pyPipe3D dossier

YGK

Context

The objective of this file is to provide a comprehensive summary of Yllari's current knowledge and doubts in the context of full spectral fitting and, particularly, pyPipe3D and pyFIT3D. As such, python code will be avoided and, instead, a reference to the file and line will be provided whenever necessary
(e.g., `pyFIT3D/modelling/stellar.py 351`).
The main objectives will be:

- Obtaining precise star formation histories (SFHs) in the analysed galaxies, and doing so as rigorously as possible
- Correctly calculate and interpret the associated uncertainties
- Full reproducibility of results in a user-friendly manner

Analysis on **SDSS** and **CAVITY** data.

Stellar population models

Stellar population models (SPs)

Three possibilities are considered for this work:

Model	Pros	Cons
maStar	fully-consistent, empirical	variable lin. res. , $\sigma_{SDSS} \sim \sigma_{\text{maStar}}$
`gsd156`	not consistent	fixed lin. res., high resolution
MILES full	consistent	no young pop., high number of models

None of them is perfect for my purposes. The default star population templates provided are the GRANADA + MILES (`gsd156`) files.

Questions:

- How should the maStar SPs be used for SDSS and CAVITY data? $\sigma_{\text{inst}} = 0$ for SDSS? Variable resolution can create many problems
- From SFS's experience, what is the best combination of SPs models (both SDSS central and CAVITY resolved)?
- Are we correctly calculating SFHs? Show a comparison between Jesus' results and MILES. They are peaked as opposed to pPXF or STECKMAP because no regularization is performed. Are they reliable? Did SFS do any tests in this regard?

Doubts pyFIT3D

Calculation

- How come emission line fits can report negative integrated flux values when both the LM and RND method are bounded by the introduced limits?
- Even if large errors are given for the flux for a range of values, the program fits it either way (show example). The spectral analysis is performed with a weighted least squares, so one would expect for those values to be ignored. Not very important, but problematic for galaxies with missing fluxes.
- The coefficients for the stellar continuum fit are not the mean of the Monte Carlo iterations, as reported in the paper; a latter search is done with random noise (`modelling/stellar.py 3020`). Why? What are the implications?
- In mass and light-to-mass calculation, no cosmology or dust correction is taken into account, right? Also, the mass is calculated with $3500 \times M/L \times F$ (`modelling/stellar.py 3532`), where F is the median flux, M/L the mass to light ratio. Where does the 3500 factor come from? In pyPipe3D this factor does not appear (`common/tools.py 2667`).
- Resampling can lead to underestimating errors. What can be done about it? Show some plots. Including covariance matrices would be too difficult.
- Mention the optimisation carried out on `stats.py`, out-of-the box $\sim \times 2, 3$ increase in speed.
- What about the `niter=2` configuration, i.e, the gas fitting that SFS does afterwards. What is the best way to do this? No big improvement, but worth commenting.
- Do the min, max values have any effect?
- Are fluxes in `flux_elines` normalized? Is this what `redefine_max` is?

Uncertainties

Why is only half of the flux error used for the MC iteration? i.e, the added noise is gaussian-distributed with $\sigma(f_\lambda)/2$ (`modelling/gas.py 1189`). The full profile is used for the SPs fitting.

How come the error in the flux for the emission lines does not depend on the number of pixels in the fitting? Particularly in the continuum. This is from the formula $\sigma = \sqrt{\sigma_{MC}^2 + \sigma_c^2}$ (`modelling/stellar.py`) where $\sigma_c = FWHM \cdot \sigma_{\text{cont}}$ and σ_{cont} is the standard deviation of the residuals for the gas-only spectrum

What is the best consistent way to calculate SFH uncertainties? Right now, even if in the article it is reported that the given value for the SP coefficient χ_j is $\sigma(\chi_i)$, the uncertainty given by the program is $\sigma(\chi_i)/N_{\text{MC}}$ (`modelling/stellar.py 3085`), which may be an underestimated uncertainty.

Also, the age and metallicity averages and STDs could be computed directly from the MC iterations, which could improve the estimation. As opposed to the linear propagation done in the current version (`modelling/stellar.py 3013-3018`).

The uncertainty in the extinction is calculated twice, first in the non-linear fit, and then as $0.43 \sum \sigma(\chi_i) A_v^2$. Where does this formula come from? (`modelling/stellar.py 3018`)

