Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **509,** 1303–1322 (2022)
Advance Access publication 2021 October 29

https://doi.org/10.1093/mnras/stab3082

The 2D metallicity distribution and mixing scales of nearby galaxies

Thomas G. Williams [1★] Kathryn Kreckel [2] Francesco Belfiore [3] Brent Groves [4] Karin Sandstrom,[5] Francesco Santoro,[1] Guillermo A. Blanc,[6,7] Frank Bigiel,[8] Médéric Boquien,[9] Mélanie Chevance [10] Enrico Congiu [7] Eric Emsellem,[11,12] Simon C. O. Glover [13] Kathryn Grasha [14] Ralf S. Klessen [13,15] Eric Koch [16] J. M. Diederik Kruijssen [10] Adam K. Leroy,[17] Daizhong Liu,[18] Sharon Meidt,[19] Hsi-An Pan,[1,20] Miguel Querejeta,[21] Erik Rosolowsky [22] Toshiki Saito,[23,24] Patricia Sánchez-Blázquez,[25] Eva Schinnerer,[1] Andreas Schruba[18] and Elizabeth J. Watkins[2]

Affiliations are listed at the end of the paper

Accepted 2021 October 20. Received 2021 October 15; in original form 2021 August 27

ABSTRACT

Understanding the spatial distribution of metals within galaxies allows us to study the processes of chemical enrichment and mixing in the interstellar medium. In this work, we map the 2D distribution of metals using a Gaussian Process Regression (GPR) for 19 star-forming galaxies observed with the Very Large Telescope/Multi Unit Spectroscopic Explorer (VLT–MUSE) as a part of the PHANGS–MUSE survey. We find that 12 of our 19 galaxies show significant 2D metallicity variation. Those without significant variations typically have fewer metallicity measurements, indicating this is due to the dearth of H II regions in these galaxies, rather than a lack of higher-order variation. After subtracting a linear radial gradient, we see no enrichment in the spiral arms versus the disc. We measure the 50 per cent correlation scale from the two-point correlation function of these radially subtracted maps, finding it to typically be an order of magnitude smaller than the fitted GPR kernel scale length. We study the dependence of the two-point correlation scale length with a number of global galaxy properties. We find no relationship between the 50 per cent correlation scale and the overall gas turbulence, in tension with existing theoretical models. We also find more actively star-forming galaxies, and earlier type galaxies have a larger 50 per cent correlation scale. The size and stellar mass surface density do not appear to correlate with the 50 per cent correlation scale, indicating that perhaps the evolutionary state of the galaxy and its current star formation activity is the strongest indicator of the homogeneity of the metal distribution.

Key words: ISM: abundances – ISM: evolution – galaxies: general – galaxies: ISM.

1 INTRODUCTION

The chemical composition of the interstellar medium (ISM) plays a fundamental role in galaxy evolution. Within galaxies, stars enrich the surrounding ISM, causing chemical changes. On large scales, 'inside-out' galaxy growth (Boissier & Prantzos 1999) naturally leads to a negative gradient in the gas-phase chemical composition (hereafter 'metallicity') of a galaxy (e.g. Searle 1971; Martin & Roy 1992; Belfiore et al. 2019). On small scales, however, individual star-forming regions (the sites of individual enrichment events) are not independent and isolated, but are embedded within dynamically active galaxies. The dynamical conditions of galaxies can act to mix metals from different star-forming regions, leading to an increase in homogeneity between neighbouring regions. For instance, interstellar turbulence (e.g. Klessen & Lin 2003; Krumholz & Ting 2018), and gravitational instability (e.g. Yang & Krumholz 2012) can drive mixing and increase homogeneity. Conversely, numerical simulations predict azimuthal variations driven by spiral arms and bars (e.g. Di Matteo et al. 2013; Fragkoudi et al. 2018, 2020). Recently, these variations have been observed in nearby galaxies (e.g. Ho et al. 2017; Kreckel et al. 2019; Sánchez-Menguiano et al. 2020; Li et al. 2021), and an important measurement of the homogeneity is a 'chemical mixing scale,' which describes the spatial scales over which neighbouring regions have highly correlated metallicity values (e.g. Krumholz & Ting 2018; Kreckel et al. 2020; Li et al. 2021). As metallicity variations are driven by non-axisymmetric features, our traditional 1D view of the distribution of metals is insufficient to study these phenomena. We therefore require high-quality, 2D maps of the metallicity for entire galaxies.

High angular resolution is required to accurately map the metallicity in galaxies, due to contamination from gas emission lines in many density regimes, where metallicity prescriptions may not be valid. In particular, we see line emission arising from diffuse ionized gas (DIG), which typically has lower density and higher temperatures than the more dense H II regions (Haffner et al. 2009; Belfiore et al. 2021). The DIG also appears to exist in two regimes – the DIG around H II regions has line ratios consistent with that of the H II regions themselves, but DIG further away from these regions is dominated by ionization from very different sources (Belfiore et al. 2021). Whether the DIG line emission can be used to infer metallicity is unclear, although attempts have been made to measure metallicities within the DIG (Kumari et al. 2019), or remove the DIG contribution from the emission (Kaplan et al. 2016; Poetrodjojo et al. 2019; Vale

★ E-mail: williams@mpia.de

© The Author(s) 2021.
Published by Oxford University Press on behalf of Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/mnras/article/509/1/1303/6414012 by guest on 01 July 2022

Asari et al. 2019). Until a robust method for measuring metallicities is developed that accounts for variations in the ionizing spectrum, our best effort is to measure metallicities in H II regions where metallicity calibrations are valid, and then interpolate between them.

With the recent advent of optical integral field units (IFUs), such as the Multi Unit Spectroscopic Explorer (MUSE; Bacon et al. 2010) on the *Very Large Telescope* (VLT), it is now possible to identify and characterize thousands of H II regions across the faces of nearby galaxy discs (e.g. Rosales-Ortega et al. 2011; Sánchez et al. 2015; Ho et al. 2017; Poetrodjojo et al. 2018; Sarzi et al. 2018; Erroz-Ferrer et al. 2019; Gadotti et al. 2019; Ho et al. 2019; Kreckel et al. 2020). This gives us an access to the chemical conditions of individual H II regions, allowing us to build a picture of the overall metallicity distribution. However, to move beyond simple radial gradients, we require hundreds or even thousands of metallicity measurements across a galactic disc (e.g. Clark et al. 2019; Kreckel et al. 2020). Given that H II regions are not located uniformly across the disc of a galaxy, and can also be sparsely distributed (Santoro et al. 2021), we require advanced statistical techniques to model the underlying metallicity distribution of the galactic disc.

Producing a 'filled-in' map for metallicities that are traditionally sparsely measured in observations has a variety of uses. In particular, a complete map is particularly powerful when combining with lower resolution, but spatially complete, data, for example *Herschel* IR photometry to study the dust-to-metals ratio (Chiang et al. 2021) or dust mass absorption coefficient (Clark et al. 2019). These maps can also be used in conjunction with the CO maps from the Atacama Large Millimeter/submillimeter Array (ALMA) of the same galaxies (Leroy et al. 2021), to study, e.g. the metallicity dependence of molecular gas properties. Simple nearest-neighbour or bilinear interpolation performs poorly over the sparse, irregular sampling that metallicity measurements typically have, and so a more unbiased interpolation method is required.

In this work, we apply a Gaussian Process Regression (GPR) technique (for a general introduction to this technique, see Rasmussen & Williams 2006) to map the smooth, higher-order (i.e. the residual after subtracting a radial gradient) metallicity variation in a number of nearby galaxies. González-Gaitán et al. (2019) have recently applied a similar technique to that which we use in this work to a number of IFU cubes, finding significantly improved results over the naïve nearest neighbour interpolation approach. The galaxies we analyse have all been mapped as a part of the Physics at High Angular resolution in Nearby GalaxieS[1] (PHANGS) survey. We use data from the PHANGS–MUSE survey (Emsellem et al. 2021), which observed the star-forming discs of a sample of 19 nearby, main sequence galaxies at homogeneous sensitivity and resolution. With this large set of galaxies, we can search for trends in the typical ISM mixing scales as a function of various galaxy parameters. With the high spatial resolution that MUSE achieves within these galaxies ($\sim$100 pc), we can also search for variations in metal enrichment within particular galactic environments (e.g. the centres of galaxies, or their spiral arms).

The layout of this paper is as follows. In Section 2, we present an overview of the data used in this study. In Section 3, we describe our procedure for modelling the 2D metallicity distribution to obtain kernel length scales over which metals are homogeneously mixed, and compare these kernel length scales to correlation scales measured from a two-point correlation function. We search for environmental dependence on metallicity enrichment in Section 4. In Section 5, we investigate the relationship between the two-point correlation scales with various galaxy parameters. Finally, we conclude in Section 6. These maps are made publicly available at the links provided in the Data Availability Section.

2 DATA

In this study, we make use of MUSE IFU data for 19 nearby galaxies taken as part of the PHANGS–MUSE survey (Emsellem et al. 2021). Full details of the data reduction and processing are given in that work, and so we give only a short summary here. The data reduction is performed using standard MUSE recipes (e.g. wavelength and flux calibration, cosmic ray rejection, mosaicking) developed by the MUSE consortium and carried out by the MUSE data reduction pipeline (Weilbacher et al. 2020), available to the community via the ESOREX package (ESO CPL Development Team 2015) and accessed using a PYTHON wrapper PYMUSEPIPE.[2] These reduced products are then processed through a data analysis pipeline (DAP[3]), which runs in three stages: first, stellar kinematics are measured (the stellar velocity and higher-order moments). Next, the properties of stellar populations are estimated (e.g. stellar age, mass, and metallicity). Both of these stages are performed on Voronoi binned data to a stellar continuum signal-to-noise ratio (S/N) of 35, to maximize reliability. The fit is performed via PPXF (Cappellari & Emsellem 2004; Cappellari 2017), and makes use of E-MILES (Vazdekis et al. 2016) single stellar population models of eight ages (0.03–14 Gyr, logarithmically spaced in steps of 0.22 dex) and four metallicities ($[Z/H] = [-1.5, -0.35, 0.06, 0.4]$). Only the wavelength range 4850–7000 Å is used in the fit (of the full 4750–9350 Å range), in order to avoid strong sky residuals in the reddest part of the MUSE spectral range. Finally, for individual spaxels the properties of emission lines are measured (fluxes and kinematics), via a simultaneous fit of continuum and emission lines also performed via PPXF. The final DAP products are 2D maps of the intensity, velocity, and higher order moments of these emission lines (and associated uncertainties), as well as the properties of the stellar populations.

We make use of the DAP products produced as a part of the first public data release, and use the 'optimally convolved' (COPT) products. Because each galaxy is observed with multiple MUSE pointings and mosaicked together, there are variations in seeing across the complete mosaic. The COPT mosaic accounts for this, by convolving each individual pointing to a common coarsest angular resolution both in wavelength, and across the entire galaxy before the spectral fitting is carried out. At this point, the shape of the point-spread function (PSF) is convolved from a Moffat to a Gaussian, which allows us to immediately convolve to other Gaussian resolutions via Gaussian convolution kernels in the following subsection. In Table 1, we present the global properties of our sample.

2.1 Convolution

Although the COPT DAP products represent a homogeneous angular resolution across the galaxy (and wavelength range), with the different distances and seeing conditions for the 19 galaxies, this does not represent a homogeneous spatial scale *across* our sample of galaxies. We therefore convolve each intensity and (intensity-weighted) velocity dispersion map, along with associated error maps, to a fixed spatial resolution. We choose a common spatial resolution

[1] http://www.phangs.org

[2] https://github.com/emsellem/pymusepipe

[3] https://gitlab.com/francbelf/ifu-pipeline

Downloaded from https://academic.oup.com/mnras/article/509/1/1303/6414012 by guest on 01 July 2022

Table 1. Key parameters for the 19 galaxies. Galaxies with a bar are marked with an asterisk. Distances are taken from Anand et al. (2021), PA and inclination from Lang et al. (2020), and sizes, global stellar masses, and global SFRs from Leroy et al. (2021). The uncertainty in stellar mass is dominated by calibration uncertainty, so is always 0.11 dex for these galaxies. We also include the number of pixels in our convolved and regridded maps, and the number of pixels that have metallicity measurements after various cuts (see Section 2.2).

Galaxy	Dist (Mpc)	PA (deg)	i (deg)	r_{25} (arcmin)	R_e (arcmin)	$\log_{10}(M_*[\mathrm{M}_\odot])$	SFR ($\mathrm{M}_\odot\,\mathrm{yr}^{-1}$)	N_{pix}	N_Z
IC5332	9.01	74.4	26.90	3.03	1.38	9.67 ± 0.11	0.41 ± 0.11	8239	362
NGC 0628	9.84	20.7	8.90	4.94	1.36	10.34 ± 0.11	1.75 ± 0.45	23028	2193
NGC 1087*	15.85	359.1	42.90	1.49	0.70	9.93 ± 0.11	1.31 ± 0.34	32592	3133
NGC 1300*	18.99	278.0	31.80	2.97	1.18	10.62 ± 0.11	1.17 ± 0.30	94322	1587
NGC 1365*	19.57	201.1	55.40	6.01	0.49	10.99 ± 0.11	16.90 ± 4.38	108960	4541
NGC 1385	17.22	181.3	44.00	1.70	0.67	9.98 ± 0.11	2.09 ± 0.54	32169	3634
NGC 1433*	18.63	199.7	28.60	3.10	0.79	10.87 ± 0.11	1.13 ± 0.29	113824	1434
NGC 1512*	18.83	261.9	42.50	4.22	0.87	10.71 ± 0.11	1.28 ± 0.33	70615	1457
NGC 1566*	17.69	214.7	29.50	3.61	0.62	10.78 ± 0.11	4.54 ± 1.17	54142	5044
NGC 1672*	19.4	134.3	42.60	3.08	0.60	10.73 ± 0.11	7.60 ± 1.97	66011	4806
NGC 2835*	12.22	1.0	41.30	3.21	0.93	10.00 ± 0.11	1.24 ± 0.32	22212	2234
NGC 3351*	9.96	193.2	45.10	3.61	1.05	10.36 ± 0.11	1.32 ± 0.34	18746	364
NGC 3627*	11.32	173.1	57.30	5.14	1.10	10.83 ± 0.11	3.84 ± 1.00	21696	3079
NGC 4254	13.1	68.1	34.40	2.52	0.63	10.42 ± 0.11	3.07 ± 0.79	44203	7652
NGC 4303*	16.99	312.4	23.50	3.44	0.69	10.52 ± 0.11	5.33 ± 1.38	56444	8198
NGC 4321*	15.21	156.2	38.50	3.05	1.24	10.75 ± 0.11	3.56 ± 0.92	55510	4175
NGC 4535*	15.77	179.7	44.70	4.07	1.36	10.53 ± 0.11	2.16 ± 0.56	32153	1135
NGC 5068*	5.2	342.4	35.70	3.74	1.30	9.40 ± 0.11	0.28 ± 0.07	5224	1047
NGC 7496*	18.72	193.7	35.90	1.67	0.70	10.00 ± 0.11	2.26 ± 0.59	23140	1245

of 120 pc (and pixel size of 60 pc), which is advantageous for two reasons: first, the native PHANGS–MUSE resolution is 25–70 pc, so we can include all data at 120 pc resolution. Secondly, at this resolution, a typical H II region will be unresolved (Hunt & Hirashita 2009; Kreckel et al. 2019; Barnes et al. 2021; Santoro et al. 2021). As metallicity diagnostics are calibrated on entire H II regions, it should be applied to pixels that contain a complete H II region, and this pixel size ensures that. Given the typical region separation length of 100–300 pc (Chevance et al. 2020), we expect each pixel to correspond to approximately a single H II region.

