Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **526,** 4520–4528 (2023) https://doi.org/10.1093/mnras/stad2524
Advance Access publication 2023 September 4

Emulating radiative transfer with artificial neural networks

Snigdaa S. Sethuram [1,2★] Rachel K. Cochrane [2] Christopher C. Hayward [2] Viviana Acquaviva,[2,3] Francisco Villaescusa-Navarro,[2,4] Gergö Popping [5] and John H. Wise[1]

[1]*Center for Relativistic Astrophysics, Georgia Institute of Technology, North Avenue, Atlanta, GA 30332, USA*
[2]*Center for Computational Astrophysics, Flatiron Institute, 162 5th Ave., New York, NY 10010, USA*
[3]*Department of Physics, CUNY NYC College of Technology, 300 Jay Street, Brooklyn, NY 11201, USA*
[4]*Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08544, USA*
[5]*European Southern Observatory, Karl-Schwarzschild-Str. 2, D-85748 Garching, Germany*

Accepted 2023 August 16. Received 2023 July 31; in original form 2023 June 13

ABSTRACT

Forward-modeling observables from galaxy simulations enables direct comparisons between theory and observations. To generate synthetic spectral energy distributions (SEDs) that include dust absorption, re-emission, and scattering, Monte Carlo radiative transfer is often used in post-processing on a galaxy-by-galaxy basis. However, this is computationally expensive, especially if one wants to make predictions for suites of many cosmological simulations. To alleviate this computational burden, we have developed a radiative transfer emulator using an artificial neural network (ANN), *ANNgelina*, that can reliably predict SEDs of simulated galaxies using a small number of integrated properties of the simulated galaxies: star formation rate, stellar and dust masses, and mass-weighted metallicities of all star particles and of only star particles with age <10 Myr. Here, we present the methodology and quantify the accuracy of the predictions. We train the ANN on SEDs computed for galaxies from the *IllustrisTNG* project's TNG50 cosmological magnetohydrodynamical simulation. *ANNgelina* is able to predict the SEDs of TNG50 galaxies in the ultraviolet (UV) to millimetre regime with a typical median absolute error of $\sim$7 per cent. The prediction error is the greatest in the UV, possibly due to the viewing-angle dependence being greatest in this wavelength regime. Our results demonstrate that our ANN-based emulator is a promising computationally inexpensive alternative for forward-modeling galaxy SEDs from cosmological simulations.

Key words: radiative transfer – methods: statistical – galaxies: evolution.

1 INTRODUCTION

Forward-modeling observables from galaxy formation simulations provides a means to confront theoretical models with observations directly. This approach represents an alternative to the more-traditional method of inferring physical properties such as stellar mass and star formation rate (SFR) from observations and comparing those with the corresponding quantities from simulations. One advantage of forward-modeling observables from a simulation is the avoidance of 'throwing away' information from the simulation; e.g. when predicting spectral energy distributions (SEDs), the full star formation history of a simulated galaxy is used.

One particularly common forward-modeling technique in the galaxy formation community is to predict ultraviolet-to-millimeter (UV-to-mm) SEDs of simulated galaxies, including both integrated SEDs and images in various observed bands (e.g. Jonsson 2006; Jonsson, Groves & Cox 2010; Camps & Baes 2015; see Steinacker, Baes & Gordon 2013 for a review). This calculation is normally done by performing dust radiative transfer on simulated galaxies in post-processing to compute how light from star particles propagates through the simulated galaxy's interstellar medium (ISM) and is absorbed, scattered, and re-emitted by interstellar dust. This approach has been applied to compare model predictions with observations of various classes of objects such as massive galaxies or active galactic nuclei (e.g. Hayward et al. 2012; Lanz et al. 2014; Narayanan et al. 2015; Safarzadeh, Hayward & Ferguson 2017b; Cochrane et al. 2019, 2023a, b; Baes et al. 2020; Parsotan et al. 2021; Cochrane, Hayward & Anglés-Alcázar 2022). Moreover, such synthetic observations can be used to perform ground-truth experiments to test and refine techniques for inferring physical quantities from observations (e.g. Wuyts et al. 2010; Michałowski et al. 2014; Hayward & Smith 2015; Smith & Hayward 2018; McKinney et al. 2021; Cochrane et al. 2022).

Unfortunately, this forward-modeling step can incur significant additional computational expense and requires detailed outputs from simulations (the full 3D density distribution of stars and dust, for example). This required detailed information is not always available, in particular for coarse-resolution simulations. In such scenarios, the Monte Carlo radiative transfer (MCRT) calculations must make various assumptions (e.g. regarding sub-grid dust clumpiness) that can affect the robustness of the predicted observables. It is thus desirable to find a computationally inexpensive, robust method for making such predictions from a limited amount of data, such as integrated galaxy properties. Having a fast method available to run radiative transfer with varying input parameters would allow predictions to be made for different input parameters, which will ultimately allow us to marginalize over the uncertainties associated with these parameters.

★ E-mail: snigdaa.ram@gatech.edu

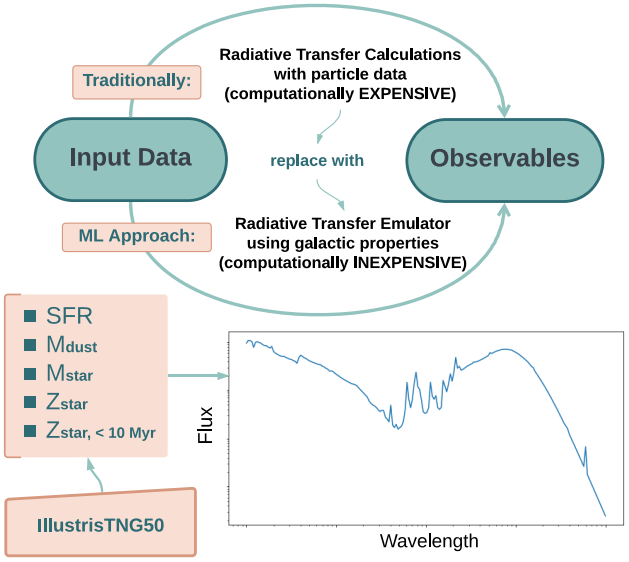


Figure 1. *ANNgelina* project workflow. In the traditional approach, computationally expensive MCRT calculations are needed to forward-model SEDs and images from simulated galaxies. In this work, we demonstrate how an artificial neural network can be used to predict SEDs of simulated galaxies, accelerating this process by at least seven orders of magnitude.

