A Comparative Study of Mid-infrared Star Formation Rate Tracers and Their Metallicity Dependence

C. M. Whitcomb1, K. Sandstrom1, E. J. Murphy2, and S. Linden3

1 Center for Astrophysics and Space Sciences, Department of Physics, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA
coryw777@gmail.com
2 National Radio Astronomy Observatory, 520 Edgemont Road, Charlottesville, VA 22903, USA
3 Department of Astronomy, University of Virginia, 530 McCormick Road, Charlottesville, VA 22904, USA

Received 2020 May 31; revised 2020 July 31; accepted 2020 August 11; published 2020 September 21

Abstract

We present a comparative study of a set of star formation rate (SFR) tracers based on mid-infrared emission in the 12.81 μm [Ne II] line, the 15.56 μm [Ne III] line, and emission features from polycyclic aromatic hydrocarbons (PAHs) between 5.2 and 14.7 μm. We calibrate our tracers with the thermal component of the radio continuum emission at 33 GHz from 33 extranuclear star-forming regions observed in nearby galaxies. Correlations between mid-IR emission features and thermal 33 GHz SFRs show significant metallicity-dependent scatter and offsets. We find similar metallicity-dependent trends in commonly used SFR tracers such as Hα and 24 μm. As seen in previous studies, PAH emission alone is a poor SFR tracer owing to a strong metallicity dependence: lower-metallicity regions show decreased PAH emission relative to their SFR compared to higher-metallicity regions. We construct combinations of PAH bands, neon emission lines, and their respective ratios to minimize metallicity trends. The calibrations that most accurately trace SFR with minimal metallicity dependence involve the sum of the integrated intensities of the 12.81 μm [Ne II] line and the 15.56 μm [Ne III] line combined with any major PAH feature normalized by dust continuum emission. This mid-IR calibration is a useful tool for measuring SFR, as it is minimally sensitive to variations in metallicity and is composed of bright, ubiquitous emission features. The Mid-Infrared Instrument (MIRI) on the James Webb Space Telescope will detect these features from galaxies as far as redshift z ~ 1. We also investigate the behavior of the PAH band ratios and find that subtracting the local background surrounding a star-forming region decreases the ratio of PAH 11.3 μm to 7.7 μm emission. This implies that PAHs are more ionized in star-forming regions relative to their surroundings.

Unified Astronomy Thesaurus concepts: Star formation (1569); Star forming regions (1565); Polycyclic aromatic hydrocarbons (1280); Metallicity (1031); Radio continuum emission (1340)

1. Introduction

The rate at which new stars are formed is a fundamental property that influences how a galaxy evolves. Many methods to measure the star formation rate (SFR) in galaxies at low and high redshifts have been employed with various complications present in each (for a review, see Kennicutt & Evans 2012). One of the main complications is interstellar dust that pervades star-forming regions. This dust protects molecular gas from ionizing or photodissociating radiation, but it also prevents such radiation from leaving the cloud. Absorbed ultraviolet and optical light is reemitted in the infrared by a complex population of dust grains, revealing their physics and that of their environment (e.g., Draine 2003; Tielens 2008, and references therein). This often necessitates hybrid, multiwavelength SFR tracers that include infrared emission from dust in addition to the more direct UV and optical tracers of massive star formation (SF; e.g., Buat 1992; Meurer et al. 1995; Calzetti et al. 2007; Kennicutt et al. 2007; Cortese et al. 2008; Kennicutt et al. 2009; Murphy et al. 2011; Leroy et al. 2012). Hybrid tracers that use both UV/optical and infrared emission such as Hα and 24 μm require observations on multiple telescopes. It is therefore of significant interest to find a tracer or multiple tracers that can predict SFR from a single spectral observation; for this purpose, emission features present in the mid-IR may be ideal.

The mid-infrared spectrum hosts many strong emission lines and broad emission bands that have the potential to be SFR tracers. Among the brightest of these features are the forbidden lines [Ne II] at 12.81 μm and [Ne III] at 15.56 μm. These lines contribute a significant portion of the cooling in H II regions as relatively abundant ionized species with lines that can be easily collisionally excited (Burbidge et al. 1963; Osterbrock 1965). Both emission lines have a low-energy spacing and a high critical density, so their emissivity varies little in typical H II regions. The ionization potentials of Ne I → Ne II and Ne II → Ne III are greater than that of hydrogen at 21.56 and 40.96 eV, respectively, so both [Ne II] emission and [Ne III] emission trace highly ionized gas. Ne III can be the dominant ion at low densities and in harder radiation fields (e.g., Gieou et al. 2002). Ne IV is not abundant in H II regions since it requires more energetic photons than O and B stars can produce. Combining [Ne II] and [Ne III] emission from H II regions thus traces the total emission from ionized neon, which traces some fraction of the overall cooling luminosity. Given a steady-state balance between heating and cooling, the neon emission should therefore trace SFR to some degree. The efficacy of SFR tracers based on ionized neon emission has been studied in several previous publications (Ho & Keto 2007; Treyer et al. 2010; Zhuang et al. 2019).

Also seen in the mid-IR spectra of galaxies are broad features that originate from a class of molecules called polycyclic aromatic hydrocarbons (PAHs; e.g., Roche et al. 1991; Helou et al. 2000; Peeters et al. 2002; Lu et al. 2003; Smith et al. 2007). PAHs emit strong features relative to the infrared thermal continuum between 3.3 and 17 μm owing to
vibrations in their carbon and hydrogen bonds when exposed to UV/optical light (Leger & Puget 1984; Sellgren 1984; Allamandola et al. 1985; Draine & Anderson 1985; Allamandola et al. 1989; Schutte et al. 1993; Draine & Li 2001). The efficiency with which PAHs absorb and reradiate UV/optical light suggests that they are ideal candidates for components of an SFR tracer, and many such calibrations exist (e.g., Roussel et al. 2001; Förster Schreiber et al. 2004; Peeters et al. 2004; Wu et al. 2005; Shipley et al. 2016; Cluver et al. 2017; Xie & Ho 2019).

However, PAH emission traces only the light that has been absorbed, so a PAH-based tracer would not trace the ionized gas and young stars directly and would include potential biases if the attenuation of UV/optical light varies with environment. In particular, it has been demonstrated that the properties of PAH emission are dependent on several environmental factors such as metallicity and radiation field hardness (Engelbracht et al. 2005; Madden et al. 2006; O’Halloran et al. 2006; Engelbracht et al. 2008), which must be accounted for to produce a reliable SFR tracer (e.g., Calzetti et al. 2007; Treister et al. 2010; Wen et al. 2014; Shipley et al. 2016; Shivaei et al. 2017; Xie & Ho 2019). In accounting for environmental variations, ratios of PAH bands and PAH-to-continuum ratios may be useful tools. Ratios of PAH bands such as $11.3\mu m$ to $7.7\mu m$ reveal properties of the local PAH population such as size and ionization (e.g., Bakes et al. 2001; Peeters et al. 2002, 2004; Bauschlicher et al. 2008). Similarly, ratios of PAH bands to mid-IR dust continuum measure the amount of emission from PAHs relative to that of the entire dust population, tracing their abundance. For instance, photometry at $8\mu m$ (dominated by the PAH feature) and the continuum measured at $24\mu m$ have often been used to trace PAH abundance (Engelbracht et al. 2005, 2008; Draine et al. 2007).

To fully escape the complications related to dust attenuation in an SFR tracer, one must move to long wavelengths. Thermal 33 GHz radio continuum (T33) originates from free–free emission and is directly proportional to the amount of ionizing radiation from young stars (e.g., Condon 1992; Murphy et al. 2011; Kennicutt & Evans 2012). This combination of independence from the presence of interstellar dust and direct correlation with radiation from young stars makes T33 emission an excellent standard of reference for SFR. Emission at 33 GHz, however, is faint even in local sources and difficult to detect in distant galaxies. Systematic observations of the 3–33 GHz continuum have been made as part of the Star Formation in Radio Survey (SFRS; Murphy et al. 2018; Linden et al. 2020) for 50 nearby galaxies in the Spitzer Infrared Nearby Galaxies Survey (SINGS; Kennicutt et al. 2003) and Key Insights on Nearby Galaxies: A Far-Infrared Survey with Herschel (KINGFISH; Kennicutt et al. 2011).

In the following, we use T33 as a standard of reference to calibrate SFR tracers with emission in the mid-infrared from neon and PAHs. We investigate the correlation between the strengths of these mid-IR features and T33 to find a tracer insensitive to absorption by interstellar dust (in normal galaxies with low mid-IR optical depth) that will be easily detected with the instruments on the James Webb Space Telescope (JWST). The $15.56\mu m$ [Ne II] line has the longest wavelength of the mid-IR features considered in this work. Thus, the Mid-Infrared Instrument (MIRI) on JWST will detect the same emission lines as the SINGS spectra described in Section 2.3 in galaxies out to redshift $z \approx 0.85$. Restricting the maximum wavelength further to include only emission below the $12.81\mu m$ [Ne II] line allows observations at $z \approx 1.25$ with MIRI.

An SFR tracer that is insensitive to metallicity is crucial for application to galaxies at higher redshifts, as the average metallicity of galaxies changes with redshift (Gallazzi et al. 2008; Madau & Dickinson 2014). In addition, the optical diagnostic lines needed to measure metallicity may or may not be available for higher-redshift targets. Therefore, to construct an SFR calibration applicable to a range of redshifts observed with JWST, we aim to ensure that any metallicity dependence is accounted for. To this end, we directly study the metallicity dependence of each SFR calibration using local measurements of the metallicity. The local metallicity of each region is determined from metallicity gradients as a function of galactocentric radius measured from H II region spectra (e.g., Moustakas et al. 2010, in our case).

The paper is organized as follows: Section 2 outlines the SINGS, SFRS, and metallicity data used in this work and the methods of our analysis. Sections 3.1 and 3.2 present the results of our correlations between the SINGS mid-IR features and SFRS T33 measurements. Section 3.3 presents a series of mid-infrared SFR tracers and comparisons of our methods against established calibrations, including the T33 tracer and the $H\alpha–24\mu m$ tracer from Murphy et al. (2011), the PAH tracer from Shipley et al. (2016), and the neon-metallicity tracer from Zhuang et al. (2019). Section 4 describes the implications of these results and novel patterns seen therein. Section 5 summarizes our key conclusions and explores potential directions for future research.

2. Observations and Analysis

2.1. Sample

The basis of our work consists of 33 GHz measurements from the National Science Foundation’s Karl G. Jansky Very Large Array (VLA) as part of SFRS (Murphy et al. 2018; Linden et al. 2020). Our full sample is the overlap between SFRS 7″-diameter aperture 33 GHz observations and regions observed in SINGS SL and SH spectral maps. We find that 56 regions from SFRS have full 7″ aperture coverage in SINGS SL data, and 52 of these were found to have full-aperture coverage in SH data.