For each galaxy, we calculate a Gaussian convolution kernel from the COPT beam to our fixed spatial resolution, where the FWHM of this kernel is given by

$$\mathrm{FWHM}_{\mathrm{conv}} = \sqrt{\mathrm{FWHM}_{120\,\mathrm{pc}}{}^2 - \mathrm{FWHM}_{\mathrm{COPT}}{}^2}. \quad (1)$$

We convolve our intensity and (intensity-weighted) velocity dispersions with a PSF of full width at half-maximum $(\mathrm{FWHM})_{\mathrm{conv}}$. For the error maps, due to earlier convolution to produce the COPT maps, values are covariant with each other. We account for this by estimating the 'per-pixel' noise that results in the measured error after smoothing. Taking a Gaussian convolution in two-dimensions, the relationship between the measured variance and the per-pixel variance is

$$\sigma^2_{\mathrm{measured}} = \sigma^2_{\mathrm{pix}} \sum_{j=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} \left(\frac{1}{2\pi \times w_{\mathrm{COPT}}{}^2} \mathrm{e}^{\frac{-(i^2+j^2)}{2\pi \times w_{\mathrm{COPT}}{}^2}} \right)^2. \quad (2)$$

i and j are the pixels to be summed over, $\sigma_{\mathrm{measured}}$ and σ_{pix} are the measured and per-pixel error maps, respectively, and w_{COPT} is the standard deviation (Gaussian width, in pixels) of the COPT PSF, where we use w instead of the usual σ in the Gaussian equation to avoid confusion between this and the error maps. The summation approximates to

$$\sigma^2_{\mathrm{measured}} \simeq \sigma^2_{\mathrm{pix}} \frac{1}{4\pi \times w_{\mathrm{COPT}}{}^2}, \quad (3)$$

and so after rearranging we arrive at

$$\sigma_{\mathrm{pix}} \simeq \sigma_{\mathrm{measured}} \times 2\sqrt{\pi} \times w_{\mathrm{COPT}}. \quad (4)$$

The final error map is then

$$\sigma_{\mathrm{conv}} = \sqrt{\sigma^2_{\mathrm{pix}} * \mathrm{PSF}^2_{120\,\mathrm{pc}}}, \quad (5)$$

where $*$ indicates a convolution and $\mathrm{PSF}_{120\,\mathrm{pc}}$ is a Gaussian kernel with FWHM = 120 pc. Following this, we Nyquist sample our maps to have two pixels across the $\mathrm{FWHM}_{120\mathrm{pc}}$, and remove any pixels within one PSF of the map edge (i.e. a two-pixel border), to avoid convolution artefacts. All reprojections are performed using PYTHON's surface brightness conserving `reproject` algorithm (as the MUSE maps are presented in units of surface brightness).

2.2 Pixel-by-pixel metallicity calculation

2.2.1 Data preparation

We calculated a gas-phase metallicity, $12 + \log_{10}(\mathrm{O/H})$, using our convolved line maps (we will use this oxygen abundance synonymously with metallicity throughout this work). We do not correct for Milky Way (Galactic) extinction, as this is already performed in the MUSE DAP. We correct fluxes for internal extinction. The internal extinction is calculated via the Balmer decrement,

$$C(\mathrm{H}\beta) = \frac{\log_{10}\left(\frac{\mathrm{H}\alpha}{\mathrm{H}\beta}\right)_{\mathrm{theo}} - \log_{10}\left(\frac{\mathrm{H}\alpha}{\mathrm{H}\beta}\right)_{\mathrm{obs}}}{0.4\left[k\left(\lambda_{\mathrm{H}\alpha}\right) - k\left(\lambda_{\mathrm{H}\beta}\right)\right]}, \quad (6)$$

where $(\frac{\mathrm{H}\alpha}{\mathrm{H}\beta})_{\mathrm{obs}}$ is the observed ratio between Hα and Hβ, and $(\frac{\mathrm{H}\alpha}{\mathrm{H}\beta})_{\mathrm{theo}}$ is the theoretically expected ratio of 2.86 (assuming case B recombination, an electron density of 100 cm^{-3}, and an electron temperature of 10^4 K; Osterbrock & Ferland 2006). This value is relatively insensitive to the assumed electron temperature, with variations of a factor of ∼2 causing changes of ∼0.1 mag in V-band extinction. The extinction coefficient at a given wavelenngth λ is $k(\lambda) = A(\lambda)/E(B-V)$. We assume the standard Milky Way value for the total to selective extinction, $R_V = 3.1$. We use the extinction

Downloaded from https://academic.oup.com/mnras/article/509/1/1303/6414012 by guest on 01 July 2022

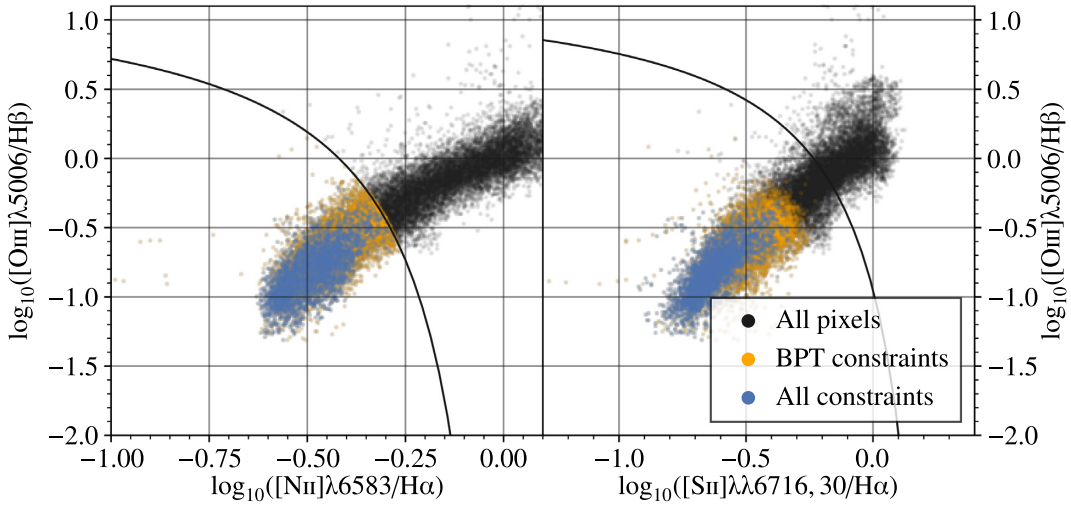


Figure 1. BPT diagnostics for pixel metallicity measurements NGC 3627. In each case, the star formation demarcation line is shown in black, with values below this line being consistent with photoionization. *Left:* Kauffmann et al. (2003) diagnostic in the [O III] λ5006/[N II] λ6583 plane. *Right:* Kewley et al. (2001) diagnostic in the [O III] λ5006/[S II] λλ6716,30 plane. In each case, we show all pixels in the convolved NGC 3627 map as a black cloud, those that satisfy both BPT constraints (i.e. are below both black lines) in yellow, and those that satisfy all our data constraints (S/N > 5, velocity dispersion <100, within morphologically defined H II region, see Section 2.2.1; i.e. those that form our final pixel map) in blue.

curve of O'Donnell (1994), calculated using PYNEB[4] (Luridiana, Morisset & Shaw 2015). We also tested the THEMIS (Köhler, Jones & Ysard 2014; Jones et al. 2017) extinction curve, which leads to <0.01 dex differences in the metallicities (using the same metallicity calibration). We therefore expect the choice of extinction curve will not qualitatively change our results. Values of $C(\mathrm{H}\,\beta) < 0$ typically have low S/N, and so we set them to 0. This has a negligible impact on our final maps, as these values typically fail our later S/N cuts. The corrected fluxes are then given by

$$I_{\mathrm{corr}}(\lambda) = I_{\mathrm{obs}}(\lambda) \times 10^{0.4\,C(\mathrm{H}\,\beta)k(\lambda)}. \quad (7)$$

Next, we make a number of cuts to ensure we select only star-forming regions, where metallicity prescriptions are appropriate. First, for the strong lines used in this work (H α, H β, [O III] λ5006, [N II] λ6583, and [S II] λλ6716,30), we require a S/N > 5. Secondly, we remove any regions with velocity dispersions > 100 km s^{-1} following Kreckel et al. (2019), to ensure we remove supernova remnants. Thirdly, we use the following two Baldwin–Phillips–Terlevich (BPT; Baldwin, Phillips & Terlevich 1981) cuts to select star-forming regions. The first is the constraint from Kauffmann et al. (2003) for the [N II] λ6583 diagram (left-hand panel of Fig. 1):

$$\log_{10}\left(\frac{[\mathrm{O\,III}]\,\lambda5006}{\mathrm{H}\,\beta}\right) < \frac{0.61}{\log_{10}\left(\frac{[\mathrm{N\,II}]\,\lambda6583}{\mathrm{H}\,\alpha}\right) - 0.05} + 1.3. \quad (8)$$

We also use the constraint from Kewley et al. (2001) for the [S II] λλ6716,30 diagram (right-hand panel of Fig. 1):

$$\log_{10}\left(\frac{[\mathrm{O\,III}]\,\lambda5006}{\mathrm{H}\,\beta}\right) < \frac{0.72}{\log_{10}\left(\frac{[\mathrm{S\,II}]\,\lambda\lambda6716,30}{\mathrm{H}\,\alpha}\right) - 0.32} + 1.3. \quad (9)$$

Both constraints must be satisfied for a pixel to be included in the final map.

Because much of the DIG has line ratios consistent with photoionization by massive stars (∼60 per cent; Belfiore et al. 2021), these cuts will not remove this diffuse component. We therefore use region masks from Santoro et al. (2021) to isolate pixels dominated by H II regions. This work gives a full account of how regions are defined, but briefly these regions are morphologically identified from MUSE H α images, using HIIPHOT (Thilker, Braun & Walterbos 2000). Following this, integrated spectra within the regions are extracted and fitted using the DAP, and bona-fide H II regions selected using similar criteria to ours described above (primarily velocity dispersion and BPT cuts). We only include these regions, and this has the effect of removing ∼50 per cent of our pixels, mostly from just outside H II regions (i.e. radiation leaking from the region). Based on visual inspection of the masks, the pixels removed in this step typically form rings around H II regions, rather than small circular regions that would indicate our per-pixel cuts are detecting H II regions that the morphological classification does not. Conversely, we also do not see pixels within the H II masks that do not make our earlier selection criteria. This indicates a good level of agreement for the earlier selection criteria between per-pixel measurements and integrated spectra of these H II regions. This masking step is applied to our convolved and regridded maps, simply by regridding to our larger pixel scale and world coordinate system. We will later use the pixel maps with all constraints except H II region masking, to assess how including the DIG component affects our results (Section 3.2). This will allow us a baseline comparison to future studies with potentially poorer spatial resolution, where removing the DIG component may not be possible.

2.2.2 *Metallicity Calibration*

Our fiducial metallicity calibration is the Pilyugin & Grebel (2016) S-calibration (hereafter Scal). This calibration uses three standard line diagnostics:

$$N_2 = I_{[\mathrm{N\,II}]\,\lambda\lambda6548,83}/I_{\mathrm{H}\beta},$$
$$S_2 = I_{[\mathrm{S\,II}]\,\lambda\lambda6716,30}/I_{\mathrm{H}\beta},$$
$$R_3 = I_{[\mathrm{O\,III}]\,\lambda\lambda4958,5006}/I_{\mathrm{H}\beta}. \quad (10)$$

[4]https://pypi.org/project/PyNeb/

Downloaded from https://academic.oup.com/mnras/article/509/1/1303/6414012 by guest on 01 July 2022

Note that for [O III] λλ4958,5006 and [N II] λλ6548,83, we use only the flux of the stronger of the two lines, and we adopt a fixed theoretical ratio of 3:1. The Scal prescription takes the form

$$12 + \log_{10}(\mathrm{O/H}) = a_1 + a_2 \log_{10}(R_3/S_2) + a_3 \log_{10}(N_2) + [a_4 + a_5 \log_{10}(R_3/S_2) + a_6 \log_{10}(N_2)] \times \log_{10}(S_2), \quad (11)$$

where the coefficients a_i are defined separately for the upper ($\log_{10}(N_2) \geq 0.6$) and lower ($\log_{10}(N_2) < 0.6$) branch. These are $a_i = [8.424, 0.030, 0.751, -0.349, 0.182, 0.508]$ for the upper branch and $a_i = [8.072, 0.789, 0.726, 1.069, -0.170, 0.022]$ for the lower branch. We use this as our fiducial metallicity calibration, as it shows small intrinsic scatter compared to direct metallicity measurements from auroral lines (Pilyugin & Grebel 2016; Ho 2019). We repeat the analysis using the Dopita et al. (2016) calibration (see Section 3.2), to assess the quantitative effect of using a different metallicity calibration. This calibration is given by

$$y = \log_{10}\left(\frac{[\mathrm{N\,II}]\,\lambda6583}{[\mathrm{S\,II}]\,\lambda\lambda6716,30}\right) + 0.264 \log_{10}\left(\frac{[\mathrm{N\,II}]\,\lambda6583}{\mathrm{H}\,\alpha}\right), \quad (12)$$

and then

$$12 + \log_{10}(\mathrm{O/H}) = 8.77 + y + 0.45(y + 0.3)^5. \quad (13)$$

For each pixel that satisfies all of our cuts in Section 2.2.1, we apply this metallicity calibration, to obtain a 'pixel metallicity map'. We propagate through the uncertainties using the PYTHON `uncertainties` package, which rigorously accounts for errors in extinction correction and fluxes in the line diagnostics. We found some values to be anomalously low (this is also seen in the catalogues of Santoro et al. 2021), and so remove outlier metallicities below $12 + \log_{10}(\mathrm{O/H}) < 7.8$.

As an additional check of the robustness of our metallicity measurements, we test against the metallicities obtained for integrated H II regions, using the catalogue from Santoro et al. (2021) and the same metallicity calibration. We reproject each of these masks on to our convolved, regridded maps, and for each region as defined in this catalogue, we take the H α-flux weighted mean metallicity of all our pixels falling within the spatial extent of the H II region. We find a mode of one pixel per H II region, and a median of three, indicating that our choice to convolve to 120 pc resolution does indeed lead to, on average, a single H II region per pixel. This comparison is shown in Fig. 2. We see an excellent agreement between these two approaches, with no systematic offset and a small scatter of 0.01 dex, so we are confident in this 'per-pixel' metallicity method going forwards.

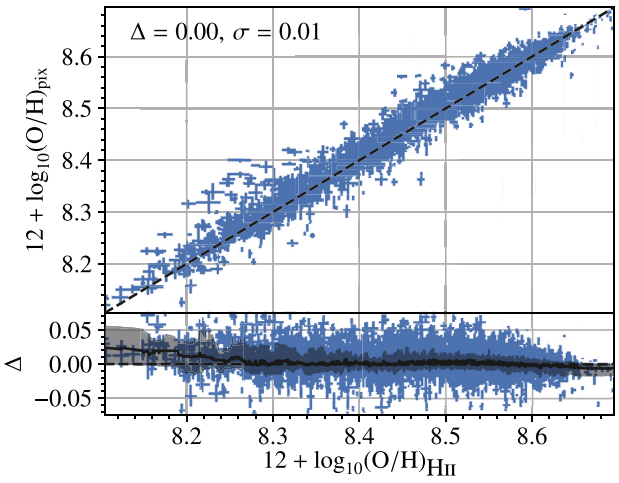


Figure 2. H α-flux weighted mean metallicities compared to integrated H II region metallicites across our full galaxy sample. We see there is no systematic offset between these two measurements, and an extremely small scatter of 0.01 dex. The lower panel shows the difference between each value, with the grey line indicating the rolling median and the shaded region the 16th and 84th percentiles of the distribution. There is a small deviation at the lowest metallicities which is slightly higher (and slightly lower at the highest metallicities), but these two regimes are dominated by low number statistics.

