arXiv:2606.19447

Reimaginig SED Fitting with Cosmological Simulations and Machine Learning

Dhruv T. Zimmerman, Desika Narayanan

1. Introduction

SED fitting

- common technique for deriving physical properties from photometry
- The inference rely on numerous assumptions (IMF, stellar isochrones, spectral libraries, SFH shape, dust law)
    - Assumed SFH shape has been shown to systematically bias the inferred stellar mass
    - Uniform 'dust screen' model (often used) may bring error
    - Stellar models have many systematic uncertainties
- The quoted uncertainties from SED fitting may not always represent actual error in the inferred properties

Machine Learning (ML)

- ML has powerful potential to solve inverse problems
- Several studies have developed their own versions of inferences using ML, but
    - Some based on observational data
    - These are designed to use for specific survey or individual cases

Gilda et al (2021)

- The first steps towards a general ML based inference
- Training data : a cosmological simulation and 3D radiative transfer
- Sampled galaxies are limited to z=0
- It requires full photometry set, i.e. it cannot deal with lacking data

This study aims to construct a ML-based inference named "Phot-Gal" that can be used for more general cases

- lacking photometric bands
- wide range of redshift
- estimate actual error more accurately.

2. Method

Data

- CAMELS 1P suites with IllustrisTNG and SIMBA
- Astrophysical parameters are varied one-at-a-time
- $25\,h^{-1}\,\mathrm{Mpc}$ boxes, 256^3 particles, 'galaxies' are selected when > 24 particles for each integer redshift snapshot $z = 0, 1, \ldots, 6.$
- Mass resolution : $2.3 \times 10^6\ M_\odot$ for gas, $1.2 \times 10^7\ M_\odot$ for DM
- total $\sim 300{,}000$ synthetic SEDs
- Photometric bands : GALEX, HST, JWST, SDSS, WISE, Spitzer, Herschel

Forward Modeling

CAMEL galaxies
(star/gas/dust amount and position, age, Z)

↓

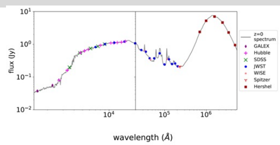
FSPS stellar spectra
Assumptions :
• each star particle $\rightarrow$ SSP
• MIST isochrones
• MILES spectra library
• Chabrier IMF

↓

POWDERDAY, HYPERION
(3D radiative transfer)
Assumptions:
• Dust attenuation law : Weingartner and Draine (2001)
• ILLUSTRIS TNG : fixed dust to metal ratio = 0.4

↓

redshift(discrete $\rightarrow$ continuous)
$z = z_0 + 0.1n + \delta, \quad |\delta| < 0.05$
Integrate through filters



(Backward) inference

Monte Carlo(MC) sampling
$(F_{\mathrm{input},i} \sim \mathcal{N}(F_i, \sigma_i)$ for many times)

↓

K-nearest neighbors (KNN) imputation

↓

NG BOOST
Photo-z, Physical property estimation

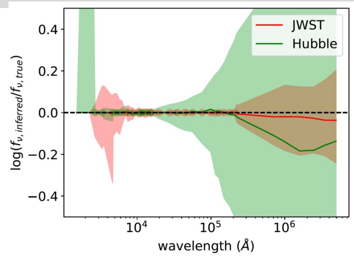


Fig4: Accuracy of the imputation when using only JWST, or only HST

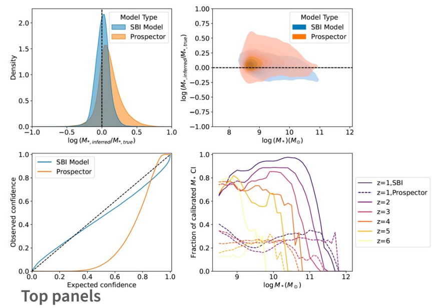
- Only JWST : overall, successfully recovered the true spectrum , with a slight bias at longest wavelength
- Only HST : much larger dispersion and bias in FIR
    - UV/optical would not offer constraints on the SED in the FIR
    - higher z $\rightarrow$ out of the range of HST $\rightarrow$ fewer available photometry

4. Model Evaluation

When compared to a traditional SED fitting code PROSPECTOR with respect to following indices,
NRMSE, NMAE, NBE, ACE, IS
(ACE, IS are the indices that reflect the difference between inferred uncertainty and true error)
Phot-Gal outperforms PROSPECTOR in every property and situation of photometry.