Our key goals in this study require knowledge of the thermal component of the 33 GHz emission and a local measurement of the metallicity. There are 43 of the 56 regions in the SFRS-SINGS overlap that have measured thermal fractions. Of those, 33 have local metallicity measurements (see Section 2.4). These 33 regions form the basis of our SFR calibrations in Sections 3.1 and 3.2. In the comparisons to literature SFR calibrations in Section 3.3, we use the largest possible subset of the data, which will be noted in the text. In Section 3.4 we investigate the PAH band ratios separately from the SFR calibrations, and therefore we use a subset of 49 regions on which the only requirement is full 11″ aperture coverage in SINGS SL mapping.

---

4 The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.
2.2. SFRS Photometry

SFRS observed 33 GHz radio continuum emission with the VLA at ~2″ resolution in 50 local galaxies (<30 Mpc) from the SINGS (Kennicutt et al. 2003) and KINGFISH (Kennicutt et al. 2011) surveys. The resulting maps were then convolved to 7″ Gaussian resolution to match the resolution of SINGS 24 μm images. The 33 GHz measurements were extracted from 7″-diameter apertures in these convolved maps and are listed in Table 1. This table also lists the location of each aperture, its galactocentric radius, and emission from Hα and 24 μm that we adopt from Murphy et al. (2018). Hα and 24 μm emission is converted to 10⁻²⁵ erg s⁻¹ cm⁻² sr⁻¹ assuming a 7″ aperture in order to facilitate comparison with mid-IR emission results described in Section 2.6. We convert 24 μm continuum to M_L units by multiplying by 12.5 THz (i.e., 24 μm in terms of frequency) for ease of comparison with line-integrated intensities. The thermal component of 33 GHz emission (T33) for each region was obtained from Linden et al. (2020), where fractions were determined by fitting radio spectra with a synchrotron emission component (nonthermal) and a free–free emission component (thermal). There are 43 of the 56 regions in our SFRS-SINGS overlap sample that have thermal fractions at 33 GHz available. We convert the T33 flux densities into surface brightness units by dividing by the solid angle of the 7″-diameter aperture.

2.3. SINGS Spectroscopy

The mid-infrared data used in this work are from the fifth release of the Spitzer Infrared Nearby Galaxies Survey (SINGS DR5; Kennicutt et al. 2003; Dale et al. 2006), which mapped nuclear and extranuclear regions of nearby galaxies with the Infrared Spectrometer (IRS; Houck et al. 2004) on the Spitzer Space Telescope (Werner et al. 2004). The SINGS DR5 spectral flux and uncertainty FITS files for extranuclear regions were downloaded from the Infrared Science Archive (IRSA) hosted by the Infrared Processing and Analysis Center (IPAC). Spectra in each common aperture between the SINGS and SFRS data were extracted from the short–low SL1 and SL2 (5.2–14.7 μm) modules. Additional spectra were extracted from SINGS data in the IRS short–high (SH) module (9.9–19.4 μm) for the 52 regions in our sample where observations are available. The extracted spectra are shown in Figure 1 ordered by their 12.81 μm [Ne II] line strength.

2.4. Ancillary Data

Metallicities for the regions in this sample were determined using data from Moustakas et al. (2010). This work presents metallicity gradients and central values in both the KK04 (Kobulnicky & Kewley 2004) and PT05 (Pilyugin & Thuan 2005) strong-line calibrations for galaxies in SINGS. These values are shown in Table 2 for the galaxies in our sample. Our investigations into the metallicity dependence of mid-IR emission features rely on having a local measurement of the metallicity at the galactocentric radius of each star-forming region. For this purpose, we select the galaxies that have a measured H II region metallicity gradient or a reported average metallicity if the galaxy is a dwarf (<10⁹ M⊙). Of the 56 SFRS-SINGS overlap regions, 49 have local metallicities. For dwarf galaxies, which are expected to have small or negligible gradients, we adopt the characteristic oxygen abundance if no gradient is measured, and for those without either a gradient or characteristic abundance we use values calculated from the luminosity–metallicity (L–Z) relation. These regions are used in our investigation into metallicity correlations in Appendix B. In general, we use the KK04 calibration throughout; changing to the PT05 calibration minimally affects the values of some of our fits and does not alter the qualitative results of our study. Table 2 also lists the 25th-magnitude isophotal radii (R25) from Moustakas et al. (2010), distances from Murphy et al. (2018), heliocentric velocities from NED,5 and masses from Kennicutt et al. (2011) or other sources that we adopt for these galaxies.

2.5. Matched Aperture Extraction

Spectra from the SINGS SL1, SL2, and SH spectral cubes were extracted in matched 7″-diameter apertures identical to those used for the 33 GHz measurements (Murphy et al. 2018). We remove regions where the 7″ aperture does not completely fall within the coverage of the SINGS data. The aperture photometry module of the Python astropy photutils package was used at each wavelength of the Spitzer SL and SH cubes to extract mid-IR spectra for each region (Astropy Collaboration et al. 2013, 2018). This function measures the flux density within the complete circular aperture by using exact fractions for partially enclosed pixels, and we repeat this for each wavelength in the data cube. We then divide by the total solid angle of the aperture to obtain spectra in units of MJy sr⁻¹. Spectral uncertainties are also propagated by the aperture photometry function from the corresponding SINGS uncertainty cube.

We do not perform any point-spread function (PSF) matching across wavelength in the IRS cubes prior to aperture photometry. In the wavelength range of interest for our study, the PSF of IRS spectral mapping observations is found to oscillate around ~3″–4″ FWHM for SL and ~4″–5″ FWHM for SH (Pereira-Santaella et al. 2010). The FWHM of the PSF is always smaller than our extraction aperture. Given the small observed variations in FWHM as a function of wavelength, PSF matching would not greatly affect our measurements, and given the uncertainties in the convolution kernels for IRS, we chose to proceed without performing any PSF matching.

We produced a second SL data set with the local background removed by subtracting the average background level determined in an annulus surrounding the circular aperture. At aperture scales of radius 3.5″ these regions are <360 pc in physical radius, and the majority of the emission from each star-forming region is enclosed. We investigate the effect of local background subtraction for these regions using the surface brightness in an annulus between radius 3.5″ and 5.5″ surrounding the original 7″-diameter aperture. The surface brightness at each wavelength in the background annulus is then averaged and subtracted from that of the inner aperture. We exclude regions that do not have full SL data coverage in the background annulus, leaving 49 of the 56 regions in the background-subtracted data set. In Section 3.4 we use this additional data set to study the effect of local background subtraction on PAH band ratios.

5 The NASA/IPAC Extragalactic Database (NED) is funded by the National Aeronautics and Space Administration and operated by the California Institute of Technology.
2.6. Measuring Integrated Intensity of PAH Bands and Emission Lines

We analyzed the SINGS SL spectra using the IDL program PAHFIT to determine the strengths and uncertainties of emission features from PAHs and ions (Smith et al. 2007).

Spectral lines from ions such as $12.81 \mu m$ [Ne II] and pure rotational lines from molecular hydrogen such as $12.28 \mu m$ H$_2$(S(2)) were fit with Gaussian profiles, while PAH features were modeled with the Drude profile of a classical damped harmonic oscillator. We input the redshift of the galaxies to...
obtain the most accurate PAHFIT measurements using values listed in Table 2. We also allowed a fit to the optical depth of the silicate absorption feature at 9.7 μm but find no significant detection in any region. A selection of PAHFIT results are shown in Table 3. We obtained measurements for some emission lines and PAH features that are not included in this

Figure 1. Extracted spectra from SL wavelengths 5.2–14.7 μm (left) and SH wavelengths 9.9–19.4 μm (right). Shown are the 28 brightest of the 56 spectra in order of 12.81 μm [Ne II] emission measurements from PAHFIT on the SL spectra. Regions with no SH data are shown as a horizontal line in the right panel.
NGC 2403c 8.89
NGC 5194c 9.18
Holmberg IIa 8.13
IC 2574a 8.24

Notes.

μ These include the weak PAH feature at 14.2 μm and other SFR calibrations for convenience, but is reproduced in Equation (2).
L–Z determined values where no characteristic or R = 0 value is given.
 Stellar mass values from Leroy et al. (2019).

The emission from [S IV] at 10.51 μm is similarly too weak to be measured at the resolution of SL spectra. Several other interesting emission lines (e.g., [S III]) lie outside the wavelength range of our SL and SH spectra but may be an interesting topic for future studies.

SINGS SH spectra were analyzed with the specutils package of astropy to calculate the integrated line emission of 15.56 μm [Ne III]. We extracted a spectral region containing the emission feature between 14.9 and 16.3 μm to include ~0.4 μm of continuum around the emission line. The continuum that underlies the emission feature was fit by taking the average surface brightness between 14.9 and 15.3 μm and another average between 15.9 and 16.3 μm. We determined the linear fit that connects these two averages to model the continuum and then subtracted this continuum from the extracted spectrum.

We set the maximum surface brightness measurement between 14.9 and 16.3 μm as an initial estimate of the peak of the 15.56 μm [Ne III] emission line. We then fit a Gaussian to the continuum-subtracted data using the emission-line parameter modeling functions of specutils to determine the integrated intensity. Uncertainties were propagated by the Monte Carlo method, where spectral data were randomly offset according to their corresponding uncertainty in each wavelength bin and the Gaussian was refit. We repeated 500 trials in this way, using the standard deviation of the results to calculate the uncertainty on the integrated intensity of the 15.56 μm [Ne III] emission. The mean and standard deviation of these 500 trials are our [Ne III] line strength and its respective uncertainty. These measurements for regions with SH spectra are shown in Table 3.

Table 3 shows the surface brightness of each H II region at 10 μm in the SL spectra in μJy units. We calculate this value as the average over three wavelength bins at ~9.9, 10.0, and 10.1 μm to obtain a measure of the mid-IR continuum. This average is then multiplied by 30 THz (i.e., 10 μm in terms of frequency). This unit conversion ensures that ratios of [Ne III] emission and PAHFIT results with 10 μm continuum values are unitless. We note that 10 μm emission will be a poor tracer of the underlying dust emission in the case where there is strong silicate absorption. We caution the reader against using any SFR calibrations that involve 10 μm continuum in that case.