3 2D METALLICITY MAPS

3.1 Fitting the pixel metallicities

Using the pixel metallicity maps obtained in Section 2.2.2, we use a simple model to describe the underlying galactic metallicity distribution. We first assume the bulk of the variation is driven by a radial dependence, which has been typically employed in the literature (e.g. Searle 1971; Ho et al. 2015). Throughout this, we will fit in the coordinate frame of the galaxy, by calculating a deprojected galactocentric radius to each pixel (using the inclination and position angles listed in Table 1). We note that this will lead to a slightly noncircular beam in the deprojected frame, but as our galaxies are relatively face-on, this will be negligible. We fit a simple model taking into account the measured uncertainties on each metallicity, as well as an intrinsic scatter which captures the scatter between points that is not explained by the statistical uncertainties in the metallicity measurements, and is typically on the order of 0.03 dex. The (ln-)likelihood function of this model is (following Hogg, Bovy & Lang 2010)

$$\ln(\mathcal{L}) = -\frac{1}{2}\sum_n \left[\frac{(Z_n - m \times R/r_{25,n} - Z_0)^2}{s_n^2} + \ln\left(2\pi s_n^2\right)\right], \quad (14)$$

where Z_n and $R/r_{25,n}$ are the metallicity and galactocentric radius for each pixel, respectively, m is the radial metallicity gradient, Z_0 the metallicity at $R = 0$, and $s_n^2 = \sigma_{n,\mathrm{meas}}^2 + \sigma_{\mathrm{int}}^2$, the quadratic sum of the measured error and intrinsic scatter of the metallicity. We maximize this likelihood function using EMCEE (Foreman-Mackey et al. 2013), with 500 steps and 500 walkers. We remove the first half of these steps as 'burn-in', and calculate our best-fitting parameters from the remaining samples. An example radial fit is shown in Fig. 3, for NGC 3627. Clearly, many measurements lie significantly off this simple radial model (see also Fig. 5), motivating the need for higher-order fitting terms.

To model this 2D variation, we use a GPR technique (for a mathematical introduction to this process, we refer the reader to Rasmussen & Williams 2006). GPR is a probabilistic interpolation that models the covariance between neighbouring points using a covariance kernel, and thus is well-suited to the task of modelling the metallicity distribution of galaxies, as H II regions have highly correlated metallicities over relatively small spatial scales (on the order of 100 pc; Kreckel et al. 2020). GPR has been shown to recover the underlying distributions in a minimally biased manner (González-Gaitán et al. 2019). As GPR is a Bayesian method, it produces a posterior probability distribution function at each position, allowing us to calculate an uncertainty for each interpolated value. This technique has previously been used to produce metallicity maps by Clark et al. (2019), who found the GPR to reliably recover metallicity values (see their appendix C). GPR is a commonly applied machine learning technique, and has recently seen an increase in the number of applications in other astronomical contexts, particularly to model the light curves of transiting exoplanets (e.g. Prsa & Hambleton 2017;

Downloaded from https://academic.oup.com/mnras/article/509/1/1303/6414012 by guest on 01 July 2022

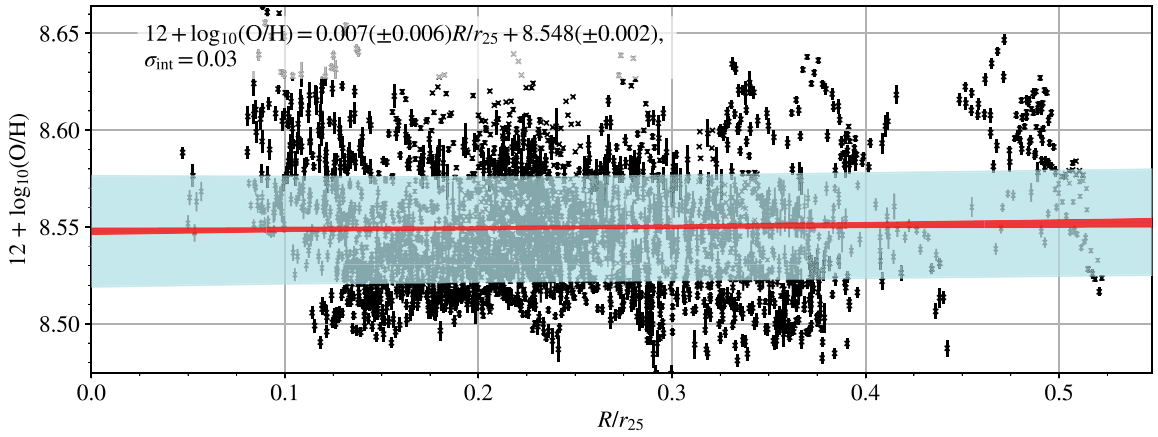


Figure 3. Radial fit for pixel metallicities in NGC 3627. The parameters for this fit are given in the top left. The shaded blue region highlights the intrinsic scatter. The shaded red region shows the uncertainty in the regression, but many points lie significantly beyond this region and have clear structure, highlighting the presence of higher-order variations with respect to the 1D radial trend.

Espinoza, Kossakowski & Brahm 2019), as well as 3D modelling of dust in the Milky Way (Green et al. 2019).

Our covariance kernel is the Matérn kernel, which is a standard choice for GPR modelling of 2D data (Rasmussen & Williams 2006). The choice of kernel will have an impact on the final map (see especially figs. 4.1–4.4 in Rasmussen & Williams 2006), but our choice of kernel here allows for variation from quite granular to smooth maps, in an attempt to cover the potential metal distributions within these galaxies. The Matérn kernel appears similar to a Gaussian, but with broader tails (sensitive to covariance over larger scales), and a narrower peak (sensitive to covariance on short distances). It takes the form

$$k(r_i, r_j) = \frac{1}{\Gamma(\nu)2^{\nu-1}} \left(\frac{\sqrt{2\nu}}{\sigma_{l,\rm obs}} d(r_i, r_j) \right)^{\nu} \mathrm{K}_{\nu} \left(\frac{\sqrt{2\nu}}{\sigma_{l,\rm obs}} d(r_i, r_j) \right), \quad (15)$$

where r_i, r_j are the two-points being considered, d the distance between them, K_{ν} a modified Bessel function, and $\Gamma(\nu)$ the gamma function. The kernel therefore only has two parameters, ν, which determines the smoothness of the kernel, and $\sigma_{l,\rm obs}$, the kernel length scale. This kernel length scale is related to the metal mixing scale (although, we will later show the absolute values differ by a significant factor). We fix ν to a value of 1.5, which makes the computation significantly more efficient, as equation (15) then simplifies to

$$k(r_i, r_j) = \left(1 + \frac{\sqrt{3} d(r_i, r_j)}{\sigma_{l,\rm obs}} \right) \exp \left(- \frac{\sqrt{3} d(r_i, r_j)}{\sigma_{l,\rm obs}} \right). \quad (16)$$

This approach is practically effective (Rasmussen & Williams 2006), and has been used to create 2D metallicity maps in previous work (Clark et al. 2019). This choice of ν imposes a level of smoothness in the final maps, but we find that experimenting with $\nu = 0.5$ leads to an extremely noisy map, and the kernel length scale is not constrained (indeed, this is noted in section 4.2 of Rasmussen & Williams 2006). We therefore proceed using the standard $\nu = 1.5$, but note that this choice of smoothness parameter may lead us to miss some variation on very small spatial scales.

We perform the fitting to the radially subtracted metallicity maps, using the `GaussianProcessRegressor` in `scikit-learn`, a PYTHON package for machine learning. We set prior bounds of $[0.001, 5]\, r_{25}$ on $\sigma_{l,\rm obs}$ (where r_{25} is the 25$^{\rm th}$ mag isophotal contour, a measure of the galaxy size; see Tab 1), to allow freedom from extremely granular higher-order metallicity maps (this range includes the resolution of the maps) to extremely smooth distributions. The MUSE field of view (FOV) typically extends to around $1r_{25}$, and so a kernel length scale larger than this may not be recoverable. We performed some simple 1D tests sampling from a Gaussian process with a number of kernel length scales (from 0 up to 10 times the FOV of the data), adding some random noise at the 1 per cent level (similar to our data). We found the kernel length scale was accurately recovered up to around five times the data FOV, and above this tended to underestimate the true value, motivating our choice of upper prior bound limit. For a kernel length scale hitting the upper prior bound, this indicates the GPR finds no significant covariance between points in the map, and not that the kernel length scale has been constrained. As the MUSE maps do not cover the entirety of the galaxy, they will not be sensitive to the typically higher scatter at much higher galactocentric radii (e.g. Pastorello et al. 2014). As such, this kernel scale length should be interpreted as an average of any variations in the mixing scaling within our FOV, and may not be representative of a mixing scale in the outskirts of galaxies. Our final metallicity map is the linear combination of the radial fit and GPR fit. This means that in regions with few measurements, the metallicity map will tend to the radial gradient, rather than $12 + \log_{10}(\mathrm{O/H}) = 0$. Our uncertainties combine the individual uncertainties from the radial and GPR fit added in quadrature. An example is shown in Fig. 4, for NGC 3627, and in Appendix A we show the equivalent plot for all 19 galaxies in the PHANGS–MUSE sample. Fig. 5 shows the improvement by including the higher-order terms in this fitting, also for NGC 3627.

We use two figures of merit to judge whether including this higher-order term improves the quality of the fit to the observed metallicities. The first is the standard χ^2 metric, to quantify how well the model describes the data, and the second is the Bayesian Information Criterion (BIC), which accounts for the number of parameters in the models, and penalizes more complex models, even if they do describe the data better. We show (a) the ratio of χ^2 values and (b) the difference in BIC (ΔBIC) statistics in Fig. 6, to show which model is preferred by each statistic. In all cases, including the 2D information leads to a fit where the χ^2 ratio prefers including the GPR fit. This is also the case for the majority of galaxies when using ΔBIC. The BIC statistic strongly favours including the GPR fitting

Downloaded from https://academic.oup.com/mnras/article/509/1/1303/6414012 by guest on 01 July 2022

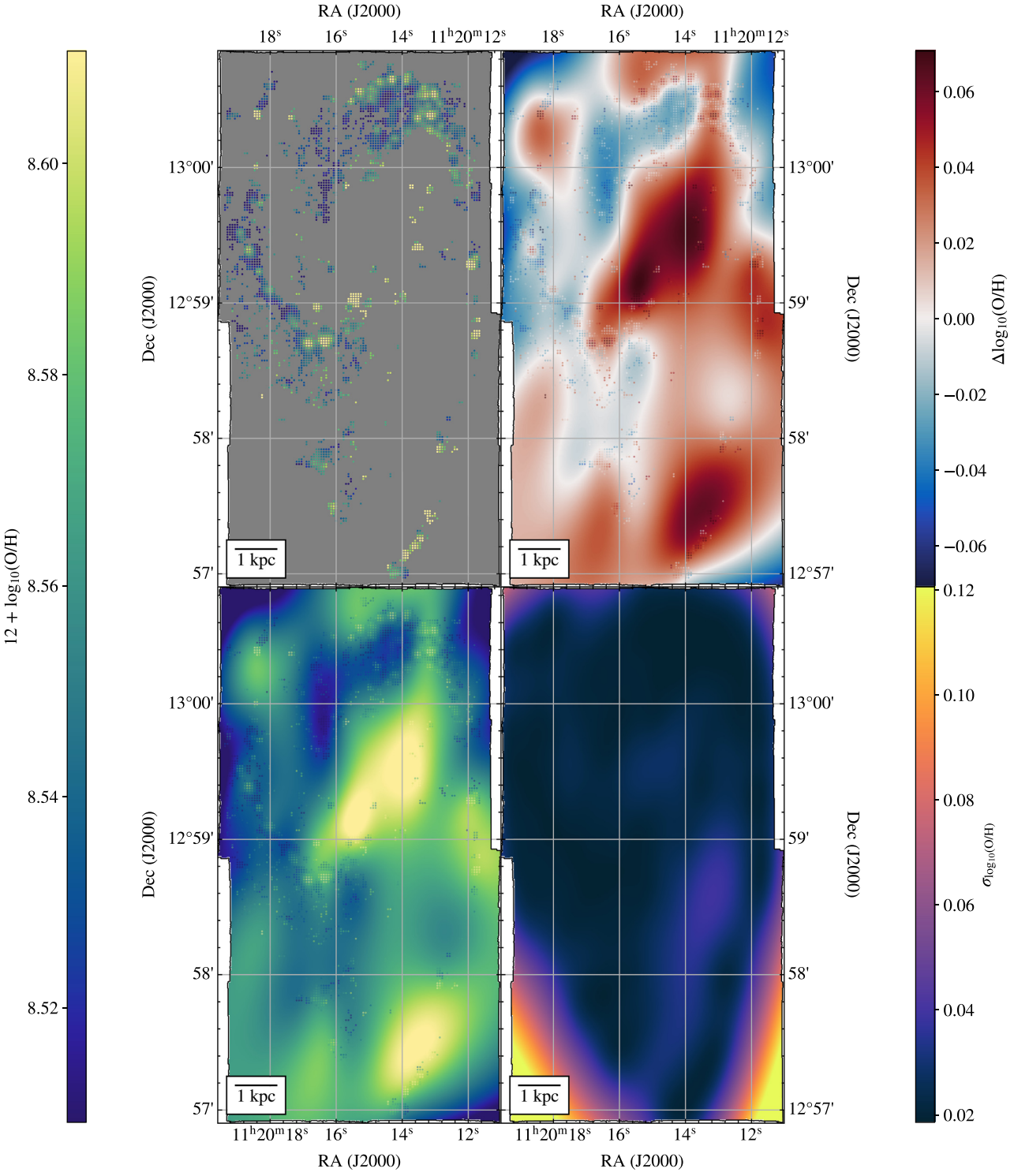


Figure 4. *Top left:* pixel metallicities for NGC 3627. This galaxy has a metallicity filling factor (Section 3.2) of 14 per cent, meaning its H II region density is slightly higher than most of our galaxy sample. *Top right:* residual metallicities after radial gradient is subtracted, overlaid on the GPR model for these points. *Bottom left:* final 2D metallicity map, showing clear higher-order variation. *Bottom right*: associated metallicity error map, combining both uncertainties in the radial fit, as well as the higher-order GPR.

for 14 of the galaxies (generally, a ΔBIC > 10 is considered strong evidence to prefer one model over another). However, there are five galaxies where the radial-only model is preferred. Of these, four strongly prefer the radial-only model over including the GPR fit, and NGC 4535 is a marginal case. In Section 3.2, we will describe how we determine if the higher-order component is deemed significant.

