How well can we determine ages and chemical abundances from spectral fitting of integrated light spectra?

Geraldo Gonçalves1*, Paula Coelho1, Ricardo Schiavon2, Christopher Usher2

1 Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, São Paulo, SP, Brazil
2 Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT
The pixel-to-pixel spectral fitting technique is often used in studies of stellar populations. It enables the user to infer several parameters from integrated light spectra such as ages and chemical abundances. In this paper, we examine the question of how the inferred parameters change with the choice of wavelength range used. We have employed two different libraries of integrated light spectra of globular clusters (GCs) from the literature and fitted them to stellar population models using the code STARLIGHT. We performed tests using different regions of the spectra to infer reddening, ages, [Fe/H] and [\(\alpha/Fe\)]. Comparing our results to age values obtained from isochrone fitting and chemical abundances from high resolution spectroscopy, we find that: (1) The inferred parameters change with the wavelength range used; (2) The method in general retrieves good reddening estimates, specially when a wider wavelength range is fitted; (3) The ideal spectral regions for determination of age, [Fe/H], and [\(\alpha/Fe\)] are: 4170-5540Å, 5280-7020Å, and 4828-5364Å, respectively; (4) The retrieved age values for old metal-poor objects can be several Gyr younger than those resulting from isochrone fitting. We conclude that, depending on the parameter of interest and the accuracy requirements, fitting the largest possible wavelength range may not necessarily be the best strategy.

Key words: stars: fundamental parameters – Galaxy: globular clusters – abundances – techniques: spectroscopic – galaxies: star clusters: stellar content

1 INTRODUCTION
Determining realistic values for ages and chemical abundances of stars is a very important step to understand how stellar systems such as galaxies and stellar clusters form and evolve. For the Milky Way, as we are able to observe stars individually, we can use isochrone fitting on the Color-Magnitude Diagram (CMD) of stellar clusters (e.g. Dotter et al. 2010) or calibrations of asteroseismologic masses (e.g. Ness et al. 2016) to determine ages, and high resolution spectroscopy to derive abundances (e.g. Barbary et al. 2018). But for more distant systems, resolving stars become increasingly more difficult and we have to work with integrated light.

There is a range of techniques that can be used to infer these parameters from integrated light, from the widely used historical line indices (e.g. Worthey et al. 1994; Trager et al. 2000; Thomas et al. 2003; Schiavon 2007) to updated approaches such as SED fitting (e.g. Walcher et al. 2011), Principal Component Analysis (e.g. Li et al. 2001), Bayesian inference (e.g. Sharma 2017), pixel-to-pixel spectral fitting (e.g. Cid Fernandes 2007; Koleva et al. 2008; Walcher et al. 2009; Conroy et al. 2014), among others.

In this paper, we are concerned with spectral fitting. One aspect which has not been studied in detail is to quantify to what extent the choice of the wavelength range affects the results of the spectral fitting. The literature reports conflicting results.

Koleva et al. (2007) estimated the star formation history of galaxies using models computed with PEGASE.HR (Le Borgne et al. 2004) and a parametric procedure. They concluded that the wavelength range used in the fit does not affect significantly the precision of derived kinematics parameters and metallicities. However, the precision of ages determinations became worse when smaller wavelength ranges were used.

In Koleva et al. (2008), the authors used two different algorithms – STECKMAP (Ocvirk et al. 2006) and NBURSTS (Chilingarian et al. 2007) – and three libraries of Simple Stellar Populations (hereafter SSP) – ELODIE 3.1 (Prugniel et al. 2007) and GALEXEV (Bruzual & Charlot 2003) – and MILES (Vazdekis 1999, Vazdekis et al. 2003) – to fit GCs from Schiavon et al. (2005)}
and the integrated spectrum of M67 from Schiavon et al. (2004a). They performed the fit in the 4000-5700 Å region, successfully retrieving the age and metallicities parameters for 36 out of 41 clusters. But they briefly reported avoiding the bluest region of the spectra in the fit (in particular the H&K lines), otherwise the inferred ages would be systematically lower by \(\sim 1\) Gyr.

Walcher et al. (2009) used the sedfit algorithm (Walcher et al. 2006) to fit spectra of NGC 6528 and NGC 6553 from Schiavon et al. (2005). They tested five different wavelength ranges and found that 4828–5364 Å is the one that results in [Fe/H], \([\alpha/Fe]\) and age values closer to the reference values from literature. The other four ranges seem more affected by the age-[Fe/H] degeneracy.

Cezario et al. (2013) used two different samples of globular clusters: M31 GCs from Alves-Brito et al. (2009) and Galactic GCs from Schiavon et al. (2005). Using the ULySS algorithm (Koleva et al. 2009) and SSP models from Vazdekis et al. (2010), they tested four different wavelength ranges with the Galactic GCs. They concluded that this choice has little influence on the metallicities, while the ages can vary a lot: a third of the spectra were fitted with intermediate ages in some cases. They favour 4000-5400 Å, being the one that best reproduced CMD ages, and applied the technique to the GCs in M31.