Full photometry , known redshift case

Stellar mass



Top panels

- PROSPECTOR exhibits a minor deviation and broader dispersion
- Photo-Gal has a slight bias at high/low mass end, which might be because of the scarcity

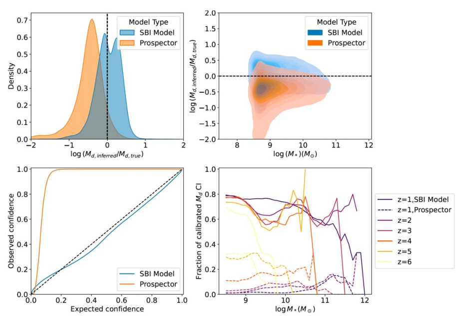
Lower left panel

- Error calibration plot often used in ML studies
- Horizontal and vertical axes are $F(M_{\mathrm{true}})$, where $F(M)$ is posterior (Cumulative) distribution function.
- PROSPECTOR's estimate is typically underconfident

Lower right panel

- Vertical axis is the fraction of galaxies at given (mass, z) bin whose true Mass is in the 68% confidence interval $\rightarrow$ 0.68 is the ideal value
- PROSPECTOR's CIs are not so much calibrated

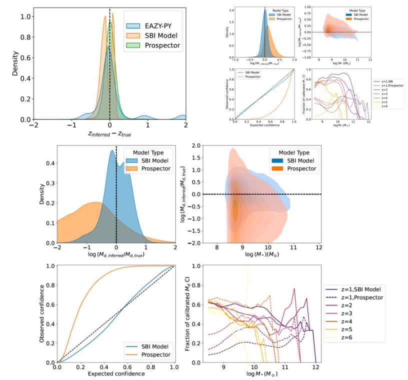
Dust mass



- Theoretically, Phot-Gal has an advantage for dust mass estimate, given the diversity of dust geometry
- In fact, Phot-Gal does not exhibit systematic bias (But a bimodal feature exists), which is apparently superior to PROSPECTOR.
- Since PROSPECTOR adopts uniform "screen" model, it tends to underestimate the dust mass

SFR : Both have little bias, but PROSPECTOR exhibit broader distribution.

JWST photometry only , unknown redshift case



5. Model Exploration

SHAP feature value : Index of how much each feature contributes to the prediction

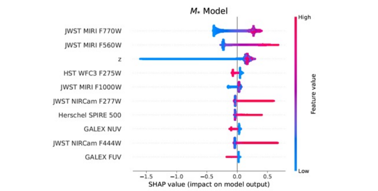
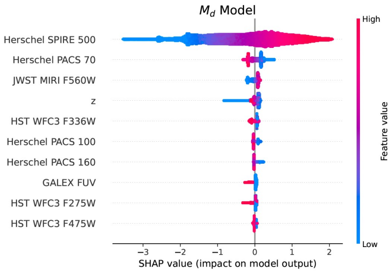


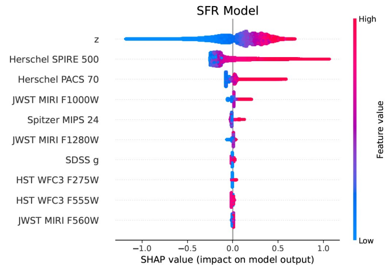
Fig. 12.— **A SHAP value plot for the NGBoost stellar mass inference model in Phot-Gal indicates that the rest-frame K band is primarily what Phot-Gal uses to predict stellar mass.** Input features are ordered from highest to lowest for how much the ML model relies on the feature to recover the stellar mass. High absolute SHAP values mean that the property is more important to the output prediction. The shading for any individual feature reflects whether the SHAP values of that normalization correspond to the higher or lower values of that feature in the sample.