2.7. Correlation Methodology

We use T33 surface brightness as the reference SFR indicator for our calibrations. Converting T33 to SFR surface density ΣSFR by Equation (11) from Murphy et al. (2011) requires a multiplicative factor, which we omit from the calibrations for convenience, but is reproduced in Equation (2). This relation also has a dependence on the local electron temperature, and we assume a constant value of T_e = 10^4 K that is typical for H II regions in nearby galaxies. Inherent in the multiplicative factor from Murphy et al. (2011) is the assumption of the initial mass function (IMF), which other SFR calibrations may not share. The SFR calibration for neon and
metallicity from Zhuang et al. (2019), for example, assumes a Salpeter IMF, which differs from the others that assume a Kroupa IMF. As shown in Murphy et al. (2011), calibrations based on the Salpeter IMF result in SFR about 50% greater than those that use the Kroupa IMF. To compare each model equally in Section 3.3, we take this offset into account.
Figure 2. Extracted spectra from SL wavelengths 5.2–14.7 μm (left) and SH wavelengths 9.9–19.4 μm (right). Shown are the 28 dimmest of the 56 spectra in order of 12.81 μm [Ne II] emission measurements from PAHFIT on the SL spectra. Regions with no SH data are shown as a horizontal line in the right panel.
We assume that all emission features can be treated in the same manner and that the functional form of the T33 calibration is a power law with the general form

$$\log_{10} \left[ \frac{T_{33}}{\text{mJy sr}^{-1}} \right] = \alpha \log_{10} \left[ \frac{X}{10^{-4} \text{ erg s}^{-1} \text{cm}^{-2} \text{ sr}^{-1}} \right] + \beta \log_{10}[Y] + \phi \log_{10}[Z] + \gamma,$$  \hspace{1cm} (1)

where \(X\), \(Y\), and \(Z\) are any emission features and \(\alpha\), \(\beta\), \(\phi\), and \(\gamma\) are constants we fit by ordinary least squares. Parameters \(Y\) and \(Z\) are typically unitless band ratios or metallicity values (converted to \([\text{O}/\text{H}]\)), but the integrated intensities of emission lines or PAH features have the units of \(X\). We first investigate single-feature models where \(\beta\) and \(\phi\) in Equation (1) are zero. All emission features in Table 3 have been converted to \(10^{-4} \text{ erg s}^{-1} \text{cm}^{-2} \text{ sr}^{-1}\) to simplify use in Equation (1). In Section 3.3 we convert to SFR using distances listed in Table 2.

We also use Equation (11) from Murphy et al. (2011) to write our Equation (1) as a relation for determining \(\Sigma_{\text{SFR}}\):

$$\log_{10} \left[ \frac{\Sigma_{\text{SFR}}}{M_{\odot} \text{ yr}^{-1} \text{ kpc}^{-2}} \right] = \alpha \log_{10} \left[ \frac{X}{10^{-4} \text{ erg s}^{-1} \text{cm}^{-2} \text{ sr}^{-1}} \right] + \gamma - 1.08,$$  \hspace{1cm} (2)

where notation is preserved from Equation (1) and parameters \(\log_{10}[Y]\) and \(\log_{10}[Z]\) are not shown for brevity but can be included following the functional form of Equation (1). We attempted to further improve our single-feature model by including a second feature \(Y\), setting only \(\phi\) to zero. These regressions are run independently of the single-feature correlations so that the constants \(\alpha\) and \(\gamma\) in Equation (1) are not necessarily equal for a given feature \(X\) after including another feature \(Y\). If KK04 or PT05 metallicity values are used for \(\log_{10}[Y]\), we set \(Y\) in terms of \([\text{O}/\text{H}]\). Band ratios and 10 \(\mu\)m normalized bands used for \(Y\) are unitless. Finally, we attempted to improve our double-feature models by including a third feature \(Z\). As before, these trials are run independently of the previous correlations so that no constants are necessarily preserved from single- or double-feature calibrations.

In our calibrations to Equation (1), we employed a Monte Carlo approach to propagate uncertainties in both the independent and dependent variables. In each trial, random offsets were added to each point according to their corresponding measurement uncertainties, assuming a Gaussian distribution. We also quantified the dependence of our results on the specific regions considered using the bootstrap technique. This approach measures the uncertainty that results from the small sample size and non-Gaussian scatter of our data set. From our set of 33 regions, we create a new set of the same size where each value is randomly selected from the 33 regions. This allows some regions to contribute to the fit more than once in a given trial, or not at all. The resulting set is then used to fit the constants of Equation (1) by the ordinary least-squares method.

We measured the quality of single- and multifeature models found by Equation (1) using the Spearman rank correlation coefficient between the predicted T33 and the observed values. In this way we uniformly quantify the strength of each correlation regardless of how many parameters are used to predict T33 by Equation (1). We then determined the residuals as the difference between the predicted and observed T33 values for each region. From the residuals we calculated the standard deviation \(\sigma\) and quantified the presence of any metallicity-dependent offset, as follows. The set of regions was divided into two groups based on the median metallicity value (12 + \log[O/H] = 8.98_{KK04} = 8.36_{PT05}). We calculated the median of the residuals in the two metallicity bins, and the difference between the high- and low-metallicity residual medians is what we define as the metallicity offset \(\Omega_Z\).

We perform \(10^4\) bootstrap/Monte Carlo trials and calculate the mean and standard deviation of all values to quantify the uncertainty. We find that the uncertainty from bootstrap variation exceeds that returned by the Monte Carlo method. We apply both methods in each trial: measurement uncertainties are propagated via the Monte Carlo technique, and then sampling uncertainties are included by applying the bootstrap method.

We confirm the significance of our \(\Omega_Z\) parameter using the two-sample Kolmogorov–Smirnov (K-S) test and the two-sample Anderson–Darling (A-D) test. Each of these tests yields a probability that the two samples are drawn from the same distribution. For each correlation, we divide the regions based on their metallicity as described previously for determining \(\Omega_Z\), and then we run the K-S and A-D tests on these two portions of the sample. We find that both tests return similar probabilities. We consider \(\Omega_Z\) values significant when these tests find \(>90\%\) confidence that the two metallicity subsets are not drawn from the same distribution. We note that the magnitude of \(\Omega_Z\) does not necessarily indicate a statistically significant difference between the two metallicity subsets. In some cases, a given tracer has a poor correlation with T33 for reasons not tied to metallicity, leading to a large scatter in the residuals. In such a case, a relatively large offset between the median residual in the upper and lower metallicity bin may be statistically insignificant. On the opposite extreme, for tracers that are well correlated with T33, a small \(\Omega_Z\) can be significant when the intrinsic scatter in the relationship is small.

A perfect correlation would have \(\rho_s = 1\), \(\Omega_Z = 0\), and \(\sigma = 0\). However, our T33 data have nonnegligible measurement uncertainties that limit the practical maximum value of \(\rho_s\) and minimum values of \(\Omega_Z\) and \(\sigma\) when used in our combined bootstrap and Monte Carlo perturbation method (Curran 2014). We attempted to quantify this limit by using T33 and its associated uncertainty as feature \(X\) in bootstrap Monte Carlo trials using Equation (1) with \(\beta\) and \(\phi\) set to zero. We expect the mean of these trials to show a perfect correlation if the random offsets in each trial were negligible, but we find a practical maximum \(\rho_s\) of 0.92 and no minimum for \(\Omega_Z\) or \(\sigma\).

On the other end of the scale, we investigated the null hypothesis for \(\rho_s\) where two data sets are considered statistically indistinguishable from random, uncorrelated data sets. Correlations with a Spearman coefficient less than or equal to a critical value \(\rho_c\) are consistent with this null hypothesis. This \(\rho_c\) is determined by producing two random sets of 33 points to match the size of our calibration set. In each trial, two sets of 33 random points are created and \(\rho_c\) is calculated between them. We repeat \(10^4\) such trials and find that the \(\rho_c\) values converge on zero with a standard deviation that defines

\(^{6}\) We find that the separation based on metallicity is the same in KK04 and PT05, except for two regions, which makes a negligible difference in the results.
the 1σ null range: \( \rho_L = 0.18 \). Data sets that have \( |\rho_L| \) within this range are considered uncorrelated. We also note that, due to the properties of our sample, metallicity is correlated with T33 with \( \rho_L = 0.3 \). This correlation likely results from a sparse sampling of low-metallicity regions below about 12 + log(O/H)_KK04 \( \sim 8.8 \), with the exception of two dwarf galaxy regions at about 8.2. We anticipate that this could introduce spurious correlations with T33 at the level of \( \rho_L = 0.3 \) for highly metallicity-dependent emission features.

Figure 3 shows an example of a single-feature correlation with T33 using Equation (1), specifically PAH 7.7 \( \mu m \) versus T33. To emphasize the relative strength of mid-IR bands, we have converted T33 to \( n_L \) units by multiplying the observed T33 by its frequency to compare with the PAH 7.7 \( \mu m \) feature only in Figure 3(a). This figure also shows the fit from each regression trial and their mean in black. We also repeat this for the high- and low-metallicity groups individually, shown in red and blue lines, respectively. Figure 3(b) summarizes our statistics on the residuals: \( \Omega_Z \) and \( \sigma \). In this case, the correlation of T33 with PAH 7.7 \( \mu m \) is weak, with \( \rho_L = 0.27 \), significant scatter \( \sigma = 0.26 \), and a large metallicity offset \( \Omega_Z = 0.26 \) (with \(< 10\% \) chance that the metallicity subsets of the residuals come from the same distribution). In Figure 3(b) we illustrate our \( \Omega_Z \) statistic as the difference between the red and blue vertical lines, and \( \pm \sigma \) (the standard deviation) is shown as the length of the horizontal black line. Table 4 lists \( \rho_L \), \( \Omega_Z \), and \( \sigma \) for the correlations discussed in Section 3, and Appendix C contains histograms in the form of Figure 3(b) for these correlations.

There are a few measurements with notable characteristics indicated in Figure 3(a). These are the dwarf galaxy regions...
Holmberg II-0 and IC 2574b (the two lowest-metallicity points in our sample) and the known anomalous microwave emission (AME) source designated NGC 6946 Enuc.4a (Murphy et al. 2010). The AME source is plotted but not included in our calculations of $\rho_s$, $\Omega_Z$, and $\sigma$ for T33 correlations owing to the observed excess 33 GHz emission relative to its SFR. We also indicate the region NGC 2403 Enuc.3, which is a more significant outlier than the known AME source. This region has recently been proposed as a potential AME source as well (Linden et al. 2020), which likely explains the significant excess of T33 relative to PAH 7.7 $\mu$m emission observed.

3. Results

We measured the properties of various correlations of mid-IR emission features with T33 and their metallicity dependence. For each correlation we measured $\rho_s$, $\Omega_Z$, $\sigma$, and their uncertainties. We focused primarily on maximizing $\rho_s$ in the following sections since this generally minimizes $\sigma$ as well. In cases of similar $\rho_s$ we then looked for minimum $\Omega_Z$ to find tracers that are more independent of metallicity. These parameters describing the quality of our results are summarized in Figure 4, where metallicity offset $\Omega_Z$ is plotted against Spearman rank correlation coefficient $\rho_s$. In this figure the features that have a statistically significant $\Omega_Z$, as defined in Section 2.7, are indicated by a black circle surrounding the symbol. Table 4 lists our main results for the following model quality statistics for several tracers: Spearman rank correlation coefficient $\rho_s$, metallicity offset $\Omega_Z$, standard deviation of residuals $\sigma$, wavelength range $\Delta \lambda$, and maximum wavelength $\lambda_{max}$. Wavelength range $\Delta \lambda$ and maxima $\lambda_{max}$ for these tracers are calculated from the peak location of constituent emission features, but additional spectral coverage will be necessary to separate most PAH and neon features from continuum. The following sections present details on trends and features of interest in this table and Figure 4. In Appendix A, we also describe the optimal correlations obtained in restricted wavelength ranges, which may be of interest for higher-redshift observations.

