3700–7300 Å in two overlapping setups: V500 (from 3745 Å to 7300 Å with a resolution of  $\sim 850$ ) and V1200 (from 3700 Å to 4750 Å with a resolution of  $\sim 1650$ ) as shown in Fig. 1.

The CALIFA data used for this work are based on the products generated by the CALIFA pipeline (version 1.3c, see Husemann et al. 2013). Our data fulfill the foreseen quality control requirements, with a spectrophotometric accuracy better than 10% everywhere within the considered wavelength range, both absolute and relative, and a depth that allows us to detect emission lines in individual H II regions as weak as  $\sim 10^{-17}$  erg s<sup>-1</sup> cm<sup>-2</sup> with a  $S/N \sim 3-5$ . For more details on the sample selection, observation strategy, data reduction procedures implemented in the pipeline, and data quality control, the reader is referred to Sánchez et al. (2012a). The galaxies under study have been selected from the CALIFA observed sample. Since CALIFA is an ongoing survey, whose observations are scheduled on a monthly basis (on dark nights), the list of objects increases regularly. The current results are based on the 150 galaxies observed until July 2012.

We used HIIEXPLORER<sup>1</sup> to extract the spectroscopic properties of our CALIFA catalog of H II regions sample from the V500-setup datacubes. HIIEXPLORER is an automatic H II region detection code described in Sánchez et al. (2012b) that allows us to select H II regions based on the contrast of the H $\alpha$  line intensity. This catalog has been recently used to study the mass-metallicity relation, and the possible dependence on the star formation rate (Sánchez et al. 2013). In a companion article we will describe the procedure for this particular dataset, summarizing the main properties of the ionized regions across these galaxies (Sánchez et al., in prep.). We present here just a brief summary

<sup>1</sup> [http://www.caha.es/sanchez/HII\\_explorer/](http://www.caha.es/sanchez/HII_explorer/)



**Fig. 1.** Representative PPAK V500 (*main window*) and V1200 (*inset window*) gratings optical spectra for H II region UGC 12494-001. The brightest optical emission lines of the CALIFA V500 spectral range are labeled. In the inset window a section of V1200 spectrum around the [O III]  $\lambda$  4363 line (used to compute the electron temperature) is shown. For both gratings the fluxes are in units of  $10^{-16}$  erg  $\text{cm}^{-2}$   $\text{s}^{-1}$ .

of the different steps included in the overall process. First we create a narrow-band image of 120 Å width, centered on the wavelength of H $\alpha$  shifted at the redshift of the targets. The narrow-band image is properly corrected for the contamination of the adjacent continuum image is used as input for HIIEXPLORER. The code provides a segmentation map that identifies each detected ionized region. Then, it extracts the integrated spectra corresponding to each segmented region.

A total of 4942 individual H II regions are selected from the datacubes of 150 galaxies, most of them part of the 1st CALIFA Data Release (Husemann et al. 2013), with the rest part of the whole CALIFA mother sample (Walcher et al., in prep.). This sample comprises galaxies of any type, mostly spirals (both early and late type), with and without bars, and with different inclinations. In addition to the extraction of the V500 spectra, on which the analysis of the indirect metallicity measurements from the CALIFA data is based, we have also extracted the V1200 spectra for those same regions and segmentation maps in order to improve the detectability of the weak [O III]  $\lambda$  4363 line in the case of the  $T_e$ -based abundance measurements. To understand the physical properties of the H II regions we need to investigate only the emission features in our spectra. For this reason, (i) each extracted spectrum was decontaminated by the underlying stellar continuum using the multi-SSP (simple stellar populations) fitting routines included in FIT3D software tool (Sánchez et al. 2006, 2011); (ii) then each emission line within the considered wavelength range was fitted with a Gaussian function to determine the line intensities and ratios; (iii) these line ratios were used to distinguish between different ionization conditions and to derive the oxygen abundances for each particular H II region/complex.

Finally, out of the  $\sim$ 5000 H II regions detected with HIIEXPLORER, we kept a total of 3423 regions where a blue

underlying continuum was clearly detected. This allowed us to exclude regions where the ionized-gas emission is not associated with massive star formation. Quantitatively speaking we required that at least 20% of the light at 5000 Å be arising from a population younger than 500 Myr, according to the multi-SSP spectral fitting provided by FIT3D (see more details in Cid Fernandes et al. 2013; Sánchez et al. 2013). Although this limit could also exclude low-burst-strength regions in the bulges of our galaxies, it will ensure that only ionized-gas emission induced by massive star formation is considered hereafter; i.e., it is a conservative limit.

Once the H II regions were selected, all the stellar-decontaminated spectra were analyzed and a final catalog of their spectroscopic properties was created for each galaxy. This catalog<sup>2</sup> includes the flux intensity and estimated error for the most prominent emission lines within the considered wavelength range, following the scheme presented in Appendix 1 of Sánchez et al. (2012b), and includes [O II]  $\lambda$  3727, [O III]  $\lambda$  4363, [O III]  $\lambda\lambda$  4959, 5007, H $\beta$ , [N II]  $\lambda$  6548, H $\alpha$ , [N II]  $\lambda$  6583 and [S II]  $\lambda\lambda$  6717, 6731. Finally, we also corrected the emission fluxes for interstellar reddening using a value of 2.86 for the theoretical H $\alpha$ /H $\beta$  ratio, assuming Case B with electron temperature  $\sim$ 10 000 K and density  $\sim$ 100  $\text{cm}^{-3}$  (OF06).

### 2.3. The CALIFA $T_e$ -based H II regions

As pointed out above, we also analyzed the V1200 extracted spectra for those regions where the [O III]  $\lambda$  4363 line was tentatively detected. The high spectral resolution data from CALIFA are important for improving the detection rate of weak emission

<sup>2</sup> An example of a similar catalog is described in Sánchez et al. (2012b).

**Table 2.** Dereddened emission line ratios relative to H $\beta$  fluxes of the CALIFA  $T_e$  sample along with their errors.

ID	[O II] $\lambda$ 3726,29	[O III] $\lambda$ 4363	[O III] $\lambda$ 5007	[N II] $\lambda$ 6583	$F_{H\beta}$	$A_V$
MCG-01-54-016-001	3.83 $\pm$ 0.06	0.041 $\pm$ 0.006	3.11 $\pm$ 0.05	0.12 $\pm$ 0.02	139.78 $\pm$ 0.34	0.63
NGC 3991-001	2.88 $\pm$ 0.12	0.018 $\pm$ 0.009	2.73 $\pm$ 0.11	0.24 $\pm$ 0.04	1859.3 $\pm$ 2.9	0.58
NGC 3991-002	3.98 $\pm$ 0.11	0.016 $\pm$ 0.007	2.18 $\pm$ 0.08	0.30 $\pm$ 0.04	1019.1 $\pm$ 1.7	0.68
NGC 3991-007	5.18 $\pm$ 0.07	0.013 $\pm$ 0.004	1.77 $\pm$ 0.04	0.37 $\pm$ 0.03	233.58 $\pm$ 0.46	1.30
NGC 7489-002	3.35 $\pm$ 0.09	0.043 $\pm$ 0.010	3.19 $\pm$ 0.07	0.26 $\pm$ 0.02	87.13 $\pm$ 0.37	0.88
NGC 7489-007	3.87 $\pm$ 0.13	0.080 $\pm$ 0.016	3.87 $\pm$ 0.10	0.26 $\pm$ 0.03	80.43 $\pm$ 0.47	0.98
UGC 00312-001	3.51 $\pm$ 0.04	0.015 $\pm$ 0.003	2.93 $\pm$ 0.04	0.23 $\pm$ 0.02	512.98 $\pm$ 0.53	1.36
UGC 00312-004	3.07 $\pm$ 0.04	0.016 $\pm$ 0.003	2.68 $\pm$ 0.04	0.25 $\pm$ 0.02	296.63 $\pm$ 0.39	0.91
UGC 00312-005	3.97 $\pm$ 0.04	0.018 $\pm$ 0.003	2.38 $\pm$ 0.03	0.26 $\pm$ 0.02	155.21 $\pm$ 0.24	0.90
UGC 00312-007	3.65 $\pm$ 0.04	0.018 $\pm$ 0.003	2.82 $\pm$ 0.03	0.24 $\pm$ 0.02	155.22 $\pm$ 0.24	0.99
UGC 00312-008	3.00 $\pm$ 0.05	0.034 $\pm$ 0.005	3.12 $\pm$ 0.04	0.19 $\pm$ 0.02	59.91 $\pm$ 0.17	0.60
UGC 00312-013	4.33 $\pm$ 0.06	0.033 $\pm$ 0.006	1.92 $\pm$ 0.03	0.29 $\pm$ 0.02	39.53 $\pm$ 0.13	0.91
UGC 08733-004	2.44 $\pm$ 0.05	0.021 $\pm$ 0.006	2.25 $\pm$ 0.04	0.26 $\pm$ 0.03	46.46 $\pm$ 0.16	0.68
UGC 10331-004	5.78 $\pm$ 0.07	0.019 $\pm$ 0.004	2.12 $\pm$ 0.04	0.39 $\pm$ 0.02	98.87 $\pm$ 0.26	1.79
UGC 10796-001	3.14 $\pm$ 0.04	0.032 $\pm$ 0.004	1.91 $\pm$ 0.03	0.32 $\pm$ 0.03	94.54 $\pm$ 0.18	0.84
UGC 12494-001	2.51 $\pm$ 0.11	0.064 $\pm$ 0.015	4.67 $\pm$ 0.12	0.12 $\pm$ 0.03	74.13 $\pm$ 0.47	0.25