In this work, we examine this question by performing assessing the dependence of the retrieved stellar population parameter with the wavelength range used. We aim at finding what would be the ideal configuration (if any) to infer physically consistent parameters, for a fixed choice of spectral fitting code and SSP models. Thus, we hope this work can contribute to studies of integrated light, its properties and methods in order to infer accurate parameters.

This paper is organized as follows: Section 2 describes the data from literature we used, section 3 describes the methodology, section 4 presents our results, that are discussed in detail in section 5. Finally, section 6 presents our conclusions.

2 THE OBSERVATIONAL DATA

In this study, we have used data from the WiFes Atlas of Galactic Globular clusters Spectra (WAGGS) project (Usher et al. 2017). The WAGGS library contains objects from the Milky Way, the Magellanic Clouds and the Fornax galaxy, thus covering a wide range in ages and chemical abundances (see Figure 1).

The data was taken with the WiFeS integral field spectrograph (Dopita et al. 2007, 2010) on the Australian National University 2.3m telescope. The spectra were taken in four different arms, using different grating setups: U7000, covering 3270–4350 Å range; B7000, covering 4170–5540 Å; R7000 covering 5280–7020 Å, and; I7000 covering 6800–9050 Å. All four gratings give spectral resolutions of \(\delta\lambda/\lambda \sim 6800\). The data is publicly available at the WAGGS project website\(^1\).

The error spectra are provided, as computed from the pipeline PYWIFES (Childress et al. 2014). The median SNR for each arm are 12, 61 and 124 for the U, B and R arms respectively\(^2\). The distribution of the SNR values per arm are shown in Figure 2. Given the natural difficulties of calibrating IFU data, these errors are likely overestimated at low SNR and underestimated at high SNR, and therefore may be biased. In particular, at high SNR, the dominant source of uncertainty is effects stochastic due to the finite stellar mass within the observed field of view.

For some additional analysis, we also used GC integrated spectra from Schiavon et al. (2005). These data have a different coverage range (3350–6430Å), with a resolution of 3.1 Å (FWHM). Each observation was taken by drifting the spectrograph slit across the core diameter of the object in exposures of 15 minutes. The data are publicly available at the National Optical Astronomical Observatory website\(^3\).

In order to obtain the reference values for ages and chemical abundances for the globular clusters, we have used the Table 1 in Usher et al. (2017) as a starting point, and performed an extensive search in literature. Searching object by object, we gathered information from literature prioritising works that provide \([\alpha/Fe]\) (here represented by [Mg/Fe]) as well as [Fe/H] and their uncertainties. This compilation is shown in Table 1.

\(^1\) http://www.astro.ljmu.ac.uk/~astcushe/waggs/

\(^2\) In this work we have not used the I7000 arm because the SSP models employed do not cover its spectral region

\(^3\) http://www.noao.edu/gcclib
How well spectral fitting performs?

Figure 2. Distribution of signal-to-noise ratios (SNRs) of the spectra in the WAGGS library. The three panels show histograms of SNRs for the U7000, B7000 and R7000, respectively. It is noticeable that U7000 show the lowest values of SNR of the sample.

Table 1. [Abridged; full table is available as online-only material] Globular clusters present in the WAGGS database and the reference values for ages and chemical abundances. The symbol '*' indicates values that were not found in our search.

| ID   | Galaxy | RA  [°] | Dec  [°] | [Fe/H]  | [Mg/Fe] | Reference for abundances | Age (Gyrs) | Reference for age |
|------|--------|---------|----------|---------|---------|---------------------------|------------|------------------|
| NGC0104 | MW     | 6.024   | -72.081  | -0.768 ± 0.031 | 0.52 ± 0.03 | Carretta et al. (2009) | 12.75 ± 0.5 | Dotter et al. (2010) |
| NGC0121 | SMC    | 6.701   | -71.536  | -1.41 ± 0.07 | * ± *  | Johnson et al. (2004) | 10.5 ± 0.5 | Glatt et al. (2008)  |
| NGC0330 | SMC    | 14.074  | -72.463  | -0.82 ± 0.11 | -0.1 ± 0.18 | Hill (1999) | 0.03 ± 0.002 | Sirianni et al. (2002) |
| NGC0362 | MW     | 15.809  | -70.849  | -1.17 ± 0.05 | 0.33 ± 0.04 | Carretta et al. (2013) | 11.5 ± 0.5 | Dotter et al. (2010) |

3 METHODOLOGY

3.1 Pre-processing of the data

Prior to the stellar population analysis, the data had to be processed to be in the rest-frame and the spectral resolution adjusted.

We measured the radial velocities of the observations using the task fxcor in IRAF.

We applied a gaussian pixel-to-pixel convolution on the spectra using the functions gaussian and convolve from Python’s scipy package. We used the equations 1 to 3 to match the data to a resolution of FWHM = 2.51 Å (Falcón-Barroso et al. 2011).

In order to fit the spectra in the RM range, we concatenated the spectra, prioritizing the arm with greater SNR (i.e. using the data with greater SNR in the regions where the arms overlap). We also excluded the 8 first and the 8 last pixels of each arm to avoid edge effects.