Downloaded from https://academic.oup.com/mnras/article/509/1/1303/6414012 by guest on 01 July 2022

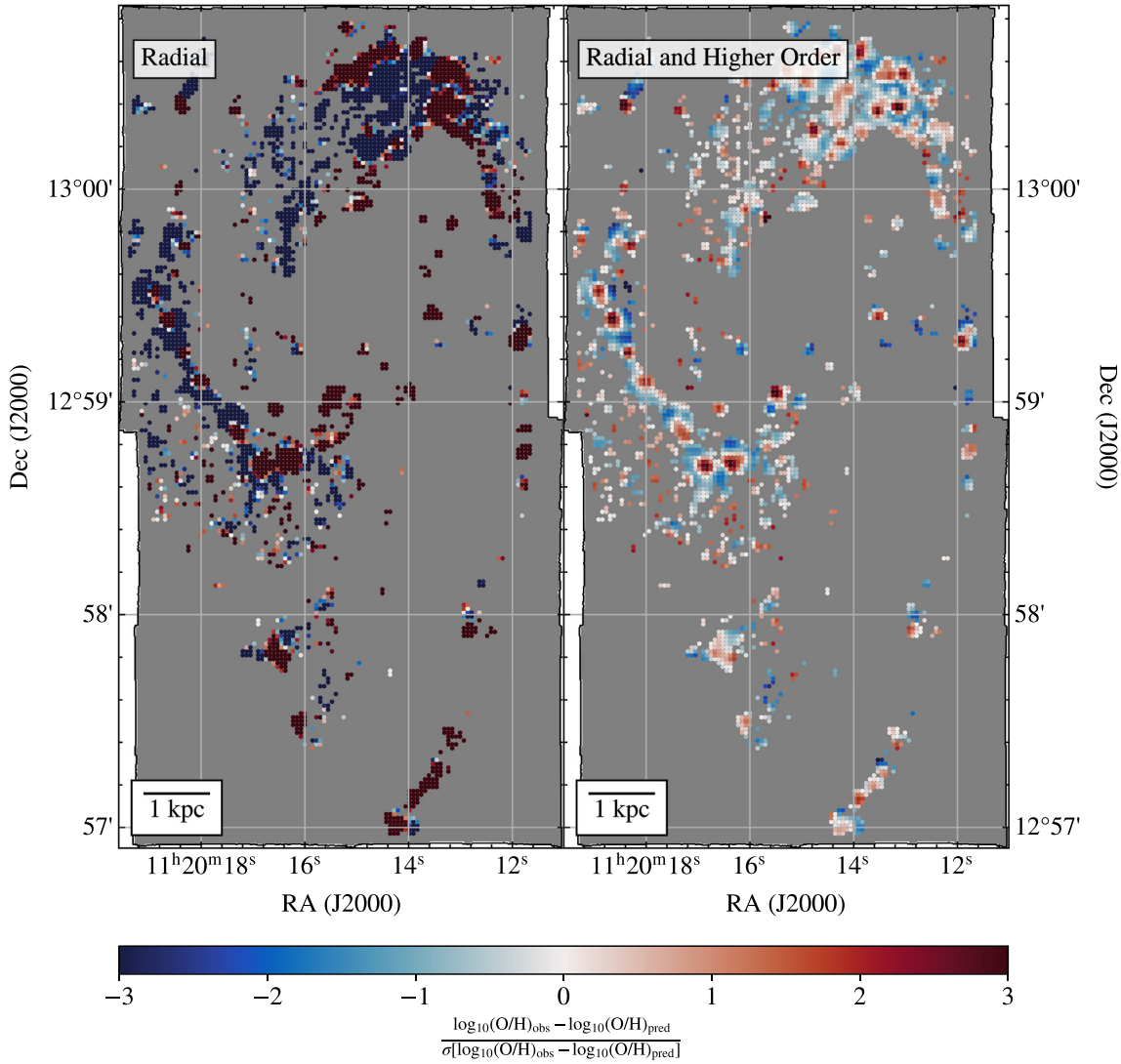


Figure 5. Map of the residual metallicity (i.e. $12 + \log_{10}(\mathrm{O/H})_{\mathrm{obs}} - 12 + \log_{10}(\mathrm{O/H})_{\mathrm{pred}}$) normalized by the combined errors in the observations and model fits of NGC 3627 for *left*: a radial gradient and *right*: our full 2D model. White indicates where the model and observations agree – including the higher-order variations produces a better fit to the data, reducing many highly significant outliers to less than 3σ.

3.2 Determining the significance of higher-order variations

As we do not have a metallicity measurement for every pixel in our map, the covariance kernel length scale may be somewhat overestimated (with fewer values, the GPR is unable to reliably model the covariance between points, and so tends towards larger kernel length scales, indicating smoother maps). To correct for this bias, we perform a nested Monte Carlo Jackknife approach ('McJack' for short). We retain a certain fraction of pixels (f_{pix} from our original pixel metallicity maps, from 0 per cent to 100 per cent, where 100 per cent corresponds to N_Z in Table 1) in steps of 10 per cent and refit the GPR. In this sense, we can extrapolate beyond $f_{\mathrm{pix}} = 1$, as this is set by the maximum number of metallicity measurements in our pixel map. This is similar to cross-validation in machine learning models, with a subset of the data being held back to verify the robustness of the output parameters. For each fraction, we also perturb the metallicities of the remaining pixels by their associated error, to estimate an uncertainty on the measured kernel length scale. For each fraction, we perform these perturbations a total of 100 times. The change in σ_l with f_{pix} appears to be well-described by a negative exponential, of the form

$$\sigma_{l,\mathrm{obs}} = A \exp(-\lambda f_{\mathrm{pix}}) + \sigma_l, \tag{17}$$

where A is the amplitude of the function at $f_{\mathrm{pix}} = 0$, λ is an exponential scale length, and σ_l the corrected kernel length scale (the extrapolation of this model to $N_{\mathrm{all}} \gg N_Z (f_{\mathrm{pix}} \rightarrow \infty)$, an estimate of the case where we have a metallicity for every pixel in the map). The model choice here is motivated by the shape of the data, rather than physically; we experimented with a number of declining functional forms (e.g. half-normal, half-student T), and found that the exponential best describes the trends seen in all cases. We could instead use f_{pix} to refer to the total number of pixels in the map (N_{pix} in Table 1); this would change λ but keep σ_l identical. We note that we only fit points that are not at (or consistent with) the upper prior bound of $5r_{25}$, and where we have more than three points to fit (so we have at least 1 degree of freedom). An example of this fit is shown in Fig. 7, for NGC 3627.

Having corrected our kernel length scale, we next turn to the question of determining whether the determined kernel length scale

Downloaded from https://academic.oup.com/mnras/article/509/1/1303/6414012 by guest on 01 July 2022

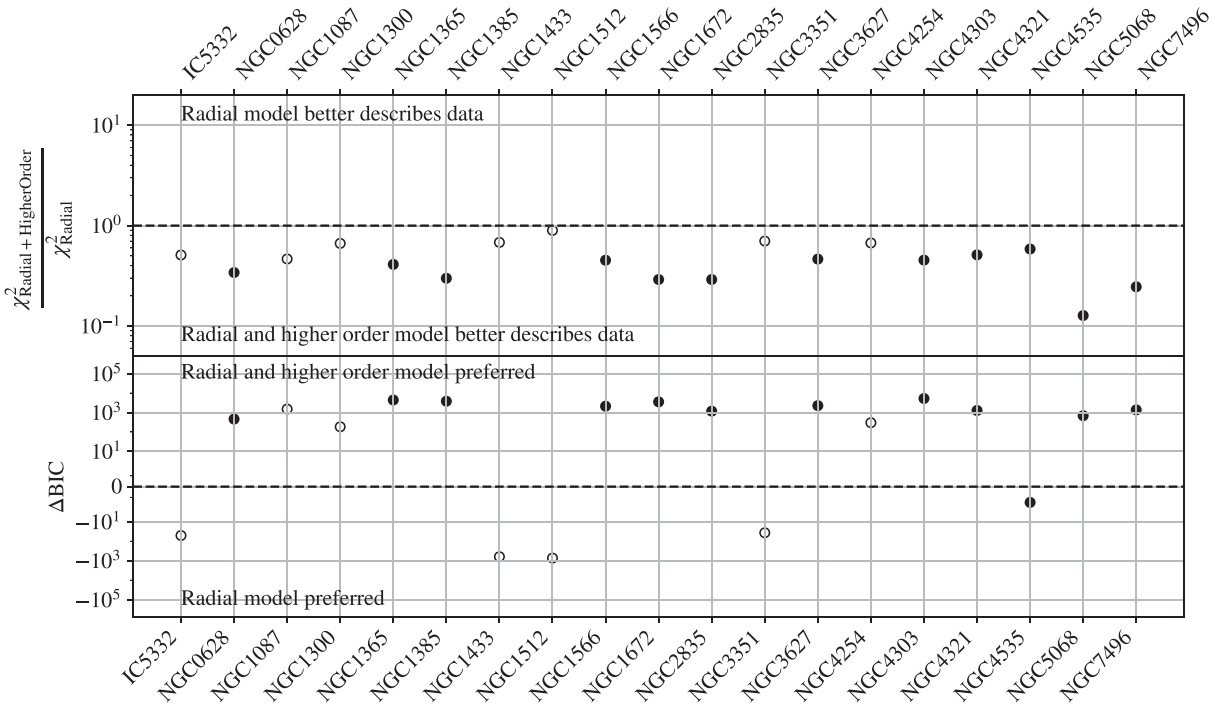


Figure 6. Figures of merit for the fitting procedures. *Top*: ratio of χ^2 values for the GPR (radial + higher-order) and radial-only model. Above one, the radial-only fit is preferred, and below one, the combined fit preferred. *Bottom*: difference in BIC statistics for the two models. Above zero, the combined model is preferred, below zero, the radial-only model is preferred. Points are filled if they pass our later significance testing (Section 3.2), and unfilled if they do not. The dashed black line indicates where the models are equally preferred, in each case.

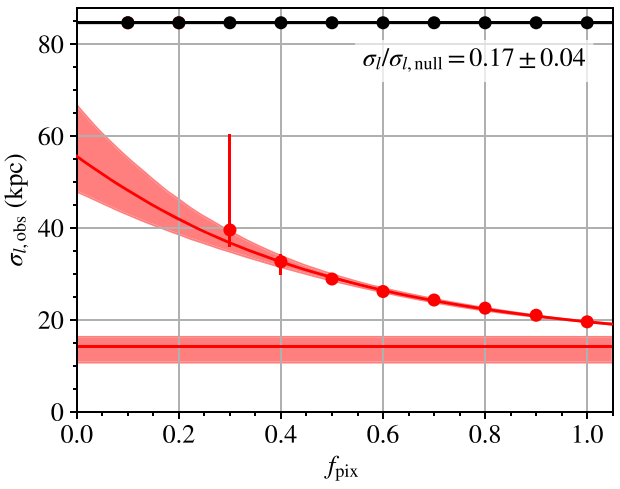


Figure 7. Monte Carlo Jackknife (McJack) method (Section 3.2) to correct the measured metallicity kernel length scale in NGC 3627. The horizontal red line shows the extrapolation of the line fitted to the real data (red points) out to infinity (our corrected kernel length scale, σ_l), the black one to our null hypothesis ($\sigma_{l,\mathrm{null}}$, which is always $5r_{25}$). Clearly, the data show a significantly different kernel length scale to the null hypothesis (highlighted by the text in the top right), and so we deem this higher-order variation to be significant.

is statistically significant with respect to a map with no correlation between neighbouring pixels. For this, we perform a null hypothesis test which is essentially identical to the McJack described in the previous paragraph, but before any pixels are removed we randomly shuffle them around the map after radial gradient subtraction. In this case, the GPR always hits the upper bound of the prior, indicating a smooth map with no real structure. If σ_l is significantly different to this null hypothesis (above the 1σ level), then we conclude we have measured statistically significant higher-order variations. Of our 19 galaxies, we find that 12 have significant higher-order variations, and that typically those that do not have significant higher-order variations are those that do not have large changes in their χ^2 values (see Fig. 6).

We make two checks on these kernel length scales to see how some of our assumptions may affect them. In the first, we repeat this entire procedure using the Dopita et al. (2016) metallicity calibration, see equation (13), and we show this in Fig. 8 for galaxies where the σ_l values measured from both metallicity calibrations are deemed to be significant. The correlation between the kernel length scales derived with these two different calibrations has Kendall's τ correlation coefficient of 0.21 $\pm$ 0.09. This is similar to the correlations seen by Li et al. (2021) between scale lengths measured from different metallicity calibrations. Li et al. (2021) use the Pearson correlation coefficient, but repeating this exercise with that correlation coefficient yields nearly identical results. However, the kernel length scales measured are typically quite different, with the Scal values being significantly higher. This reflects the increased scatter of metallicities between neighbouring H II regions in the Dopita et al. (2016) calibration, which the GPR models as significant variation over small scales. We adopt Scal as our fiducial metallicity calibration due to its low scatter with respect to direct metallicity measurements in an attempt to avoid this issue, but we note that maps derived from different calibrations will produce (sometimes significantly) different looking maps (see also appendix C of Kreckel et al. 2019).

Downloaded from https://academic.oup.com/mnras/article/509/1/1303/6414012 by guest on 01 July 2022

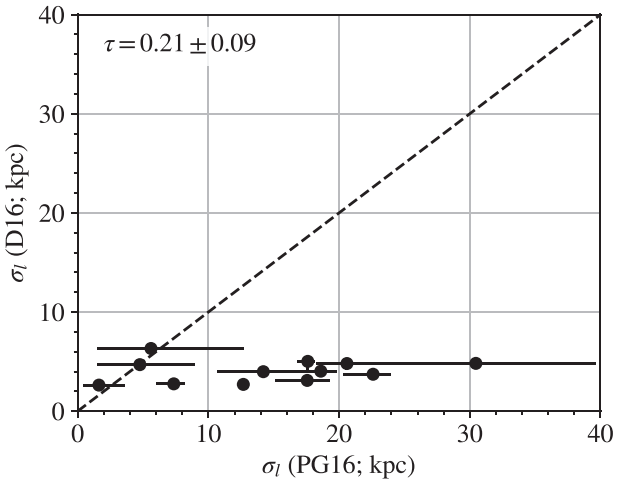


Figure 8. Comparison between kernel length scales derived using the Dopita et al. (2016) or the Pilyugin & Grebel (2016) Scal metallicity calibration. The 1:1 relationship is shown as a dashed black line. Kendall's τ correlation coefficient is shown in the top left. Typically, the kernel length scales derived from the Dopita et al. (2016) are significantly shorter, and reflect the higher scatter in this metallicity calibration, which the GPR models as significant variation over short scales.

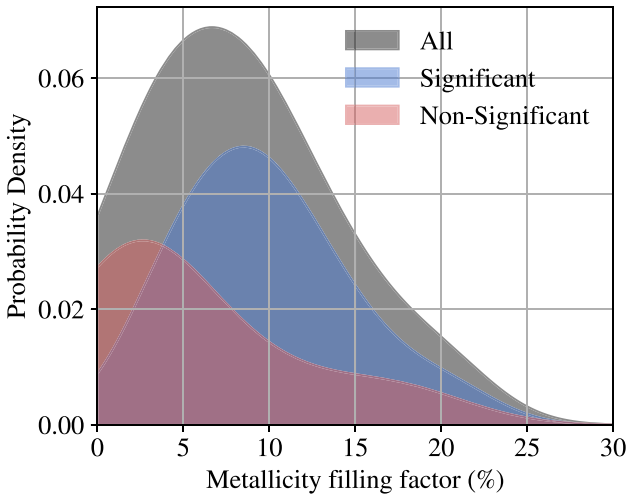


Figure 10. Distribution of metallicity filling factors for galaxies in which we measure a significant higher-order variation (blue) and galaxies where we do not (red). The overall distribution for all galaxies is shown in grey in the background. Due to the relatively small number of galaxies, we show the distribution as a Kernel Density Estimate (KDE), to aid visualization. We typically see a lower filling factor for galaxies where we do not measure significant higher-order variations.