**Notes.** Fluxes are measured in units of  $10^{-16}$  erg cm $^{-2}$  s $^{-1}$ . The observed lines are dereddened with a [Cardelli et al. \(1989\)](#) extinction law.  $A_V$  is calculated from the Balmer decrement ( $H\alpha/H\beta$ ) adopting an MW extinction law ( $A_V/A_{H\beta} = 1.164$ , [Cardelli et al. 1989](#)).

lines, such as this one, since they are strongly affected by spectral beam dilution. After imposing a cut on S/N ( $>3\sigma$ ) at the peak of this line and on the relative total line flux error ( $\leq 20\%$ ), and visually inspecting the spectra for potential spurious detections, we identified a total of 16 CALIFA H II complexes where we can reliably derive O $^{++}$  zone electron temperatures and from these, the oxygen abundance, using the formulation of P10. In Table 2 we present the most important emission line fluxes measured (relative to the H $\beta$  fluxes) for our CALIFA- $T_e$  H II regions. The optical spectra of one of these targets (UGC 12494-001) is shown in Fig. 1 where we plot the V500 grating in the main window and a section of the V1200 grating in the inset panel.

As extensively described in Sect. 1, our aim is to use the sample listed above to calibrate the most widely used indicators (O3N2, N2) versus (O/H) using both *direct* (bona fide calibrators that, despite the dramatic improvement compared to previous calibration efforts, are still somewhat limited in coverage and statistics) and *indirect* (based on multiple strong-line ratios but with extensive metallicity coverage and good statistics) methods to derive the gas metallicities. The following sections describe the details of the analysis and methods used along with our results.

### 3. Analysis and results

A number of relations have been proposed in the literature to derive metal abundances and temperatures from the metallicity-sensitive emission line ratios (e.g., [Dopita & Evans 1986](#); [Zaritsky et al. 1994](#); [Vilchez & Esteban 1996](#); [Díaz & Pérez-Montero 2000](#); [Pilyugin 2000, 2001a](#); [Pettini & Pagel 2004](#); [Tremonti et al. 2004](#); [Pilyugin & Thuan 2005](#); [Liang et al. 2006](#); [Stasińska 2006](#); [Pérez-Montero et al. 2007](#); [Thuan et al. 2010](#)). In this regard, we highlight the recent comparative study of LS12 and similar previous studies by [Pérez-Montero & Díaz \(2005\)](#), [Kewley & Ellison \(2008\)](#) and [López-Sánchez & Esteban \(2010, hereafter LSE10\)](#). Oxygen abundances anchored to  $T_e$ -based calibrations were derived using different methods for each of the two types of data described in Sects. 2.1 and 2.2. For this reason we investigated which kind of *indirect* empirical calibrators based on single line ratios could be most accurate in deriving the metallicity of H II regions when only shallow spectra (of high-redshift galaxies, for example) are available. We also

computed oxygen abundance using both *direct* ( $T_e$ -based) methods and different indirect derivations of the metallicity given in the literature and used the strength of both methods to best characterize the applicability and associated uncertainties of single strong-line ratios to infer oxygen abundances.

On one hand, the electron temperatures and the metallicity were computed from relations similar to the ones proposed by [Campbell et al. \(1986\)](#), [Pérez-Montero & Díaz \(2003\)](#) and [Pilyugin \(2007\)](#) and to the  $T_e$ -method (see, e.g., P10). In particular we recalculated the oxygen abundance using the equations proposed in Sect. 2.2 of P10 and of P12 for all the targets of our  $T_e$ -sample. By combining the auroral and the nebular lines we were able to compute confident electron temperatures  $t_{3,O}$  from the [O III]  $\lambda\lambda$  4959+5007/[O III]  $\lambda$  4363 ratio,  $t_{2,N}$  from [N II]  $\lambda\lambda$  6548+6584/[N II]  $\lambda$  5755 ratio and  $t_{2,O}$  using the [O II]  $\lambda$  3727/[O II]  $\lambda\lambda$  7320+7330 ratio that lead to oxygen abundances with a typical error of 0.12 dex. The electron temperature  $t_{3,S}$  can be estimated from the [S III]  $\lambda\lambda$  9069+9532/[S III]  $\lambda$  6312 ratio by using Eq. (3) proposed in P12. We note that even in the case of robust detections of the [O II]  $\lambda\lambda$  7320, 30, doublet we did not make use of these measurements because of the uncertainty expected in the resulting oxygen abundances (larger than the  $t_2$ - $t_3$  relation, see [Kennicutt et al. 2003](#), and P12).

There are also some H II regions in our compilation for which it is possible to obtain different measurements for the electron temperature (from both [O III]  $\lambda$  4363 and [N II]  $\lambda$  5755). In these cases we determine different abundance values for each method/auroral line as done in P12. Finally, in Table 3 the new results derived for the CALIFA- $T_e$  sample are listed. The errors in the  $T_e$ -based oxygen abundances were obtained using Monte Carlo (MC) simulations taking the errors on the different line ratios involved in their determination into account. We assumed a Gaussian distribution for the errors and that the errors among the different line ratios were not correlated for the MC. This can be considered as an upper limit for the errors since some of the contributors to these errors (e.g., uncertainties in the stellar continuum subtraction) might indeed be partly correlated. As a sanity check, we also derived the electron temperature of our CALIFA- $T_e$  sample using *PyNeb* ([Luridiana et al. 2012](#)), a python tool based on the n-level atom model able to solve the equilibrium equations and obtain the physical conditions of the

**Table 3.** Temperatures and oxygen abundances derived for the CALIFA  $T_e$  sample using the [O III] $\lambda$ 4363 auroral line.

ID	$S/N_{[\text{O III}]\lambda 4363}$	$T_3$ [K]	$12+\log(\text{O}/\text{H})$
MCG-01-54-016-001	8.2	$12\,750 \pm 70$	$8.09 \pm 0.07$
NGC 3991-001	10.7	$10\,070 \pm 330$	$8.30 \pm 0.05$
NGC 3991-002	4.9	$10\,430 \pm 360$	$8.30 \pm 0.05$
NGC 3991-007	3.2	$10\,340 \pm 460$	$8.36 \pm 0.06$
NGC 7489-002	4.0	$12\,920 \pm 330$	$8.05 \pm 0.03$
NGC 7489-007	4.1	$15\,470 \pm 930$	$7.92 \pm 0.10$
UGC 00312-001	8.5	$9360 \pm 120$	$8.47 \pm 0.02$
UGC 00312-004	7.5	$9670 \pm 180$	$8.37 \pm 0.03$
UGC 00312-005	5.6	$10\,450 \pm 370$	$8.31 \pm 0.05$
UGC 00312-007	5.6	$9940 \pm 240$	$8.38 \pm 0.03$
UGC 00312-008	5.7	$11\,920 \pm 230$	$8.12 \pm 0.03$
UGC 00312-013	6.9	$14\,340 \pm 200$	$7.91 \pm 0.03$
UGC 08733-004	4.0	$11\,160 \pm 710$	$8.09 \pm 0.09$
UGC 10331-004	3.3	$10\,990 \pm 340$	$8.34 \pm 0.04$
UGC 10796-001	6.2	$14\,090 \pm 930$	$7.84 \pm 0.06$
UGC 12494-001	16.4	$12\,970 \pm 120$	$8.08 \pm 0.02$

nebulae from the emission-line ratios. Small differences ( $<7\%$  in individual measurements) are found, since they are compatible with our error estimations.