Several previous works (Hayward et al. 2011; Safarzadeh et al. 2016; Lovell et al. 2021; Cochrane et al. 2023b) have demonstrated that it is possible to predict observed-frame far-IR (FIR) fluxes of simulated galaxies using only a small number of integrated properties of the galaxies, such as SFR and dust mass. The success of these approaches suggests that geometric variations amongst simulated galaxies play a subdominant role in determining their FIR SEDs and gives confidence that one can employ such simple approaches as an alternative to full MCRT, as has been done in a handful of works (e.g. Hayward et al. 2013, 2021; Miller et al. 2015; Safarzadeh, Lu & Hayward 2017a; Popping et al. 2020; Cochrane et al. 2023b). However, these works have focused on the thermal dust emission; it is likely more difficult to predict full UV-mm SEDs because geometry plays a greater role in the UV-optical due to the non-isotropic nature of dust attenuation on galaxy scales. Nevertheless, in this work, we aim to predict UV-mm SEDs using only a small number of integrated properties of simulated galaxies, without specifying any information about the galaxy geometry. The general workflow of the project is shown in Fig. 1.

Using machine learning (ML) techniques to emulate computationally intensive calculations in simulations has shown substantial promise in astrophysics (Buchner 2019; Kasim et al. 2021; Bird et al. 2022) as well as other computationally intensive STEM fields, e.g. climate modeling (Weber et al. 2020). Motivated by such works and the demonstrated possibility of predicting FIR SEDs from a small number of galaxy properties, in this work, we develop an artificial neural network (ANN) emulator for MCRT calculations on simulated galaxies. Neural networks (NNs) are deep learning algorithms, a type of ML that employs multiple layers of computation to learn trends and correlations in data. NNs are robust at understanding complex relationships between input and output data due to their iterative calculations that optimize the data fit in pieces. ML algorithms take into account not only the non-linear mapping of input to output data but also the distribution of properties across the full training set (Berner, Elbrächter & Grohs 2019).

In this paper, we explore the possibility of an ANN reproducing non-linear relationships between galaxy properties and galaxy SEDs. Our work addresses the following question: 'given a set of 3D MCRT calculations with fixed assumptions (regarding e.g. the stellar initial mass function, single-age stellar population SED templates, and dust model) performed on galaxies selected from a single cosmological simulation, how well can an ANN emulate the MCRT calculations to predict integrated UV-mm SEDs of the simulated galaxies'? The inverse workflow has been attempted, including using ML to derive galaxy properties from a given SED (Gilda, Lower & Narayanan 2021) and using ML to derive star formation histories for galaxies in the *Illustris* and EAGLE simulations (Lovell et al. 2019), but to the best of our knowledge, using an NN to predict galaxy SEDs has not been previously attempted. We use SEDs from the TNG50 cosmological magnetohydrodynamical simulation (Nelson et al. 2019; Pillepich et al. 2019), which is the highest-resolution simulation in the *IllustrisTNG* suite (TNG) (Nelson et al. 2019; Pillepich et al. 2019), to train our ANN, *ANNgelina*.[1]

The structure of this paper is as follows. In Section 3, we discuss the methods used to create *ANNgelina* and present an overview of the TNG50 data set used for training. In Section 4, we present our preliminary results, touching on some of the interpretation that went into analyzing *ANNgelina*'s performance. We discuss applications and limitations of *ANNgelina*, as well as improvements to be made in future work. We draw conclusions in Section 5.

2 DATA

The version of *ANNgelina* released with this paper (ANNGELINA_V1.0) was trained on synthetic SEDs obtained by running a radiative transfer code on galaxies in the TNG50 (Nelson et al. 2019; Pillepich et al. 2019) simulation. Here, we provide an overview of the TNG simulations, the sample of galaxies selected from TNG50, and the methods used to model their SEDs.

2.1 IllustrisTNG simulations

The *IllustrisTNG* project[2] is a suite of large-volume cosmological magnetohydrodynamical simulations that model the formation of galaxies from early times to $z = 0$. The numerical methods and galaxy formation model used for the simulations are described in detail in Pillepich et al. (2017), Weinberger et al. (2018), and Pillepich et al. (2018), so we will only briefly summarize them here.

The simulations were run with AREPO[3] (Springel 2010), an unstructured moving-mesh hydrodynamics code. Star formation is implemented by stochastically spawning stellar particles at a rate set by a volume density-dependent Kennicutt–Schmidt law (Schmidt 1959; Kennicutt 1998); a density threshold of 0.13 cm^{-3} is employed. Stellar populations are evolved self-consistently, including chemical enrichment and gas recycling, resulting in mass loss. An effective equation of state is used to approximate how supernovae (SNe) heat the ISM, and stellar feedback-driven winds are implemented by stochastically isotropically kicking and temporarily hydrodynamically decoupling gas cells (Springel & Hernquist 2003). Black hole accretion is modelled as Eddington-limited modified Bondi–Hoyle accretion. A two-mode active galactic nucleus (AGN) feedback model is employed: a relatively efficient kinetic channel at low

[1] https://github.com/snigdaa/runANNgelina
[2] https://www.tng-project.org
[3] https://gitlab.mpcdf.mpg.de/vrs/arepo

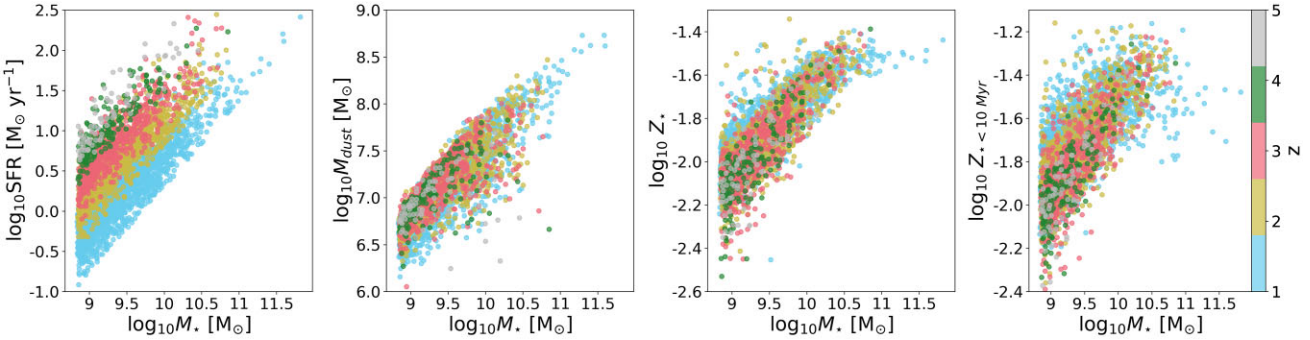


Figure 2. Correlations between various integrated properties and stellar mass for all galaxies from TNG50 with $M_\star > 10^9$ M$_\odot$, illustrating the region of parameter space spanned by our simulated galaxy sample: SFR (*left-most*), dust mass (*middle-left*), and mass-weighted metallicities of all stars (*middle-right*) and young (<10 Myr-old) stars (*right-most*). In each panel, the points are colour-coded by redshift. All properties have been calculated within twice the stellar half-mass radius of each galaxy. As expected, all of the considered quantities correlate with total stellar mass. The normalization of the SFR-$M_\star$ relation increases with increasing redshift, whereas the $M_{\rm dust}$–$M_\star$ correlation is independent of redshift for the range considered.