For our second test, we relax our constraint on limiting to pixels within H II regions as defined from their H α morphology by Santoro et al. (2021). We repeat our analysis, and show the derived kernel length scales from this exercise compared to our fiducial assumptions in Fig. 9. There is a reasonable agreement, with Kendall's τ of 0.30 ± 0.12, indicating that including pixels that are dominated by DIG emission does not bias our results in a sample-wide sense, but including DIG emission can lead to significant differences in individual kernel length scales. We therefore urge caution in interpreting length scales measured on data where morphological classification of H II regions is not possible.

The lack of significant higher-order variations in seven galaxies may be due to two reasons. The first is that these galaxies truly only possess a radial metallicity gradient, or second that our data are insufficient in number for the fitting algorithm to measure this variation. To test this, we calculate a 'metallicity filling factor', which is simply the fractional area on the sky occupied by H II regions. We show the distribution of these filling factors in Fig. 10. There may be local variations of the filling factor in radius and azimuth, which will manifest as larger uncertainties in the GPR fit. Clearly, for galaxies where we do not deem the higher-order component to be significant, we have far fewer points to fit the GPR to. This lack

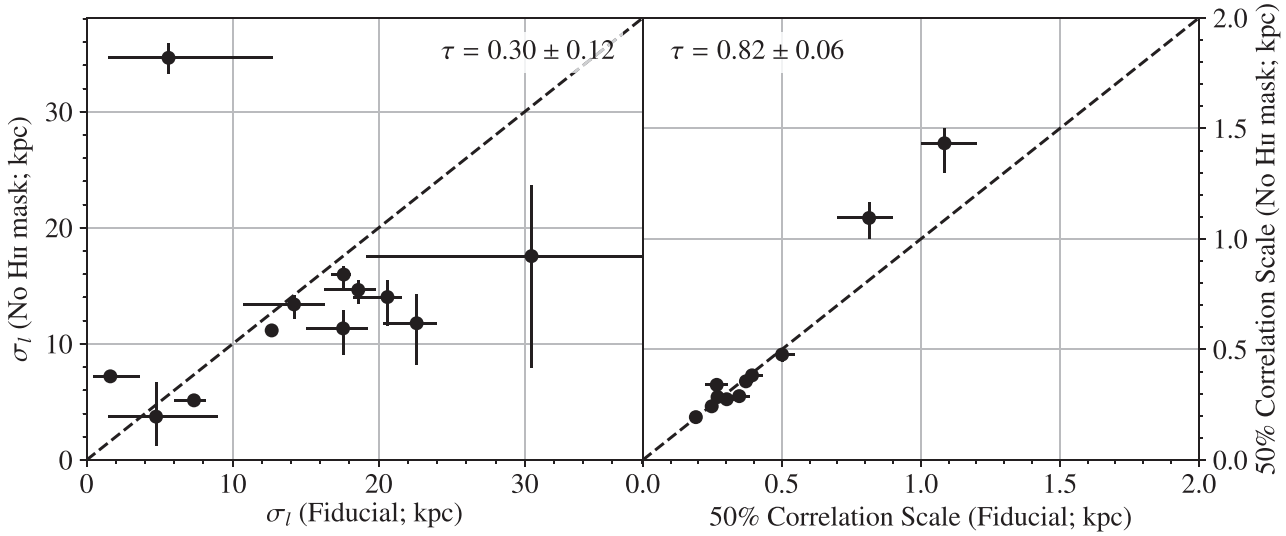


Figure 9. *Left:* kernel length scales calculated from our fiducial assumptions on which pixels to include, compared to when we relax our restriction to only include pixels within H II regions as defined by Santoro et al. (2021). *Right:* the same exercise but for the 50 per cent correlation scales of the two-point correlation function. In each case, the 1:1 line is shown as a dashed black line, and we show Kendall's τ correlation coefficient.

Downloaded from https://academic.oup.com/mnras/article/509/1/1303/6414012 by guest on 01 July 2022

Table 2. Corrected scale lengths (σ_l), and 50 per cent two-point correlation scale for the 12 galaxies with significant azimuthal variation.

Galaxy	σ_l (kpc)	50 per cent Correlation Scale (kpc)
NGC 0628	$1.63^{+2.00}_{-1.19}$	$0.19^{+0.00}_{-0.01}$
NGC 1365	$5.61^{+7.13}_{-4.11}$	$4.12^{+0.08}_{-0.12}$
NGC 1385	$4.76^{+4.19}_{-3.28}$	$0.30^{+0.04}_{-0.01}$
NGC 1566	$17.56^{+1.71}_{-2.47}$	$0.39^{+0.04}_{-0.02}$
NGC 1672	$22.60^{+1.39}_{-2.26}$	$0.50^{+0.05}_{-0.02}$
NGC 2835	$18.60^{+1.18}_{-2.36}$	$0.81^{+0.09}_{-0.11}$
NGC 3627	$14.20^{+2.10}_{-3.49}$	$0.25^{+0.02}_{-0.02}$
NGC 4303	$12.67^{+0.34}_{-0.41}$	$0.35^{+0.04}_{-0.05}$
NGC 4321	$20.60^{+0.95}_{-2.36}$	$0.27^{+0.01}_{-0.01}$
NGC 4535	$30.47^{+9.18}_{-11.34}$	$0.37^{+0.02}_{-0.02}$
NGC 5068	$7.35^{+0.85}_{-1.32}$	$0.27^{+0.04}_{-0.04}$
NGC 7496	$17.61^{+0.49}_{-0.84}$	$1.08^{+0.12}_{-0.08}$

of higher-order structure is likely driven by the lack of a sufficient number of H II regions within these galaxies, and not necessarily by a real lack of higher-order variations in the galaxy itself. Since these galaxies are the ones in which the GPR does not provide a significant improvement over simply a radial metallicity gradient, we will not use these in our later analysis. Our corrected kernel length scales are given in Table 2. Our kernel length scales vary from 1.5 to 30 kpc, with a median value of 15 kpc (0.1–1.9r_{25}, median 1.1r_{25}). We find that there is a slight trend with the MUSE FOV size (Kendall's τ of 0.15 ± 0.12), but no trend with the galaxy inclination (Kendall's τ of 0.12 ± 0.12).

3.3 Two-point correlation function

Kreckel et al. (2020) performed an analysis of the two-point correlation function of metals for a subset of eight galaxies of the PHANGS–MUSE sample, applied to measurements of individual H II regions. They found a high correlation in metallicity over small scales, with the exception of IC 5332 (a galaxy where we also do not detect significant higher-order variations). They quantify the mixing scale via a percentage correlation scale in the two-point correlation function (their sect. 4). We apply the same measurement to our maps here, to see how comparable our results are to those applied only to H II regions. The two-point correlation function at a spatial scale r is given as

$$\xi(r) = \left\langle \frac{\overline{Z(\mathbf{r_1})Z(\mathbf{r_2})} - \overline{Z}^2}{\overline{(Z-\overline{Z})^2}} \right\rangle, \quad (18)$$

where Z represents a metallicity measurement, and $|\mathbf{r_1} - \mathbf{r_2}| \leq r$. Horizontal lines indicate averaging over all H II regions in the galaxy, while angle brackets indicate averaging over all choices of $\mathbf{r_1}$. At $r = 0$, each H II region correlates perfectly with itself and a 100 per cent correlation is recovered. We apply this to our radially-subtracted, GPR fitted maps, taking a small percentage (1 per cent, leaving us a minimum of ∼3000 pixels to calculate this statistic) of the total pixels in our maps each time, to reduce computation time and allow us to determine jackknife errors. Of the remaining pixels, we also perturb these by the measured errors. We calculate the two-point correlation for a number of scales (from 0 to 5 kpc), and calculate the 50 per cent correlation scale. We show the relation of the kernel length scales and 50 per cent correlation scale in the left panel of Fig. 11, and list these values in Table 2. The kernel length scale and the 50 per cent correlation scale are positively correlated with each other, but the kernel length scale is often more than an order of magnitude larger in absolute value. The difference in these two values reflects the fact that the 50 per cent correlation function is a measure of small-scale, highly local correlation between values, whilst the kernel scale length is sensitive more to large-scale, smoother variations between more separated measurements. The two quantities are not monotonically related, and there is no simple transformation between them. NGC 1365 is an outlier in this comparison, with a 50 per cent correlation scale of over 4 kpc. We do not include this in our plot, or calculation of the Kendall's τ coefficient, which is $\tau = 0.38 \pm 0.07$. We also compare how the 50 per cent correlation scales changes based on whether we use H II region masks to define pixels to fit or not. This is shown in the right panel of Fig. 9, and there is an excellent agreement between the values calculated using these two different assumptions, with a $\tau = 0.82 \pm 0.06$ and most points clustered along the 1: 1 line. Whilst we have earlier motivated our choice to remove pixels outside H II regions before fitting the GPR, this result highlights this choice will not bias our results significantly.

We also apply the two-point correlation function to the full PHANGS nebulae catalogues (Santoro et al. 2021, i.e. only using H II regions) in Appendix B for all 19 galaxies in the PHANGS–MUSE sample, and we compare the values calculated from our interpolated radially-subtracted GPR maps to those calculated from the H II region catalogues in the right panel of Fig. 11, for those where we deem the measured kernel length scales to be significant, and excluding NGC 1365, as this also has a very high 50 per cent correlation scale measured from the H II regions. This likely indicates the scale length is primarily driven by H II regions along the bar of this galaxy, which fills most of the MUSE FOV, rather than some inherent limitation in either the GPR fitting or applying a two-point correlation statistic to H II regions. We see a good agreement between the 50 per cent two-point correlation scales between the GPR map and from the H II region catalogues, with many following the 1: 1 relationship, and determine a Kendall's correlation coefficient $\tau = 0.24 \pm 0.07$. We are therefore confident that our maps reflect a similar covariance between neighbouring regions as those found when considering morphologically selected H II regions. Furthermore, since the 50 per cent correlation scales are similiar between these two methods, we are confident these different analyses characterize the metal mixing in a similar way. We will therefore use the 50 per cent correlation scale from the two-point correlation function in our analysis going forwards, as this has been previously employed in the literature (Kreckel et al. 2020), and shown to produce similar correlation scales to the model of Krumholz & Ting (2018), as shown by Li et al. (2021). However, as we have shown with our comparison between the kernel scale length and 50 per cent correlation scale, using different measures of 'scale lengths' can result is significantly different absolute values. In Section 5 we will search for trends between this correlation scale and various global parameters.

4 ENVIRONMENTAL DEPENDENCE OF METAL ENRICHMENT

With an interpolated metallicity map, we can now study how the metallicity varies in different environments of the galaxy. Abundance variations have been predicted from numerical galaxy simulations (e.g. Di Matteo et al. 2013; Grand et al. 2016), but observational works show conflicting results. Some studies have reported no significant variation between the spiral arm and inter-arm regions (e.g.

Downloaded from https://academic.oup.com/mnras/article/509/1/1303/6414012 by guest on 01 July 2022

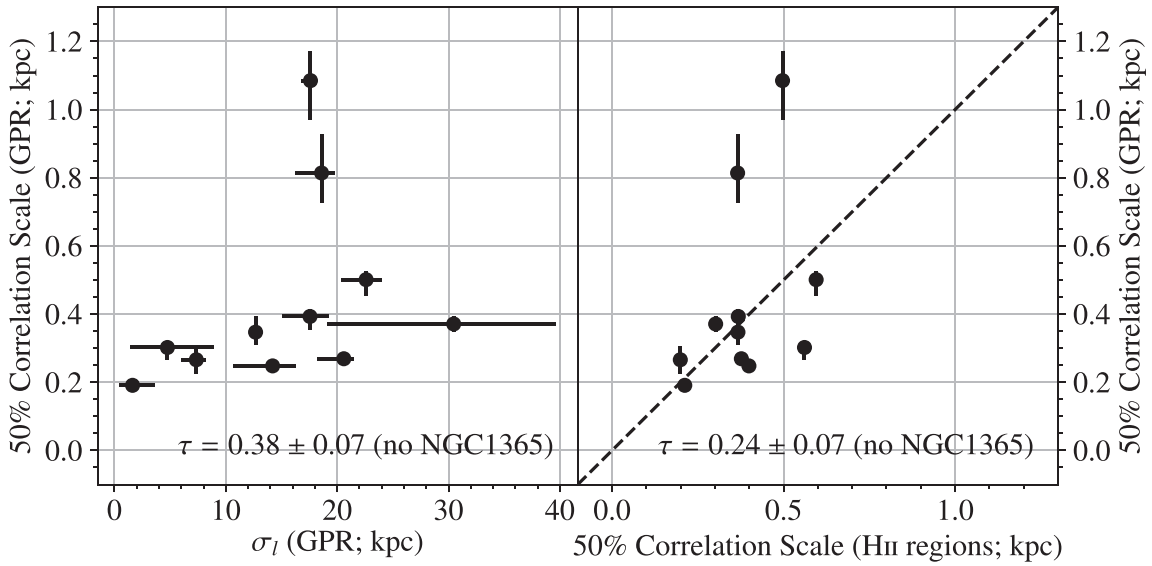


Figure 11. *Left*: Comparison between derived kernel length scales and 50 per cent correlation scale from two-point correlation function. The 50 per cent correlation scales are significantly shorter than the kernel length scales, by an order of magnitude or more. *Right*: Comparison between 50 per cent two-point correlation functions derived in this work, and for the same galaxies from the nebulae catalogues of Santoro et al. (2021, Appendix B). The dashed black line indicates the 1: 1 relation. In each case, we show Kendall's τ correlation coefficient excluding NGC 1365 (the notable outlier).

Martin & Belley 1996; Cedrés & Cepa 2002; Kreckel et al. 2016), whilst others have found evidence that metallicities are enhanced in spiral arms with respect to the inter-arm region (e.g. Sakhibov et al. 2018; Sánchez-Menguiano et al. 2020). These variations may be spatially localized, perhaps only occurring in one of the spiral arms (Kreckel et al. 2019), or could vary radially (and peaking at co-rotation; Spitoni et al. 2019). With these maps, we can now address if (and where) metallicities are enhanced with a large, homogeneous sample of high-resolution metallicity maps.