3.1. Correlations of Individual Mid-IR Emission Features with Thermal 33 GHz Emission

In this section, we present single-feature fit results for Equation (1) (with $\beta = \phi = 0$). The best-correlated mid-IR observable with T33 is the 15.56 $\mu$m [Ne III] line with $\rho_s = 0.72$. This emission line is the only single feature considered that has a statistically significant negative metallicity offset ($\Omega_Z = -0.13$ dex). Low-metallicity H II regions show increased 15.56 $\mu$m [Ne III] emission relative to their SFR compared to higher-metallicity regions. On average, a low-metallicity region is found to have about 35% more [Ne III]
emission than a typical high-metallicity region of the same T33. After [Ne III] the next best individual feature is the 12.81 μm [Ne II] line, which has a strong positive metallicity offset ($\rho_s = 0.59$ and $\Omega_Z = 0.28$ dex). In comparison to [Ne III], an average high-metallicity region is found to emit about 90% more [Ne II] emission than a typical low-metallicity region of the same T33.

All PAH bands in this study show weak correlation with T33, characterized by $\rho_s \sim 0.3$. We find significant positive metallicity offsets of $\Omega_Z \sim 0.25$ dex for each PAH band correlation with T33. Adding the PAH band emission into a total $\Sigma$PAH does not improve T33 correlations or reduce the metallicity offset. We find that our high- and low-metallicity subsamples are better correlated separately than combined because of these metallicity-dependent offsets as seen in Figure 3(a). Ratios of PAH bands are uncorrelated (i.e., $\rho_s < \rho_c$). If we normalize the PAH bands by 10 μm continuum, the magnitude of $\rho_s$ is typically unaffected, but the metallicity offset is reduced and the two metallicity samples become consistent with being drawn from the same distribution.

Following the work of Ho & Keto (2007), we combine the 12.81 μm [Ne II] and 15.56 μm [Ne III] emission-line-integrated intensities into a new “single” feature: $\Sigma$Ne. We refer to $\Sigma$Ne as an individual parameter when used in Equation (1) despite being a combination of two distinct emission features. This combined feature is our best single parameter $X$ for predicting T33 by Equation (1), with $\rho_s = 0.84$ and a small but statistically significant $\Omega_Z = 0.08$ dex (see Appendix C for histogram). We find that adding the [Ne II] and [Ne III] emission lines to create $\Sigma$Ne results in a similar $\rho_s$ within about 1σ to that obtained by using them separately as parameters $X$ and $Y$ ($\rho_s = 0.80 \pm 0.09$). We discuss this observation in Section 4.

3.2. Correlations of Multiple Observables with Thermal 33 GHz Emission

3.2.1. Improvements to $\Sigma$Ne Correlations with T33

In the previous section, we showed that single-feature correlations with T33 have metallicity-dependent residuals that can decrease the correlation coefficient. To improve the correlations, we therefore attempt to add an additional feature to the model that acts to remove metallicity offsets. Section 3.1 shows that using $\Sigma$Ne results in the highest $\rho_s$ and lowest $\Omega_Z$ values of the single mid-IR feature models. We first explicitly include metallicity as the second parameter to check that the correlation improves and to what degree. Indeed, Table 4 shows that using the KK04 metallicity in combination with $\Sigma$Ne results in $\rho_s = 0.91$, which is consistent within 1σ with the maximum 0.92 noted in Section 2.7. We also find an equivalent improvement using the PT05 metallicity values.

The KK04 and PT05 metallicity calibrations require optical emission lines, and observations may not be available for all galaxies, particularly at high redshift. To avoid requiring this additional information, we searched for observables in the mid-IR that could improve the $\Sigma$Ne correlation with T33 by acting as a proxy for metallicity. The ratio of 15.56 μm [Ne III] and 12.81 μm [Ne II] is expected and observed to correlate well with metallicity; however, we find that including it results in no significant improvement in $\rho_s$ for $\Sigma$Ne. We discuss this observation further in Section 4.

Another potential proxy for metallicity is our 10 μm continuum-normalized PAH bands, which should trace the abundance of PAHs relative to dust and correlate well with metallicity. We find that combining $\Sigma$Ne with the strongest PAH bands normalized by 10 μm continuum results in $\rho_s = 0.87$ and a statistically insignificant $\Omega_Z$. These values are within 1σ of those obtained from combining $\Sigma$Ne with KK04 or PT05 metallicities. Table 4 shows that the 12.6 μm PAH feature normalized by continuum at 10 μm can be used to remove the metallicity offset and improve the $\Sigma$Ne–T33 correlation in place of KK04 or PT05 metallicity. We find that 10 μm continuum-normalized PAH bands improve $\rho_s$ and $\Omega_Z$ for $\Sigma$Ne better than ratios of two PAH bands for every band considered. In Appendix B we investigate the correlation between continuum-normalized PAH bands and metallicity. In theory, adding another feature to fit could further improve the correlation; however, we find that including a third mid-IR observable results in no significant improvement to $\rho_s$ for the $\Sigma$Ne–12.6/10 μm tracer.

3.2.2. Improvements to Other Correlations with T33

We next investigated two-feature correlations with T33 that do not involve $\Sigma$Ne. No pair of two features was found that predicts T33 measurements better than $\Sigma$Ne paired with metallicity or a normalized PAH band. In Section 3.1 we showed that the individual emission feature that best correlates with T33 is the 15.56 μm [Ne III] line. We find that this [Ne III] correlation is improved best by including 24 μm continuum as parameter $Y$ or with [Ne II] to form $\Sigma$Ne; each combination results in the same $\rho_s$ of 0.85 within 1σ. Figure 4 summarizes this similarity between the [Ne III]–24 μm tracer and the $\Sigma$Ne tracer.

A commonly used hybrid SFR tracer involves linearly combining $H_\alpha$ emission from ionized gas and 24 μm continuum from dust (e.g., Calzetti et al. 2007). Although we use a different functional form (power-law rather than a linear combination), we tested the quality of the combination of $H_\alpha$ and 24 μm by using the Murphy et al. (2018) measurements listed in Table 1 as parameters in Equation (1). As expected, the correlation is strong with $\rho_s = 0.77$ as shown in Table 4 but with a statistically significant metallicity offset $\Omega_Z = 0.11$ dex.

We then searched for individual features that could replace either $H_\alpha$ or 24 μm to give an improved correlation with T33 based on a larger $\rho_s$ and smaller $\Omega_Z$ and $\sigma$. As mentioned previously, pairing 24 μm with 15.56 μm [Ne III] emission results in $\rho_s = 0.85$ and $\Omega_Z = 0.04 \pm 0.04$, with no statistical difference between the two metallicity bins, which is a significant improvement over pairing 24 μm with $H_\alpha$ in our power-law model Equation (1). We find no better feature to pair with $H_\alpha$ emission than 24 μm; however, pairing $H_\alpha$ with the 12.81 μm [Ne II] line results in an equivalent $\rho_s$ and $\Omega_Z$. These different combinations of $H_\alpha$, 15.56 μm [Ne III], 24 μm, and 12.81 μm [Ne II] are further discussed in Section 3.2.3. We also directly compare our power-law $H_\alpha$ and 24 μm model with the commonly used linear hybrid $H_\alpha$ and 24 μm tracer from Calzetti et al. (2007) in Section 3.3.

Despite their weak correlation with T33 individually, we investigated the full potential of PAH emission as an SFR tracer by including metallicity information. As before, the
correlations are similar for each PAH band with \( \rho_s \approx 0.6 \) after KK04 metallicity is included. We also attempted using the ratio of 15.56 \( \mu m \) [Ne III] to 12.81 \( \mu m \) [Ne II] emission again as a mid-IR proxy for metallicity. Figure C4 in Appendix C shows that combining \( \Sigma \)PAH with this neon line ratio results in a stronger correlation with T33 than if \( \Sigma \)PAH is combined with KK04 metallicity (\( \rho_s = 0.69 \) and statistically insignificant \( \Omega_Z = 0.02 \) dex compared to \( \rho_s = 0.63 \) and \( \Omega_Z = 0.13 \) dex). These results imply that emission from PAHs is a poor option for an SFR indicator even though the metallicity-corrected PAH correlation with T33 has a \( \rho_s \) similar to some single-feature tracers such as 15.56 \( \mu m \) [Ne III] emission.

### 3.2.3. Summary

We summarize our results for mid-IR/T33 correlations by plotting their \( \Omega_Z \) as a function of \( \rho_s \) in Figure 4. We have shown in Section 2.7 that the T33 measurement uncertainties limit the maximum value of \( \rho_s \) for any correlation to \( \sim 0.92 \), which is indicated by the star symbol in this figure. Figure 4 also shows the null range for \( \rho_s \) with data set size of 33 points as a gray band around zero as defined in Section 2.7. In Figure 5 we focus on some of the best-correlated tracers from Table 4, including some that are not shown in Figure 4 for clarity. Figure 5 shows that \( \rho_s \) for the tracer using 12.81 \( \mu m \) [Ne II], PAH 12.6 \( \mu m \), and 10 \( \mu m \) continuum is equivalent to the H\( \alpha \) and 24 \( \mu m \) tracer but with a smaller \( \Omega_Z \). The \( \Sigma \)Ne and PAH 12.6/10 \( \mu m \) tracer has a statistically insignificant \( \Omega_Z \) similar to the \( \Sigma \)Ne and KK04 metallicity tracer and with \( \rho_s \) within the 1\( \sigma \) uncertainty.

In addition to each individual feature, we have included several pair tracers of interest in Figure 4. The strong similarities between the \( \Sigma \)Ne tracer and the [Ne III]–24 \( \mu m \) pair tracer noted in Section 3.2.2 are evident in their near overlap on this plot. Figure 4 reveals a trend that combining two emission features with opposite metallicity dependence results in a stronger correlation with T33. For example, combining 15.56 \( \mu m \) [Ne III] (negative \( \Omega_Z \)) and any PAH feature (positive \( \Omega_Z \)) results in a strong correlation with T33 and negligible \( \Omega_Z \). The frequently used hybrid tracer based on H\( \alpha \) and 24 \( \mu m \) is found to follow a similar pattern; however, the small, negative \( \Omega_Z \) determined for H\( \alpha \) emission is not statistically significant.

This trend is further supported by considering our previous observation that pairing 24 \( \mu m \) with [Ne III] performs better than pairing with H\( \alpha \) since the value of \( \Omega_Z \) for [Ne III] is statistically significant while H\( \alpha \) has a negligible \( \Omega_Z \). Similarly, pairing [Ne II] with H\( \alpha \) performs as well as 24 \( \mu m \) with H\( \alpha \). We also find that pairing features with \( \Omega_Z \) of the same sign such as [Ne II] with 24 \( \mu m \) or [Ne III] with H\( \alpha \) does not improve \( \rho_s \), \( \Omega_Z \), or \( \sigma \), based on calculated uncertainties.