Eddington ratio ('radio-mode') and a less-efficient thermal channel at high Eddington ratio ('quasar mode'; see Weinberger et al. 2018 for details). The free parameters in the subgrid models employed were tuned to match the galaxy stellar mass function and stellar mass-to-halo mass relation at $z \sim 0$, in addition to the cosmic SFR history at $z \lesssim 10$. The black hole mass–stellar mass relation, halo gas fractions and stellar half-mass radii of galaxies were also considered (see section 3.2 of Pillepich et al. 2017). The TNG simulations adopt a cosmology consistent with the Planck 2015 results (Ade et al. 2016): $\Omega_m = 0.31$, $\Omega_\Lambda = 0.69$, $\Omega_b = 0.0486$, $h = 0.677$, $\sigma_8 = 0.8159$, and $n_s = 0.97$ (Nelson et al. 2019).

2.1.1 *IllustrisTNG50 data set*

We use the highest-resolution simulation from the *IllustrisTNG* suite, TNG50 (Nelson et al. 2019), which has a volume of $(35\,h^{-1}\,\mathrm{Mpc})^3$. The mass resolution is 8.5×10^4 M$_\odot$ for baryonic particles and 4.5×10^5 M$_\odot$ for dark matter particles. TNG50 employs a collisionless softening length of 0.3 kpc at $z = 0$ and an adaptive gas softening length with a minimum of 74 comoving pc.

We draw our sample of TNG50 galaxies from Popping et al. (2022), who modelled integrated SEDs for approximately 3800 star-forming galaxies[4] on or above the star formation main sequence with $M_\star > 10^{8.7}$ M$_\odot$. Our sample is comprised of 1709 (central and satellite) galaxies at $z = 1$, 1149 galaxies at $z = 2$, 620 galaxies at $z = 3$, 184 galaxies at $z = 4$, and 76 galaxies at $z = 5$. There are fewer SEDs available for the higher-redshift snapshots, since fewer galaxies meet the redshift-independent stellar mass criterion. Fig. 2 shows the distribution of the parameter space covered by our TNG50 data set.

2.2 Generation of synthetic SEDs

Synthetic galaxy SEDs were generated by Popping et al. (2022) using the SKIRT MCRT code (Camps & Baes 2015).[5] SKIRT takes the 3D dust and stellar density distributions from a galaxy simulation and propagates photon packets from radiation sources (i.e. star particles; AGN emission was not modeled) through the simulated galaxies' ISM using a Monte Carlo approach to model dust absorption, scattering, and re-emission. Popping et al. (2022) followed the methods described by Schulz et al. (2020) to calculate the SEDs for TNG50 galaxies. We briefly summarize their methods here.

Gas and stellar particles within 7.5-times the stellar half-mass radius were extracted from subhalos. Star particles with ages greater than 10 Myr were assigned model template SEDs based on their age and metallicity, according to the Bruzual & Charlot (2003) single-age stellar population model. Star particles with ages ≤10 Myr were assigned SEDs from Groves et al. (2008), which include a sub-grid model for H II and photodissociation regions. Dust was modelled using the gas-phase metal density distribution, with a constant dust-to-metals mass ratio of 0.4 (e.g. Dwek 1998; James et al. 2002). Gas cells with $T > 7.5 \times 10^4$ K were assumed to contain no dust due to dust destruction via sputtering. The Weingartner & Draine (2001) Milky Way dust model was employed, with a mix of graphite, silicate, and polycyclic aromatic hydrocarbon (PAH) grains. Self-absorption of dust grains was not taken into account, but this is unlikely to be significant for the bulk of the galaxy population considered here (e.g. Hayward et al. 2011).

The modeled SEDs span rest-frame $0.1 - 1000$ μm and are sampled at 200 uniformly log-spaced wavelengths. The simulated galaxies were 'observed' 'face-on' (i.e. the detector was placed such that the line of sight was along the angular momentum vector of the simulated galaxy, corresponding to face-on projections for simulated disc galaxies).

2.2.1 *Modeled SEDs for TNG50 galaxies*

In Fig. 3, we show examples of modeled SEDs for TNG50 galaxies in several stellar mass bins. The upper panel shows SEDs colour-coded by SFR. The lower panel shows SEDs colour-coded by redshift. We use this visualization to understand trends in our data. As expected, higher SFRs generally yield higher fluxes. There are interesting trends at short wavelengths, though, with some high-SFR SEDs displaying particularly low fluxes at $\lambda \lesssim 1$ μm, owing to high levels of dust attenuation. We also note that SFR correlates with the shape of the FIR SED: the FIR SED peaks at shorter wavelengths for galaxies with higher SFRs. For a given stellar mass bin, higher-redshift (i.e. earlier-forming) galaxies are also brighter at all wavelengths (note that these SEDs are generated in the rest frame). This likely reflects

[4]Here, a single halo is considered a distinct galaxy in each of the snapshots. Given the large redshift spacing between snapshots, this assumption is reasonable.
[5]https://skirt.ugent.be/root/_home.html