To do this, we take the environmental masks from Querejeta et al. (2021). To avoid being too granular in defining environments, we use the 'simple' masks as defined in that work, which divide galaxies up into 'centres', 'bars', 'spiral arms', and 'discs' (which includes the outer disc, the entire disc of galaxies without strong spiral arms, as well as inter-arm and inter-bar regions). We reproject these masks on to our radially subtracted metallicity maps, and calculate a Kernel Density Estimate (KDE) of the radially subtracted metallicities ($\Delta\log_{10}$(O/H); i.e. the top right panel of (Fig. 4) for each pixel grouped by its environment. Because some environments may have more sparse metallicity measurements (see e.g. along the bar in NGC 3627 in Fig. 4), we also perturb the values by their associated errors before calculating the KDE. The error reported by the GPR represents the sparsity of measurements, along with the uncertainty of each given measurement, and so this perturbation will account for environments with fewer metallicity measurements (e.g. along bars). For the KDE, we use a Silverman B. W. (1986) bandwidth (see also this reference for an introduction to KDE). To avoid giving equal weight to more uncertain metallicity values in fainter regions, the KDE is calculated using weighting based on the Hα flux, and refer to this distribution as $\Delta\log_{10}(\mathrm{O/H})_{\mathrm{H}\alpha}$. Because the bar may also promote efficient mixing and reduces local abundance variations (Di Matteo et al. 2013), we split the centres of galaxies up further into a barred and unbarred sample. For our 12 galaxies, this is shown in Fig. 12. The distributions for all environments look reasonably Gaussian and (except the centres of barred galaxies, which tend to be slightly higher) peak around a $\Delta\log_{10}(\mathrm{O/H})_{\mathrm{H}\alpha}$ of 0.01 dex,

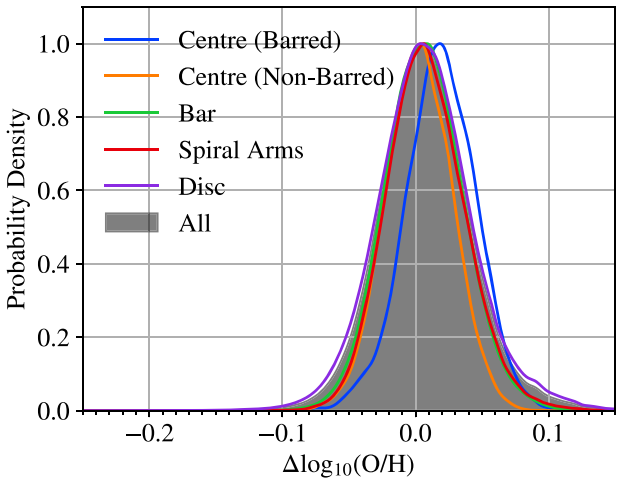


Figure 12. $\Delta\log_{10}(\mathrm{O/H})_{\mathrm{H}\alpha}$ for each environment for all 12 galaxies with significant higher-order variations. The overall distribution is shown as a shaded grey region, and the different environments with differently coloured lines. Each distribution is normalized to have a peak of 1. We separate the centres of barred and non-barred galaxies, as we may expect a different trend in abundance variation between barred and unbarred galaxies.

indicating that, when averaged across the entire sample, there are no clear overall enhancements within each environment (i.e. that the amplitude of the fitted GPR averages to zero over these larger scales, and not that the residual metallicity distribution is necessarily flat within each environment).

However, by averaging across the whole sample, we may be washing out variations between galaxies. We therefore repeat this exercise for each galaxy individually, calculating the KDE distribution of the radially subtracted metallicities for each environment, and weighting each measurement by the Hα intensity. We use the median of these distributions (i.e. the median of each environmental

Downloaded from https://academic.oup.com/mnras/article/509/1/1303/6414012 by guest on 01 July 2022

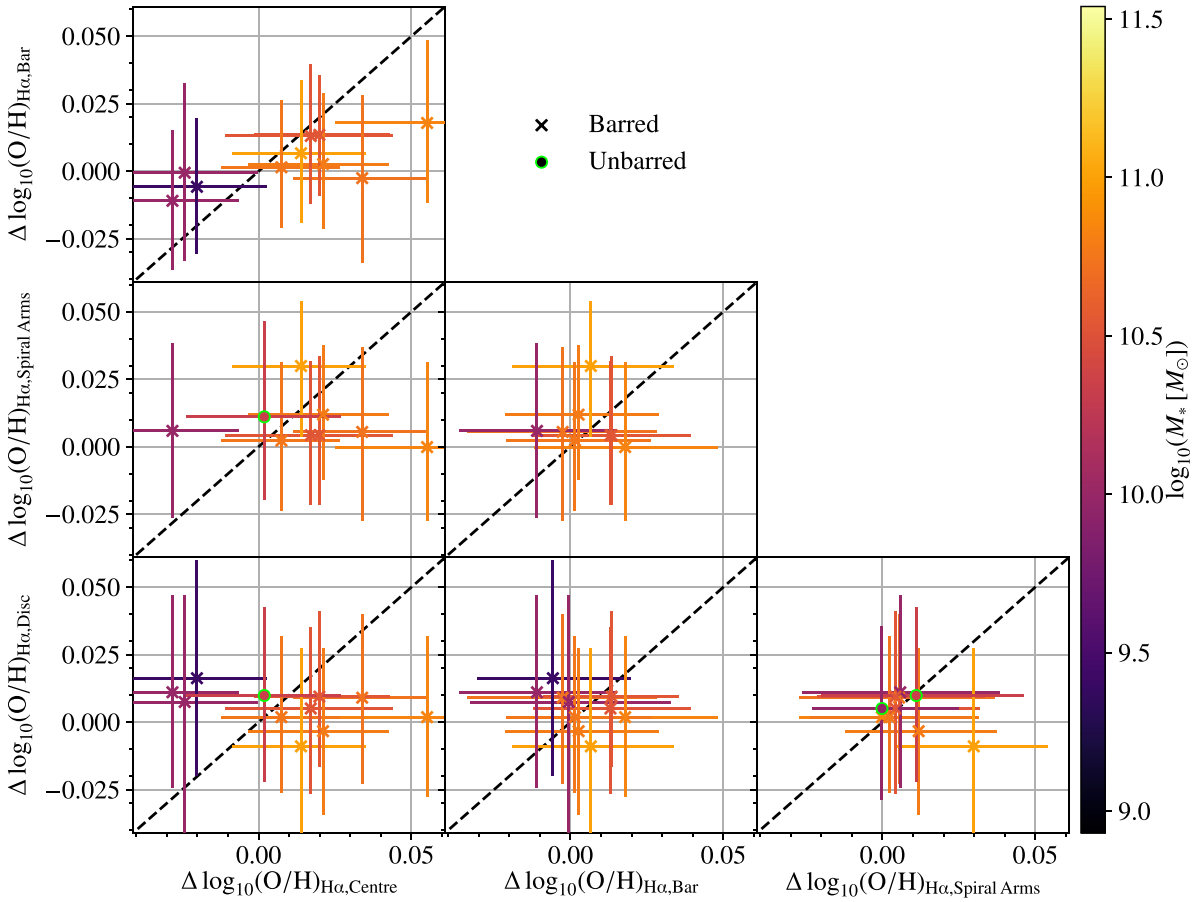


Figure 13. $\Delta\log_{10}(O/H)_{H\alpha}$ for various environments for the 12 galaxies with significant higher-order variations. The error bars represent the 16th and 84th percentiles of the KDE distribution, and the point marks the median. In many cases, the spread in values in galaxy centres is very narrow, and so the error bars are comparable to the size of the point. In each case, the point is coloured by the stellar mass of the galaxy, and we indicated barred galaxies with crosses, and unbarred with green-outlined circles. The 1:1 line is shown as a dashed-black line.

distribution in Fig. 12 for each galaxy) to indicate the average abundance offset in that particular environment, and the 16th and 84th percentiles to show the spread of values. We show this for each galaxy in Fig. 13. In this formalism, points that lie above the 1: 1 line have enhanced abundances in the *y*-axis environment with respect to the *x*-axis environment (and vice versa). We see that generally the $\Delta\log_{10}(O/H)_{H\alpha}$ values are consistent with 0, showing little indication of abundance variations within galaxies. In particular, the spiral arm and disc $\Delta\log_{10}(O/H)_{H\alpha}$ values are very close to zero (apart from NGC 1365 towards the bottom right of this subplot), contrary to the results of Sakhibov et al. (2018) and Sánchez-Menguiano et al. (2020), using similar definitions for the spiral arms and disc. Note that, although we include the outer disc, the FOV of the MUSE observations typically does not extend far into this regime. The only environment where we see clear abundance variations are in the centres of galaxies, which typically show enhanced metallicity with respect to the rest of the galactic environments, even after subtraction of the bulk radial metallicity gradient. This effect becomes more pronounced with stellar mass (the correlation between $\Delta\log_{10}(O/H)_{H\alpha,Centre}$ with $\log_{10}(M_*)$ is $\tau = 0.42 \pm 0.29$), with more massive galaxies having a more metal-enriched centre with respect to the rest of the galaxy than expected from their overall (linear) radial metallicity gradient. The fact that we see this is not surprising, as the metallicity gradient typically appears to be steeper towards the centres of galaxies (e.g. Sánchez et al. 2014; Sánchez-Menguiano et al. 2016), and can also be seen in the lower than expected CO-to-H_2 conversion factor seen in the centres of galaxies (Sandstrom et al. 2013).

5 WHAT DRIVES VARIATION IN THE ISM MIXING SCALE?

We may expect the global properties of the galaxy to encode some information about their mixing scale. For instance, more massive or earlier type galaxies may have more homogeneous metal distributions, and so may display a larger scale length (Li et al. 2021), perhaps because the majority of enrichment happened much earlier in the life cycle of the galaxy. Turbulence in the ISM can be injected via cosmological gas accretion (Klessen & Hennebelle 2010) or increased star formation (Krumholz et al. 2018), and this increased turbulence may increase metal mixing. We show the relationship between the two-point correlation scale and measures of the evolutionary state, star formation activity, and gas turbulence in Fig. 14. As in Section 3.3, we exclude the anomalously high value for NGC 1365. We normalize mass quantities by πR_e^2 (the effective radius that contains half the stellar mass of the galaxy), to give a disc-averaged surface density.

Downloaded from https://academic.oup.com/mnras/article/509/1/1303/6414012 by guest on 01 July 2022

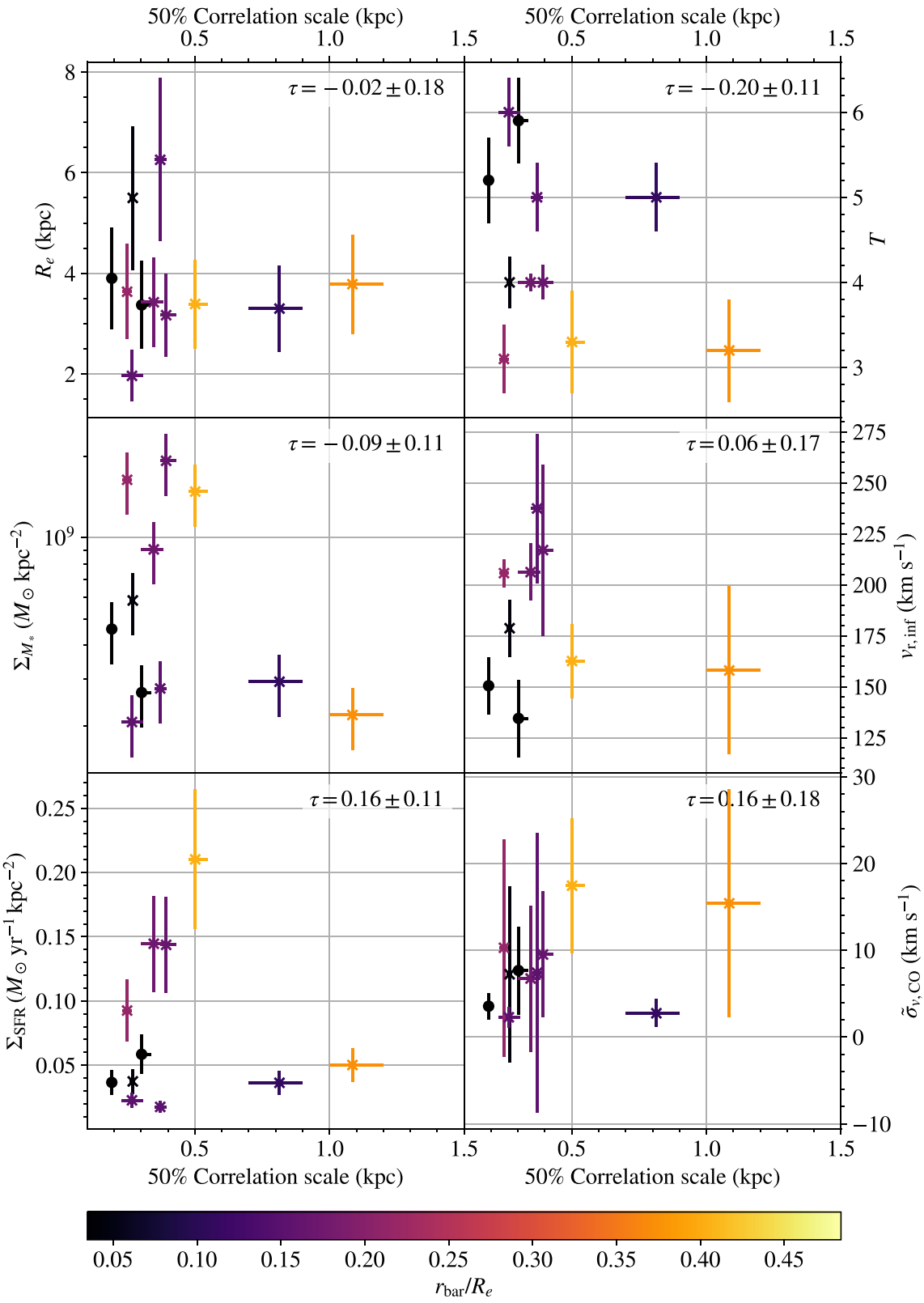


Figure 14. From top left to bottom right, the dependence on the 50 per cent two-point correlation scale with R_e, total stellar surface density (stellar mass normalized by πR_e^2), total star formation rate surface density, Hubble type T, asymptotic rotation velocity, and CO intensity-weighted CO velocity dispersion. In each case, Kendall's τ correlation coefficient is shown in the top right. Barred galaxies are indicated as crosses, and coloured according to their (normalized) bar length. Unbarred galaxies are black circles.

Downloaded from https://academic.oup.com/mnras/article/509/1/1303/6414012 by guest on 01 July 2022

We see a number of correlations with these quantities. First, galaxies with a higher Hubble type T typically have a lower 50 per cent correlation scale. This indicates that earlier type galaxies have a more homogeneous metal distribution. However, this is not echoed in the galaxy size, stellar mass surface density or total mass (from the asymptotic rotation velocity calculated by Lang et al. 2020). In Li et al. (2021), trends were found with the galaxy size, and we do not find this here (either with R_e, or r_{25} as used in their work). However, we do find the same trends with the stellar mass surface density as in Li et al. (2021). We also find a trend with the star formation rate surface density, unlike this earlier work. These results seem to indicate the correlation scales seen are driven by a combination of factors – first, the stage of a galaxy in its lifecycle, from the morphological type, and secondly its current star formation activity.

We also study how the correlation scale depends on the gas velocity dispersion. The stochastically forced diffusion model of Krumholz & Ting (2018) suggest that the correlation scale should strongly correlate with the gas velocity dispersion. We therefore compare our correlation scale to the CO velocity dispersions derived from the PHANGS–ALMA data (Leroy et al. 2021, we use the strict moment 0 map for intensity, and moment 2 map for velocity dispersion). We calculate a CO intensity-weighted CO velocity dispersion measured at ∼100 pc resolution across the whole CO map as an average measure of the cold gas velocity dispersion, and the 16th and 84th percentiles as a measure of the spread. Given the spectral resolution of the MUSE instrument, in many cases the Hα velocity dispersion is barely resolved and so we do not include it here, but with the higher spectral resolution of the ∼2.5 km s^{-1} ALMA data, we expect this to be less of an issue. This is shown in the bottom right panel of Fig. 14. We see no clear trend with gas velocity dispersion, which was also seen in Li et al. (2021) for a larger sample of galaxies. We also note that repeating this exercise using the coarser velocity resolution MUSE data gives identical results, and replacing 50 per cent correlation scales with σ_l show the same trends. These results are thus in contradiction with the Krumholz & Ting (2018) model. However, as we are averaging over the entire galaxy here, we may be missing some local effects. Indeed, Kreckel et al. (2020) found that in 2 kpc annular rings, the two-point correlation scales are strongly correlated with the gas velocity dispersion within that ring. Our analysis may therefore be washing out these local effects. Our results suggest the single largest global drivers of mixing are the morphological type T of the galaxy, and star formation activity, with mass, size, and gas turbulence playing little role in predicting the global metal mixing efficiency.