3.2 Spectral fitting

In this work, we have used the software STARLIGHT (Cid Fernandes et al. 2005), a full spectral fitting code widely

4 http://ast.noao.edu/data/software

5 We have not used spectra observed in the I7000 arm in the present work, but report the obtained relations for completeness.
used in galaxy studies. It creates a model $M$ fitting an observed spectrum as a combination of SSP models, as shown in Equation 5.

$$M = \sum_j x_j \text{SSP}_j \otimes G(v_\ast, \sigma_\ast) \times 10^{-0.4A_V} \quad (5)$$

where $x_j$ is the population vector (it corresponds to the light contribution in percentage of each SSP $j$ in the model), $G$ is the kinematical filter (where $v_\ast$ is the radial velocity and $\sigma_\ast$ is the velocity dispersion) and $A_V$ is the reddening (in magnitudes in the V band).

Fitting this model allows us to infer light-weighted and mass-weighted ages and chemical abundances, as shown in Equations 6, 7 and 8, alongside reddening values and the kinematical parameters. In the present work, we fixed the kinematics as we already have the spectra in the rest frame and with the broadening matching that of the SSP models (see Section 3.1) Although the clusters have their own velocity dispersion (about 5 to 15 km/s, Dalgleish et al. 2020), these are significant smaller than the resolution of the MILES library.

$$\langle \log(\text{age}) \rangle = \sum_j x_j \cdot \log(\text{age})_j \quad (6)$$

$$\langle \text{[Fe/H]} \rangle = \sum_j x_j \cdot \text{[Fe/H]}_j \quad (7)$$

$$\langle \alpha/\text{Fe} \rangle = \sum_j x_j \cdot \alpha/\text{Fe}_j \quad (8)$$

When it comes to mass weighted parameters $x_j$ is replaced by $M_{\text{corr},j}$, which is the mass-weight of the SSP $j$ in the model, corrected for mass lost during its evolution.

STARLIGHT also returns the $\chi^2$ of each SSP fitted on the data, so we are able to work both with multi-population and simple-population fits.

The SSPs come from the MILES library (Vazdekis et al. 2015), which uses the BaSTI isochrones (Pietrinferni et al. 2004b, 2006b). Each model is defined by three values: age (varying from 0.03 to 14.00 Gyrs), [Fe/H] (-2.27 to 0.26 dex) and $\alpha/\text{Fe}$ (0.0 and 0.4 dex). The models with $[\alpha/\text{Fe}] = 0.0$ adopt the solar abundances from Grevesse & Sauval (1998), scaled according to [Fe/H]. The $\alpha$-enhanced models adopt abundances of $[\text{X}/\text{Fe}] = 0.4$ for O, Ne, Mg, Si, S, Ca and Ti, and solar values for the other elements.

The library contains more than one thousand models, but we have selected for use 324 models, those with ages: 0.03, 0.06, 0.09, 0.20, 0.35, 0.50, 0.80, 1.25, 2.00, 2.75, 3.50, 4.50, 6.00, 7.50, 9.00, 10.50, 13.00 and 13.50 Gyr (one out of every three ages in the complete library). The spectral range goes from 3540.5 to 7409.6 Å at a resolution of FWHM = 2.51 Å (Falcón-Barroso et al. 2011).

Given that the lowest age in the library is of 30 Myr, we did not include the GCs NGC2004 and NGC2100 in the analysis, as they are reported in literature to be younger than this limit (Table 1).

3.3 A note on the definition of ages

As reference values, we adopt ages derived from isochrone fitting to high quality CMD, compiled from literature (see Table 1). But when inferring ages from integrated spectra, several definitions of an age can be used. Below we shortly specify the different definitions used in this work.

Isochrone age $t_{\text{iso}}$: The age inferred from isochrone fitting, obtained from fitting stellar evolution isochrones to observed CMD. This is the value we adopt throughout this work as reference value for age.

Light-weighted age $t_{\text{light}}$: A multi-population mean age, computed by weighting the age of each SSP $j$ by its light contribution $x_j$, as obtained from STARLIGHT. This is the definition shown in Equation 6.

Mass-weighted age $t_{\text{mass}}$: A multi-population mean age, computed by weighting the age of each SSP $j$ by its stellar mass. The mass fraction in stars for each SSP needs to be provided to STARLIGHT to transform the light contribution $x_j$ into mass contribution $M_{\text{corr},j}$.

SSP-equivalent age $t_{\text{SSP}}$: Age of the SSP that best fits the observation (SSP model with the smallest $\chi^2$).

Age of the top contributor to the integrated light $t_{\text{max},x_j}$: The age of the SSP which most contribute to the integrated light (max($x_j$)).

Equivalent definitions are valid for [Fe/H] and $[\alpha/\text{Fe}]$, but our tests show that variations among the definitions are larger in the case of ages, as discussed in Section 4.

4 RESULTS

For each cluster, we run STARLIGHT in five different wavelength ranges, as listed in Table 2. In Table 3, we list some of the main spectral features in each range. They include the four strongest lines of the Balmer series (Hβ, Hβ, Hγ, Hδ), the Balmer break, and strong the optical absorption features (Worthey et al. 1994; Worthey & Ottaviani 1997; Poggianti & Barbaro 1997).

