Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **512,** 378–400 (2022) https://doi.org/10.1093/mnras/stac471
Advance Access publication 2022 February 23

Strong CO absorption features in massive ETGs

Elham Eftekhari [1,2]★ Francesco La Barbera [3] Alexandre Vazdekis [1,2] Carlos Allende Prieto [1,2] and Adam Thomas Knowles [4]

[1]*Instituto de Astrofísica de Canarias, E-38200 La Laguna, Tenerife, Spain*
[2]*Departamento de Astrofísica, Universidad de La Laguna, E-38205 La Laguna, Tenerife, Spain*
[3]*INAF-Osservatorio Astronomico di Capodimonte, sal. Moiariello 16, I-80131 Napoli, Italy*
[4]*Jeremiah Horrocks Institute, School of Physical Sciences and Computing, University of Central Lancashire, Preston PR1 2HE, UK*

Accepted 2022 February 17. Received 2022 February 17; in original form 2021 July 14

ABSTRACT

Massive Early-Type Galaxies (ETGs) in the local Universe are believed to be the most mature stage of galaxy evolution. Their stellar population content reveals the evolutionary history of these galaxies. However, while state-of-the-art Stellar Population Synthesis (SPS) models provide an accurate description of observed galaxy spectra in the optical range, the modelling in the Near-Infrared (NIR) is still in its infancy. Here, we focus on NIR CO absorption features to show, in a systematic and comprehensive manner, that for massive ETGs, all CO indices, from *H* through to *K* band, are significantly stronger than currently predicted by SPS models. We explore and discuss several possible explanations of this 'CO mismatch', including the effect of intermediate-age, asymptotic-giant-branch-dominated, stellar populations, high-metallicity populations, non-solar abundance ratios, and the initial mass function. While none of these effects is able to reconcile models and observations, we show that ad hoc 'empirical' corrections, taking into account the effect of CO-strong giant stars in the low-temperature regime, provide model predictions that are closer to the observations. Our analysis points to the effect of carbon abundance as the most likely explanation of NIR CO line-strengths, indicating possible routes for improving the SPS models in the NIR.

Key words: galaxies: stellar content – infrared: galaxies.

1 INTRODUCTION

The study of the stellar content of galaxies, both in the local and early Universe, is fundamental to understand how they shape up over cosmic time, as it provides key constraints on star formation rates, total stellar masses, chemical enrichment, and the stellar initial mass function (IMF). Since Stellar Population Synthesis (SPS) models are an essential tool to constrain the stellar content of galaxies, a detailed assessment of their validity and limitations is crucial to determine the physical and evolutionary properties of these systems. While spectral synthesis modelling at optical wavelengths is now a mature field of research and optical galaxy spectra can be accurately matched with SPS models, there is still a long road ahead for Near-Infrared (NIR) SPS models to consistently agree with observations (Riffel et al. 2019; Eftekhari, Vazdekis & La Barbera 2021). For example, only in the last decade, the problem of matching strong sodium absorption lines of massive Early-Type Galaxies (ETGs) has been scrutinized in the NIR. The strength of NIR sodium features in massive ETGs is much stronger than would be expected from an old stellar population with a Milky Way (MW)-like IMF and with solar elemental abundance ratios. A combination of a bottom-heavy IMF and a highly enhanced sodium abundance can reconcile the tension between observations and NIR models (Smith et al. 2015; La Barbera et al. 2017; Röck et al. 2017); however, the finding of massive ETGs with MW-like IMFs – as derived by strong gravitational lensing analyses – and strong sodium line-strengths at 1.14 μm calls for caution in interpreting the NIR sodium line-strengths (Smith & Lucey 2013; Smith et al. 2015). Another disagreement between observations and stellar population models in the NIR arises from CO absorption features, which are prominent in the *H*- and *K*-band spectral regions, and have always been a puzzle.