### 3.3. Comparisons with Previous SFR Calibrations

Our calibrations for \( \Sigma_{\text{SFR}} \) tracers based on T33 were performed assuming that a power-law functional form (Equation (1)) can be applied to all emission features considered. We compare our tracers built using this assumption against other calibrations from the literature. These other calibrations differ from this work in functional form, and few publications have directly studied the metallicity dependence of SFR tracers using local metallicities. We used the largest applicable subset of the 56 regions for any given comparison.

Observed T33 surface brightnesses were used to derive SFR by Equation (11) from Murphy et al. (2011). We refer to the SFR calculated using observed T33 as the reference model. We combined our Equation (1) for predicting T33 with the SFR–T33 relation in Equation (11) from Murphy et al. (2011) such that

\[
\left( \frac{\text{SFR}}{M_\odot \text{ yr}^{-1}} \right) = 7.8 \times 10^{-4} \left( \frac{T_e}{10^4 \text{ K}} \right)^{-0.45} \left( \frac{D}{\text{Mpc}} \right)^2 \times \left( \frac{X}{10^{-4} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ sr}^{-1}} \right) \gamma (Z)^\varphi,
\]

where notation is preserved from Equation (1), \( T_e \) is the local electron temperature, and \( D \) represents the distance to the host galaxy. Units for \( Y \) and \( Z \) vary from \((10^{-4} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ sr}^{-1})\) for emission features to unitless [O/H] for KK04 and PT05 metallicities, while band ratios and 10 \( \mu m \) normalized bands are also unitless. SFRs derived using Equation (3) are referred to as the power-law model. The corresponding constants for use in Equation (3) are listed in Table 5. As noted in this table, these same constants can be used to model T33 in Equation (1) if 9.04 (i.e., the logarithm of the reciprocal solid angle of a 7″ aperture in steradians) is added to the constant \( \gamma \).

Lack constraints on the electron temperature, we assume a value of \( 10^4 \) K, which is typical of H\( \Pi \) regions. Changing this assumption by a factor of two results in a difference of \( \sim 37\% \). Converting T33 from Equation (1) to SFR in Equation (3) requires distance as an additional parameter. Therefore, we expect that statistics such as the correlation coefficient or the standard deviation of residuals will not be preserved from Table 4. The Spearman correlation coefficient between SFR derived using our power-law relation and a comparison tracer is referred to as \( \rho_{\text{SFR}} \) to differentiate. We also calculate the median value of the residuals as \( \Delta \) to measure any remaining offset between the two SFR tracers. We also list the standard deviation of the residuals when we report \( \Delta \).
We compare our power-law model using Hα and 24 μm emission in Equation (3) with Equation (7) from Murphy et al. (2011), which depends only on these same parameters. The constant in this relation is updated slightly from the primary reference, Calzetti et al. (2007), but remains very similar. We refer to the relation from Murphy et al. (2011) as the “linear model” owing to its functional form. Neither model has an explicit metallicity, 15.56 μm [Ne III], or T33 dependence, so the full data set of 56 regions with Hα and 24 μm emission is used in Figure 6(c). Panels (a) and (b) of this figure show that our power-law model predicts T33-derived SFR about as well as the commonly used linear model for Hα and 24 μm emission. Figure 6(c) shows that our power-law model is also as accurate as the linear model for the 23 H II regions outside the calibration subset of 33 regions.

We also compare ΣNe and metallicity power-law model to the model presented as Equation (6) in Zhuang et al. (2019; hereafter ZHS model). Both models depend only on the metallicity and functions of the 12.81 μm [Ne II] and 15.56 μm [Ne III] bands. As noted in Section 2.7, an offset is expected between the ZHS model, which is based on the Salpeter IMF, and the T33 model, which is based on the Kroupa IMF. We corrected for this by including a factor of 1.5 decrease (−0.18 dex) to SFR from the ZHS model (Murphy et al. 2011). We calculate SFR for the 46 regions with SH data and metallicity measurements when comparing our power-law model to the ZHS model in Figure 7(c) and the 33 regions in the calibration subset when comparing our power-law model to the ZHS model in Figure 7. We find that SFR predicted by the ZHS model is well correlated with the T33-derived SFR, but with a significant average residual offset. Figure 7(a) shows that the IMF-corrected SFR from the ZHS model are greater than SFR from T33 by about an average of 0.13 dex. This discrepancy will be further discussed in Section 4. This observed offset is propagated into the comparison between our power-law model and the ZHS model in Figure 7(c) since our calibration is based on T33 emission; however, the correlation is very strong as indicated by ρSFR.

Finally, we apply the PAH-based SFR tracers presented by Shipton et al. (2016). These tracers were calibrated for entire regions.

Table 5

| Feature | X | Y | Z | Constants |
|---------|---|---|---|-----------|
| ΣNe    |   |   |   | α = 0.87 ± 0.09, β = −0.48 ± 0.14, φ = 0.09 ± 0.06, γ = −2.17 ± 0.42 |
| ΣNe    |   | 12.6/10 |   | α = 0.84 ± 0.10, β = −0.40 ± 0.18, φ = 0.09 ± 0.06, γ = −1.14 ± 0.19 |
| [Ne III] |   | 24 μm |   | α = 0.38 ± 0.06, β = 0.38 ± 0.09, φ = 0.09 ± 0.06, γ = −1.14 ± 0.17 |
| ΣNe    |   |   |   | α = 0.76 ± 0.15, β = 0.09 ± 0.15, φ = 0.09 ± 0.06, γ = −1.14 ± 0.17 |
| [Ne III] | 13.6 |   |   | α = 0.48 ± 0.06, β = 0.26 ± 0.09, φ = 0.09 ± 0.06, γ = −1.14 ± 0.17 |
| [Ne II] | 12.6 | 12.6/10 |   | α = −0.47 ± 0.24, β = 1.10 ± 0.25, φ = −0.86 ± 0.35, γ = −1.27 ± 0.41 |
| Hα 24 μm |   |   |   | α = 0.25 ± 0.10, β = 0.43 ± 0.11, φ = 0.09 ± 0.06, γ = −1.34 ± 0.20 |
| [Ne III] |   |   |   | α = 0.48 ± 0.08, β = 0.26 ± 0.09, φ = 0.09 ± 0.06, γ = −1.34 ± 0.20 |
| [Ne II] | 12.6 |   |   | α = 1.09 ± 0.35, β = −0.76 ± 0.25, φ = 0.09 ± 0.06, γ = −1.34 ± 0.20 |
| [Ne II] |   |   |   | α = 0.36 ± 0.16, β = 0.26 ± 0.09, φ = 0.09 ± 0.06, γ = −1.34 ± 0.20 |
| PAH 11.3 μm | 12.0/10 |   |   | α = 0.44 ± 0.21, β = −0.84 ± 0.45, φ = 0.09 ± 0.06, γ = −1.95 ± 0.70 |
| PAH 7.7 μm | 7.7/10 |   |   | α = 0.48 ± 0.19, β = −0.89 ± 0.50, φ = 0.09 ± 0.06, γ = −1.20 ± 0.26 |

Note.

ς Constant γ as listed gives T33 in mJy; units of Equation (1) (mJy sr⁻¹) are obtained by adding a constant 9.54 to γ.
galaxies with a strict lower limit on metallicity, approximately equivalent to the median metallicity value used in this work. Therefore, we expect the tracers to be most accurate for the regions we have previously defined as high-metallicity. For fair comparison with the work of Shipley et al., we restrict our calculations of statistics such as the Spearman correlation coefficient \( \rho \) and the median of the residuals \( \hat{\Delta} \) to high-metallicity regions \((12 + \log(O/H))_{\text{KK04}} > 8.98\). The correlations with T33 are indeed much stronger among the high-metallicity points. We find that each tracer is statistically equivalent in both \( \rho_{\text{SFR}} \) and \( \sigma \). Figure 8 shows a significant offset of about \(-0.2\) dex in the PAH 7.7 \( \mu \)m SFR and about \(-0.3\) dex in the PAH 11.3 \( \mu \)m SFR compared to that from T33. However, we find no significant offset between the T33 SFR and the PAH 6.2 \( \mu \)m SFR relation.

3.4. PAH Band Ratios in H II Regions

In Section 2.5 we described our method for producing a set of background-subtracted star-forming region spectra. This data set was used to compare changes in the PAH spectrum in star-forming regions after background subtraction; the ratio of PAH features at 11.3 \( \mu \)m and 7.7 \( \mu \)m is used as an indicator of the relative number of neutral to ionized PAH molecules, and the ratio of PAH features at 6.2 \( \mu \)m and 7.7 \( \mu \)m is used as an indicator of the relative number of small to large PAH molecules (e.g., Draine & Li 2001). We find a decrease in the weighted mean of the 11.3/7.7 ratio after the local background subtraction and a less significant increase in the weighted mean of the 6.2/7.7 ratio. Figure 9 shows that the 6.2/7.7 PAH-size ratio increases by 3.6% and the 11.3/7.7 PAH-ionization ratio decreases by 15.6%. Also indicated in this figure are significant outliers such as the dwarf galaxy regions Holmberg II-0 and IC 2574b and the known AME source designated NGC 6946 Enuc.4a (Murphy et al. 2010). These outliers are found to be regions where the aperture is not centered on the peak of PAH emission or PAH emission is sufficiently extended such that subtracting the surrounding annulus removes a significant portion of the total aperture brightness. We find that this form of annulus background subtraction results in no significant effect on T33 correlations.

4. Discussion

4.1. Properties of Effective Mid-IR SF Tracers

Our study has revealed a novel recipe for creating effective SFR tracers by considering their dependence on metallicity. Common hybrid SFR tracers use UV/optical emission that originates from young stars and ionized gas, such as H\(\alpha\), and combine it with a correction factor for the light that was attenuated, which is assumed to be proportional to the 24 \( \mu \)m continuum (see review by Kennicutt & Evans 2012). Mid-infrared emission is largely insensitive to dust attenuation, and our study finds that the best SFR tracers based on these features instead pair an ionized gas tracer with a metallicity-dependent correction term. Our mid-IR tracers do not rely on two features to trace separate portions of emission from a star-forming region and instead employ a feature such as \( \Sigma Ne \) to trace the emission from ionized gas and a second parameter that acts to further improve the correlation, typically by correcting for variations due to metallicity.

In Section 3.2.3 we noted a pattern among SFR tracers where combining features with opposite metallicity offset \( \Omega_Z \) improves correlations with T33. For example, our 15.56 \( \mu \)m [Ne II]–24 \( \mu \)m tracer is constructed of 24 \( \mu \)m, which we found to have a positive metallicity offset, while [Ne III] has a large negative metallicity offset. The direction of the [Ne III] metallicity offset implies underestimated SFR in low-metallicity relative to high-metallicity regions for a given [Ne III] surface brightness, and vice versa for 24 \( \mu \)m emission. Pair tracers that combine features with these inverse metallicity dependencies in T33 correlations are found to best predict T33-based SFR.