In this section we present the results of the spectral fits, and their implications are discussed further in Section 5.

4.1 Quality of the spectral fits

The quality of the spectral fitting as performed by STARLIGHT is given in terms of $\chi^2$ and the mean relative difference between model and observation $\Delta_\lambda$.

$$\Delta_\lambda = \frac{1}{N} \sum_\lambda \left| \frac{f_{\text{model}}(\lambda) - f_{\text{obs}}(\lambda)}{f_{\text{obs}}(\lambda)} \right| \quad (9)$$

where $N$ is the number of pixels, $f_{\text{model}}$ is the fitted spectrum and $f_{\text{obs}}$ is the observed spectrum.

A good fit is illustrated in Figure 3, for the case of NGC104 (47 Tuc), as observed in the B7000 arm in the WAGGS library ($\chi^2 \sim 2.6$, $\Delta_\lambda \sim 0.4$).
In Figure 4 we illustrate the histograms of $\chi^2$ and $\Delta_\lambda$ for all the fits performed. We only show $\Delta_\lambda$ values up to 5, but for a few fits, they can go as high as $\sim$18% for R7000 and $\sim$500% for U7000 (both cases for NGC2004 spectra, which are very noisy with SN: $\gtrsim$ 1.0 in both cases ). All the cases where $\Delta_\lambda \gtrsim 5$ are listed in the caption of Figure 4. Overall, most spectra in the U arm are fitted within 2% of the flux, and most spectra of B and R-arms are fitted within 1%.

Another aspect of the fit quality has been discussed in Cid Fernandes (2018). The article discusses a way to identify if the algorithm performed the fit well or if the Markov-Chain did not converge. To verify this, the user can compare the $\chi^2_{\text{SSP}}$ (the $\chi^2$ of fitting each SSP in the library) with the global $\chi^2$ of the multi-population fit. According to the author, in the case of a good fit $\chi^2_2$ should be lower than any of the $\chi^2_{\text{SSP}}$ values. Otherwise, the fit should not be trusted. Looking at these values, we found that only three of our fits (Fornax3–U7000, Fornax5–R7000 and NGC7099–B7000) do not meet this quality requirement.

The cases which fall into the criteria above ($\chi^2 > \min(\chi^2_{\text{SSP}})$) and/or were fitted with $\Delta_\lambda > 5$ have been removed from the analysis hereafter.

### 4.2 Reddening

Ages and reddening are the two parameters which primarily affect the continuum shape of the fitting process. As such, they are the ones likely to be more affected by uncertainties in flux calibration.

Figure 5 shows the reddening values $A_V$ inferred by Starlight compared to the reference values from literature (Harris 1996; McLaughlin & van der Marel 2005). For the majority of the fits, the correspondence is close to the 1-to-1 line. This good agreement between the fitted values and the literature values would seem to indicate that the relative WAGGS flux calibration is reliable. The average difference between Starlight and reference values are 0.10 for $R_M$, -0.25 for $R_U$, 0.12 for $R_B$, 0.31 for $R_R$ and 0.14 for $R_W$. For spectra in the B and R arms, the retrieved $A_V$ of systems with $A_V \gtrsim 2$ show a tendency to be overestimated.

In the fitting runs we allow negative values for reddening down to -1. Gallazzi et al. (2005) and Mateus et al. (2006) discuss some reasons (mainly related to inconsistencies in the base of SSP models) why this should be allowed (clearly the results would have no physical meaning in these cases). Most of these cases with $A_V \gtrsim 0$ appear in the U7000 arm panel, which are the spectra with lower SNR (see Figure 2).

### 4.3 Ages

Figure 6 shows the age values inferred from Starlight compared to the isochrone ages, for all definitions listed in Section 3.3. For most of the young and intermediate-age objects, ages are generally well retrieved in spectra from the U7000 and B7000 arms, regardless of the age definition adopted. Values obtained from the R7000 arm spectra show larger dispersion, which is expected given that age features are clustered in the blue bands – in the wavelength interval $R_R$ (see Table 2) basically only $H_\alpha$ is available as an age feature.

In the case of old systems, a range of ages is retrieved, with a tendency for the metal-poor objects to be modelled with ages younger than the isochrones ages. This is better illustrated in Figure 7 showing $t_{\text{light}}$ and $t_{\text{mass}}$ in detail for the old systems. The difference between the spectroscopic age and the reference age can reach up to 10 Gyr. We further discuss these findings in Section 5.

### 4.4 Chemical abundances

Figure 8 shows the retrieved values of [Fe/H] compared to the reference from literature. It is noticeable that, despite a small systematic shift ($\sim$ 0.2 dex), the light-weighted and mass-weighted values show a good agreement with the literature. [Fe/H]$_{\text{SSP}}$ and the [Fe/H]$_{\text{max}}$ $x_j$, on the other hand, show a bigger scatter, up to 2.0 dex.