The appearance of the first overtone of CO in *K* band, at ∼2.3 μm, in the spectra of galaxies was discussed by many authors in the 1970s and 1990s (Baldwin, Frogel & Persson 1973b; Frogel et al. 1975, 1978; Aaronson et al. 1978; Frogel, Persson & Cohen 1980; Oliva et al. 1995; Mobasher & James 1996; James & Mobasher 1999). CO absorption originates in the atmospheres of red giants and supergiants, which tend to have deeper CO absorptions than dwarf stars (Baldwin et al. 1973a). Faber (1972) opened up the discussion that optical data could not be used to uniquely determine the proportion of M dwarfs and M giants in the galactic nucleus of M31, showing that while a model with enhanced M dwarfs in synthesized models would match the Na doublet at 8190 Å, a model dominated by M giants is required to explain the *K*-band CO strength. Since then, several authors have analysed the *K*-band CO absorption of galaxies, alone or in combination with other indices, by comparing the observed strengths with those of stellar spectra (Baldwin et al. 1973a, b; Frogel et al. 1975, 1978; Oliva et al. 1995; Mobasher & James 1996; James & Mobasher 1999). All of these studies found that line-strengths of the *K*-band CO absorption lie in the range of giant stars, concluding that most of the light emitted from galaxies in the CO spectral region comes from these stars.

Using NIR observations of globular clusters, Aaronson et al. (1978) showed that the 2.3 μm CO index strength is strongly correlated with metallicity. They constructed SPS models and compared

★ E-mail: elhamea@iac.es

© 2022 The Author(s)
Published by Oxford University Press on behalf of Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article/512/1/378/6534922 by University of Central Lancashire user on 24 March 2022

them with observations of the central regions of ETGs, claiming that metal-rich models with a Salpeter IMF adequately fit the CO index of the brightest ellipticals. Frogel et al. (1978) also found that any significant increase in the number of late-type dwarfs beyond those already contained in the SPS models drives the *K*-band CO index to unacceptably low values,[1] concluding that the changes observed in the CO indices of galaxies are due to variations in the mean metallicity of their stellar populations. However, Frogel et al. (1980) attributed the differences between various colours and *K*-band CO index of ETGs with respect to those of globular clusters and stellar synthesis models to a population of low-temperature luminous stars present neither in the clusters nor in the models. They hypothesized giant branch stars with higher metallicity than the Sun and/or asymptotic giant branch (AGB) stars above the first red giant tip as two candidates for such a population.

Separation of the relative contributions to the *K*-band CO line-strength from young supergiants in a burst population (a few Myr) and giants in an older stellar system (∼1 Gyr) has also been a subject of debate; the NIR CO features are mainly sensitive to effective temperature but are also somewhat shallower in giants than supergiants of similar temperatures (Kleinmann & Hall 1986; Origlia, Moorwood & Oliva 1993; Oliva et al. 1995). However, metal-rich red giant stars can have CO absorptions that are as strong as the red supergiants found in starbursts; in other words, cold giants and slightly warmer supergiants can have equally strong CO line-strengths (Origlia et al. 1993; Oliva et al. 1995). This hampers the interpretation of CO absorptions in galaxies alone, in the absence of independent measurement of the stellar temperature.

ETGs are known to host old stellar populations with little contribution, if any, from recently formed stellar populations. Indeed, since Frogel et al. (1980), the CO (2.3 μm) absorption has been used to possibly infer the presence of young stars (red giants and supergiants) in ETGs. In particular, Mobasher & James (2000) found that the CO line-strength is stronger for ellipticals in the outskirts of the Coma cluster than in the core, interpreting this as an evidence for younger populations in galaxies inhabiting low-density environments (see also James & Mobasher 1999). Mobasher & James (1996) and Mármol-Queraltó et al. (2009) also interpreted the observed higher value of the 2.3 μm CO line-strength of galaxies in the field with respect to those in denser environments of clusters as due to relatively more recent episodes of star formation in field galaxies. Unfortunately, most of these analyses has hitherto been based on a direct comparison of CO indices in galaxies to those for stellar spectra, with no detailed comparison to predictions of SPS models. Recently, Baldwin et al. (2018) measured the 2.3 μm CO line-strength for 12 nearby ETGs, comparing to predictions from different sets of SPS models. They found that all models systematically underpredict the strength of the *K*-band CO.