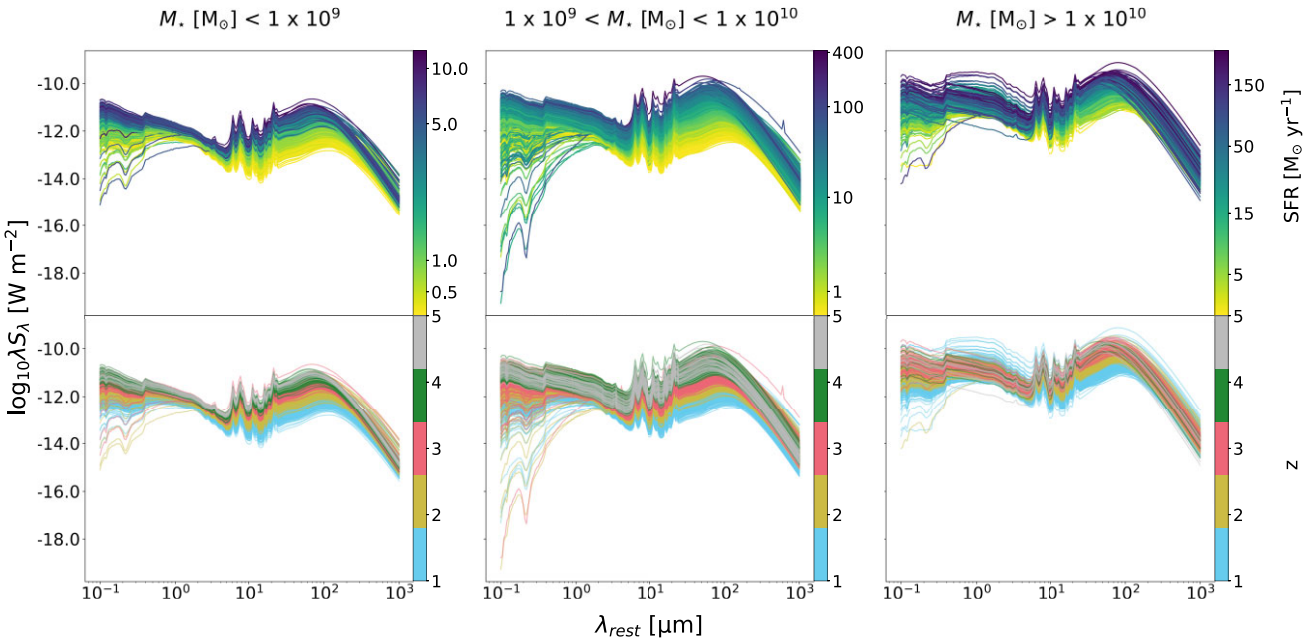


Figure 3. TNG50 SEDs used to train our ANN. The top row shows the SEDs colour-coded by SFR, and the bottom row shows the SEDs colour-coded by redshift; the columns show different stellar mass bins, as indicated at the tops of the figures. In a given mass bin, galaxies with higher SFRs tend to have higher IR fluxes, and their FIR SEDs are hotter. In the two lower-mass bins, the UV-optical fluxes also increase with SFR. This is less evident in the highest-mass bin due to increasing dust attenuation. In all mass bins, there is a subset of galaxies with high SFRs and heavily reddened UV-optical SEDs. Higher-redshift galaxies tend to have larger fluxes due to the evolution of the star formation main sequence (at fixed mass, SFR increases with increasing redshift).

the evolution of the 'main-sequence': at a given stellar mass, higher-redshift star-forming galaxies have higher SFRs and hence higher intrinsic (pre-dust-attenuated) fluxes.

3 METHODS

We design an ANN, *ANNgelina*, to emulate SED generation of TNG50 galaxies and hence bypass the radiative transfer procedure described in Section 2.2. An ANN architecture consists of an input layer with several 'features'; one or more hidden layers, in which the features are manipulated; and an output layer, which contains the desired outputs, or 'labels'. The structure of *ANNgelina* is illustrated in Fig. 4. We use a fully connected network, where each neuron talks to all neurons in the previous and following layers. Each neurons in the input layer corresponds to a galaxy property, such as stellar mass or SFR, and each neuron in the output later corresponds to the flux at one wavelength. *ANNgelina* is built with *PyTorch* (Paszke et al. 2019).[6] In the following sections, we describe the architecture and parameter choices in more detail and outline the training procedure.

3.1 Data normalization and partition

Each input feature, X, is normalized as follows:

$$X_{\rm norm} = \frac{\log_{10} X - \mu(\log_{10} X)}{\sigma(\log_{10} X)}. \quad (1)$$

For any features that had zero values, we added a negligible nonzero value to that feature before normalizing to avoid taking the logarithm of zero. This normalization procedure accelerates the training of the network.

[6] https://pytorch.org/

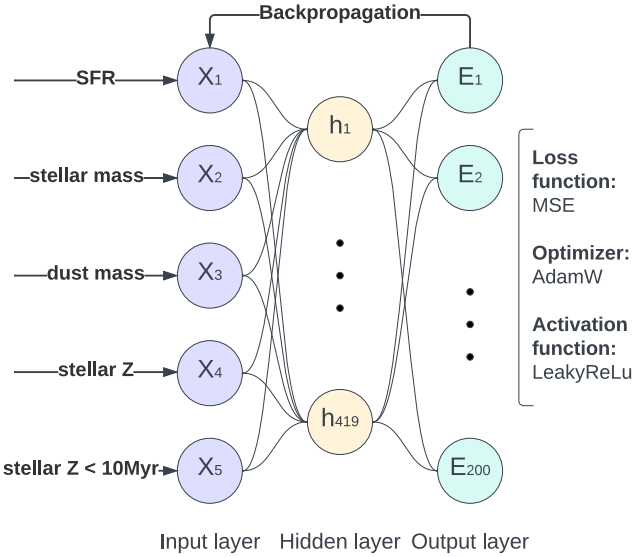


Figure 4. The architecture of *ANNgelina*, with input features labeled $X_1 - X_5$ (purple), neurons in the hidden layer labeled $h_1 - h_{419}$ (orange), and output neurons labeled $E_1 - E_{200}$ (green). Not all neurons are shown. The following galaxy properties are used as features: SFR, stellar mass, dust mass, and stellar metallicities of both all and young (<10 Myr old) stars. The output comprises rest-frame fluxes at 200 wavelengths, uniformly spaced in log-space. We note the loss function, optimizer and activation functions used in training.

The flux is input as the irradiance, E, defined as $E = \lambda S_\lambda/(\mathrm{W\,m^{-2}})$. We use $\log_{10}E$ as the labels in the network. Data are then randomized and partitioned into training (70 per cent), validation (15 per cent), and test (15 per cent) sets.

3.2 Training the ANN

Here, we provide a broad overview of the training process and various hyper-parameter and function choices. During a 'forward pass', the ANN pushes the training data through the hidden layers by using the current values of the weights for each feature. After each forward pass, the model updates itself through a 'backpropagation' phase, in which it determines a loss value (based on the difference between the predicted and true outputs) and updates the weights according to an optimization function, usually a type of gradient descent. The number of times the forward pass and backpropagation cycle is executed depends on the pre-defined batch size and number of epochs.

Several parameters define the workflow of the network. Hyper-parameters such as the number of hidden layers and neurons in each layer, dropout fraction, weight decay, and activation function are related to the structure of the network. These hyper-parameters were optimized as described in Section 3.3. The learning rate, number of epochs, and batch size affect the training procedure; appropriate choices ensure generalizability of the model and improved efficiency while training. The batch size (128) and number of epochs (1500) were tuned manually. The batch size refers to the number of samples to be propagated through the network in one full pass. One epoch is completed once all training samples have been propagated through the network. The initial learning rate, dropout fraction, weight decay, number of hidden layers, and number of neurons per hidden layer were determined using *Optuna* (Akiba et al. 2019), as described in Section 3.3.

We use 'dropout', a simple and efficient regularization technique that consists of removing a random subset of weights at every training epoch, thereby ensuring that the learned weights are more robust and preventing over-fitting. The dropout fraction determines what percentage of weights are removed. Weight decay works in tandem with dropout and is another regularization technique applied to all weights in a network. There is an extra term added to the loss function to reduce the variance in the network weights.

We adopt the following optimizer, activation and loss functions:

Optimizer function: We chose the optimizer function AdamW, which is a modified version of the Adam optimizer (Kingma & Ba 2014) in which the learning rate and the weight decay of the model are optimized separately. It takes an initial learning rate and weight decay value and updates them for each feature's weight individually while performing gradient descent. It performs well at handling non-smooth data. The learning rate controls how much the model changes in response to predicted error during each forward pass.