Due to the smaller dynamical range of the [α/Fe] values,
we opted to show the results as the distribution of $\Delta[\alpha/Fe] = [\alpha/Fe]_{\text{Starlight}} - [\alpha/Fe]_{\text{Reference}}$. We can see in Figure 9 and in Table 5 that the retrieved values are closer to the reference values when the fit is performed using the $R_W$ range, not showing significant differences between the light-weighted and the mass-weighted values.

4.5 Evaluating uncertainties in the retrieved parameters

The sources of uncertainties are multiple: in the reference values, from the models, from the observations and from the method; and not straightforward to access.

The uncertainties in the reference values reported in literature are typically small (see Table 1). In order to evaluate the uncertainties in our results due to the quality of the observations (SNR), we used a similar method to the one presented in Cid Fernandes et al. (2005). Using the flux uncertainties in the second extensions of the publicly available FITS files for WAGGS library, we created twenty perturbed spectra of each observation, perturbing each flux value $\text{flux}_j$ with a random value inside its error bar $\text{err}_j$ (Equation 10). We then fitted all 20 perturbed observed spectra using the same method and models described above. The uncertainty $\sigma_P$ of each parameter $P$ is given by the standard deviation of the values obtained from the 20 realisations (Equation 11).

\begin{equation}
\text{flux}_j \rightarrow \text{flux}_j + \text{random}[-1, 1] \times \text{err}_j
\end{equation}

\begin{equation}
\sigma_P = \sqrt{\frac{\sum_{i=1}^{20} (P_i - \langle P \rangle)^2}{20}}
\end{equation}

To optimise computing time, we evaluated these uncertainties performing the spectral fits in the interval $R_W$ (4828–5364 Å), favoured by Walcher et al. (2009). We computed the standard deviation of the parameters retrieved from the 20 realisations, and report their median values in Table 4. These values can be seen as typical observational uncertainties due to the SNR of the WAGGS sample, while
How well spectral fitting performs?

Figure 4. Distributions of $\chi^2$ (upper line) and $\Delta_\lambda$ (lower line) for the spectra of the WAGGS library. The columns show the histograms for the five different ranges fitted (see Table 2). In the case of $\Delta_\lambda$, we show only the values up to 5. For the following cases, $\Delta_\lambda$ are not shown in the figure: NGC1846-$R_M$ ($\Delta_\lambda \sim 28.3$), NGC6440-$R_M$ ($\Delta_\lambda \sim 31.5$), NGC6553-$R_M$ ($\Delta_\lambda \sim 24.1$), NGC6528-$R_M$ ($\Delta_\lambda \sim 31.9$), NGC2009-$R_U$ ($\Delta_\lambda \sim 21$), NGC6553-$R_U$ ($\Delta_\lambda \sim 51$), NGC6440-$R_U$ ($\Delta_\lambda \sim 55$), NGC1846-$R_U$ ($\Delta_\lambda \sim 98$) and NGC6528-$R_U$ ($\Delta_\lambda \sim 115$).

Figure 5. $A_V$ values from Starlight versus reference values. Each panel shows the results for one of the wavelength ranges tested in this work (see Table 2). The identity line $x=y$ is shown in solid black to guide the eye. Markers are colored by host galaxy, as indicated in the panel in the right.

Individual uncertainties will vary from spectrum to spectrum. In any case, these simulations show that the typical observational errors are small.

Another source of statistical uncertainty not accounted for in the previous estimation is the stochastic sampling of short-lived stellar evolutionary phases, due to the finite stellar mass inside the field of view (e.g. da Silva et al. 2012). In a general sense, one should trust more the results for clusters with larger enclosed mass, as given in column 15 of Table 1 from Usher et al. (2017). We investigated if there is a relation between the residuals of the retrieved parameters and the mass enclosed, but no correlation was found.
4.6 Uncertainties coming from different observations of the same clusters

We repeated the spectral fitting in the range $R_W$ using instead the integrated globular spectra from Schiavon et al. (2005), for the 35 globular clusters in common with WAGGS. Figure 10 compares the parameters obtained from the two datasets. It is noticeable that the values obtained from the two sets correlate well, with most of the fits showing discrepancy values smaller than 0.1 for log(ages), 0.3 for [Fe/H] and 0.1 for [$\alpha$/Fe].

The source of the deviations can be two-fold: the two data sets have been observed with different observing strategies and data reduction pipelines; and the differences in resolutions, Schiavon et al. (2005) spectra having lower resolution (FWHM = 3.1Å) compared to WAGGS convolved to MILES resolution (FWHM = 2.5Å).

To verify these possibilities, we convolved WAGGS spectra to FWHM = 3.1Å and run new fits in the $R_W$ range. We evaluate that the differences due to different resolutions are: $-0.015 \pm 0.057$ in log($t_{\text{mass}}$), $-0.014 \pm 0.063$
How well spectral fitting performs?

Figure 7. Retrieved values of age versus isochrone ages, in linear scale. Upper line shows the results of light-weighted ages, while the lower line shows the results of mass-weighted ages (see Section 3.3). Each column represents one of the wavelength ranges tested in this work (see Table 2). The identity line x=y is shown in solid black to guide the eye. Markers are colored by \([\text{Fe/H}]\).