While the CO bandhead in the *K* band has been extensively analysed in the literature, not much effort has been put so far into studying other NIR CO lines, which are prominent in galaxy spectra, especially in the *H* band. This is because low-temperature and heavily obscured stars are brighter in the *K* band than the *H* band, and, perhaps more importantly, severe contamination of the *H*-band spectral range from sky emission lines has prevented its exploitation in the work of the late 1900s. However, nowadays, due to the high resolution of NIR spectrographs and new sky-subtraction techniques, the *H* band is fully accessible to detailed stellar population studies.

CO absorptions in the *K* band arise from transitions between two vibrational states with a difference ($\Delta\nu$) in quantum number ν of 2, while the $\Delta\nu$ of CO absorptions in the *H* band is equal to 3 (see table 10 of Rayner, Cushing & Vacca 2009). This results in a lower (∼2 orders of magnitude) optical depth of CO lines in the *H* band than the *K* band and therefore, the CO lines in the *H* band saturate in cooler stars than those in the *K* band. Hence the strength of the CO lines in the *H* and *K* bands is expected to behave differently with spectral type and luminosity (Origlia et al. 1993). Indeed, performing a simultaneous analysis of different features arising from the same chemical species is of paramount importance, as it helps in breaking degeneracies among relevant stellar population parameters (e.g. Conroy & van Dokkum 2012). The only effort in this direction has been put by Riffel et al. (2019), who have analysed CO line-strengths in both the *H* and *K* bands, for a sample of nearby ETGs, covering a range of galaxy mass, as well as star-forming galaxies. They found that while some CO lines are matched by the models, others seem to exhibit a significant disagreement.

In this paper, we perform a detailed analysis of a whole battery of CO absorption features that are found in the NIR spectra of ETGs, focusing on a homogeneous, high-quality sample of very massive nearby galaxies, with a velocity dispersion of $\sim$300 km s^{-1} (i.e. the high-mass end of the galaxy population), as well as on other galaxy samples collected from previous works. We show that, indeed, all CO features, besides the well-studied 2.3 μm CO bandhead, are systematically underestimated by the models for the very massive ETGs. We scrutinize several possible explanations of this 'CO mismatch' problem, including the effect of a non-universal IMF, a contribution from young and intermediate-age populations, high-metallicity stars, as well as the effect of non-solar abundance ratios. We also present ad hoc empirical SPS models that might help to solve the problem, by taking advantage of the scatter of stars in the available stellar libraries used to construct the models.

The paper is organized as follows: In Sections 2 and 3, we describe our samples of ETGs, as well as the stellar libraries and synthesis models used in this work. In Section 4, we show the CO mismatch problem for NIR absorption features, by comparing models and observations. Different experiments are presented in Section 5 in order to search for a possible solution of the CO puzzle. Our empirical approach to address the tension is explained in Section 6. Section 7 provides a discussion of our results. The overall conclusions are summarized in Section 8.

2 SAMPLES

We used different samples of ETGs drawn from the literature. Our main galaxy sample is that of La Barbera et al. (2019) (hereafter LB19), consisting of exquisite-quality, high-resolution, optical and NIR spectra for a sample of very massive ETGs at $z \sim 0$, collected with the X-SHOOTER spectrograph (Vernet et al. 2011) at the ESO Very Large Telescope (VLT). Other samples of ETGs were used wherever the quality of data and wavelength coverage were suitable for our analysis. Although these samples are far from being homogeneous (as they were observed with different instruments), they encompass a wide range in galaxy stellar mass and stellar population parameters, allowing for a comprehensive study of NIR CO features. In particular, we have included two samples of ETGs residing in varying environments (see below), namely in the field (Mármol-

[1] Kroupa & Gilmore (1994) obtained a similar result by simulating the *K*-band spectrum of the cooling-flow ellipticals, i.e elliptical galaxies with a low-mass star accretion population. They used the spectral library of low-mass stars from Arnaud, Gilmore & Collier Cameron (1989) and showed that by spectroscopy around the *K*-band CO feature, an overabundance of low-mass stars in these galaxies can be estimated.