On the other hand, we compared the results obtained from different empirical calibrations, such as the ON-calibration, the ONS-calibration (P10-ON and P10-ONS), the C-method (P12), the N2O2 method (Kewley & Dopita 2002), and the method described in Peña-Guerrero et al. (2012, hereafter PG12). The details of this comparison are beyond the scope of this article and will be presented elsewhere. For the 3423 H II regions provided by CALIFA we found the P10-ONS calibrator most appropriate, because it was already anchored to oxygen abundances derived using electron temperatures. In this regard, we only summarize some important aspects of the comparison made between the different calibrations.

As a summary, the aforementioned comparison concludes that there is a good correlation between the different methods once the following considerations are taken into account: (1) the P10-ON and the C-method abundances are well correlated with P10-ONS in the entire abundance range; (2) the PG12 calibration (as the authors themselves suggest), overestimates the abundance values by  $\sim 0.2$  dex, due to their suggested treatment on the temperature and ionizing structure of the nebula and the depletion of oxygen onto dust grains; (3) the N2O2 method is appropriate for  $12 + \log(\text{O}/\text{H}) > 8.6$  but overestimates the abundance by 0.2–0.3 dex since it is based on a combination of photoionization and stellar population synthesis models. Thus, in the case of our CALIFA H II regions, we populate the abundance sequence using the P10-ONS method (hereafter CALIFA-ONS). It correlates best with the  $T_e$ -based measurements in the whole range of metallicities, showing good agreement with the other methods in the range of applicability of the latter and yields less outliers.

It is important to emphasize here that our indirect approach, based on strong lines of oxygen, nitrogen, and sulfur (for most of the abundance range covered in this article, the ONS-calibration is similar to the  $R_{23}$  calibration regarding its information content<sup>3</sup>) is particularly important because of the wide range of ionization conditions covered by the 3423 H II regions in CALIFA (in terms of the N/O relative abundances, ionization conditions

and the electron densities) compared to the regions where auroral lines are detected.

The compilation of H II regions with electron temperatures and strong emission lines provides a more straightforward calibration, but without the statistical significance, due to its more reduced coverage of the space of physical parameters, poorer statistics, and inhomogeneity compared to the CALIFA catalog. In that regard, for the 16  $T_e$ -based CALIFA regions we made use of the P10- $T_e$  method that was based on the derivation of  $t_3$  temperature from the ratio of [O III] auroral to nebular lines, as in the case of the oxygen abundance measurements included in the P12 dataset.

In the following sections we want to investigate the behavior of the widely used single-parameter indicators (O3N2 and N2) with the oxygen abundance computed both from CALIFA-ONS and  $T_e$ -based measurements.

### 3.1. O3N2 index

The O3N2 index depends on two (strong) emission line ratios and was firstly introduced by Alloin et al. (1979) as

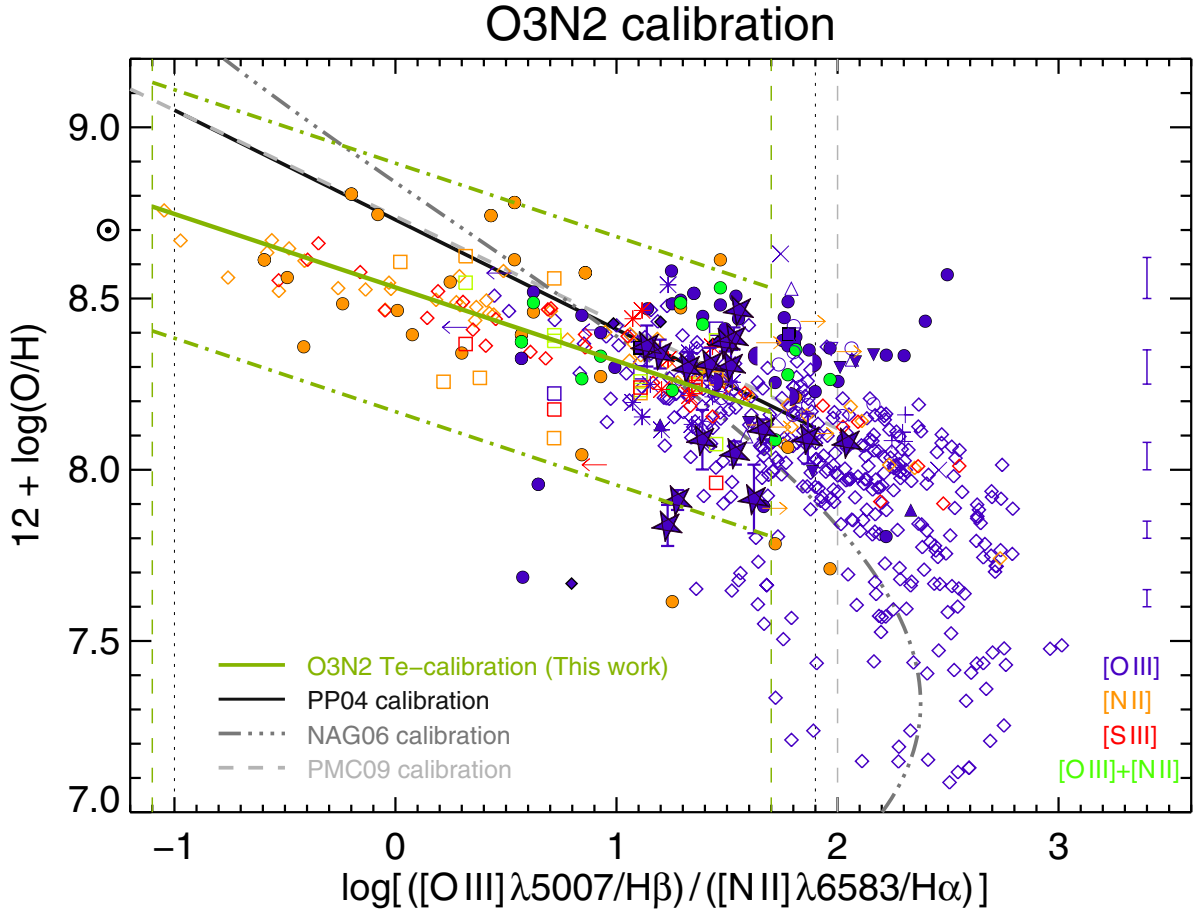
$$\text{O3N2} = \log \left( \frac{[\text{O III}]\lambda 5007}{\text{H}\beta} \times \frac{\text{H}\alpha}{[\text{N II}]\lambda 6583} \right). \quad (1)$$

Although O3N2 formally includes two line ratios in its definition, given the way the [O III] $\lambda$ 5007 and [N II] $\lambda$ 6584 fluxes are corrected for dust attenuation, this index is only sensitive to the extinction-corrected [O III] $\lambda$ 5007/[N II] $\lambda$ 6584 ratio. Two of the most popular calibrations (among many) that relate the oxygen abundance and the O3N2 index with a simple linear regression are the ones proposed by PP04 where  $12 + \log(\text{O}/\text{H}) = 8.73 - 0.32 \times \text{O3N2}$  and by PMC09 where  $12 + \log(\text{O}/\text{H}) = 8.74 - 0.31 \times \text{O3N2}$ . Here we also explore the calibration obtained by NAG06 using the third-order polynomial function presented in their Table 6. It is worth emphasizing that this oxygen abundance estimator is widely used in the high-metallicity regime up to the solar value, where the N2 index saturates. In Fig. 2 we show the variation in the oxygen abundance with the O3N2 index for 603 H II calibrating regions. The sample of  $T_e$ -based H II regions is represented (see the caption for details) and the 16 CALIFA H II regions are plotted with the respective errors. All these data span a range of metallicity  $7.0 < 12 + \log(\text{O}/\text{H}) < 8.9$ . We also plot the PP04 calibration, the PMC09 calibration, and the NAG06 one. We do a robust fit (least absolute deviation method) to our  $T_e$ -based H II regions data in the range of O3N2 between  $-1.1$  and  $1.7$  in O3N2 and we find that

$$12 + \log(\text{O}/\text{H}) = 8.533[\pm 0.012] - 0.214[\pm 0.012] \times \text{O3N2}. \quad (2)$$