Table 4. Typical standard deviations of the parameters obtained from perturbing the observed spectrum according to its error spectrum.

| Parameter | median(σ) |
|-----------|-----------|
| log\((t_{\text{light}})\) | 0.002 |
| log\((t_{\text{mass}})\) | 0.002 |
| \([\text{Fe/H}]_{\text{light}}\) | 0.02 |
| \([\text{Fe/H}]_{\text{mass}}\) | 0.03 |
| \([\alpha/\text{Fe}]_{\text{light}}\) | 0.02 |
| \([\alpha/\text{Fe}]_{\text{mass}}\) | 0.02 |

in \([\text{Fe/H}]_{\text{mass}}\), and 0.015 ± 0.057 in \([\alpha/\text{Fe}]_{\text{mass}}\). In comparison, the differences in parameters between WAGGS (at FWHM = 3.1Å) and Schiavon data are: –0.005 ± 0.051 in log\((t_{\text{mass}})\), –0.025 ± 0.167 in \([\text{Fe/H}]_{\text{mass}}\), and 0.035 ± 0.078 in \([\alpha/\text{Fe}]_{\text{mass}}\).

We therefore conclude that different observing strategies and/or data reduction process are comparable to the effect of changing the spectral resolution. Still, the different observations seem to be dominant effect on the dispersion of abundance parameters seen in Fig. 10.

4.7 Summary of results

Table 5 shows the median ± the interquartile range (IQR) of each \(\Delta P = P_{\text{Starlight}} - P_{\text{Reference}}\), where \(P\) represents any of the retrieved parameters. We see that reddening values are better constrained when more information is used (in our case, the \(R_M\) range). When it comes to ages, the mass-weighted-\(R_B\) values are the ones with less deviation from the isochrone (reference) values. The closest match to the reference metallicities is reached in the mass-weighted case, using the \(R_W\) range in the fit. For \([\alpha/\text{Fe}]\), \(R_W\) returns the more consistent results, not showing a significant variation between light and mass weighted values.

The values shown in Table 5 can guide the user in the choice of the wavelength range to use for his/her own application. This choice will be dependent on each science case: e.g., if the main goal is to obtain reliable \([\alpha/\text{Fe}]\) values and a larger uncertainties in ages are manageable, the interval \(R_W\) will be favoured; conversely, if maximising the accuracy of \(\Delta \log(t_{\text{mass}})\) is the goal, we suggest \(R_B\). Unfortunately we cannot suggest a single wavelength range for all cases.

The ranges suggested here are admittedly based on the WAGGS spectra, but for a more general application one could inspect Table 3 to verify which spectral features are driving the results. By looking at Table 3 alone one would expect \(R_U\) to be the range most sensitive to age, but we interpret that the lower SNR in this region causes the best age-determination range to shift to \(R_B\). The range \(R_W\) is a narrow one around the MgB triplet, so it is not surprising that it performed best for deriving \([\alpha/\text{Fe}]\) values. Regarding metallicity, the range which performed best on average is \(R_R\), basically devoid of age features (other than Hα), even though with a relatively large IQR.

5 DISCUSSIONS

On blue HB stars and blue stragglers Arguably, the most striking failure in retrieving the parameters from our exercise in spectral fitting are the ages of the low metallicity systems (figure 7). We better highlight the dependence of the problem with metallicity in Figure 11, which shows
Figure 8. Retrieved values of [Fe/H] versus reference values. Each line shows the results for one of the different definitions of [Fe/H] (described in Section 3.3). Each column represents one of the wavelength ranges tested in this work (see Table 2). The identity line x=y is shown in solid black to guide the eye. Markers are colored by isochrone age.

Table 5. Median values of ∆parameters for each given wavelength range. $\Delta P = P_{\text{Starlight}} - P_{\text{Reference}}$ where P represents each parameter. Columns 2 – 6 indicate the fitted wavelength range, as defined in table 2.

| Parameter | $R_M$ | $R_U$ | $R_B$ | $R_R$ | $R_W$ |
|-----------|-------|-------|-------|-------|-------|
| $\Delta \log(t_{\text{light}})$ | -0.44 ± 0.45 | -0.27 ± 0.23 | -0.40 ± 0.41 | -0.56 ± 0.41 | -0.19 ± 0.29 |
| $\Delta \log(t_{\text{mass}})$ | -0.19 ± 0.34 | -0.19 ± 0.16 | -0.06 ± 0.17 | -0.17 ± 0.18 | -0.14 ± 0.22 |
| $\Delta [\text{Fe/H}]_{\text{light}}$ | 0.33 ± 0.28 | 0.34 ± 0.21 | 0.24 ± 0.20 | 0.25 ± 0.27 | 0.18 ± 0.24 |
| $\Delta [\text{Fe/H}]_{\text{mass}}$ | 0.23 ± 0.32 | 0.39 ± 0.29 | 0.34 ± 0.26 | 0.09 ± 0.39 | 0.22 ± 0.27 |
| $\Delta \alpha/\text{Fe}_{\text{light}}$ | -0.10 ± 0.19 | -0.05 ± 0.22 | -0.05 ± 0.23 | -0.17 ± 0.22 | -0.05 ± 0.17 |
| $\Delta \alpha/\text{Fe}_{\text{mass}}$ | -0.08 ± 0.20 | -0.06 ± 0.22 | -0.02 ± 0.22 | -0.14 ± 0.23 | -0.05 ± 0.17 |
| $\Delta A_V$ | 0.10 ± 0.30 | -0.25 ± 0.50 | 0.12 ± 0.28 | 0.31 ± 0.40 | 0.14 ± 0.38 |
How well spectral fitting performs?