Downloaded from https://academic.oup.com/mnras/article/512/1/378/6534922 by University of Central Lancashire user on 24 March 2022

Queraltó et al. 2009), and in the Fornax cluster (Silva, Kuntschner & Lyubenova 2008), whose *K*-band spectra were obtained with the same observational set-up. These two samples allow us to explore the effect of the environment on the CO strengths.

The main properties of our galaxy samples are summarized as follows:

(i) La Barbera et al. (2019) (hereafter XSGs): The seven massive ETGs of this sample are at redshift $z \sim 0.05$, and span a velocity dispersion (σ) range of ∼300–360 km s^{-1}. Five galaxies are centrals of galaxy groups, while two systems are satellites (see LB19 for details). The galaxies were observed using the X-SHOOTER three-arm echelle spectrograph, mounted on UT2 at the ESO VLT. The wavelength coverage of the spectra in the NIR arm ranges from 9800 to 25 000 Å with a final resolution of ∼5500 [full width at half-maximum (FWHM)]. The spectra used for this work were extracted for all galaxies within an aperture of radius 1.5 arcsec, and have a high signal-to-noise ratio (SNR), above 170 Å^{-1}. In addition to individual spectra, we also used a stacked spectrum in our analysis, to characterize the average behaviour of the sample. Using optical (Hβo, H$_{\gamma F}$, TiO1, TiO2$_{SDSS}$, aTiO, Mg4780, [MgFe]', NaD, Na I8190) and NIR (Na I1.14, Na I2.21) spectral indices, LB19 showed that stellar populations in the centre of these galaxies are old (∼11 Gyr), metal-rich (∼+0.26 dex), enhanced in α-elements ([α/Fe]∼+0.4 dex), and with bottom-heavy IMF (with logarithmic slope $\Gamma_b \sim 3$ for the upper segment of a low-mass tappered IMF, often regarded as 'bimodal').

(ii) François et al. (2019): The 12 nearby ($z < 0.016$) and bright ($B_t \sim$ 11–13 mag) galaxies of this sample span ∼35–335 km s^{-1} in velocity dispersion and are distributed along the Hubble sequence from ellipticals to spirals. They have been observed with the X-SHOOTER spectrograph at the VLT. NIR spectra were obtained with a 1.2 arcsec slit, providing a resolving power of 4300. The one-dimensional spectra were extracted by sampling the same spatial region for all galaxies. Using optical indices (<Fe>, [MgFe]', Mg$_2$, Mg$_b$, and H$_\beta$), François et al. (2019) showed that stellar populations in the centre of these galaxies span a wide range of values in age and metallicity ($0.8 \leq$ age (Gyr) ≤ 15 and $-0.4 \leq$ [Z/H] ≤ 0.5).

(iii) Baldwin et al. (2018): They obtained *JHK*-band spectra for 12 nearby ETGs from the ATLAS^{3D} sample (Cappellari et al. 2011), with a velocity dispersion range of $\sigma = 80$–120 km s^{-1}, using GNIRS, the Gemini Near-Infrared Spectrograph at the 8m Gemini North telescope in Hawaii. The galaxies span a broad range in age from 1 to 15 Gyr (optically derived SSP-equivalent ages) at approximately solar metallicity. They used a 0.3 arcsec slit, with a spectral resolution of R ∼1700. One-dimensional spectra were extracted within an aperture of $\pm\frac{1}{8}$R$_{\rm eff}$ except for one galaxy (whose spectrum was extracted within $\pm\frac{1}{12}$R$_{\rm eff}$). The SNR of the spectra is in the range 50–200.