6 CONCLUSIONS

In this work, we have mapped the 2D variations of metals across the discs of 19 nearby galaxies. We have done this by calculating a 'per-pixel' metallicity at a common worst resolution of 120 pc, subtracting the dominant linear radial metallicity gradient, and then performing an interpolation using GPR. The key parameter that we extract from this fitting is σ_l, a characteristic kernel length scale that is indicative of the distance over which neighbouring measurements are highly correlated. We have performed a McJack procedure to correct these kernel length scales for incomplete coverage, and establish whether they are statistically significant. We find that, in our sample, 12 of our 19 galaxies show significant higher-order variations. Those that do not typically have many fewer metallicity measurements, and so this may be limited to the lack of H II regions in the galaxy, and not reflect a real lack of higher-order variations in the galaxy. Our measured kernel length scales range from 1.5 to 30 kpc, with a median value of 15 kpc. We compared our kernel length scales to the 50 per cent correlation scale from a two-point correlation function, and find them to be related, although the kernel length scales measured in this work are typically around an order of magnitude larger. The two-point correlation functions measured from our maps are similar to those measured for H II region catalogues, indicating the GPR fitting is sensitive to the same small-scale features as the two-point correlation function.

With these 12 galaxies, we have investigated how the residual (i.e. radially subtracted) metal enrichment varies with galactic environment (e.g. spiral arms, and bars). We have divided each galaxy up into centres, bars, spiral arms, and discs using the environmental masks from Querejeta et al. (2021). We see no clear signs of enrichment in any particular environment (e.g. spiral arms, and disc) when taking all galaxies as a whole, but see that centres are typically enriched (up to 0.05 dex higher, dependent on the total stellar mass of the galaxy) with respect to the linear radial metallicity gradient. We find no evidence that spiral arms are enriched compared to the disc, unlike recent work from Sánchez-Menguiano et al. (2020). Abundance variations within a particular environment that we fit with the GPR typically average out when considering the entire environment. However, azimuthal variations (Kreckel et al. 2019), or fluctuations that vary radially (Spitoni et al. 2019), would be missed by our analysis.

We have also looked at how the 50 per cent correlation scale from the two-point correlation function varies with different global galaxy parameters. Higher star formation rate surface density and lower Hubble type T have larger correlation scales. This may indicate that galaxies that are more evolved and with higher levels of star formation activity have mixed (or are mixing) their metals more efficiently. Unlike predictions from the model of Krumholz & Ting (2018), we find no significant correlation between the scale length and global gas velocity dispersion, in agreement with the findings of Li et al. (2021).

Whilst generally a second-order effect compared to the dominant radial metallicity gradient in galaxies, higher-order variations appear to be ubiquitous and non-negligible (with variations of up to 0.05 dex from the radial gradient, see Fig. 12) in star-forming spiral galaxies. With advanced statistical techniques and high-quality data, we have demonstrated that it is possible to measure and map these variations in a statistically robust way. These kinds of models will be suitable for comparison to the outputs of simulations, where metallicities are known locally, and can be compared to the effects of large-scale dynamical processes, such as bar mixing (e.g. Grand et al. 2016). Our interpolated metallicity maps also provide a minimally biased way to combine metallicities measurements from one observatory with observations at other wavelengths and resolutions, which probe different galactic properties, for resolved studies of, e.g. the dust-to-metals ratio (e.g. De Vis et al. 2019; Chiang et al. 2021), or studying the metallicity dependence on the CO-to-H_2 conversion factor, which is critical in obtaining reliable H_2 masses from CO data (e.g. Sandstrom et al. 2013). In moving from 1D first-order radial gradients to 2D higher-order variations, this work provides a stepping stone towards more realistic models of metal variations within galaxies.

ACKNOWLEDGEMENTS

This work has been carried out as part of the PHANGS collaboration. The authors would like to thank the anonymous reviewer, for comments and suggestions that have improved the quality of this paper.

Downloaded from https://academic.oup.com/mnras/article/509/1/1303/6414012 by guest on 01 July 2022

This work has made use of ASTROPY (Astropy Collaboration 2013; Price-Whelan et al. 2018), MATPLOTLIB (Hunter 2007), NUMBY (Harris et al. 2020), SCIPY (Virtanen et al. 2020), SCIKIT-LEARN (Pedregosa et al. 2011), and SEABORN (Waskom et al. 2017).

This work is based on observations collected at the European Southern Observatory under ESO programmes 1100.B-0651, 095.C-0473, and 094.C-0623, 094.B-0321, 099.B0242, 0100.B-0116, 098.B-0551, and 097.B-0640.

This paper makes use of the following ALMA data, which have been processed as a part of the PHANGS–ALMA CO($J = 2 - 1$) survey:

ADS/JAO.ALMA#2012.1.00650.S,
ADS/JAO.ALMA#2013.1.00803.S,
ADS/JAO.ALMA#2013.1.01161.S,
ADS/JAO.ALMA#2015.1.00121.S,
ADS/JAO.ALMA#2015.1.00782.S,
ADS/JAO.ALMA#2015.1.00925.S,
ADS/JAO.ALMA#2015.1.00956.S,
ADS/JAO.ALMA#2016.1.00386.S,
ADS/JAO.ALMA#2017.1.00392.S,
ADS/JAO.ALMA#2017.1.00766.S,
ADS/JAO.ALMA#2017.1.00886.L,
ADS/JAO.ALMA#2018.1.01321.S,
ADS/JAO.ALMA#2018.1.01651.S,
ADS/JAO.ALMA#2018.A.00062.S,
ADS/JAO.ALMA#2019.1.01235.S,
ADS/JAO.ALMA#2019.2.00129.S,

ALMA is a partnership of ESO (representing its member states), NSF (USA), and NINS (Japan), together with NRC (Canada), NSC, and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO, and NAOJ. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under a cooperative agreement by Associated Universities, Inc.

TGW, FS, H-AP, and ES acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 694343). KK gratefully acknowledges funding from the German Research Foundation (DFG) in the form of an Emmy Noether Research Group (grant number KR4598/2-1, PI Kreckel). FB acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 726384/Empire). MB acknowledges FONDECYT regular grant 1170618. MC and JMDK gratefully acknowledge funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through an Emmy Noether Research Group (grant number KR4801/1-1), as well as from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme via the ERC starting grant MUSTANG (grant agreement number 714907). EC acknowledges support from ANID project Basal AFB-170002. The work of AKL was partially supported by the National Science Foundation (NSF) under grant no. 1615105 and 1653300. ER acknowledges the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), funding reference number RGPIN-2017-03987. SCOG, EJW, and RSK acknowledge support from the DFG via the collaborative research center (SFB 881, Project-ID 138713538) 'The Milky Way System' (subprojects A1, B1, B2, B8, and P2). SCOG and RSK additionally acknowledge support from the Heidelberg Cluster of Excellence 'STRUCTURES' in the framework of Germany's Excellence Strategy (grant EXC-2181/1, Project-ID 390900948) and from the ERC via the ERC Synergy Grant 'ECOGAL' (grant 855130). MQ acknowledges support from the research project PID2019-106027GA-C44 from the Spanish Ministerio de Ciencia e Innovación. KS acknowledges funding support from National Science Foundation award no. 1816462. EWK acknowledges support from the Smithsonian Institution as a Submillimeter Array (SMA) Fellow.

DATA AVAILABILITY

The MUSE data used in this paper are presented in Emsellem et al. (2021). The ALMA data are presented in Leroy et al. (2021), and available online at https://www.canfar.net/storage/list/phangs/RELEASES/PHANGS-ALMA/. The code used in this paper is available at https://github.com/thomaswilliamsastro/metallicity_gpr. The metallicity maps (and associated error maps) can also be found in this repository.

REFERENCES

Anand G. S. et al., 2021, MNRAS, 501, 3621
Astropy Collaboration, 2013, A&A, 558, A33
Bacon R. et al., 2010, in McLean I. S., Ramsay S. K., Takami H., eds, SPIE Conf. Ser., Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III. SPIE, Bellingham, p. 773508
Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5
Barnes A. T. et al., 2021, MNRAS, submitted
Belfiore F., Vincenzo F., Maiolino R., Matteucci F., 2019, MNRAS, 487, 456
Belfiore F. et al., 2021, A&A, submitted
Boissier S., Prantzos N., 1999, MNRAS, 307, 857
Cappellari M., 2017, MNRAS, 466, 798
Cappellari M., Emsellem E., 2004, PASP, 116, 138
Cedrés B., Cepa J., 2002, A&A, 391, 809
Chevance M. et al., 2020, MNRAS, 493, 2872
Chiang I.-D. et al., 2021, ApJ, 907, 29
Clark C. J. R. et al., 2019, MNRAS, 489, 5256
De Vis P. et al., 2019, A&A, 623, A5
Di Matteo P., Haywood M., Combes F., Semelin B., Snaith O. N., 2013, A&A, 553, A102
Dopita M. A., Kewley L. J., Sutherland R. S., Nicholls D. C., 2016, Ap&SS, 361, 61
Emsellem E. et al., 2021, A&A, submitted, preprint (arXiv:2110.03708)
Erroz-Ferrer S. et al., 2019, MNRAS, 484, 5009
ESO CPL Development Team, 2015, Astrophysics Source Code Library. record ascl:1504.003
Espinoza N., Kossakowski D., Brahm R., 2019, MNRAS, 490, 2262
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Fragkoudi F., Di Matteo P., Haywood M., Schultheis M., Khoperskov S., Gómez A., Combes F., 2018, A&A, 616, A180
Fragkoudi F. et al., 2020, MNRAS, 494, 5936
Gadotti D. A. et al., 2019, MNRAS, 482, 506
González-Gaitán S., de Souza R. S., Krone-Martins A., Cameron E., Coelho P., Galbany L., Ishida E. E. O., COIN Collaboration, 2019, MNRAS, 482, 3880
Grand R. J. J. et al., 2016, MNRAS, 460, L94
Green G. M., Schlafly E., Zucker C., Speagle J. S., Finkbeiner D., 2019, ApJ, 887, 93
Haffner L. M. et al., 2009, Rev. Mod. Phys., 81, 969
Harris C. R. et al., 2020, Nature, 585, 357
Ho I. T., 2019, MNRAS, 485, 3569
Ho I. T., Kudritzki R.-P., Kewley L. J., Zahid H. J., Dopita M. A., Bresolin F., Rupke D. S. N., 2015, MNRAS, 448, 2030
Ho I. T. et al., 2017, ApJ, 846, 39
Ho I. T. et al., 2019, ApJ, 885, L31
Hogg D. W., Bovy J., Lang D., 2010, preprint (arXiv:1008.4686)
Hunt L. K., Hirashita H., 2009, A&A, 507, 1327

Downloaded from https://academic.oup.com/mnras/article/509/1/1303/6414012 by guest on 01 July 2022

Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Jones A. P., Köhler M., Ysard N., Bocchio M., Verstraete L., 2017, A&A, 602, A46
Kaplan K. F. et al., 2016, MNRAS, 462, 1642
Kauffmann G. et al., 2003, MNRAS, 346, 1055
Kewley L. J., Heisler C. A., Dopita M. A., Lumsden S., 2001, ApJS, 132, 37
Klessen R. S., Hennebelle P., 2010, A&A, 520, A17
Klessen R. S., Lin D. N., 2003, Phys. Rev. E, 67, 046311
Köhler M., Jones A., Ysard N., 2014, A&A, 565, L9
Kreckel K., Blanc G. A., Schinnerer E., Groves B., Adamo A., Hughes A., Meidt S., 2016, ApJ, 827, 103
Kreckel K. et al., 2019, ApJ, 887, 80
Kreckel K. et al., 2020, MNRAS, 499, 193
Krumholz M. R., Ting Y.-S., 2018, MNRAS, 475, 2236
Krumholz M. R., Burkhart B., Forbes J. C., Crocker R. M., 2018, MNRAS, 477, 2716
Kumari N., Maiolino R., Belfiore F., Curti M., 2019, MNRAS, 485, 367
Lang P. et al., 2020, ApJ, 897, 122
Leroy A. K. et al., 2021, ApJS accepted (arXiv:2104.07739)
Li Z., Krumholz M. R., Wisnioski E., Mendel J. T., Kewley L. J., Sánchez S. F., Galbany L., 2021, MNRAS, 504, 5496
Luridiana V., Morisset C., Shaw R. A., 2015, A&A, 573, A42
Martin P., Belley J., 1996, ApJ, 468, 598
Martin P., Roy J.-R., 1992, ApJ, 397, 463
O'Donnell J. E., 1994, ApJ, 422, 158
Osterbrock D. E., Ferland G. J., 2006, Astrophysics of gaseous nebulae and active galactic nuclei. University Science Books
Pastorello N., Forbes D. A., Foster C., Brodie J. P., Usher C., Romanowsky A. J., Strader J., Arnold J. A., 2014, MNRAS, 442, 1003
Pedregosa F. et al., 2011, J. Mach. Learn. Res., 12, 2825
Pilyugin L. S., Grebel E. K., 2016, MNRAS, 457, 3678
Poetrodjojo H. et al., 2018, MNRAS, 479, 5235
Poetrodjojo H., D'Agostino J. J., Groves B., Kewley L., Ho I. T., Rich J., Madore B. F., Seibert M., 2019, MNRAS, 487, 79
Price-Whelan A. M. et al., 2018, AJ, 156, 123
Prsa A., Hambleton K. M., 2017, in American Astronomical Society Meeting Abstracts, Vol. 229. p. 344.09
Querejeta M., et al., 2021, A&A, accepted
Rasmussen C., Williams C., 2006, Gaussian Processes for Machine Learning, Adaptive Computation and Machine Learning. MIT Press, Cambridge, MA, USA, p. 248
Rosales-Ortega F. F., Díaz A. I., Kennicutt R. C., Sánchez S. F., 2011, MNRAS, 415, 2439
Sakhibov F., Zinchenko I. A., Pilyugin L. S., Grebel E. K., Just A., Vílchez J. M., 2018, MNRAS, 474, 1657
Sánchez S. F. et al., 2014, A&A, 563, A49
Sánchez S. F. et al., 2015, A&A, 573, A105
Sánchez-Menguiano L. et al., 2016, A&A, 587, A70
Sánchez-Menguiano L., Sánchez S. F., Pérez I., Ruiz-Lara T., Galbany L., Anderson J. P., Kuncarayakti H., 2020, MNRAS, 492, 4149
Sandstrom K. M. et al., 2013, ApJ, 777, 5
Santoro F. et al., 2021, A&A, submitted
Sarzi M. et al., 2018, A&A, 616, A121
Searle L., 1971, ApJ, 168, 327
Silverman B. W., 1986, Density Estimation For Statistics And Data Analysis / B. W. Silverman,..., Monographs on statistics and applied probability. Chapman and Hall CRC Press, London Glasgow Weinheim [etc.], Boca Raton
Spitoni E., Cescutti G., Minchev I., Matteucci F., Silva Aguirre V., Martig M., Bono G., Chiappini C., 2019, A&A, 628, A38
Thilker D. A., Braun R., Walterbos R. A. M., 2000, AJ, 120, 3070
Vale Asari N., Couto G. S., Cid Fernandes R., Stasińska G., de Amorim A. L., Ruschel-Dutra D., Werle A., Florido T. Z., 2019, MNRAS, 489, 4721
Vazdekis A., Koleva M., Ricciardelli E., Röck B., Falcón-Barroso J., 2016, MNRAS, 463, 3409
Virtanen P. et al., 2020, Nat. Methods, 17, 261
Waskom M. et al., 2017, mwaskom/seaborn: v0.8.1 (september 2017)
Weilbacher P. M. et al., 2020, A&A, 641, A28
Yang C.-C., Krumholz M., 2012, ApJ, 758, 48

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Appendix A. GPR Fits for all PHANGS–MUSE Galaxies.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

APPENDIX A: GPR FITS FOR ALL PHANGS–MUSE GALAXIES

Here, we show the equivalent of Fig. 4 for all 19 galaxies in our sample.