\[ \Delta t_i = t_i - t_{iso} \] (being \( t_i \) the different ages defined in section 3.3) vs. [Fe/H], for the range \( R_W \). The classical interpretation of blue-light excess in Galactic old populations is the presence of extended HB morphologies or blue-stragglers, when unaccounted for in SSP models. The "excess" of blue light coming from these stars could be interpreted by the inferred ages of the integrated spectra. These effects have been discussed in a myriad of papers (e.g. de Freitas Pacheco & Barbuy 1995; Lee et al. 2000; Maraston & Thomas 2000; Peterson et al. 2003; Schiavon et al. 2004b).

In particular in the context of spectral fitting, we refer the reader to the work of Koleva et al. (2008) and Ocvirk (2010). Koleva et al. (2008) were able to reproduce CMD ages for 35 out of 40 clusters, using hot stars together with the SSPs when needed, to mimic the effect of the HB morphology (or blue stragglers). Ocvirk (2010) proposes that any young starburst superimposed on an old stellar population in this range [Fe/H] = [−2, −1.2] could be regarded as a modelling artefact, if it weighs less than 12% of the optical light. The work by Conroy et al. (2018) illustrates a similar result (see their Figure 15), in a sample with [Fe/H] between [−1.5, +0.3]. It remains to be investigated in a future work if this effect can explain all the deviant cases found in this work.

Quality of the SSP model at low metallicities The SSP models used in this work (Vazdekis et al. 2015) are built upon the empirical stellar library MILES (Sánchez-Blázquez et al. 2006). The coverage of the Hertzsprung–Russell diagram (HRD hereafter) in empirical stellar libraries in general is known to be poorer towards lower metallicities, due to observational constraints (see e.g. Coelho 2009 and Section 3.2 in Vazdekis et al. 2010).

On the one hand, according to the analyses in Vazdekis et al. (2010), old models (>~ 10 Gyr) can be safely used also at low metallicity. In this aspect the poorer coverage at low metallicities should not be a reason for the discrepant ages observed here. On the other hand, the recent work by Coelho et al. (2019) investigated the effect that the HRD coverage has on the SSP model predictions. In their figure 15 and table 7, there is evidence that the sparse coverage of HRD imply in an underestimation of the inferred ages. The authors observe a median effect of \( \Delta \log(t) \approx 0.11 \), observed in systems with metallicity between [Fe/H] = −1 and 0. It remains an open possibility, then, if this effect could partly account for the age discrepancies observed in this work.

The choice of isochrones Ideally, one would want the SSP models to adopt the same isochrones as the ones used to derive the reference ages. Or the other way around, obtain from literature age estimations using the same isochrones as the ones used in the SSP models. This is not the case here, since the SSP models are using isochrones from Pietrinferni et al. (2004a, 2006a), while the reference ages come from mixed sources (see table 1).

The review from Gallart et al. (2005) compares several stellar evolutionary libraries from literature, and quantify how they (dis)agree when used to infer age, metallicity, or distance of a population. Ages predicted by different models agree within \( \pm 0.01 \) Gyr at young ages (~0.1 Gyr), and differ by up to 1 Gyr in the intermediate and old age regimes. We therefore assume that different isochrones can account for a spread of 1 Gyr in inferred ages, which may be important to intermediate-age populations.
On the other hand, one may argue that using the same isochrones does not necessarily help: the isochrones used to measure the age from CMDs could make a good prediction of the main-sequence turn-off (MSTO) but not of later stellar evolution phases (such as the horizontal branch or red giant branch), which may dominate over the MSTO in the integrated light.

Elusive $\alpha$/Fe The $\alpha$/Fe is an important chemical marker in stellar populations, working as a cosmic clock for the history of star formation (e.g. Matteucci 2003). It has been measured in galaxies for more than a decade using spectral indices (e.g. Thomas et al. 2005), but our results seem to indicate that it can be elusive to measure from spectral fitting. Qualitatively similar results are shown in Figure 15 in Conroy et al. (2018), where measuring individual abundances other than [Fe/H] show considerable scatter. We could partly attribute the difficulties to the dynamical range – considerably smaller than the other parameters – from $\sim -0.2$ to $\sim +0.5$ in the observations, and only two values available in the models (0.0 and +0.4). In any case, the results in Table 5 and Figure 9 indicate that uncertainties similar or larger than 0.2 dex are to be expected, and that fitting a narrow wavelength range is more indicated (at the expense of worsening the other results).