(iv) Mármol-Queraltó et al. (2009): They observed 12 ETGs in the field in the velocity dispersion range 59 km s^{-1} $< \sigma <$ 305 km s^{-1} with the ISAAC NIR imaging spectrometer, mounted on UT1 at the VLT. NIR spectra were obtained with a 1 arcsec slit, providing a resolving power of 7.1 Å at 2.3 μm. The wavelength coverage of the spectra is short (from 2.20 to 2.29 μm), including Na I, Fe I, Ca I, and the first CO absorptions at the red end of the *K* band (see below). They extracted galaxy spectra within a radius corresponding to $\frac{1}{8}$R$_{\rm eff}$. The ages of these galaxies were determined by Sánchez-Blázquez et al. (2006b), using the H$_\beta$ index and a preliminary version of the SPS models of Vazdekis et al. (2010).

(v) Silva et al. (2008): This sample consists of eight ETGs in the Fornax cluster with $\sigma = 70$–360 km s^{-1}, all observed with the ISAAC NIR imaging spectrometer. A 1 arcsec slit was used during the observations, giving a spectral resolution of $R \approx 2900$ (7.7 Å FWHM). The central spectra were extracted within a radius corresponding to $\frac{1}{8}$R$_{\rm eff}$ with an SNR of 48–280, and covering a wavelength range of 2.12–2.37 μm. The ages of these galaxies are drawn from Kuntschner & Davies (1998), and were computed with Worthey et al. (1994) SPS models.

Note that, for all spectra, we measured CO spectral indices (see Section 4), when they lie within the available spectral range and are considered to be safe for the stellar population analysis, according to the criteria given in Eftekhari et al. (2021) (see their section 4.8).

3 STELLAR LIBRARIES AND STELLAR POPULATION MODELS

We compare observed CO index strengths with predictions of various SPS models which differ in a number of ingredients, such as the adopted isochrones, stellar libraries, IMF shape, age and metallicity coverage, as well as the prescription to implement the AGB phase. The latter is one of the most important aspects when studying the NIR spectral range, as it is the main source of differences seen among models. We also compare CO indices observed in our galaxy samples with individual stars, from theoretical and empirical stellar libraries. The main features of the models and libraries are summarized next.

3.1 Stellar population models

(i) E-MILES: We used two model versions: base E-MILES (Vazdekis et al. 2016) and α-enhanced E-MILES (an updated version of Na-enhanced models described in La Barbera et al. 2017). The models are available for two sets of isochrones: BaSTI (Pietrinferni et al. 2004) and Padova00 (Girardi et al. 2000). These isochrones provide templates for wide range of ages, from 1–14 Gyr in BaSTI and 1–17.78 Gyr for Padova00, and metallicities, from −0.35 to +0.26 dex in BaSTI and −0.4 to +0.22 dex in Padova00. The base model is computed for different IMF shapes – Kroupa universal, revised Kroupa, Chabrier, unimodal, and bimodal – while the α-enhanced model is only available for bimodal IMF distributions (see Vazdekis et al. 1996, 1997 for a description of different IMF parametrizations). E-MILES models cover a wide range in wavelength from ultraviolet (1680 Å) to infrared (50 000 Å) and are based on NGSL stellar library (Gregg et al. 2006) in the ultraviolet, MILES (Sánchez-Blázquez et al. 2006a), Indo-US (Valdes et al. 2004), and CaT (Cenarro et al. 2001) stellar libraries in the optical and 180 stars of IRTF stellar library (see below) in the infrared. For this study, we utilized the BaSTI-based model predictions with bimodal IMF shape with slopes $\Gamma_b = 1.3$ (representative of the Kroupa-like IMF) and $\Gamma_b = 3.0$ (representative of a bottom-heavy IMF), and two metallicity values, around solar (+0.06) and metal-rich (+0.26). Note that the BaSTI models have cooler temperatures for low-mass stars and use simple synthetic prescriptions to include the AGB regime.