Activation function: We use the LeakyReLu (Leaky Rectified Linear Unit) activation function (Maas et al. 2013), which is defined as

$$y = \begin{cases} x & (x \geq 0), \\ 0.01x & (x < 0). \end{cases} \tag{2}$$

Loss function: NN losses in supervised learning models are determined by quantifying the difference between the predicted and true outputs using a loss function. We adopt the mean square error (MSE) as our loss function. The MSE guarantees that the prediction represents the posterior mean without assuming any shape for the posterior distribution. During the training phase, the network minimizes the MSE averaged over each batch. The MSE for a batch of size b is given by

$$\mathrm{MSE} = \frac{1}{200 \cdot b} \sum_{i=1}^{b} \sum_{\lambda=1}^{200} \left[\log_{10}(E_{\lambda,\, i,\, \mathrm{pred}}/E_{\lambda,\, i,\, \mathrm{true}})\right]^2 \tag{3}$$

Table 1. Correlation matrix for TNG50 galaxy properties.

Feature	M_{dust}	$M_\star$	$Z_\star$	$Z_{\star<10\,\mathrm{Myr}}$	SFR
M_{dust}	1.0	0.88	0.41	0.28	0.48
$M_\star$	0.88	1.0	0.44	0.29	0.48
$Z_\star$	0.41	0.44	1.0	0.9	0.37
$Z_{\star<10\,\mathrm{Myr}}$	0.28	0.29	0.9	1.0	0.25
SFR	0.48	0.48	0.37	0.25	1.0

where λ represents the wavelength index of the SED, given 200 wavelengths sampled between 0.1 and 1000 μm, and $E_{\lambda,i}$ is the irradiance at a given wavelength, λ, for a given galaxy in the batch.

3.3 Hyperparameter optimization

We use the *Optuna* hyperparameter tuning package (Akiba et al. 2019) to tune the follwing model parameters: learning rate, dropout fraction, weight decay, number of hidden layers, and number of neurons per hidden layer. After testing with up to 11 hidden layers, each with up to 1000 neurons, optimal performance was achieved with a shallow, one hidden layer architecture with 419 neurons, as shown in Fig. 4. All hyperparameter values for our optimized model can be found in the public repository listed in Section 2.

3.4 Galaxy feature selection tests

This version of *ANNgelina* uses just five galaxy properties (SFR, M_{dust}, $M_\star$, $Z_\star$, and $Z_{\star<10\,\mathrm{Myr}}$) to predict SEDs. While designing *ANNgelina*, we tested five additional galaxy properties: gas mass (M_{gas}), redshift (z), half-SFR radius, half-stellar-mass radius, and half-dust-mass radius.

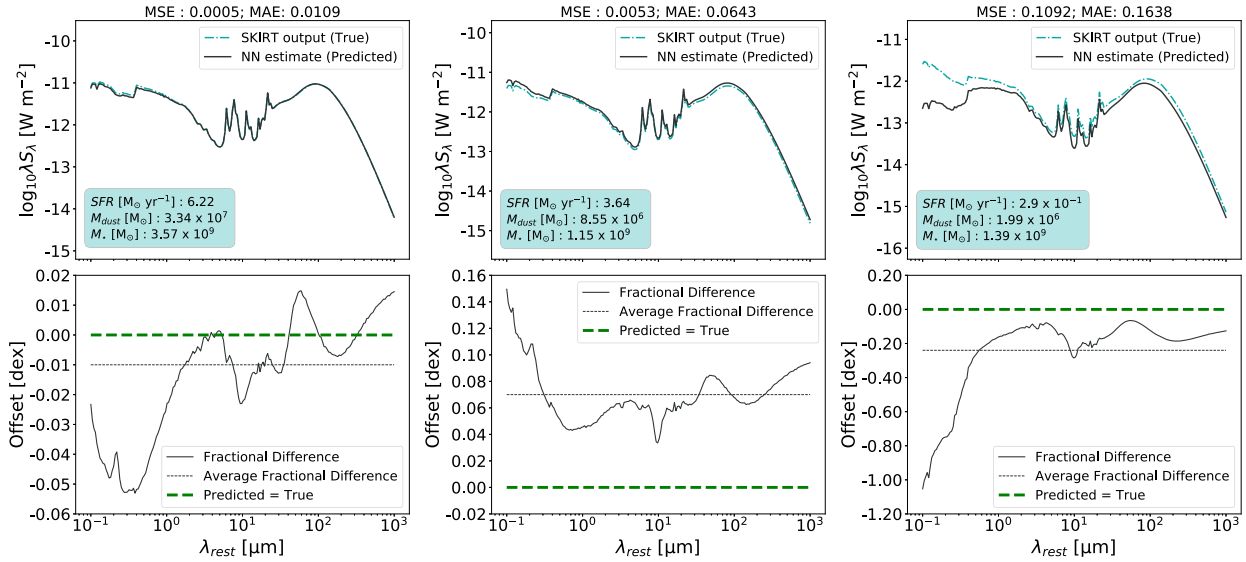
We first tuned an ANN using all ten features ('full' model). We then trained ten further nine-feature ANNs, using the same architecture but with a different feature removed. In all cases, all input features were normalized, as described in equation (1). We compared the loss values of each of the nine-feature ANN to that of the 'full' model. We eliminated features that did not significantly impact the loss values of the model. It is reassuring that the features that were determined to significantly affect the model are all properties that should matter according to physical intuition (e.g. SFR, dust mass). Redshift is not important because we are predicting SEDs in the rest frame.

In Table 1, we show the correlation matrix of the galaxy properties used in *ANNgelina*. Correlations are calculated as the covariance of each variable pair divided by the product of their standard deviations. A correlation with an absolute magnitude close to 1.0 indicates very strong correlation, while an absolute magnitude closer to 0.0 indicates a weak correlation. While some properties are strongly correlated (most notably stellar and dust masses), *ANNgelina*'s performance is impacted negatively if any of the listed properties are omitted from the input feature set, indicating that even strongly correlated features contribute non-redundant information.

4 RESULTS AND DISCUSSION

4.1 Characterizing *ANNgelina*'s overall performance

We assess the performance of *ANNgelina* using a test set of TNG50 galaxies that were not included in the training set. Here, we define the metrics used to characterize how well SEDs are predicted. We define the offset of the predicted (log-space) SED from the target



(a) Among the best predictions in the test set, with an MSE corresponding to the 2nd percentile of the sample's distribution.

(b) A typical example (50th percentile).

(c) One of the least successful predictions (98th percentile).

Figure 5. Three example SEDs predicted using *ANNgelina*, with MSE increasing from left to right. In each column, the top panel shows the true SED computed using SKIRT (cyan, dot–dashed) and that predicted using *ANNgelina* (black). The bottom panels show the logarithm of the absolute fractional difference (black lines). The average value across all wavelengths is indicated via the horizontal short-dashed black lines; the green dashed horizontal lines indicate zero prediction error. The labels at the top of each column indicate the MSE and MAE (i.e. median absolute fractional difference) values. The typical absolute fractional difference between true and predicted SEDs is in the range from 0.03 to 0.07. The average fractional difference lines indicate that our model has low bias, even for the worst example shown (right column). Note that the UV emission is more poorly predicted than any other part of the SED in all cases.