The limits in the fit range is mainly come from the large dispersion of the  $T_e$ -based data in the low-metallicity regime ( $\text{O3N2} > 1.8$ ). Our new empirical calibration is thus found by using 309 (of 603)  $T_e$ -based H II regions and in Fig. 3 is shown with 95 per cent (68 per cent) of the measurements within  $\pm 0.36$  dex ( $\pm 0.18$  dex). The standard errors in the zero point and in the slope are reported. These errors were computed with  $10^5$  bootstrap repetitions for all the robust fits presented in this work. In Fig. 3 the CALIFA-ONS H II regions are drawn as a density contour plot, and also in this case we computed an independent robust fit in the range of  $-1.1 < \text{O3N2} < 1.7$  to the 3423 CALIFA-ONS H II regions that yields  $12 + \log(\text{O}/\text{H}) = 8.505[\pm 0.001] - 0.221[\pm 0.004] \times \text{O3N2}$  with  $\sigma = 0.08$ . This fit is shown in Fig. 3. The large number of regions used in deriving this calibration allow us to also compute the prediction intervals for

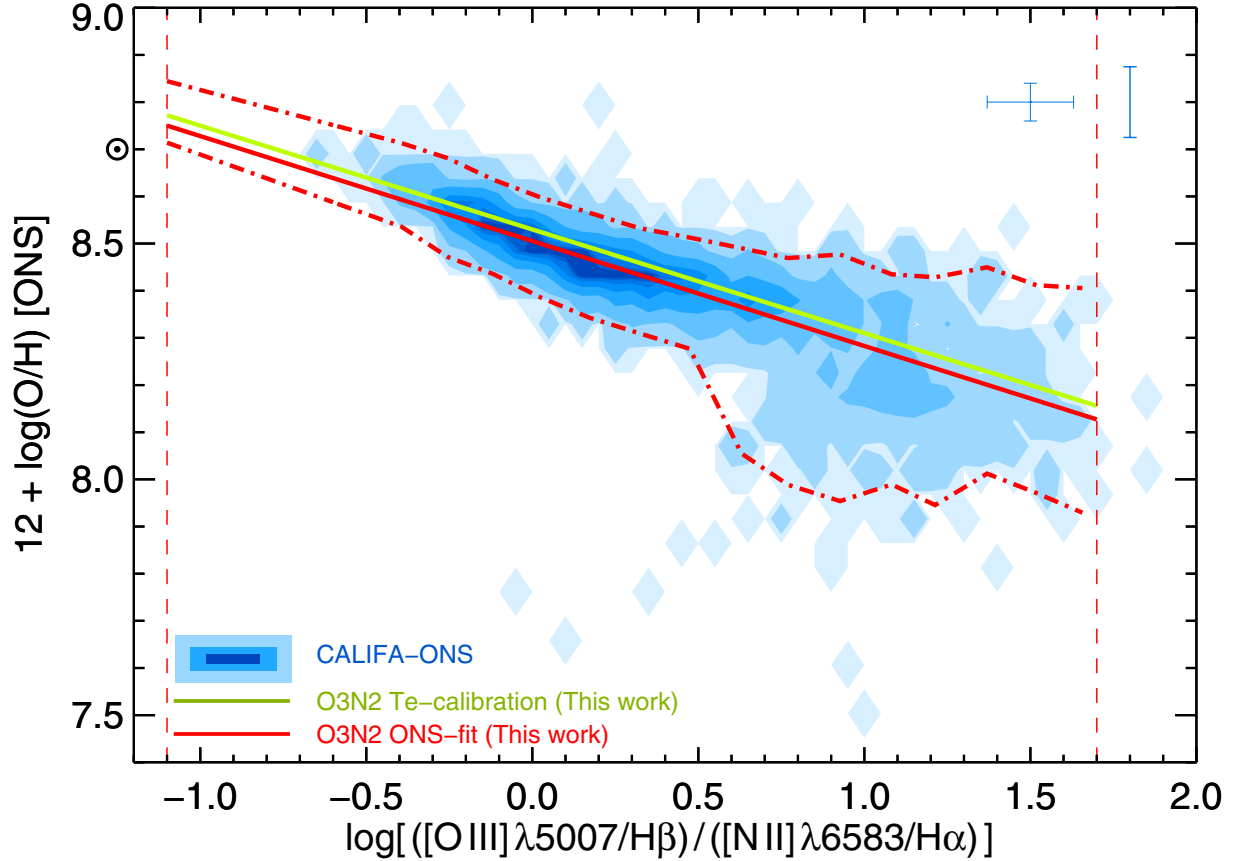
<sup>3</sup>  $R_{23} = ([\text{O II}]\lambda\lambda 3727+3729 + [\text{O III}]\lambda\lambda 4959+5007)/\text{H}\beta$ , (Pagel et al. 1979).



**Fig. 2.** Oxygen abundance versus the O3N2 index for  $T_e$ -based HII regions abundances. The  $T_e$ -based HII regions are shown using different symbols and colors depending on the work they came from and the auroral line used for computing the electron temperature. The meaning of the symbols is the following: filled squares for Berg et al. (2012) data, empty circles for Bresolin et al. (2012) data, filled stars for the new CALIFA- $T_e$  data, plus signs for Crowther & Bibby (2009) data, empty upside-down triangles for Crowther et al. (2009) data, filled triangles for Guseva et al. (2012) data, filled upside-down triangles for Hadfield & Crowther (2007) data, crosses for Kehrig et al. (2011) data, filled circles for Pérez-Montero & Contini (2009) data, empty diamonds for Pilyugin et al. (2012) data, filled diamonds for Sanders et al. (2012) data, asterisks for Stasińska et al. (2013) data, empty triangles for Zahid & Bresolin (2011) data, empty squares for Zurita & Bresolin (2012) data, left pointing arrows for García-Benito et al. (2010) data, right pointing arrows for Westmoquette et al. (2013) data, filled left semi-circles for Esteban et al. (2013) data and filled right semi-circles for Monreal-Ibero et al. (2012) data. In blue we show those regions for which the temperature  $t_{3,O}$  was computed using the [O III]  $\lambda$ 4363 line; in orange those regions for which we obtain the temperature  $t_{2,N}$  from the [N II]  $\lambda$ 5755 line; in red we represent those regions for which we are able to compute the temperature  $t_{3,S}$  using the [S III]  $\lambda$ 6312, and in green are shown those regions that have both [O III]  $\lambda$ 4363 and [N II]  $\lambda$ 5755 measurements, so we could compute the abundance using  $t_{3,O}$  and  $t_{2,N}$ . The PP04 calibration is plotted with a black solid line with its applicability interval (from  $-1$  to  $1.9$ ) shown with the vertical gray dotted lines. The PMC09 calibration is represented with a gray dashed line until its limit of validity O3N2 =  $2$ . The gray 3-dot-dashed line shows the NAG06 calibration. The new  $T_e$ -based calibration,  $12 + \log(O/H) = 8.533 - 0.214 \times \text{O3N2}$  is shown with a green solid line. The green dot-dashed lines encompass  $2\sigma$  ( $=\pm 0.36$  dex) of all measurements. The vertical dashed lines indicate the interval of our fit, from O3N2 =  $-1.1$  to  $1.7$ . We only plot in blue the errors associated to the new measurements of the CALIFA- $T_e$  HII regions. For the sake of clarity we show on the right the average random error associated to the computation of the oxygen abundances for different abundance bins.

each given O3N2 value, which are shown in the same figure ( $\pm 2\sigma$  in this case). These intervals can be used to derive the random error in  $12 + \log(O/H)$  when this is computed from the O3N2 index. Since for the ONS-based abundances we are using more line ratios than for the O3N2 calibration, these prediction intervals allow us to provide an estimate of the degeneracies in the O3N2 calibration for predicting the actual oxygen abundance. In that regard and until a larger number of precise high-metallicity  $T_e$ -based measurements are available, so the metallicity in the high regime can be derived by *direct* methods, the prediction intervals shown in Fig. 3 at O3N2 values below  $-0.2$  should be taken with caution. These high metallicities is where the [N II]  $\lambda$ 6584/ $H\alpha$  ratio saturates, reducing the difference in information content between the ONS calibration and the one given here for the O3N2 single-line ratio.