For example, JST F770W flux distribution has bimodality. "The group that have High SHAP values typically have high flux" indicate that high F770W flux contributes to higher M_* prediction.

- 1–$10\ \mu m$ range is most important to infer M_*

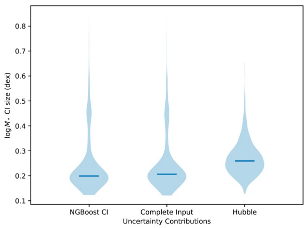


The prediction relies mostly on IR(Herschel $500\ \mu\mathrm{m}$) photometry.



Redshift and Far IR are most important
- The normalization of the SFMS is highly dependent on redshift
- Scaling relation between L_{IR} and SFR
- Much of the UV light associated with star formation in the Universe is expected to be obscured by dust

more about CIs
What does contribute to the distribution of M_* ?
- full photometry, known redshift case $\to$ only the NGBoost's distribution remains : $0.2\ \mathrm{dex}$ contribution to $\log M_*$ CI size
- Additionally considering photometric error $\to$ slight contribution

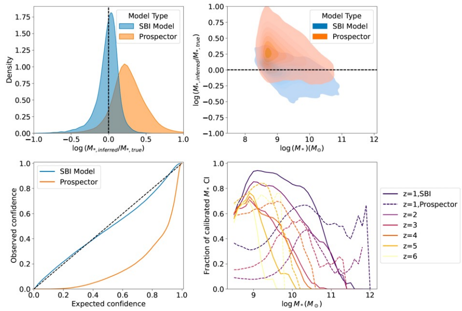


- HST only, unknown redshift $\to 0.25\ \mathrm{dex}$

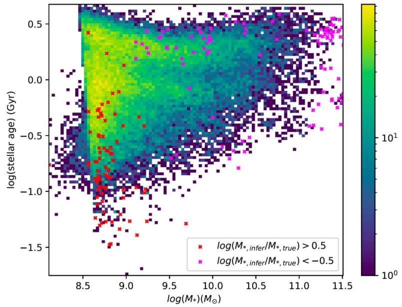
In short, CI width : $0.2\ \mathrm{dex} \to 0.25\ \mathrm{dex}$ (small change)

However, in HST only , unknown case, the fraction of true M_* in CI is significantly worsened.

Phot-Gal's **imputation step** is not adding sufficient uncertainty that reflect the loss of information.



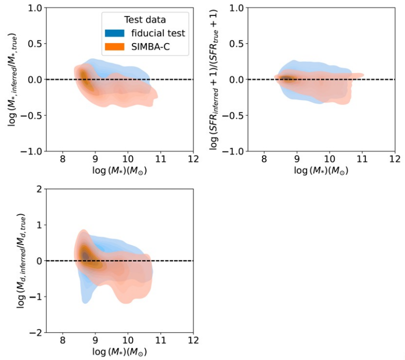
outliers (Phot-Gal has difficulty in inferring)



- Phot-Gal tends to over/underestimate M_* when M_* is high/low or stellar age is young/old
- Possible reason
    - Poor training samples
    - irregularly high or low mass-to-light ratio

Epistemic Uncertainty
Apply Phot-Gal to SIMBA-C data set (, which is not included in training)

- Systematic bias in M_* estimate



6. (additional) Limitations

- Training samples are dominated by $M_* \in [10^{8.5},\ 10^{10.5}]\ M_\odot$ galaxies. $\to$ applying Phot-Gal to massive, dwarf, or rare-history galaxies is problematic
- At non-integer redshift, the set of spectra was not produced from self-consistent SFH
- Model dependence : In forward modelling , they assume certain IMF, dust extinction law, and SPS assumption
- AGN emission , emission from HII regions are not included
- Problems of estimator
    - KNN imputation : computation costs increase as filters or training sets increase
    - NGBoost : only Gaussian distribution