We can also interpret the results for the often used H\(\alpha\)–24 \( \mu \)m calibration as following a similar trend of pairing strong SF tracers with opposite metallicity dependence, leading to the best correlation with SF. H\(\alpha\) has a small negative metallicity offset, although the residuals in the two metallicity bins are consistent with being drawn from the same sample, and a moderate correlation with T33, while 24 \( \mu \)m has a strong positive metallicity offset and similar correlation with T33. The metallicity offset is expected for the dust-attenuation-dependent emission from 24 \( \mu \)m, as found by Relaño et al. (2007) in a similar study of star-forming regions in nearby galaxies. The
weak metallicity dependence of Hα with respect to T33 suggests that attenuation geometry may dominate as the source of scatter between T33 and Hα. Combining Hα and 24 μm leads to an improvement in the correlation with T33 and a lower-metallicity offset than the 24 μm has alone, reinforcing the idea that pairing strong SF tracers with opposite metallicity dependence results in the best correlation with SF.

Interestingly, we found that an SFR tracer using 15.56 μm [Ne III] and 24 μm emission in Equation (3) outperforms a tracer using Hα and 24 μm in all model quality parameters of Table 4. [Ne III] and Hα each trace highly ionized gas. The negative metallicity dependence as defined in this work is found to be statistically significant for [Ne III], but not for Hα. This is likely because Hα emission exclusively traces the unobscured portion of radiation from ionized hydrogen gas and thus depends less directly on metallicity than [Ne III]. [Ne III] emission has a complex dependence on metallicity, density, and radiation field hardness, but a general trend of decreasing emission with respect to SFR as metallicity increases is expected and observed (e.g., Giveon et al. 2002; Madden et al. 2006; Zhuang et al. 2019). Observations at 15.56 μm are relatively unaffected by dust attenuation that hinders Hα emission. These results imply that an [Ne III] and 24 μm tracer outperforms the Hα and 24 μm tracer because Hα traces only the unobscured ionized gas while [Ne III] emission has a dependence on all the ionized gas, as well as the local metallicity.

Generalizing this pattern, we propose that a tracer is more effective if both component emission features have a strong correlation with SF and the residual dependence on metallicity of one is opposite that of the other. The metallicity dependence in the SFR correlation of each feature independently is...
and its weak metallicity dependence may be related to the ratio of PAH features at 6.2 \mu m and 7.7 \mu m plotted against the ratio of PAH features at 6.2 \mu m and 7.7 \mu m for 49 extranuclear star-forming regions. Vertical and horizontal lines indicate that the weighted mean and theoretical tracks obtained from Draine & Li (2001) depicted in gray squares and black stars represent expected PAH emission for interstellar radiation field strength $U_{\text{ISM}}$.

Figure 9. Ratio of PAH features at 11.3 \mu m and 7.7 \mu m plotted against the ratio of PAH features at 6.2 \mu m and 7.7 \mu m for 49 extranuclear star-forming regions. Vertical and horizontal lines indicate that the weighted mean and theoretical tracks obtained from Draine & Li (2001) depicted in gray squares and black stars represent expected PAH emission for interstellar radiation field strength $U_{\text{ISM}}$.

Mitigated after they are combined, strengthening the SFR dependence.

4.2. SF Tracers From $\Sigma$Ne

In Section 3.2.1 we noted that combining $\Sigma$Ne and KK04 metallicity in Equation (1) results in a correlation coefficient consistent with the maximum allowed by the uncertainties of our T33 measurements. We also found that $\Sigma$Ne alone better traces T33 than combining H\alpha and 24 \mu m. The total emission from the mid-IR ionized neon lines depends on the total abundance of neon, so an ionized neon tracer could, in theory, have a complex dependence on metallicity and the radiation field hardness. These effects may cancel each other on average, as high-metallicity regions will have more neon in total, but less of it will be in an ionized state since the radiation field will be softer compared to lower-metallicity regions. Similarly, the low-metallicity regions will have less neon in total, but more of it will be in the Ne II and Ne III states owing to the harder radiation field of stars at low metallicity. Because both are $\alpha$-elements (Woosley & Weaver 1995), the total neon abundance is expected and observed to correlate with oxygen abundance, so regions of the same metallicity have similar neon abundance (e.g., Berg et al. 2020). These opposing effects may be mitigated on average in the $\Sigma$Ne--T33 correlation, as seen in the significantly reduced metallicity offset compared with the correlation of either [Ne II] or [Ne III] individually. Residual scatter in the $\Sigma$Ne--T33 correlation is found to be entirely dependent on metallicity since including metallicity as a second parameter in the fit results in an optimal correlation coefficient.

Another explanation for the success of the $\Sigma$Ne calibration and its weak metallicity dependence may be related to the importance of the two mid-IR neon lines in H II region cooling. Because Ne II and Ne III are relatively abundant and have low-lying fine-structure levels that can be easily excited by collisions, the resulting 12.81 \mu m [Ne II] and 15.56 \mu m [Ne III] lines are important coolants in H II region energy balance (e.g., Burbidge et al. 1963; Osterbrock 1965). In steady state, the H II region heating, which is dominated by photoionization of hydrogen, and the cooling, which is partially from the neon lines, will be equal. Therefore, $\Sigma$Ne should trace the ionizing photon rate, modulo a potentially varying fraction of the cooling luminosity emerging in other emission lines. The lack of substantial metallicity offset in $\Sigma$Ne suggests that over the metallicity range of our study the neon lines together carry a relatively constant fraction of the overall cooling luminosity of H II regions.

In Section 3.2.1 we found that unlike local metallicity (12 + log_{10}[O/H]), the ratio of [Ne III]/[Ne II] emission does not improve $\Sigma$Ne correlations with T33. However, we also found in Appendix B that this ionized neon ratio has the strongest correlation with metallicity of all mid-IR ratios considered. The neon ratio illustrates an important point that a ratio of emission features can have a strong correlation with metallicity and not have a metallicity-dependent offset in T33 correlations. We find that any other correlation that is improved by pairing with metallicity is similarly improved by pairing with the [Ne III]/[Ne II] ratio. This indicates that the $\Sigma$Ne correlation with T33 is a special case that cannot be improved with the [Ne III]/[Ne II] ratio. It is possible that combining the 12.81 \mu m [Ne II] and 15.56 \mu m [Ne III] surface brightnesses into $\Sigma$Ne includes all the relevant metallicity information about the relationship between [Ne II] and [Ne III] with respect to SFR. We find that using metallicity itself or any other ratio that is well correlated with metallicity (such as a normalized PAH band) improves the $\Sigma$Ne correlation with T33 to approximately the maximum allowed.

4.3. Metallicity Dependence of PAH-Based SF Tracers

Individual PAH band correlations with T33 are found to all be similar within their measured uncertainties and weak: with Spearman coefficients of $\rho_s \sim 0.3$ (within about 1$\sigma$ of our calculated critical coefficient for the null hypothesis). We found that this weak relation between PAH emission and T33 is due to a strong metallicity dependence. Combining emission from multiple PAH bands also does not improve the correlation, suggesting that band ratios do not vary substantially with SFR or metallicity. At a given T33 or SFR, we observed that high-metallicity regions are approximately twice as bright in the 7.7 \mu m PAH band as low-metallicity regions. As noted in Section 2.7, KK04 metallicity itself is found to have $\rho_s \sim 0.3$ in correlations with T33 owing to the sparse sampling of low metallicities in our data set. This fact, combined with the strong metallicity dependence seen in PAH emission, implies that the significance of PAH correlations with T33 may even be slightly exaggerated by our sample choice. The scatter in PAH--T33 correlations can, however, be improved significantly, to $\rho_s \sim 0.6$ by using a metallicity tracer as a second parameter. Therefore, we conclude that the PAH correlations with T33 SFR are very weak owing to strong metallicity dependence but not uncorrelated.

In choosing the best second feature to include in a PAH-based tracer, we found that pairing any PAH band with the [Ne III]/[Ne II] ratio improves T33 correlations and metallicity offsets. These improvements are greater than those obtained from pairing PAH bands with KK04 metallicity itself. The population of PAH molecules and their emission have a dependence on both metallicity and radiation field hardness.
(e.g., Madden et al. 2006; Lebouteiller et al. 2007; Gordon et al. 2008), since PAHs can be destroyed by energetic photons. These dependencies are also expected in the ionized neon ratio [Ne III]/[Ne II] (Zhuang et al. 2019). If the metallicity dependence of PAH emission is due in part to their destruction in harder radiation fields, the use of [Ne III]/[Ne II] as a second parameter in the T33 correlation may capture that effect better than metallicity alone.

4.4. Comparison to Existing SFR Calibrations

Our investigation provides an interesting comparison with existing SFR calibrations for several reasons. We calibrate our relationship using resolved measurements for individual star-forming regions, as was done in Calzetti et al. (2007) and Kennicutt et al. (2009), for instance. Our calibration is tied to thermal 33 GHz, which has a relatively direct relationship with the ionizing flux from massive stars, subject to few of the systematic effects tied to UV/optical SF tracers (e.g., Murphy et al. 2011), infrared tracers (e.g., Leroy et al. 2012), or SED-fitting tracers (e.g., Leja et al. 2019). In addition, our resolved measurements of SF regions minimize issues caused by diffuse emission unrelated to SF (i.e., cirrus or diffuse ionized gas) that may contaminate integrated galaxy SF calibrations.

One difference between our investigation and some previous studies is the assumption of a power-law form for the calibration, defined by Equation (3). This assumption does not appear to affect the accuracy of our SF predictions, however. Section 3.3 shows that our results are consistent with previous calibrations such as the linear hybrid tracers with Hα and 24 μm (Murphy et al. 2011) and more complex SFR calibrations such as the 12.81 μm [Ne II], 15.56 μm [Ne III], and metallicity model from Zhuang et al. (2019).

In Section 3.1 we found that combining the [Ne II] with the [Ne III] surface brightness linearly into a single feature ΣNe results in an equivalent correlation with T33 compared to including [Ne II] and [Ne III] in Equation (1) as two separate parameters. This supports our assumption that all parameters can be treated similarly using a power law. As noted previously in Section 1, the sum of [Ne II] and [Ne III] will trace the total emission from ionized neon, but our study cannot definitively conclude whether this is better modeled as a linear combination or a combination of two power laws.