(SKIRT-predicted) SED at a given wavelength λ as follows:

$$\mathrm{Offset}_\lambda/\mathrm{dex} = \log_{10}(\mathrm{E}_{\lambda,\mathrm{predicted}}/\mathrm{E}_{\lambda,\mathrm{true}}). \quad (4)$$

For each galaxy in the test set, we take the median absolute offset value across all 200 wavelengths, and call this the MAE:

$$\mathrm{MAE/dex} = \mathrm{median}\,|\mathrm{Offset}_\lambda/\mathrm{dex}|\,. \quad (5)$$

We use the MAE as a single value characterizing the SED prediction for each galaxy. We note that, because of our definition, the MAE is essentially equivalent to the Median Absolute Percentage Error in the low-error regime.

In Fig. 5, we show several examples of target and predicted SEDs for galaxies in our test set. From left to right, we show predicted SEDs with MSEs at the ≈ 2nd, 50th, and 98th percentiles of the MSE distribution, where the 98th percentile refers to the worst of the three predictions, or the highest MSE. The top panels show the true SED calculated using SKIRT in cyan and that predicted by *ANNgelina* in black. The bottom panels show the logarithm of the prediction error versus wavelength. The MSE and MAE are quoted at the tops of the columns. For the second percentile example, the true and predicted SEDs are essentially indistinguishable – the SED is predicted to within ∼ 0.05 dex (∼ 12 per cent) across the full UV-mm. The MAE is 0.0109 dex (2.5 per cent). For the 50th percentile example, *ANNgelina* overpredicts the SED at both UV and FIR wavelengths, but the error is at most ∼ 0.15 dex (∼ 41 per cent). The MAE is 0.0643 dex (16 per cent). The 98th percentile example displays a catastrophic failure. The UV luminosity density is underpredicted by an order of magnitude. The MAE is 0.1638 dex (46 per cent).

We compile MAE values for all test-set galaxies in Fig. 6. *ANNgelina* performs well on the TNG50 data set we employ, with an average MAE of ∼ 0.06 dex (15 per cent), with the worst predictions having an MAE of ∼ 0.2 dex (60 per cent). In Fig. 7, we show 2D projections of the input feature space coloured by MSE. If a particular region of parameter space were problematic for *ANNgelina*, there would be local maxima in the MSE distribution. We see no such local maxima. This indicates that *ANNgelina* does not exhibit higher MSE values in any one part of the parameter space; rather, it predicts the SEDs uniformly well throughout the parameter space probed. The scatter in the predictions may be due to additional variables not included in the analysis.

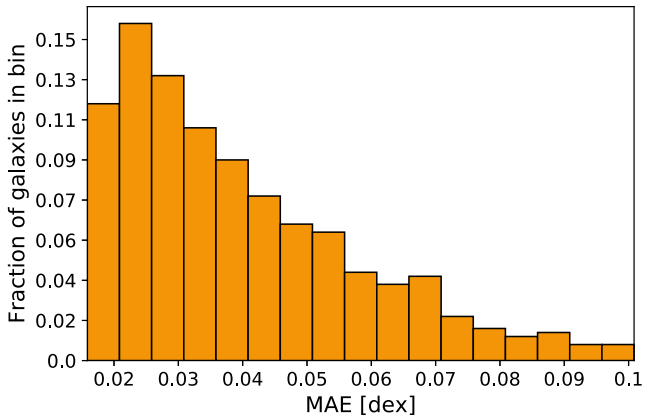


Figure 6. Distribution of the median absolute error, calculated for each galaxy using equation (5). The top 2 per cent and bottom 2 per cent of values are excluded for clarity. On average, *ANNgelina* predicts the irradiance to within 0.06 dex, but there is a large spread from galaxy to galaxy.

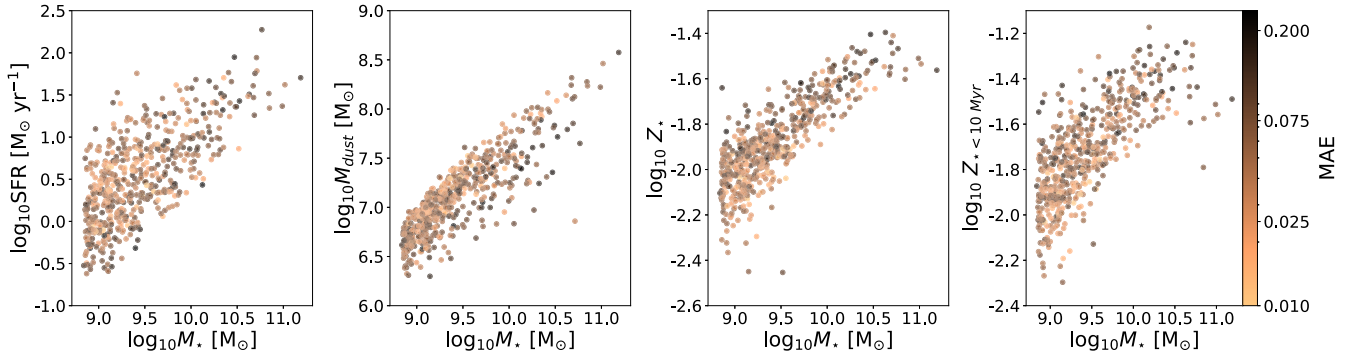


Figure 7. TNG50 galaxies in the test set, colour-coded by MAE value. The uniformity of MSE values in all feature–feature planes indicates that *ANNgelina* performs uniformly well across the parameter space probed.

4.2 *ANNgelina*'s performance in different wavelength regimes

As seen in Fig. 5, for a given SED, *ANNgelina*'s accuracy varies with wavelength. In the examples shown, *ANNgelina* appears to struggle to reproduce rest-frame UV emission more than other wavelengths. In this section, we characterize distributions of MAE values within individual wavelength ranges for the whole test sample.

We divide SEDs into the following wavelength ranges: $0.1 - 0.4\,\mu$m (UV), $0.4 - 0.7\,\mu$m (optical), $0.7 - 3\,\mu$m (near-IR; NIR), $3 - 50\,\mu$m (mid-IR; MIR), $50 - 207\,\mu$m (FIR), and $207 - 1000\,\mu$m (sub-mm). For each galaxy in the test set, we calculate the MAE within each of these wavelength ranges, following equation (5). In Fig. 8, we compile these values (note that the worst 4 per cent of values have been omitted for clarity). Vertical lines indicate the average MAE across test galaxies within each of these wavelength bands. The UV stands out from the others, with a significantly higher average MAE of 0.15 dex (41 per cent), while all other wavelength bands hover around 0.06 dex (15 per cent). This is expected, as the UV emission depends much more strongly on the geometry of the observed simulated galaxy since the UV is highly attenuated by dust grains, and the amount of attenuation can vary considerably with the line of sight.