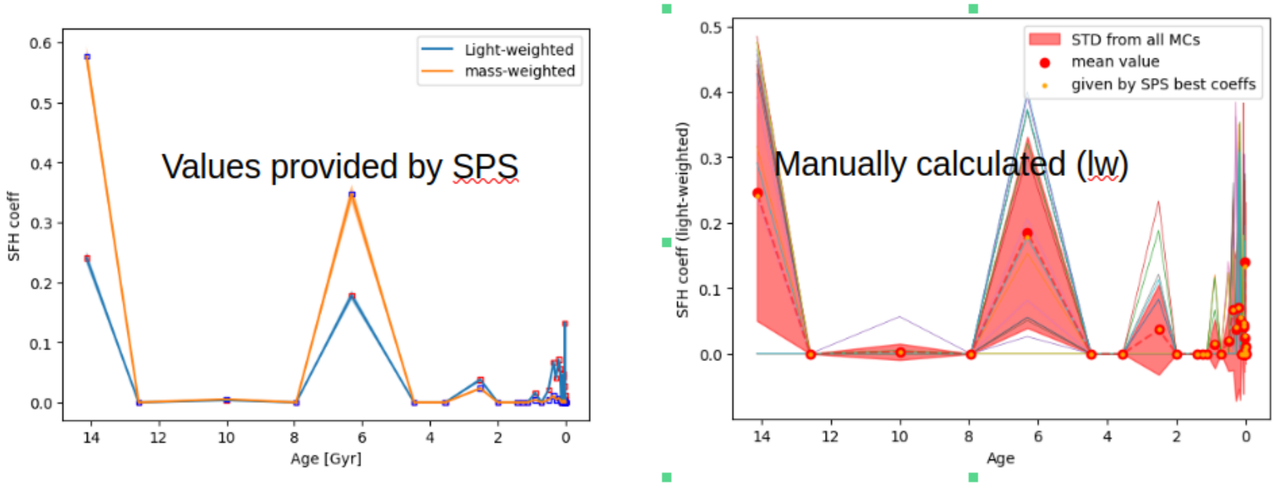

Figure 1: Comparison between the SFH provided values (left) and the manually calculated as mean and standard deviation (right).

Doubts pyPipe3D

General

- Review the pipeline making sure that I understand everything correctly. What are the optimal parameters for different steps?
- Have there been any changes in the pipeline since the latest version on gitlab?
- What is the proper way to determine the center of the galaxy (x_0, y_0)? Just using the pseudo-V band mean flux as in the code?
- Why does the instrumental dispersion change between the different steps of the procedure? What sigma should I use in each one? **This is probably the quadratic difference and the proper inst. disp., but to be completely sure**
- What to do with extremely large (bad fitting) values for emission line fluxes? These are probably being filtered directly by Sebastián (show cube). These appear because the fitting wavelength range given by the config and the one provided or derived in the non-linear fit are different **This I probably could fix on my own**. I did by changing the default redshift from around zero to the guided/central one. There are still some issues with OII 3727, probably because the values fall outside of the range.
- The instrumental sigmas are not being corrected for redshift, right? See Cappellari 2017 MNRAS 466, 798–811

The dust-corrected stellar mass density is calculated as:

$$\log(M_{\text{dust}}) = \log(M) + \log_{10}(e)A_v \tag{1}$$

but, in the formula, the preceding factor for A_v should be 0.4, from the correction $F_0 = F \cdot 10^{0.4A_v}$ (?).

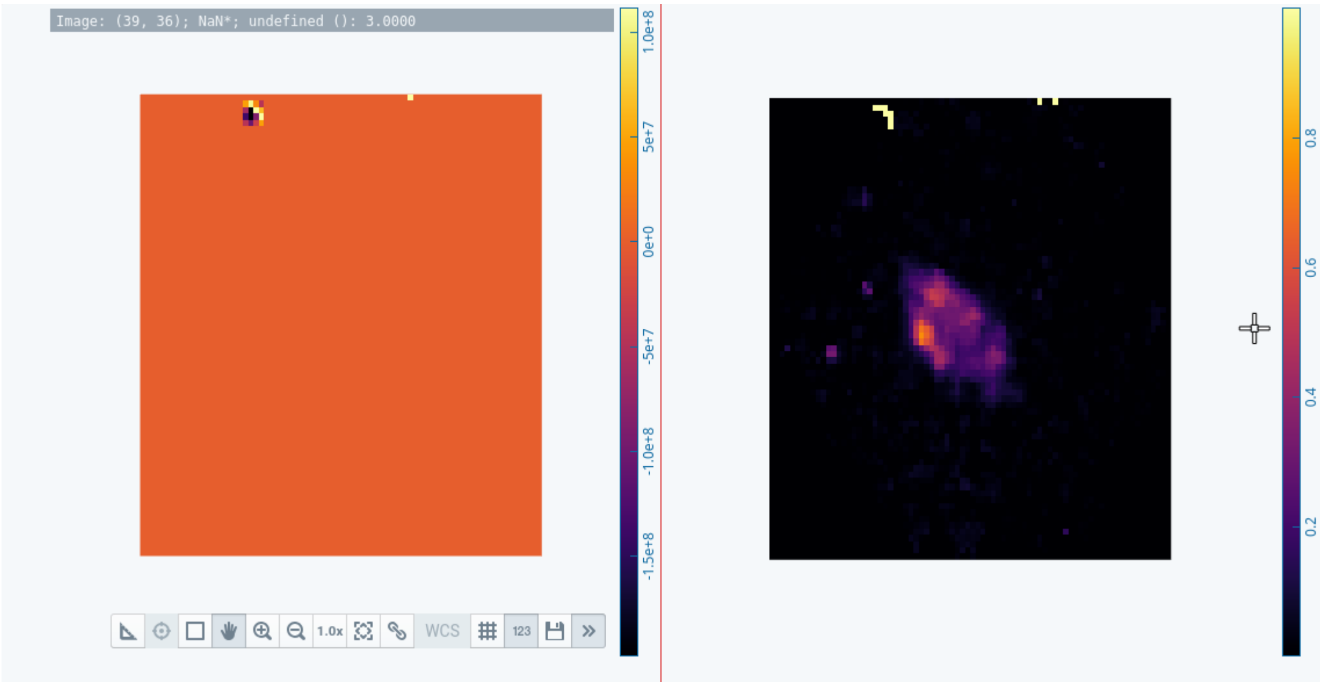

Figure 2: Comparison for cube flux without (left) and with filtering for extremely large values (right).

Uncertainties

Since all the derived values come from pyFIT3D, not many doubts here:

- Why are there no uncertainties in the dust-corrected mass? Neither stellar mass uncertainties are given in CAVITY data sometimes.
- Could we provide uncertainties for high order dataproducts?

What else?

High-order data products? Perhaps he is also interested in comparing SFHs with different SPs?