We compared our ΣNe SFR calibration with the Zhuang et al. (2019) SFR calibration based on 12.81 μm [Ne II], 15.56 μm [Ne III], and metallicity and found that the ZHS model overestimates SFR by about twofold (0.31 dex) before correcting for IMF assumption. In Section 2.7 we noted that the difference in IMF assumption between the ZHS model and the other SFR calibrations considered in this study is known to account for about 0.18 dex of this offset (Murphy et al. 2011). The remaining 0.13 dex offset may result from the assumed value for solar metallicity, as the SFR from the ZHS model has a direct dependence on this value. Reducing our assumed KK04 solar metallicity by this 0.13 dex (from 12 + log10(O/H) = 8.8 to 8.67) removes the offset. The H II regions considered in this work have KK04 values ranging from about 8.1 to 9.2, but regions from normal spiral galaxies like the Milky Way have metallicities between 8.7 and 9.2 in this calibration as shown in Figures B1 and B2. This indicates that the zero-point of the metallicity scale could be the cause of the offset in SFR from the ZHS model. Recent studies have also found that calibrations based on spectral energy distribution (SED) fitting of entire galaxies can overestimate SFR by a similar factor to that observed in Figure 7(b) (Leja et al. 2019). We expect T33 to provide an ideal SFR that is independent of the assumptions in SED modeling, so the offset between Zhuang et al. (2019) and our calibration may be evidence of this overestimate from SED-based SFR tracers, although after correcting for the IMF we do not see this level of disagreement. We note that for distant, unresolved galaxies it will be necessary to take into account neon emission lines powered by active galactic nuclei rather than SFR in some cases, as investigated by Zhuang et al. (2019).

Our comparisons between T33-based SFR tracers and others rely on an assumed constant H II region electron temperature of 10^4 K. Changing this assumption by a factor of two results in a difference of ~37% (0.14 dex). The temperature of H II regions is metallicity dependent, so an additional correction term may be necessary to reduce the variations from this dependence, especially for observations at high redshifts.

The high-metallicity PAH band SFR tracer of Shipley et al. (2016) was found to correlate well with SFR given by T33 measurements despite the weak correlations between PAH bands and T33 seen in Section 3. This demonstrates a key difference between calibrations for SFR from luminosities and calibrations for ΣSFR from surface brightnesses. We found that the PAH 11.3 μm and 7.7 μm tracers underestimate SFR by about 60% (0.2 dex). A possible cause for this is the assumed limit that defines the high-metallicity range. Our high-metallicity group of regions is defined based on the median value of our data set, while that of Shipley et al. (2016) is a fixed value based on a different metallicity calibration. It is also likely that differences in the PAH band ratios have a role in the offset. In Section 3.4 we found that the ratio of 11.3 μm and 7.7 μm PAH bands varies within star-forming regions compared to the immediate surrounding area. This implies that a calibration such as Shipley et al. (2016) that is based on emission from a galaxy will then represent an average of such variations between PAH bands. Both this difference in the relative strength of PAH band emission and our assumed high-metallicity definition may contribute to the observed offset between the Shipley et al. (2016) and T33 SFR models.

4.5. Metallicity Dependence of PAH Properties

A number of previous studies have used photometric ratios of PAH-dominated 8 μm emission relative to continuum at 24 μm to trace the abundance of PAHs (Engelbracht et al. 2005; Draine et al. 2007) and to explore its metallicity dependence in galaxies. In our work, we can take a step further by measuring the integrated strength of individual PAH features normalized by neighboring continuum. Our results support the conclusion of these works that PAH abundance varies with metallicity. In addition, we found that normalizing PAH emission with 10 μm continuum behaves similarly to 24 μm as a function of metallicity. We have shown that all normalized PAH band correlations with metallicity have statistically equivalent Spearman coefficients ρ2 ≈ 0.5.

As an additional investigation we studied ratios of ionized to neutral PAHs after subtracting the local average in an annulus of outer radius 5.5′ surrounding the 3.5′ region. We found no significant change in the ratio of 6.2 μm to 7.7 μm emission, which is an indicator of the relative size of PAH molecules in some models (e.g., Draine & Li 2001). This suggests that in terms of size, PAH molecules are distributed in the surrounding
diffuse area similarly to how they are in the star-forming region. The background subtraction decreases the 11.3 μm to 7.7 μm PAH ratio, indicating an increase in the ionization of PAH molecules closer to the star-forming region. Such an increase could be explained by the harder and more intense radiation fields near the star-forming region compared to the diffuse interstellar medium (Bakes et al. 2001).

5. Conclusions

We have used resolved observations of thermal 33 GHz radio continuum, mid-IR emission from ionized neon and PAHs, and local metallicity to calibrate SFR tracers in nearby star-forming galaxies. We summarize our main conclusions as follows:

1. All individual infrared emission features considered show significant metallicity-dependent offsets in correlations with T33. Known hybrid tracers of SFR such as Hα and 24 μm are also found to have metallicity offsets with respect to T33.

2. ΣNe alone is found to better correlate with T33 than the commonly used hybrid SFR tracer of Hα and 24 μm emission. Residual scatter in the ΣNe–T33 correlation was found to be completely dependent on metallicity.

3. A mid-infrared SFR tracer with accuracy equivalent to thermal 33 GHz emission can be constructed using ΣNe and any PAH feature normalized by dust continuum. This accuracy is the same as using metallicity directly as the second parameter.

4. All PAH correlations with T33 are statistically similar and very weak. Including a correction for metallicity improves the correlation, but significant scatter remains.

5. Continuum-normalized PAH emission is highly metallicity dependent and follows similar trends to 8 μm/24 μm ratios, which are suggested to trace PAH abundance.

6. We found that combining tracers of opposite metallicity dependence in T33 correlations leads to a significant improvement in correlation strength. This trend is evident in combinations such as [Ne III]–24 μm, Hα–24 μm, and [Ne III]–PAH/10 μm. We proposed that this pairing of opposite metallicity dependence features shows a novel pattern for constructing SFR tracers wherein an indicator of ionized gas is corrected for variations due to metallicity by the additional indicator.

7. We found that subtracting local background emission from an H II region decreases the 11.3 μm/7.7 μm PAH ratio. This implies a decrease in the relative abundance of neutral PAH molecules closer to the star-forming region.

Our study is unique in using resolved thermal 33 GHz emission and local metallicity measurements for individual star-forming regions to calibrate SFR tracers. The mid-infrared SFR tracers presented in this study have wide applicability in analyzing future observations with the James Webb Space Telescope. The Mid-Infrared Instrument (MIRI) will detect 12.81 μm [Ne II] out to redshift z ≈ 1.2 and PAH 7.7 μm as far as z ≈ 2.5. In the 2030s, proposed telescopes such as the Space Infrared Telescope for Cosmology and Astrophysics (SPICA) and the Origins Space Telescope will have the potential to detect these features out to the epoch of reionization, enabling investigations throughout the full history of SF.

We thank the anonymous referee and the statistical referee for useful comments that improved the paper. The authors thank Daniela Calzetti, J. D. T. Smith, and Adam Leroy for useful discussions. C.W. and K.S. also thank Jeremy Chastenet, I-Da Chiang, and Petia Yanchulova Merica-Jones for their efforts over the years. We thank the SINGS team for their efforts in delivering processed data products for the community.

C.W. and K.S. acknowledge funding support from NASA ADAP grants NNX16AF48G and NNX17AF39G and National Science Foundation grant No. 1615728.

This research has made use of the NASA/IPAC Extragalactic Database (NED) and NASA/IPAC Infrared Science Archive (IRSA), which are operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. This work uses observations made with the Spitzer Space Telescope, which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. This research has made use of NASA’s Astrophysics Data System.

Appendix A
Optimal Correlations under Various Observational Restrictions

We investigated which fits with Equation (1) provided the most accurate model of observed T33 and subsequently the most accurate SFR under several data restrictions. This is useful because observations at higher redshift, for example, may exclude emission at longer wavelengths such as the 15.56 μm [Ne III] or 12.81 μm [Ne II] lines.

We find that the best SFR tracer that can be made exclusively from emission features between 5.2 and 15.6 μm uses X = ΣNe and Y = PAH 12.6 μm/10 μm. A histogram of the residual values for modeled T33 using this tracer is shown in Appendix C. Other continuum-normalized PAH bands result in statistically similar correlations (ρ ~ 1.0), so the 12.6 μm feature is chosen for observational convenience. This feature is nearest to the [Ne II] line at 12.81 μm, which must be measured to determine ΣNe and allows the range of necessary spectral observations to be reduced to about 6 μm wide from 10 to 16 μm (including continuum to measure 15.56 μm [Ne III]). SFRs can be derived from this 6 μm wide region of the mid-infrared spectrum with negligible residual metallicity dependence. The Mid-Infrared Instrument (MIRI) on the upcoming James Webb Space Telescope will detect this range of mid-IR spectrum in galaxies to redshift z < 0.8.

Observations at higher redshifts may exclude the 15.56 μm [Ne III] line, so we investigated the best tracer that can be made with emission features between 5.2 and 13.6 μm (z ≤ 1.12 for MIRI). We found in Section 3.1 that the best-correlated individual observable in this range is the 12.81 μm [Ne II] line. We find that the best tracer in this range is made by combining [Ne II] as the basis and including the nearby 12.6 μm PAH feature and its continuum-normalized form PAH 12.6/10 μm continuum, as shown in Table 4. This three-parameter tracer based on [Ne II], PAH 12.6 μm, and 10 μm continuum has ρ ~ equivalent to the commonly used Hα and 24 μm combination, about 0.78, and has a statistically insignificant Ωe, as shown in Figure 5.
We continued restricting our possible observables by excluding 12.81 μm [Ne II] to investigate the best tracer that can be made with emission features between 5.2 and 12.6 μm. In this range we are left with only PAH bands, of which there is no single feature that is better correlated with T33 than the others. For this reason we search for the best improvements to all PAH band correlations but with a focus on the strongest features such as 7.7 and 11.3 μm. We find that the best multifeature correlations in this range involve a bright PAH feature paired with a 10 μm continuum-normalized PAH band. However, the maximum value of $\rho_Z$ that can be achieved with these features is about 0.55, and the large $\Omega_Z$ of the individual PAH band tracers remains unaffected as shown in Figures A1(b) and (c). This is seen when comparing the individual tracer based on PAH 7.7 μm shown in Figure 3(b) with the PAH 7.7 μm and 7.7 μm/10 μm pair tracer shown in Figure A1(c).

In Section 3.1 we noted a trend where high-metallicity regions are significantly better correlated with T33 when considered separately from low-metallicity regions. We find that this trend is preserved when a normalized PAH feature is included as seen in Figure A1(c). This figure shows that the metallicity offset is similar to 7.7 μm alone (Figure 3(b)), but the overlap between high- and low-metallicity points is notably decreased. We find that the correlation coefficient with T33 for high-metallicity points separately from low reaches about 0.6, and likewise for low-metallicity regions separately from high.

**Appendix B**

**Correlations of Mid-IR Emission Feature Ratios with Metallicity**

We investigated the potential of a mid-IR metallicity tracer with a similar procedure to that of Section 2.7. We modify the general form of Equation (1) to solve for KK04 metallicities in terms of unitless [O/H] in place of T33 in units of mJy sr$^{-1}$. We also restrict our possible parameters to unitless quantities such as band ratios and continuum-normalized bands. These correlations use the set of 46 regions with high-confidence metallicity measurements as described in Section 2.4.