Finally, and although the ONS-P10 calibration of the oxygen abundance is anchored to  $T_e$ -based measurements (see P10), an average offset of  $\sim 0.02$  dex is still found between the two calibrations ( $T_e$ -based and ONS-based). This offset should be considered as a systematic error to be added to the random error above in cases where abundance measurements based on different indices and calibrations (e.g., from galaxies at different redshifts and/or of different data quality) are to be combined. Thanks to our unprecedented statistics, the equations given above represent the most update version for the O3N2 calibration. We thus confirm that our O3N2 index calibration is significantly more robust than PP04, NAG06, and PMC09 in the high-abundance regime. We find a good correlation between the  $T_e$ -based and the CALIFA-ONS based measurements except for a small offset that can be due in part to the still quite low number



**Fig. 3.** Oxygen abundance versus the O3N2 index for the H II complexes detected within the CALIFA galaxies. For the sake of clarity we plot the CALIFA-ONS H II regions as a blue density contour plot. Our new O3N2 calibration (see Fig. 2) is plotted with a green solid line. Red lines are used to show the CALIFA-ONS best linear fit (solid) and its  $2\sigma$  prediction intervals (dot-dashed) with its applicability interval (from  $-1.1$  to  $1.7$ , vertical dashed lines) (see text for details on these fits). In this case we obtain  $12 + \log(\text{O}/\text{H}) = 8.505 - 0.221 \times \text{O3N2}$ . The blue cross and bar in the top right corner of the plot indicate the average random errors (cross) of the CALIFA-ONS data points and the typical systematic errors (vertical bar) associated with the measurement of ionized-gas oxygen abundances.

of objects (not exactly the same as the ones used by P10) with  $T_e$  measurements in a broad range of oxygen abundances.

From an operational point of view, we recommend potential users of the O3N2 index to use the  $T_e$ -based calibration given above when oxygen abundances from different methods are to be combined. If only O3N2 data are to be analyzed (within a galaxy or potential changes from galaxy to galaxy) the calibration and prediction intervals based on the ONS calibration could be used instead, except for very low values of O3N2 ( $< -0.2$ ) where an average  $T_e$ -based uncertainty of 0.18 dex should be more reliable because in this regime the ONS abundances are heavily weighted by the value of the O3N2 index.

### 3.2. N2 index

The N2 index was studied by several groups including Storchi-Bergmann et al. (1994), van Zee et al. (1998), Raimann et al. (2000), and Denicoló et al. (2002), and it is defined as

$$N2 = \log \left( \frac{[\text{N II}]\lambda 6583}{\text{H}\alpha} \right). \quad (3)$$

Despite the saturation that the N2 suffers in the high-metallicity regime, this parameter is a very useful indicator of the oxygen abundance for different reasons: (1) it is very sensitive to the metal content of a nebulae; (2) it does not suffer from reddening correction or flux calibration issues due to the close wavelength

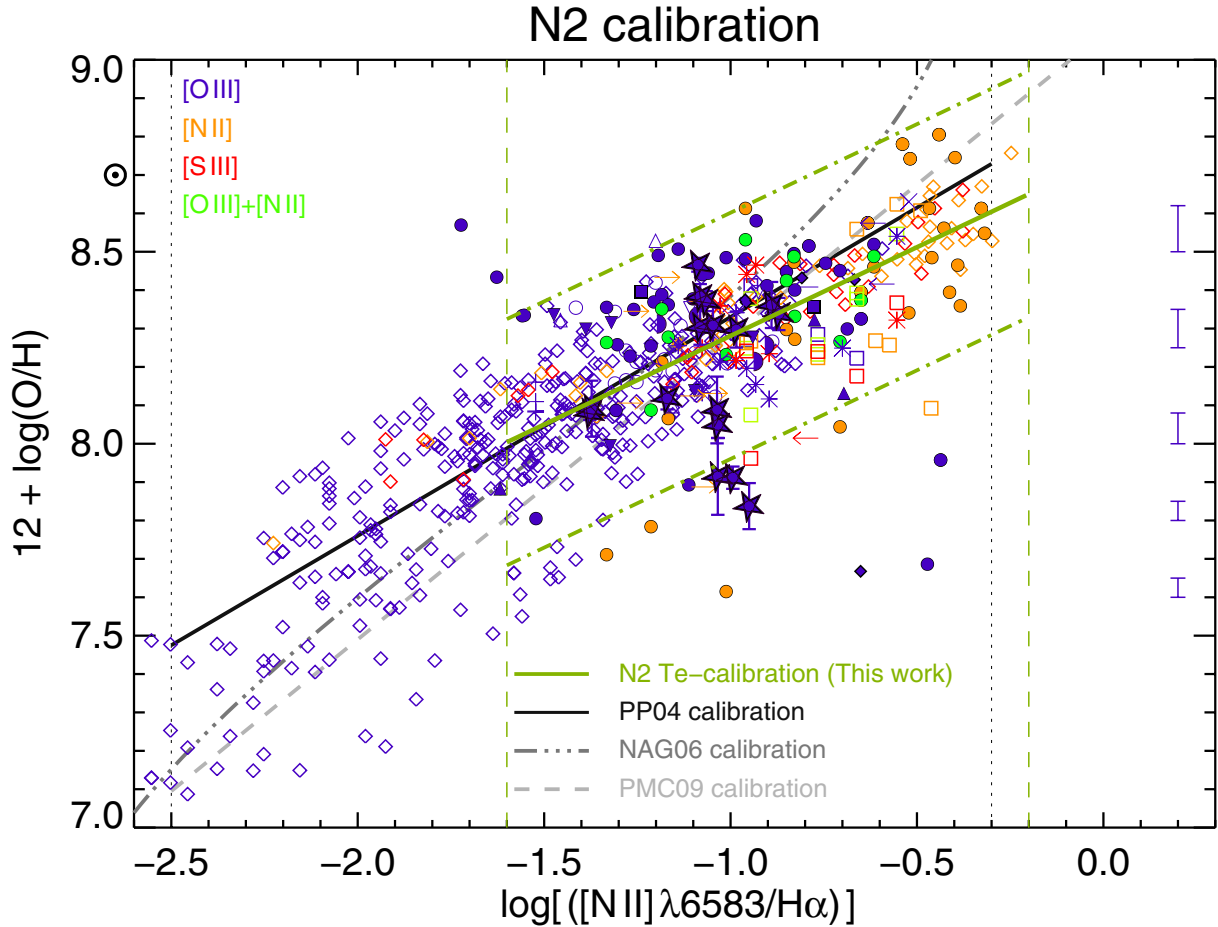
of these lines; and (3) it can be detected with new-generation near-infrared spectrographs in 8–10 m class telescopes at high redshifts. For the N2 index PP04 and PMC09 proposed the following relations:  $12 + \log(\text{O}/\text{H}) = 8.90 + 0.57 \times N2$  and  $12 + \log(\text{O}/\text{H}) = 9.07 + 0.79 \times N2$ , respectively. Our results, along with the PP04, NAG06, and PMC09 calibrations, are shown in Fig. 4. The robust regression made to 452  $T_e$ -based abundance measurements yields

$$12 + \log(\text{O}/\text{H}) = 8.743[\pm 0.027] + 0.462[\pm 0.024] \times N2, \quad (4)$$

with 95 per cent (68 per cent) of the measurements of  $\log(\text{O}/\text{H})$  found within  $\pm 0.32$  dex ( $\pm 0.16$  dex), in the interval  $-1.6 < N < -0.2$ . On the other hand, the robust fit to the CALIFA-ONS H II regions yields  $12 + \log(\text{O}/\text{H}) = 8.667[\pm 0.006] + [0.455 \pm 0.011] \times N2$  with an average rms of  $\sigma = 0.09$ .

As in the case of the O3N2 calibrations, these new calibrations provide somewhat shallower slopes than the PP04, NAG06, and PMC09 calibrations. As expected, this calibrator has a monotonic behavior with the oxygen abundance but saturates at  $\log([\text{N II}]\lambda 6583/\text{H}\alpha) \sim -0.4$ , where the piling of both  $T_e$ -based and CALIFA-ONS H II regions is evident. We also find good agreement between the trend of the  $T_e$ -based H II regions and the CALIFA-ONS H II regions, although there is a systematic offset of 0.08 dex (in the sense that  $T_e$ -based abundances are slightly larger than ONS-based ones) that should be taken into account should the ONS-based calibration be used in combination with other oxygen abundance estimates. The strength of the





**Fig. 4.** Oxygen abundance versus the N2 index for the  $T_e$ -based H II regions. The symbols and the color code are the same as for Fig. 2. The PP04 calibration is plotted with a black solid line with its applicability interval (from  $N2 = -2.5$  to  $-0.3$ ) shown with vertical gray dotted lines. The PMC09 calibration is represented with a gray dashed line, while the gray 3-dot-dashed line shows the NAG06 calibration. The new  $T_e$ -based calibration,  $12 + \log(O/H) = 8.743 + 0.462 \times N2$ , is shown with a green solid line, while the green dot-dashed lines encompass  $2\sigma$  ( $\pm 0.32$  dex) of the measurements. The vertical dashed lines indicate the interval of our fit, from  $N2 = -1.6$  to  $-0.2$ . We only plot in blue the errors associated to the new measurements of the CALIFA- $T_e$  H II regions. For the sake of clarity we show on the right the average random error associated to the computation of the oxygen abundances for different metallicity bins.

large collection of CALIFA data is clear in this case, where the prediction intervals can be safely adopted within the entire range of applicability of the N2 index, given the comparatively weak dependence of the ONS calibration on the value of the N2 line ratio itself, especially as we move to high metallicities where the latter saturates.