4.3 Limitations of our approach

There are a few considerations to highlight to potential users of *ANNgelina*. First, all SEDs used to train *ANNgelina* are for 'face-on' projections of TNG50 galaxies (i.e. along the angular momentum axis). Consequently, for galaxies with disc-like morphologies, our data set is biased toward the least-obscured lines of sight, and heavily obscured lines of sight (e.g. edge-on discs) are thus under-represented. The SEDs of simulated galaxies with exactly the same values for the input features can vary due to viewing-angle effects: dust attenuation is non-isotropic, especially in the UV. For this reason, if the 'raw' predictions from *ANNgelina* are used, the scatter in the predicted SEDs will be unrealistically low. One cannot simply perturb the predicted photometry by adding random noise to account for this effect, as the prediction errors for flux values in different bands should be correlated (e.g. for an edge-on galaxy, *ANNgelina* is likely to systematically overpredict the UV-optical flux). In future work, we will employ a larger training set with multiple lines of sight and incorporate some geometric features to attempt to account for viewing angle effects in the ANN's predictions. Since there is intrinsic scatter to be expected when predicting an SED given a set of galaxy properties, we will also attempt to use normalizing flows to sample the distribution instead of providing the model a single value as we have done here.

Secondly, *ANNgelina* was trained on TNG50 galaxies sampling a restricted region of parameter space (simulated galaxies with $M_\star > 10^{8.7}\,M_\odot$ on or above the main sequence), and it should only be applied within this region. Applying *ANNgelina* to simulated galaxies outside of this parameter space may result in high prediction error, even when these galaxies are simulated with exactly the same code (i.e. numerical method and sub-grid models) and resolution. Moreover, only a single simulation – and thus single code and single resolution – was used. It is possible that the prediction error would be significantly greater if *ANNgelina* were applied to simulations that employ the *IllustrisTNG* model but different resolution or/and simulations that employ a different numerical method or/and sub-grid models. In this proof-of-concept work, we have opted to employ only a single simulation, but in future work, we will explore how well we can cross-apply our emulator to other simulations.

Finally, we note that various assumptions are 'baked in' to the MCRT calculations and thus *ANNgelina*. For example, it is necessary to make assumptions about the stellar populations (stellar initial mass function and single-age stellar population SED templates) and dust (optical properties, dust-to-metal ratio, temperature above which dust is destroyed, and sub-grid dust distribution). These assumptions can affect the resulting SEDs, and depending on the science question of interest, they may affect the quantitative or even qualitative results. We have explored the sensitivity of our results to such assumptions in various previous works (e.g. Hayward et al. 2011; Snyder et al. 2013; Safarzadeh et al. 2017b). Exploring the impact of such assumptions is beyond the scope of this proof-of-concept work. Instead, as noted in the introduction, we have addressed the following question: 'given a set of 3D MCRT calculations with fixed assumptions (regarding e.g. the stellar initial mass function, single-age stellar population SED templates, and dust model) performed on galaxies selected from a single cosmological simulation, how well can an ANN emulate the MCRT calculations to predict integrated UV-mm SEDs of the simulated galaxies'?

It is important that users understand that the results will in general depend on the assumptions used in the MCRT calculations to generate the training set. For this reason, the reported errors *do not* include systematic errors associated with the MCRT assumptions and thus underestimate the 'true' error. These errors only represent how imperfectly the ANN can emulate MCRT calculations for this single simulation and set of MCRT assumptions. If, for example, one desires to predict SEDs assuming SMC-like dust, it is necessary to generate a training set using SMC-like dust in the MCRT calculations and

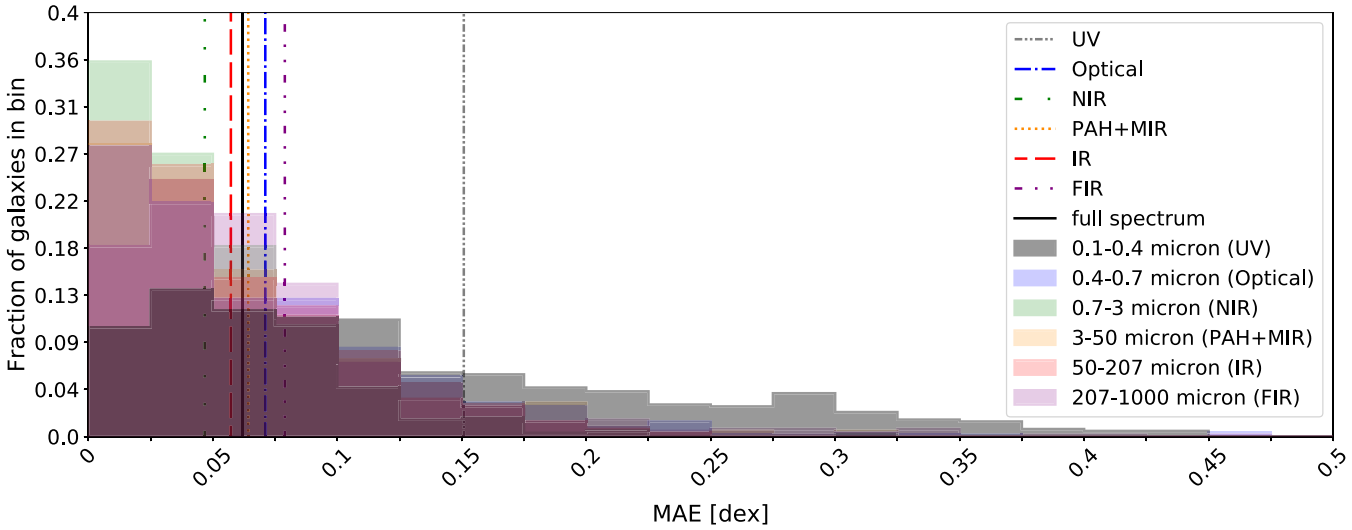


Figure 8. Distributions of the median absolute error between the true and predicted SED per wavelength band for each galaxy in the test set. The vertical lines indicate the average MAE within each wavelength bin. The UV (*grey dash–dot–dotted line*) has a significantly higher average MAE than all the other wavelength bands. This is due to the UV being most highly attenuated by dust, as dust attenuation is anisotropic. The other wavebands exhibit MAE values of ∼0.05–0.07 dex. The solid black line indicates the average MAE across the full wavelength spectrum.

then train a new ANN rather than using the ANN that we have made publicly available.

5 CONCLUSIONS

In this work, we have presented an ANN-based emulator to predict UV-mm SEDs of simulated galaxies. We trained the ANN on a sample of SEDs generated in previous work by performing dust MCRT on galaxies from the TNG50 simulation. We find that the ANN performs well at predicting the SEDs of simulated galaxies in the test set – the average MAE is 0.06 dex. For the vast majority of simulated galaxies in the test set, *ANNgelina* can predict the flux across the full UV-mm wavelength range with a maximum relative error of $\lesssim 50$ per cent. The UV dominates the prediction error, likely because the viewing angle dependence is strongest in this wavelength regime, and our input feature set includes no information about viewing angle. Our results demonstrate that our ANN-based emulator is a promising computationally inexpensive alternative to performing MCRT in order to predict UV-mm SEDs of simulated galaxies.