Example of an extreme case and uncertainties: We discussed possible uncertainties coming from the observations in Sections 4.5 and 4.6. Estimating errors in the models and method are considerably more difficult. In the case of the SSP fits, we can follow a procedure similar to the one adopted in Charlot et al. (2002, see section 4.2), namely, considering that all models within $\Delta \chi^2 < 1$ around the best $\chi^2$ could also be picked as a good fit.

In our analysis this approach is limited to the SSP fits (because STARLIGHT outputs the $\chi^2$ of all SSP fits, but not of all multi-population fits of the Markov-Chain), and should be regarded with care given the biases that we are aware to exist in the SNR estimation (see section 2).

In any case, as a test-case, we choose to apply this scheme to the results from NGC6397, which falls into the worst-case scenario of our analysis of ages: isochrone fitting returns an age of 13.5 Gyr, while the retrieved light-weighted, mass-weighted and SSP-equivalent age values are about 1.0, 5.0 and 2.0 Gyr, respectively.

If one estimates the age uncertainty from the criteria $\Delta \chi^2 < 1$, the acceptable ages for NGC6397 range from 1.25 to 7.50 Gyr for SSPs with [Fe/H] $= -2.27$, and from 1.25 to 4.50 Gyr for SSPs with [Fe/H] $= -1.79$ (bracketing the literature value of [Fe/H] = −2.0 for this cluster). Although not large enough to bring the result in agreement with the isochrones age of the cluster, the intervals are rather large, suggesting that indeed the models and/or method may not be able to clearly distinguish the age parameter. For cases less extreme than NGC6397, this could bring the results in agreement with literature.

6 CONCLUSIONS
In this work we obtained reddening $A_V$, ages, [Fe/H] and $\alpha$/Fe from spectral fitting of integrated spectra of globular clusters (Usher et al. 2017), simulating a procedure which is routinely applied to the analysis of galaxy spectra. We repeated the analyses for different wavelength ranges, and compared the retrieved parameters with reference values compiled from literature (obtained from CMD analysis and high-resolution spectroscopy), to decide which ranges are more robust. Our main conclusions are summarised below.

The intervals which best reproduced fiducial values of reddening, ages, [Fe/H] and $\alpha$/Fe are different (respectively $R_V$ 3540–7409, $R_B$ 4170–5540, $R_R$ 5280–7020 and $R_V$ 4828–5364). For all parameters, changing the wavelength range of the spectral fit changes the parameters inferred. Reddening is the only parameter which is favoured by the use of the widest possible interval. On the other limit, $\alpha$/Fe is better retrieved when the narrowest range is fitted.

Ages are poorly constrained for metal-poor objects, in accordance with previous results in literature. The correlation between age residuals (retrieved – literature) with metallicity seems to be consistent with the effect of blue HB stars or blue stragglers. Alternatively, the sparser coverage of the HRD by the models at low metallicities may also play a role.

The explanations of our findings can be multi-fold, and one can argue that it may be due to the choice of models or code employed. But we stress that our results are qualitatively similar to what other work found with different models and different fitting codes (e.g. Walcher et al. 2009; Cezario et al. 2013, which employed codes that do not use the continuum information). Admittedly, the result that different wavelength ranges will return different parameters is an uncomfortable one. Moreover, our results do not conform with the common assumption existent in the spectral fitting community where “the more pixels the better”. That drives us back to the decades-old work on defining spectral indices, where the features driving the stellar population parameters were determined almost in an artisanal manner. We make the hypothesis that to use more pixels can bring more signal but do not necessarily bring more information, in the sense that we want to maximise something similar to the SNR concept (increase the signal without increasing the noise). Using the case of $\alpha$/Fe as an example, centring the fit around Mg triplet feature may maximise the information, which otherwise would be diluted in a larger wavelength range.

In essence we conclude that a user of spectral fitting codes interpreting galaxies will not obtain the best result necessarily by using more data (wider wavelength ranges).

7 DATA AVAILABILITY
Value-added catalogues produced in this work using spectral fitting are available in the article’s online supplementary material. Instructions on how to acquire the public data herein used are described in Sec. 2.

ACKNOWLEDGEMENTS
GG acknowledges support from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). PC acknowledges support from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq 310041/2018-0) and
Figure 10. Values for mass-weighted ages and [Fe/H] (upper and middle panels, respectively) retrieved from the WAGGS library versus the values retrieved when fitting spectra from Schiavon et al. (2005)'s library. The identity line \( x=y \) is shown in solid black to guide the eye. The third panel shows the distribution of \( \Delta[\alpha/Fe] = [\alpha/Fe]_{\text{WAGGS}} - [\alpha/Fe]_{\text{Schiavon}} \).

Figure 11. \( \Delta t = t_{\text{Starlight}} - t_{\text{iso}} \) values (light-weighted in upper panel, mass weighted in lower panel) versus [Fe/H], fitting the \( R_{18} \) interval. It is noticeable that there is a clear correlation between the differences between fitted and isochrone ages with the metallicity.