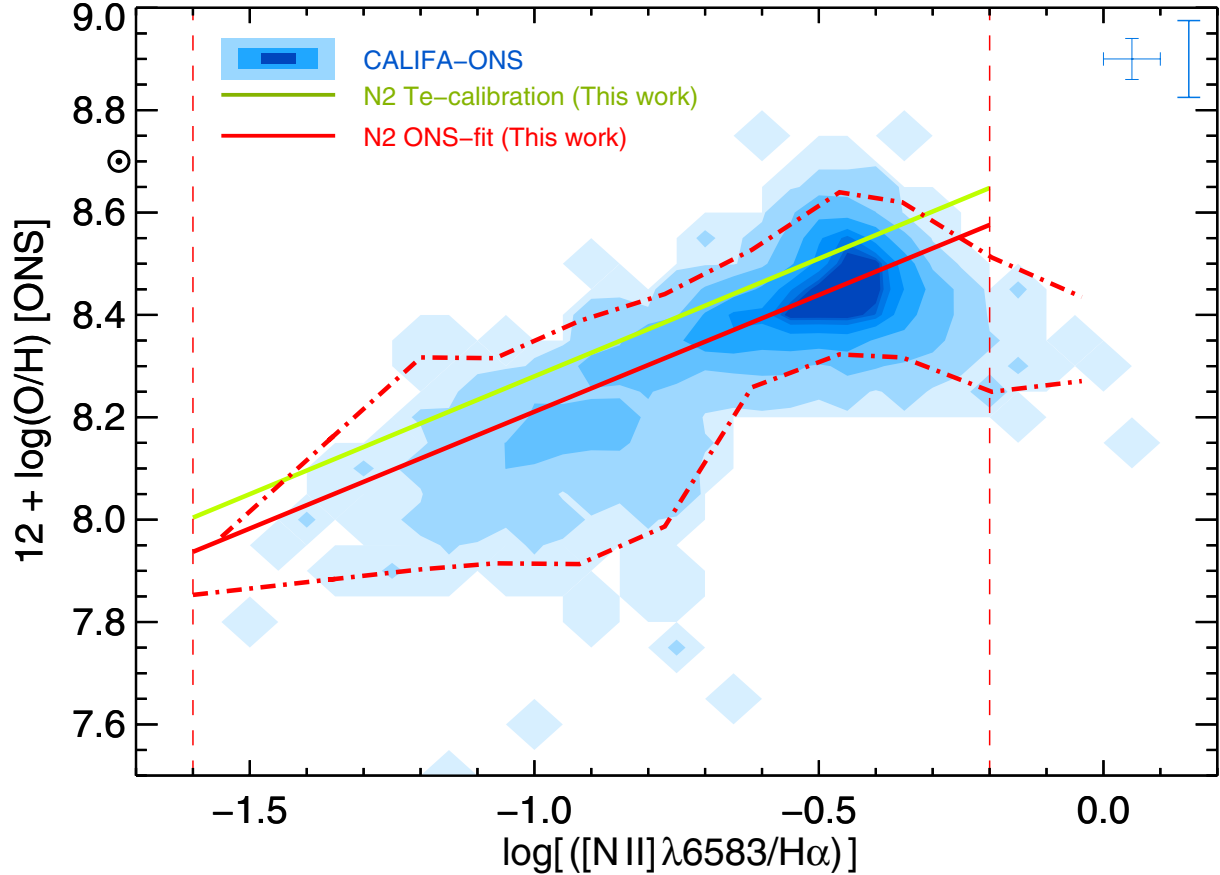
#### 4. Discussion and conclusions

In this paper we provide revisited empirical calibrations for the oxygen abundances in H II regions based on the O3N2 and N2 indicators. This work represents the most comprehensive compilation of both  $T_e$ -based and multiple strong-line (ONS-based) ionized-gas abundance measurements in external galaxies to date.

The differences found between these calibrations and those obtained by PP04, NAG06, and PMC09 recommend revising the results obtained using the latter calibration. These differences are particularly noticeable in the high-metallicity regime ( $12 + \log(O/H) > 8.2$ ), and on one hand, they are mainly due to the lack of high-quality observations of H II regions with auroral lines at the high-metallicity end at the time of the publication of the PP04 calibration, hence their justified need for including predictions from photoionization models at this regime. On

the other hand, the large differences found at high metallicity between our calibration and the one of NAG06 occur because NAG06 in this metallicity range uses a sample of galaxies (their sample C, for more details see Sect. 2.2 in Nagao et al. 2006) that have no  $T_e$ -based abundances measured. We find that a linear relation provides a good fit to the oxygen abundance as a function of the O3N2 (N2) parameters with rms values of 0.18 dex (0.16 dex). The relatively few statistics and rather heterogeneous origin of the direct abundance measurements precludes deriving reliable predicting intervals as a function of these parameters or the derived oxygen abundance. This is where the results of the analysis of large and homogeneous sets of ONS-based measurements are most useful with the caveats and limitations described in Sects. 3.1 and 3.2.

Hereafter we discuss the potential physical origin of the relations shown in Figs. 2 through 5. The ionization degree of the H II region also plays an important role when deriving a proper oxygen abundance. Many other empirical methods based on both observations (e.g. Pilyugin 2001a,b; Pilyugin & Thuan 2005) and photoionization models (e.g. McGaugh 1991; Kewley & Dopita 2002) do consider an ionization parameter – typically the  $[O III]/[O II]$  ratio – to improve the accuracy of their derived O/H values. Indeed, it has been shown that empirical calibrations that do not consider an ionization parameter



**Fig. 5.** Oxygen abundance versus the N2 index for the H II complexes within CALIFA galaxies. We plot the CALIFA-ONS H II regions as a blue density contour plot. Our new N2 calibration (see Fig. 4) is plotted with a green solid line. Red lines are used to show the CALIFA-ONS best linear fit (solid) and its  $2\sigma$  prediction intervals (dot-dashed) with its applicability interval (from  $N2 = -1.6$  to  $-0.2$ , vertical dashed lines) (see text for details on these fits). In this case we obtain  $12 + \log(O/H) = 8.667 + 0.455 \times N2$ . The blue cross and bar in the top right corner of the plot indicate the average random errors (cross) of the CALIFA-ONS data points and the typical systematic errors (vertical bar) associated with the measurement of ionized-gas oxygen abundances.

have a larger scatter (which is even larger than observational errors) than those ones that do not use it (Kobulnicky et al. 1999; López-Sánchez & Esteban 2010; Moustakas et al. 2010; López-Sánchez et al. 2012). That is particularly important when dealing with IFS data, since observations may be actually showing the ionization structure of the H II regions (e.g., Kehrig et al. 2008; López-Sánchez et al. 2011; Monreal-Ibero et al. 2010, 2011, 2012; Pérez-Montero et al. 2011; James et al. 2013a,b). This behavior is clearly seen in Fig. 13 of LS11, which compares the  $[O III]/[O II]$  ratio and oxygen abundance as provided by several empirical calibrations. Using N2 and O3N2 calibrations proposed by PP04, NAG06, and PMC09, it is clearly seen that regions with higher ionization degree tend to have lower oxygen abundances (see also Fig. 12 of PP04). Another important issue to consider is that the ionization parameter is related to the age of the most-recent star formation event. Galaxies hosting young starbursts will have higher ionization parameters than galaxies where the main star-formation event happened a long time ago. This fact introduces an extra bias to all O/H estimations based on calibrations that do not consider an ionization parameter for the gas (Stasińska 2010; López-Sánchez & Esteban 2010). In the case of objects with high N/O ratios (e.g. Pustilnik et al. 2004; López-Sánchez et al. 2007, 2011; López-Sánchez & Esteban 2010; Monreal-Ibero et al. 2010, 2012; Pérez-Montero et al. 2011; Amorín et al. 2012; Kehrig et al. 2013), the use of a  $[N II]$ -based calibration will provide O/H values that are higher

than the real ones. According to the photoionization models described in Pérez-Montero & Díaz (2005) and PMC09, all the dispersion found in the relation between both O3N2 and N2 parameters and the oxygen abundances derived from the  $T_e$  method can be explained in terms of an additional dependence on ionization parameter and on the nitrogen-to-oxygen ratio. In the case of our compilation, we checked that the grid of models presented in PMC09 covers all our observed sample by varying these two parameters.