We find strong correlations between metallicity and ratios involving 15.56 μm [Ne III] emission. Figure B1 shows KK04 metallicities as a function of [Ne III] ratios with the 12.6 μm PAH band and the 12.81 μm [Ne II] line. We find that [Ne III] ratios with [Ne II] and all PAH bands are statistically equivalent to one another based on the Spearman correlation coefficient with metallicity $\rho_Z \sim -0.76$ for each. Figure B1 also shows that these models using ratios of 15.56 μm [Ne III] fail to accurately reproduce values for the lowest-metallicity regions (12 + log(O/H) < 8.7 KK04 or <8.2 PT05). Our sample contains two regions below this metallicity that are located in the dwarf galaxies Holmberg II and IC2754. These results for KK04 correlations are found to be true for correlations with PT05 metallicities as well. We find that including a second parameter $Y$ does not improve correlations between [Ne III] ratios and metallicity.

We also find significant correlations between metallicity and continuum-normalized PAH bands. Figure B2 shows three of these normalized PAH correlations with metallicity: PAH 12.6 μm/10 μm, PAH 7.7 μm/10 μm, and PAH 7.7 μm/24 μm. These normalized bands can be used to improve correlations with T33 and reduce the metallicity offset of single-feature tracers, as seen in Table 4. We find that normalizing by the continuum at 10 μm is equivalent to that of 24 μm in metallicity correlations with $\rho_Z \sim 0.54$ within 1σ. Compared to the 15.56 μm [Ne III] ratio metallicity correlations in Figure B1, these continuum-normalized PAH bands show a weaker correlation, but they more accurately predict the lowest-metallicity points. Similar to [Ne III] ratios, we find no improvement in correlations with metallicity by adding a second parameter $Y$. A tracer with $X = [\text{Ne III}]/12.6$ and $Y = 7.7/10 \mu m$, for example, improves their individual $\rho_Z$ values, but not by greater than the 1σ uncertainty.
Figure B1. KK04 metallicity as a function of 15.56 μm [Ne III] ratios: (a) [Ne III]/12.81 μm [Ne II] and (b) [Ne III]/PAH 12.6 μm.

Figure B2. KK04 metallicity as a function of 10 μm continuum-normalized PAH bands: (a) PAH 12.6 μm/10 μm and (b) PAH 7.7 μm/10 μm, as well as (c) PAH 7.7 μm band normalized by 24 μm continuum.
Appendix C
Individual Correlation Histograms

This section presents correlation histograms for single-feature tracers (Figure C1), two-feature tracers with $\Sigma$Ne (Figure C2), two-feature tracers with 24 $\mu$m (Figure C3), and two-feature tracers with $\Sigma$PAH (Figure C4).

**Figure C1.** Residual difference in predicted T33 for calibrations with (a) 12.81 $\mu$m [Ne II], (b) 15.56 $\mu$m [Ne III], and (c) $\Sigma$Ne.
Figure C2. Residual difference in predicted T33 using Equation (1) based on $\Sigma$Ne surface brightness and (a) KK04 metallicity values and (b) 10 $\mu$m continuum-normalized PAH 12.6 $\mu$m emission (12.6/10). All other normalized PAH features improve $\rho$, $\Omega_Z$, and $\sigma$ comparably to 12.6/10.

Figure C3. Residual difference in predicted T33 based on 24 $\mu$m emission and (a) H$\alpha$ and (b) 15.56 $\mu$m [Ne III] emission.
**Figure C4.** Residual difference in predicted T33 based on ΣPAH emission and (a) KK04 metallicity and (b) the ratio of 15.56 μm [Ne III] to 12.81 μm [Ne II].

### ORCID iDs

| C. M. Whitcomb | https://orcid.org/0000-0003-2093-4452 |
| K. Sandstrom | https://orcid.org/0000-0002-4378-8534 |
| E. J. Murphy | https://orcid.org/0000-0001-7089-7325 |
| S. Linden | https://orcid.org/0000-0002-1000-6081 |

### References

Allamandola, L. J., Tielens, A. G. G. M., & Barker, J. R. 1985, ApJL, 290, L25

Allamandola, L. J., Tielens, A. G. G. M., & Barker, J. R. 1989, ApJS, 71, 733

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 18

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33

Bakes, E. L. O., Tielens, A. G. G. M., & Bauschlicher, C. W. J. 2001, ApJ, 565, 501

Bauschlicher, C. W. J., Peeters, E., & Allamandola, L. J. 2008, ApJ, 678, 316

Berg, D. A., Pogge, R. W., Skillman, E. D., et al. 2020, ApJ, 893, 96

Buat, V. 1992, A&A, 264, 444

Burbidge, G. R., Gould, R. J., & Pottasch, S. R. 1963, ApJ, 138, 945

Calzetti, D., Kennicutt, R. C., & Engelbracht, C. W., et al. 2007, ApJ, 666, 870

Cluver, M. E., Jarrett, T. H., Dale, D. A., et al. 2017, ApJ, 850, 68

Condon, J. J. 1992, ARA&A, 30, 575

Cortese, L., Boselli, A., Franzetti, P., et al. 2008, MNRAS, 386, 1157

Curran, P. A. 2014, arXiv:1411.3816

Dale, D. A., Smith, J. D. T., Armus, L., et al. 2006, ApJ, 646, 161

Dale, B. T. 2003, ARA&A, 41, 241

Dale, B. T., & Anderson, N. 1985, ApJ, 292, 494

Dale, B. T., Dale, D. A., Bendo, G., et al. 2007, ApJ, 663, 866

Dale, B. T., & Li, A. 2001, ApJ, 551, 807

Engelbracht, C. W., Gordon, K. D., Rieke, G. H., et al. 2005, ApJL, 628, L29

Engelbracht, C. W., Rieke, G. H., Gordon, K. D., et al. 2008, ApJ, 678, 804

Förster Schreiber, N. M., Galliano, F., Sauvage, M., & Charmandaris, V. 2004, A&A, 419, 501

Gallazzi, A., Brinchmann, J., Charlot, S., & White, S. D. M. 2008, MNRAS, 383, 1439

Gavazzi, G., Boselli, A., Franzetti, P., et al. 2008, MNRAS, 386, 1157

Galliano, F., Armus, L., Rieke, G. H., et al. 2005, ApJL, 630, L87

Galliano, F., Armus, L., Rieke, G. H., et al. 2006, ApJL, 638, L91

Galliano, F., Armus, L., Rieke, G. H., et al. 2007, ApJ, 663, 866

Galliano, F., Armus, L., Rieke, G. H., et al. 2008, ApJ, 682, 336

Helou, G., Lu, N. Y., Werner, M. W., Malhotra, S., & Silbermann, N. 2000, ApJL, 532, L21

Ho, L. C., & Keto, E. 2007, ApJ, 658, 314

Houck, J. R., Roellig, T. L., van Cleve, J., et al. 2004, ApJS, 154, 18

Kennicutt, R. C., Calzetti, D., Aniano, G., et al. 2011, PASP, 123, 1347

Kennicutt, R. C., & Evans, N. J. 2012, ARA&A, 50, 531

Kennicutt, R. C., Calzetti, D., Walter, F., et al. 2007, ApJ, 671, 333

Kennicutt, R. C. J., Hao, C.-N., Calzetti, D., et al. 2009, ApJ, 703, 1672

Kobulnicky, H. A., & Kewley, L. J. 2004, ApJL, 617, 240

Lebouteiller, V., Brandl, B., Bernard-Salas, J., Devost, D., & Houck, J. R. 2007, ApJ, 656, 390

Legor, A., & Pugot, J. L. 1984, A&A, 500, 279

Leja, J., Johnson, B. D., Conroy, C., et al. 2019, ApJ, 877, 140

Leroy, A. K., Bigiel, F., de Blok, W. J. G., et al. 2012, AJ, 144, 3

Leroy, A. K., Sandstrom, K. M., Lang, D., et al. 2019, ApJS, 244, 24

Linden, S. T., Murphy, E. J., Dong, D., et al. 2020, ApJS, 248, 25

Lu, N., Helou, G., Werner, M. W., et al. 2003, ApJ, 588, 199

Madau, P., & Dickinson, M. 2014, A&A&A, 52, 415

Madden, S. C., Galliano, F., Jones, A. P., & Sauvage, M. 2006, A&A, 446, 877

Meurer, G. R., Heckman, T. M., Leitherer, C., et al. 1995, AJ, 110, 2665

Moustakas, J., Jr., R. C., Tremonti, C. A., et al. 2010, ApJS, 190, 233

Murphy, E. J., Condon, J. J., Schinnerer, E., et al. 2011, ApJ, 737, 67

Murphy, E. J., Dong, D., Monajian, E., et al. 2018, ApJS, 234, 24

Murphy, E. J., Helou, G., Condon, J. J., et al. 2010, ApJL, 709, L108

O'Halloran, B., Satyapal, S., & Dudik, R. P. 2006, ApJ, 641, 795

Osterbrock, D. E. 1965, ApJ, 142, 1423

Peeters, E., Hony, S., Van Kerckhoven, C., et al. 2002, A&A, 390, 1089

Peeters, E., Spoon, H. W. W., & Tielens, A. G. G. M. 2004, ApJ, 613, 986

Pepe, E., Santos, N. C., Alibert, Y., et al. 2010, A&A, 518, 59

Pilyugin, L. S., & Thuan, T. X. 2005, ApJ, 631, 231

Relaño, M., Liskenfeld, U., Pérez-González, P. G., Vilchez, J. M., & Battaner, E. 2007, ApJL, 667, L141

Roche, P. F., Aitken, D. K., Smith, C. H., & Ward, M. J. 1991, MNRAS, 248, 606

Roussel, H., Sauvage, M., Viguoux, L., & Bosma, A. 2001, A&A, 372, 427

Sandstrom, K. M., Leroy, A. K., Walter, F., et al. 2013, ApJ, 777, 5

Schutte, W. A., Tielens, A. G. G. M., & Allamandola, L. J. 1993, ApJ, 415, 397

Sellgren, K. 1984, ApJ, 277, 623

Shapley, H. V., Papovich, C., Rieke, G. H., Brown, M. J. I., & Moustakas, J. L. 2016, ApJ, 818, 60

Shivaei, I., Reddy, N. A., Shapley, A. E., et al. 2017, ApJ, 837, 157

Smith, J. D. T., Draine, B. T., Dale, D. A., et al. 2007, ApJ, 656, 770


Tielens, A. G. G. M. 2008, ARA&A, 46, 289
Treyer, M., Schiminovich, D., Johnson, B. D., et al. 2010, ApJ, 719, 1191
Wen, X.-Q., Wu, H., Zhu, Y.-N., et al. 2014, MNRAS, 438, 97
Werner, M. W., Roellig, T. L., Low, F. J., et al. 2004, ApJS, 154, 1
Wu, H., Cao, C., Hao, C.-N., et al. 2005, ApJL, 632, L79
Xie, Y., & Ho, L. C. 2019, ApJ, 884, 136
Zhuang, M.-Y., Ho, L. C., & Shangguan, J. 2019, ApJ, 873, 103
Woosley, S. E., & Weaver, T. A. 1995, ApJS, 101, 181