APPENDIX B: HII REGION TWO-POINT CORRELATION FUNCTION

We revisit the two-point correlation function calculation from Kreckel et al. (2020), based on the H II region positions and metallicities, for the full sample of 19 galaxies. This analysis makes use of the H II region catalogue presented in Santoro et al. (2021) and metallicity measurements presented in Groves et al. (in preparation), and is thus an update on the values previously published for eight galaxies in Kreckel et al. (2020) that were based on the H II region catalogue in Kreckel et al. (2019). In addition to more than doubling the galaxy sample, the updated H II region catalogue is based on a new data reduction, which includes a more accurate sky subtraction amongst other improvements (Emsellem et al. 2021). The H II region identification algorithm has also been refined. As a result of these improvements, the newer catalogues contain 20–30 per cent more H II region detections per galaxy.

We select H II regions all nebulae that meet the following criteria:

(i) S/N $>$ 5 in all strong lines (Hβ, [O III]λ5006, [N II]λ6583, Hα, [S II]$\lambda\lambda$6716,30) that are used as diagnostics for the ionization source or the metallicity calculation;

(ii) Line ratios consistent with photoionzation in the BPT diagrams using the Kauffmann et al. (2003) diagnostic in the [N II]λ6583 diagram, the Kewley et al. (2001) diagnostic in the [S II]$\lambda\lambda$6716,30 diagram;

(iii) Spatial separation by more than the PSF from any bright foreground stars or the field edges;

(iv) Hα velocity dispersion $<$ 60 km s^{-1}, to remove supernova remnant contaminants and spurious fits.

The resulting catalogue consists of a total of 23 436 regions, and identifies between 476 and 2355 H II regions per galaxy.

Metallicities are computed by adopting the Pilyugin & Grebel (2016) Scal, and removing a linear radial gradient as fit in Santoro et al. (2021). In computing the H II region separations, distances are deprojected within each galaxy and computed assuming the position angle and inclination, as listed in Table 1.

Within each galaxy, we calculate the two-point correlation of metals (ξ) as a function of spatial scale (r) using equation 18. The resulting two-point correlation functions are shown in Fig. B1. We estimate the uncertainty in our measured function by performing 100 random samples of our uncertainties in 12 + log$_{10}$(O/H), and repeating our analysis. The 1σ distribution is determined at each spatial scale, though it is generally thinner than the line

Downloaded from https://academic.oup.com/mnras/article/509/1/1303/6414012 by guest on 01 July 2022

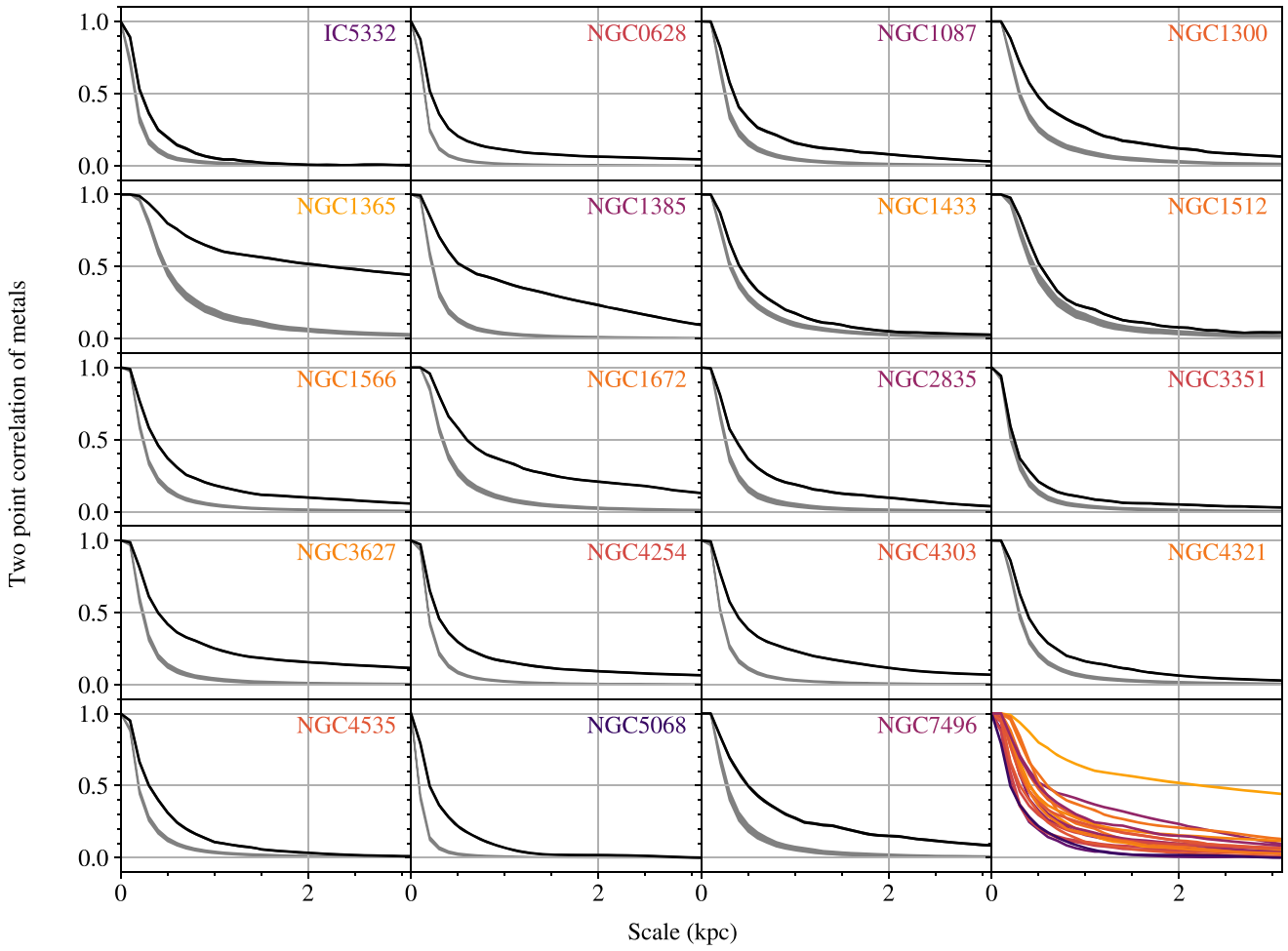


Figure B1. Two-point correlation of metals for each of the 19 PHANGS–MUSE galaxies. Galaxies colours are ordered by their total stellar mass, from low (purple) to high (orange). We estimate the uncertainty in our measured function by performing 100 random samples of our uncertainties in $12 + \log_{10}$(O/H), and repeating our analysis. A band is plotted covering the 1σ distribution at each spatial scale, though it is generally not visible as it is thinner than the width of the line (<10 pc). To determine the significance of our two-point correlation functions, we assume the null hypothesis (that all H II regions are perfectly uncorrelated) by randomly shuffling (for 100 actualizations) all measurements of $12 + \log_{10}$(O/H) across each galaxy and repeating our analysis. This is shown by the grey bands.

drawn (<10 pc). To determine the significance of our two-point correlation functions, we assume the null hypothesis (that all H II regions are perfectly uncorrelated) by randomly shuffling (for 100 actualizations) all measurements of $12 + \log_{10}$(O/H) across each galaxy and repeating our analysis. This is shown by the grey bands, and in all galaxies the two-point correlation functions are measured at $>2\sigma$ out to kpc scales.

We quantify the 30 per cent and 50 per cent correlation scales for each galaxy, along with the 1σ uncertainties, with all values listed in Table B1. A direct comparison with the values reported in Kreckel et al. (2020) are shown in Fig. B2, and reveal very good systematic agreement for the eight overlapping galaxies.

NGC 1365 shows remarkably high correlations out to the largest scale we test (5 kpc). NGC 3351 and NGC 1512 show the least significant difference from the null hypothesis, with only $\sim 2\sigma$ difference at the 30 per cent and 50 per cent correlation scales. Measurements for all other galaxies are detected at $>6\sigma$ levels. As shown in the bottom right corner of Fig. B1, only a weak trend with galaxy stellar mass (indicated by the line colours) is apparent.

Downloaded from https://academic.oup.com/mnras/article/509/1/1303/6414012 by guest on 01 July 2022

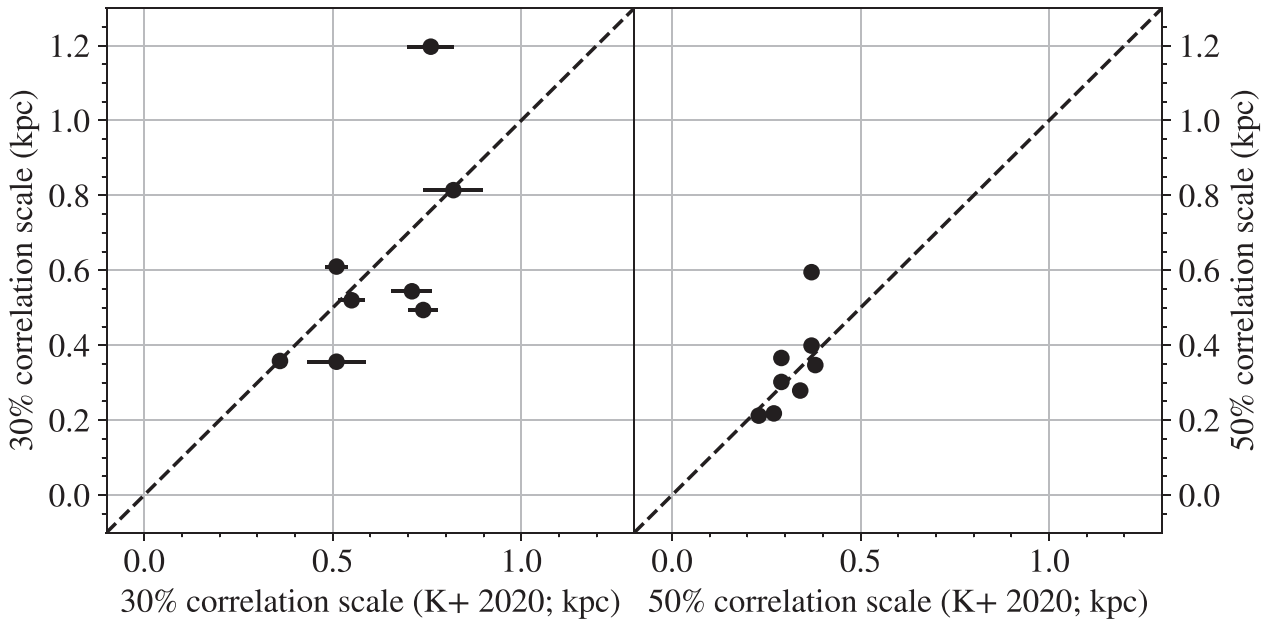


Figure B2. Comparison of the 30 per cent and 50 per cent correlation scales for the eight galaxies in Kreckel et al. (2020) with the measurements provided in this work (y-axis). Error bars reflect the 1σ uncertainties, which in most cases are smaller than the points. NGC 1672 shows the most discrepant values between the two H II region catalogues, with the remaining galaxies showing good agreement.

Table B1. Correlation scales measured for the 19 galaxies.

Galaxy	50 per cent correlation scale (pc)	30 per cent correlation scale (pc)
IC5332	218 ± 4	356 ± 5
NGC 0628	212 ± 1	358 ± 2
NGC 1087	347 ± 3	544 ± 8
NGC 1300	476 ± 5	870 ± 16
NGC 1365	2247 ± 29	>5000
NGC 1385	561 ± 10	1529 ± 17
NGC 1433	412 ± 3	666 ± 7
NGC 1512	525 ± 5	752 ± 7
NGC 1566	368 ± 3	615 ± 5
NGC 1672	595 ± 3	1197 ± 9
NGC 2835	366 ± 4	610 ± 8
NGC 3351	242 ± 2	380 ± 4
NGC 3627	399 ± 2	814 ± 6
NGC 4254	279 ± 1	494 ± 4
NGC 4303	367 ± 2	695 ± 5
NGC 4321	377 ± 1	585 ± 3
NGC 4535	302 ± 4	520 ± 5
NGC 5068	199 ± 1	379 ± 3
NGC 7496	497 ± 7	905 ± 17

[1]*Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*
[2]*Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, Mönchhofstraße 12-14, D-69120 Heidelberg, Germany*
[3]*INAF - Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, I-50157 Firenze, Italy*
[4]*International Centre for Radio Astronomy Research, University of Western Australia, 7 Fairway, Crawley 6009, WA, Australia*
[5]*Center for Astrophysics and Space Sciences, Department of Physics, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093, USA*
[6]*The Observatories of the Carnegie Institution for Science, 813 Santa Barbara Street, Pasadena, CA 91101, USA*
[7]*Departamento de Astronomía, Universidad de Chile, Camino del Observatorio 1515 Las Condes, Santiago, Chile*
[8]*Argelander-Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany*
[9]*Universidad de Antofagasta, Centro de Astronomía, Avenida Angamos 601, Antofagasta 1270300, Chile*
[10]*Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, Mönchhofstraße 12-14, D-69120 Heidelberg, Germany*
[11]*European Southern Observatory, Karl-Schwarzschild-Str. 2, D-85748 Garching, Germany*
[12]*Université Lyon 1, ENS de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69230 Saint-Genis-Laval, France*
[13]*Universität Heidelberg, Zentrum für Astronomie Heidelberg, Institut für Theoretische Astrophysik, Albert-Ueberle-Str. 2, D-69120 Heidelberg, Germany*
[14]*Research School of Astronomy and Astrophysics, Australian National University, Weston Creek, ACT 2611, Australia*
[15]*Universität Heidelberg, Interdisziplinäre Zentrum für Wissenschaftliches Rechnen, Im Neuenheimer Feld 205, D-69120 Heidelberg, Germany*
[16]*Center for Astrophysics | Harvard & Smithsonian, 60 Garden St., Cambridge, MA 02138, USA*
[17]*Department of Astronomy, The Ohio State University, 140 West 18th Avenue, Columbus, OH 43210, USA*
[18]*Max-Planck-Institut für extraterrestrische Physik, Giessenbachstraße 1, D-85748 Garching, Germany*
[19]*Sterrenkundig Observatorium, Universiteit Gent, Krijgslaan 281 S9, B-9000 Gent, Belgium*
[20]*Department of Physics, Tamkang University, Tamsui Dist., New Taipei City 251301, Taiwan*

Downloaded from https://academic.oup.com/mnras/article/509/1/1303/6414012 by guest on 01 July 2022

[21]*Observatorio Astronómico Nacional (IGN), C/Alfonso XII 3, Madrid E-28014, Spain*
[22]*Department of Physics, University of Alberta, Edmonton, Alberta T6G 2E1, Canada*
[23]*Department of Physics, General Studies, College of Engineering, Nihon University, 1 Nakagawara, Tokusada, Tamuramachi, Koriyama, Fukushima 963-8642, Japan*
[24]*National Astronomical Observatory of Japan, 221-1 Osawa, Mitaka, Tokyo 181-8588, Japan*
[25]*Departamento de Física de la Tierra y Astrofísica, Facultad de CC Físicas, Universidad Complutense de Madrid, E-28040 Madrid, Spain*

This paper has been typeset from a TEX/LATEX file prepared by the author.

Downloaded from https://academic.oup.com/mnras/article/509/1/1303/6414012 by guest on 01 July 2022