The comparison between the  $T_e$ - and ONS-based abundances in the case of the O3N2 yields a very small difference of 0.02 dex, which can be attributed to the intrinsic dispersion in ONS for a given  $T_e$ -based abundance (see P10). In the case of N2, on the other hand, we find a systematic difference of 0.08 dex that is evident as an offset between the green and red lines of Fig. 5. This cannot be due exclusively to the limitations of the ONS abundances along but mainly to differences between either the N2 or ONS measurements of the CALIFA regions compared to those where  $T_e$ -based measurements have been derived. Since for a given O3N2 we find very good agreement between the  $T_e$ - and ONS-based measurements, it is more likely that the difference arises in the N2 measurements (only 16 CALIFA  $T_e$ -based regions are not enough to directly confirm this). One possibility for explaining the differences between the trending of the ONS-fit and our  $T_e$ -based calibration in the case of N2 is that the change in N2 (for a given O3N2 and ONS) is

related to spatial-resolution effects. The reasoning behind this is that the vast majority of the  $T_e$ -based measurements come from long-slit or multi-object spectroscopy of very nearby galaxies (typically inside the Local Volume;  $d < 11$  Mpc), where a 1 arcsec resolution yields a physical resolution better than 50 pc, while the typical physical resolution of the CALIFA data is, in the best case, four times this number. In this regard, Mast et al. (in prep.) do not find a systematic increase in the N2 index with decreasing spatial resolution using simulations based on IFS data on very nearby systems from PINGS (Rosales-Ortega et al. 2010). However, since their work is not based on a large diameter-limited sample like CALIFA, we cannot completely rule out this effect. On the other hand, we should also consider that while a significant fraction of the H II regions included in our  $T_e$ -based analysis were extracted from dwarf galaxies, these are almost absent (by design) from the CALIFA sample, which is mainly constituted of disk galaxies. Therefore, any difference (for a given oxygen abundance) in the N/O ratio or the excitation conditions of the nebulae in dwarfs with respect to disk galaxies could also result in a different N2 for the CALIFA sample even for the same O3N2 and ONS measurements. Should that be the case, future analyses of the chemical abundances in galaxies that are based solely on the N2 index should pay close attention to the sample selection since systematics in the N2-based abundances could arise if galaxies of different types are combined, for example, at different redshifts.

Our new calibration also has important implications for studies that investigate the gas metallicity evolution of galaxies up to high redshifts. For example, the mass-metallicity relation (MZR) at  $z > 2$  presented by Erb et al. (2006), which is often used as a reference sample, was based on the N2 calibration of PP04. They found an evolution in MZR of  $\sim 0.3$  dex at  $z > 2$  compared to the local SDSS sample. However, the metallicity evolution is enhanced using our new N2 calibration by  $\sim 0.1$ – $0.2$  dex. Our Eq. (4) would lead to an average decrease of 0.4 dex to  $z \sim 2$ , while using the fit to the CALIFA-ONS H II regions results in an evolution of 0.5 dex. This is more in accordance with recent determinations of metallicity in distant zCOSMOS galaxies exploiting measurements of all five strong lines (Pérez-Montero et al. 2013).

Thus, we conclude that the relations given in this paper will improve our understanding of the chemical evolution of the Universe, even when only single-parameter abundance measurements are available and as long as the predicting intervals derived here are taken into account. Possible effects associated to spatial resolution and sample selection should also be adequately accounted for.

*Acknowledgements.* We would like to thank the anonymous referee for the review of our manuscript. The comments and suggestions helped to improve the content of the paper. R.A. Marino is funded by the Spanish program of International Campus of Excellence Moncloa (CEI). This study makes use of the data provided by the Calar Alto Legacy Integral Field Area (CALIFA) survey (<http://www.califa.caha.es>). Based on observations collected at the Centro Astronómico Hispano-Alemán (CAHA) at Calar Alto, operated jointly by the Max-Planck-Institut für Astronomie and the Instituto de Astrofísica de Andalucía (CSIC). CALIFA is the first legacy survey being performed at Calar Alto. The CALIFA collaboration would like to thank the IAA-CSIC and MPIA-MPG, as major partners of the observatory, and CAHA itself, for the unique access to telescope time and support in manpower and infrastructures. The CALIFA collaboration also thanks the CAHA staff for their dedication to this project. We thank the *Viabilidad, Diseño, Acceso y Mejora* funding program, ICTS-2009-10, for supporting the initial development of this project. S.F.S., F.F.R.O., and D. Mast thank the *Plan Nacional de Investigación y Desarrollo* funding programs, AYA2012-31935 of the Spanish *Ministerio de Economía y Competitividad*, for the support given to this project. S.F.S. thanks the *Ramón y Cajal* project RyC-2011-07590 of the Spanish *Ministerio de Economía y Competitividad*, for the support giving to this project. F.F.R.O.

acknowledges the Mexican National Council for Science and Technology (CONACYT) for financial support under the program Estancias Posdoctorales y Sabáticas al Extranjero para la Consolidación de Grupos de Investigación, 2010-2012. We acknowledge financial support for the ESTALLIDOS collaboration by the Spanish Ministerio de Ciencia e Innovación under grant AYA2010-21887-C04-03. BG-L also acknowledges support from the Spanish *Ministerio de Economía y Competitividad* (MINECO) under grant AYA2012-39408-C02-02. J.F.-B. acknowledges financial support from the Ramón y Cajal Program and grant AYA2010-21322-C03-02 from the Spanish Ministry of Economy and Competitiveness (MINECO), as well as to the DAGAL network from the People's Program (Marie Curie Actions) of the European Union's Seventh Framework Program FP7/2007-2013/ under REA grant agreement number PITN-GA-2011-289313. CK has been funded by project AYA2010-21887 from the Spanish PNYA. P.P. acknowledges support by the Fundação para a Ciência e a Tecnologia (FCT) under project FCOMP-01-0124-FEDER-029170 (Reference FCT PTDC/FIS-AST/3214/2012), funded by FCT-MEC (PIDDAC) and FEDER (COMPETE). R.M.G.D. and R.G.B. also acknowledge support from the Spanish *Ministerio de Economía y Competitividad* (MINECO) under grant AyA2010-15081. V.S., L.G., and A.M.M. acknowledge financial support from the Fundação para a Ciência e a Tecnologia (FCT) under program Ciência 2008 and the research grant PTDC/CTE-AST/112582/2009.