ACKNOWLEDGEMENTS

The authors thank the anonymous referee for their useful comments, which helped improve the clarity of the manuscript. SS is supported by the National Aeronautics and Space Administration (NASA) Future Investigators in NASA Earth and Space Science Technology (FINESST) fellowship award 80NSSC22K1589. SS acknowledges support from the Flatiron Institute's Center for Computational Astrophysics (CCA) Pre-doctoral Program, during which this work was initiated. The Flatiron Institute is supported by the Simons Foundation. JHW acknowledges support from NASA grants NNX17AG23G, 80NSSC20K0520, and 80NSSC21K1053 and National Science Foundation (NSF) grants OAC-1835213 and AST-2108020.

DATA AVAILABILITY

A github repository containing the tools to run our RT emulator is available at https://github.com/snigdaa/runANNgelina. Our cleaned data are available at the github repository linked above.

REFERENCES

Ade P. A. R. et al., 2016, A&A, 594, A13
Akiba T., Sano S., Yanase T., Ohta T., Koyama M., 2019, preprint (arXiv: 1907.10902)
Baes M. et al., 2020, MNRAS, 494, 2912
Berner J., Elbrächter D., Grohs P., 2019, preprint (arXiv:1905.09803)
Bird S., Ni Y., Matteo T. D., Croft R., Feng Y., Chen N., 2022, MNRAS, 512, 3703
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Buchner J., 2019, PASP, 131, 108005
Camps P., Baes M., 2015, Astron. Comput., 9, 20
Cochrane R. K. et al., 2019, MNRAS, 488, 1779
Cochrane R. K., Hayward C. C., Anglés-Alcázar D., 2022, ApJ, 939, L27
Cochrane R. K. et al., 2023a, MNRAS, 523, 2409
Cochrane R. K., Hayward C. C., Anglés-Alcázar D., Somerville R. S., 2023b, MNRAS, 518, 5522
Dwek E., 1998, ApJ, 501, 643
Gilda S., Lower S., Narayanan D., 2021, ApJ, 916, 43
Groves B., Dopita M. A., Sutherland R. S., Kewley L. J., Fischera J., Leitherer C., Brandl B., van Breugel W., 2008, ApJS, 176, 438
Hayward C. C., Smith D. J. B., 2015, MNRAS, 446, 1512
Hayward C. C., Kereš D., Jonsson P., Narayanan D., Cox T. J., Hernquist L., 2011, ApJ, 743, 159
Hayward C. C., Jonsson P., Kereš D., Magnelli B., Hernquist L., Cox T. J., 2012, MNRAS, 424, 951
Hayward C. C., Behroozi P. S., Somerville R. S., Primack J. R., Moreno J., Wechsler R. H., 2013, MNRAS
Hayward C. C. et al., 2021, MNRAS, 502, 2922
James A., Dunne L., Eales S., Edmunds M. G., 2002, MNRAS, 335, 753
Jonsson P., 2006, MNRAS, 372, 2
Jonsson P., Groves B. A., Cox T. J., 2010, MNRAS, 403, 17
Kasim M. F. et al., 2021, Mach. Learn.: Sci. Technol., 3, 015013
Kennicutt R. C., Jr., 1998, ApJ, 498, 541

Kingma D. P., Ba J., 2014, Adam: A Method for Stochastic Optimization. https://arxiv.org/abs/1412.6980
Lanz L., Hayward C. C., Zezas A., Smith H. A., Ashby M. L. N., Brassington N., Fazio G. G., Hernquist L., 2014, ApJ, 785, 39
Lovell C. C., Acquaviva V., Thomas P. A., Iyer K. G., Gawiser E., Wilkins S. M., 2019, MNRAS, 490, 5503
Lovell C. C., Geach J. E., Davé R., Narayanan D., Li Q., 2021, MNRAS, 502, 772
Maas A. L., Hannun A. Y., Ng A. Y., 2013, in DasguptaS., McAllester D., eds, Proceedings of the 30th International Conference on Machine Learning, Proceedings of Machine Learning, 28, 3
McKinney J., Hayward C. C., Rosenthal L. J., Martínez-Galarza J. R., Pope A., Sajina A., Smith H. A., 2021, ApJ, 921, 55
Michałowski M. J., Hayward C. C., Dunlop J. S., Bruce V. A., Cirasuolo M., Cullen F., Hernquist L., 2014, A&A, 571, A75
Miller T. B., Hayward C. C., Chapman S. C., Behroozi P. S., 2015, MNRAS, 452, 878
Narayanan D. et al., 2015, Nature, 525, 496
Nelson D. et al., 2019, MNRAS, 490, 3234
Parsotan T., Cochrane R. K., Hayward C. C., Anglés-Alcázar D., Feldmann R., Faucher-Giguère C. A., Wellons S., Hopkins P. F., 2021, MNRAS, 501, 1591
Paszke A. et al., 2019, in, Advances in Neural Information Processing Systems 32. Curran Associates, Inc., p. 8024, http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf Last Accessed December 2022
Pillepich A. et al., 2017, MNRAS, 473, 4077
Pillepich A. et al., 2018, MNRAS, 475, 648
Pillepich A. et al., 2019, MNRAS, 490, 3196
Popping G. et al., 2020, ApJ, 891, 135
Popping G. et al., 2022, MNRAS, 510, 3321
Safarzadeh M., Hayward C. C., Ferguson H. C., Somerville R. S., 2016, ApJ, 818, 62
Safarzadeh M., Lu Y., Hayward C. C., 2017a, MNRAS, 472, 2462
Safarzadeh M., Hayward C. C., Ferguson H. C., 2017b, ApJ, 840, 15
Schmidt M., 1959, ApJ, 129, 243
Schulz S., Popping G., Pillepich A., Nelson D., Vogelsberger M., Marinacci F., Hernquist L., 2020, MNRAS, 497, 4773
Smith D. J. B., Hayward C. C., 2018, MNRAS, 476, 1705
Snyder G. F., Hayward C. C., Sajina A., Jonsson P., Cox T. J., Hernquist L., Hopkins P. F., Yan L., 2013, ApJ, 768, 168
Springel V., 2010, MNRAS, 401, 791
Springel V., Hernquist L., 2003, MNRAS, 339, 289
Steinacker J., Baes M., Gordon K. D., 2013, ARA&A, 51, 63
Weber T., Corotan A., Hutchinson B., Kravitz B., Link R., 2020, Atmos. Chem. Phys., 20, 2303
Weinberger R. et al., 2018, MNRAS, 479, 4056
Weingartner J. C., Draine B. T., 2001, ApJ, 548, 296
Wuyts S., Cox T. J., Hayward C. C., Franx M., Hernquist L., Hopkins P. F., Jonsson P., van Dokkum P. G., 2010, ApJ, 722, 1666

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