## References

- Alloin, D., Collin-Souffrin, S., Joly, M., & Vigroux, L. 1979, *A&A*, 78, 200  
 Amorín, R., Pérez-Montero, E., Vílchez, J. M., & Papaderos, P. 2012, *ApJ*, 749, 185  
 Berg, D. A., Skillman, E. D., Marble, A. R., et al. 2012, *ApJ*, 754, 98  
 Bresolin, F., Kennicutt, R. C., & Ryan-Weber, E. 2012, *ApJ*, 750, 122  
 Campbell, A., Terlevich, R., & Melnick, J. 1986, *MNRAS*, 223, 811  
 Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *ApJ*, 345, 245  
 Cid Fernandes, R., Pérez, E., García Benito, R., et al. 2013, *A&A*, 557, A86  
 Cresci, G., Mannucci, F., Sommariva, V., et al. 2012, *MNRAS*, 421, 262  
 Crowther, P. A., & Bibby, J. L. 2009, *A&A*, 499, 455  
 Croxall, K. V., van Zee, L., Lee, H., et al. 2009, *ApJ*, 705, 723  
 Denicoló, G., Terlevich, R., & Terlevich, E. 2002, *MNRAS*, 330, 69  
 Díaz, A. I., & Pérez-Montero, E. 2000, *MNRAS*, 312, 130  
 Dopita, M. A., & Evans, I. N. 1986, *ApJ*, 307, 431  
 Erb, D. K., Shapley, A. E., Pettini, M., et al. 2006, *ApJ*, 644, 813  
 Esteban, C., Carigi, L., Copetti, M. V. F., et al. 2013, *MNRAS*, 433, 382  
 García-Benito, R., Díaz, A., Hägele, G. F., et al. 2010, *MNRAS*, 408, 2234  
 Guseva, N. G., Izotov, Y. I., Fricke, K. J., & Henkel, C. 2012, *A&A*, 541, A115  
 Hadfield, L. J., & Crowther, P. A. 2007, *MNRAS*, 381, 418  
 Husemann, B., Jahnke, K., Sánchez, S. F., et al. 2013, *A&A*, 549, A87  
 James, B. L., Tsamis, Y. G., Barlow, M. J., Walsh, J. R., & Westmoquette, M. S. 2013a, *MNRAS*, 428, 86  
 James, B. L., Tsamis, Y. G., Walsh, J. R., Barlow, M. J., & Westmoquette, M. S. 2013b, *MNRAS*, 430, 2097  
 Kehrig, C., Vílchez, J. M., Sánchez, S. F., et al. 2008, *A&A*, 477, 813  
 Kehrig, C., Oey, M. S., Crowther, P. A., et al. 2011, *A&A*, 526, A128  
 Kehrig, C., Pérez-Montero, E., Vílchez, J. M., et al. 2013, *MNRAS*, 432, 2731  
 Kelz, A., Verheijen, M. A. W., Roth, M. M., et al. 2006, *PASP*, 118, 129  
 Kennicutt, Jr., R. C., Bresolin, F., & Garnett, D. R. 2003, *ApJ*, 591, 801  
 Kewley, L. J., & Dopita, M. A. 2002, *ApJS*, 142, 35  
 Kewley, L. J., & Ellison, S. L. 2008, *ApJ*, 681, 1183  
 Kobulnicky, H. A., & Phillips, A. C. 2003, *ApJ*, 599, 1031  
 Kobulnicky, H. A., Kennicutt, Jr., R. C., & Pizagno, J. L. 1999, *ApJ*, 514, 544  
 Liang, Y. C., Yin, S. Y., Hammer, F., et al. 2006, *ApJ*, 652, 257  
 Lilly, S. J., Carollo, C. M., & Stockton, A. N. 2003, *ApJ*, 597, 730  
 López-Sánchez, Á. R., & Esteban, C. 2010, *A&A*, 517, A85  
 López-Sánchez, Á. R., Esteban, C., García-Rojas, J., Peimbert, M., & Rodríguez, M. 2007, *ApJ*, 656, 168  
 López-Sánchez, Á. R., Mesa-Delgado, A., López-Martín, L., & Esteban, C. 2011, *MNRAS*, 411, 2076  
 López-Sánchez, Á. R., Dopita, M. A., Kewley, L. J., et al. 2012, *MNRAS*, 426, 2630  
 Luridiana, V., Morisset, C., & Shaw, R. A. 2012, in *IAU Symp.*, 283, 422  
 McGaugh, S. S. 1991, *ApJ*, 380, 140  
 Monreal-Ibero, A., Vílchez, J. M., Walsh, J. R., & Muñoz-Tuñón, C. 2010, *A&A*, 517, A27  
 Monreal-Ibero, A., Relaño, M., Kehrig, C., et al. 2011, *MNRAS*, 413, 2242  
 Monreal-Ibero, A., Walsh, J. R., & Vílchez, J. M. 2012, *A&A*, 544, A60  
 Moustakas, J., Kennicutt, Jr., R. C., Tremonti, C. A., et al. 2010, *ApJS*, 190, 233  
 Nagao, T., Maiolino, R., & Marconi, A. 2006, *A&A*, 459, 85  
 Osterbrock, D. E., & Ferland, G. J. 2006, *Astrophysics of gaseous nebulae and active galactic nuclei* (University Science Books)

- Pagel, B. E. J., Edmunds, M. G., Blackwell, D. E., Chun, M. S., & Smith, G. 1979, *MNRAS*, 189, 95
- Peña-Guerrero, M. A., Peimbert, A., & Peimbert, M. 2012, *ApJ*, 756, L14
- Pérez-Montero, E., & Contini, T. 2009, *MNRAS*, 398, 949
- Pérez-Montero, E., & Díaz, A. I. 2003, *MNRAS*, 346, 105
- Pérez-Montero, E., & Díaz, A. I. 2005, *MNRAS*, 361, 1063
- Pérez-Montero, E., Hägele, G. F., Contini, T., & Díaz, Á. I. 2007, *MNRAS*, 381, 125
- Pérez-Montero, E., Vílchez, J. M., Cedrés, B., et al. 2011, *A&A*, 532, A141
- Pérez-Montero, E., Contini, T., Lamareille, F., et al. 2013, *A&A*, 549, A25
- Pettini, M., & Pagel, B. E. J. 2004, *MNRAS*, 348, L59
- Pettini, M., Shapley, A. E., Steidel, C. C., et al. 2001, *ApJ*, 554, 981
- Pilyugin, L. S. 2000, *A&A*, 362, 325
- Pilyugin, L. S. 2001a, *A&A*, 369, 594
- Pilyugin, L. S. 2001b, *A&A*, 374, 412
- Pilyugin, L. S. 2007, *MNRAS*, 375, 685
- Pilyugin, L. S., & Thuan, T. X. 2005, *ApJ*, 631, 231
- Pilyugin, L. S., Vílchez, J. M., & Thuan, T. X. 2010, *ApJ*, 720, 1738
- Pilyugin, L. S., Grebel, E. K., & Mattsson, L. 2012, *MNRAS*, 424, 2316
- Pustilnik, S., Kniazev, A., Pramskij, A., et al. 2004, *A&A*, 419, 469
- Queyrel, J., Contini, T., Kissler-Patig, M., et al. 2012, *A&A*, 539, A93
- Raimann, D., Storchi-Bergmann, T., Bica, E., Melnick, J., & Schmitt, H. 2000, *MNRAS*, 316, 559
- Rosales-Ortega, F. F., Kennicutt, R. C., Sánchez, S. F., et al. 2010, *MNRAS*, 405, 735
- Roth, M. M., Kelz, A., Fechner, T., et al. 2005, *PASP*, 117, 620
- Sánchez, S. F., García-Lorenzo, B., Jahnke, K., et al. 2006, *New Astron. Rev.*, 49, 501
- Sánchez, S. F., Rosales-Ortega, F. F., Kennicutt, R. C., et al. 2011, *MNRAS*, 410, 313
- Sánchez, S. F., Kennicutt, R. C., Gil de Paz, A., et al. 2012a, *A&A*, 538, A8
- Sánchez, S. F., Rosales-Ortega, F. F., Marino, R. A., et al. 2012b, *A&A*, 546, A2
- Sánchez, S. F., Rosales-Ortega, F. F., Jungwiert, B., et al. 2013, *A&A*, 554, A58
- Sanders, N. E., Caldwell, N., McDowell, J., & Harding, P. 2012, *ApJ*, 758, 133
- Shields, G. A. 1990, *ARA&A*, 28, 525
- Stasińska, G. 2006, *A&A*, 454, L127
- Stasińska, G. 2010, in *IAU Symp. 262*, eds. G. R. Bruzual, & S. Charlot, 93
- Stasińska, G., Peña, M., Bresolin, F., & Tsamis, Y. G. 2013, *A&A*, 552, A12
- Steidel, C. C., Shapley, A. E., Pettini, M., et al. 2004, *ApJ*, 604, 534
- Storchi-Bergmann, T., Calzetti, D., & Kinney, A. L. 1994, *ApJ*, 429, 572
- Strauss, M. A., Weinberg, D. H., Lupton, R. H., et al. 2002, *AJ*, 124, 1810
- Thuan, T. X., Pilyugin, L. S., & Zinchenko, I. A. 2010, *ApJ*, 712, 1029
- Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, *ApJ*, 613, 898
- van Zee, L., Salzer, J. J., & Haynes, M. P. 1998, *ApJ*, 497, L1
- Verheijen, M. A. W., Bershady, M. A., Andersen, D. R., et al. 2004, *Astron. Nachr.*, 325, 151
- Vílchez, J. M., & Esteban, C. 1996, *MNRAS*, 280, 720
- Westmoquette, M. S., James, B., Monreal-Ibero, A., & Walsh, J. R. 2013, *A&A*, 550, A88
- York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, *AJ*, 120, 1579
- Zahid, H. J., & Bresolin, F. 2011, *AJ*, 141, 192
- Zaritsky, D., Kennicutt, Jr., R. C., & Huchra, J. P. 1994, *ApJ*, 420, 87
- Zurita, A., & Bresolin, F. 2012, *MNRAS*, 427, 1463