(ii) Conroy et al.: We used two model versions: Conroy & van Dokkum (2012) (hereafter CvD12) and Conroy et al. (2018) (hereafter C18). The CvD12 models rely on 91 stars from the IRTF stellar library in the NIR, using different isochrones to cover different phases of stellar evolution, from the hydrogen burning limit to the end of the AGB phase, namely the Dartmouth isochrones (Dotter et al. 2008), Padova isochrones (Marigo et al. 2008), and Lyon isochrones (Chabrier & Baraffe 1997; Baraffe et al. 1998). The publicly available models of CvD12 cover either non-solar abundance templates at 13.5 Gyr, or younger ages (from 3 to 13.5 Gyr) at solar abundance

Downloaded from https://academic.oup.com/mnras/article/512/1/378/6534922 by University of Central Lancashire user on 24 March 2022

ratios and metallicity. The 13.5 Gyr model of solar metallicity and solar abundance is available for a Salpeter IMF, Chabrier IMF, two bottom-heavy IMFs with logarithmic slopes of $x = 3$ and $x = 3.5$, and a bottom-light IMF. In this paper, we utilized C-, O-, and α-enhanced models of age 13.5 Gyr in addition to solar metallicity models of ages 3 to 13.5 Gyr with a Chabrier IMF. The updated version of CvD12 models is based on the MIST isochrones (Choi et al. 2016) and utilizes the Extended IRTF Stellar Library (Villaume et al. 2017) that includes continuous wavelength coverage from 0.35 to 2.4 μm. C18 models cover a wide range in metallicity ($-1.5 \lesssim$ [Fe/H] $\lesssim 0.3$) and include new metallicity- and age-dependent response functions for 18 elements. In this paper, we used C18 models with overall metallicity [Z/H] of 0.0 and 0.2 dex, respectively, with a Kroupa-like IMF, for ages between 1 and 13 Gyr. We also used C-enhanced models ([C/Fe] $= +$ 0.15) of age 13 Gyr, solar metallicity, and a Kroupa-like IMF.

(iii) Maraston (2005) (hereafter M05): This model is based on the fuel consumption theorem of Renzini & Voli (1981) for the post-main-sequence stages of the stellar evolution (in particular TP-AGBs) and makes use of two stellar libraries: the BaSel theoretical library (Lejeune, Cuisinier & Buser 1998) in the optical and NIR and the Lançon & Wood (2000) empirical library of TP-AGB stars. The models are available for Salpeter and Kroupa IMF and for two horizontal branch (HB) morphologies: blue and red. They cover a wide range in metallicity (from −2.35 to +0.67 dex) and age (from 10^3 yr to 15 Gyr). For our analysis we have used models with solar metallicity, Kroupa IMF, and ages between 1 and 14 Gyr, with red HB morphology.

3.2 Stellar libraries

(i) IRTF (Cushing, Rayner & Vacca 2005; Rayner et al. 2009): This is an empirical spectral library of 210 cool stars covering the *J*, *H*, and *K* bands (0.8–5 μm) at a resolving power of $R \sim 2000$. The library includes late-type stars, AGB, carbon, and S stars, mostly with solar metallicity,[2] providing absolute flux-calibrated spectra. For this study, we have used a subsample of 180 IRTF stars that are also used to construct E-MILES models in the NIR.

(ii) Theoretical stars from Knowles (2019): We included a small set of theoretical stellar models, generated using the same method presented in Knowles et al. (2021). In summary, these models are computed using ATLAS9 (Kurucz 1993), with publicly available[3] opacity distribution functions described in Mészáros et al. (2012). We used the 1D and LTE mode of ASSϵT (Advanced Spectrum SynthEsis Tool, Koesterke 2009) with input ATLAS9 atmospheres to produce fully consistent synthetic spectra at air wavelengths, with abundances varied in the same way in both model atmosphere and spectral synthesis components. The models here adopt Asplund, Grevesse & Sauval (2005) solar abundances and a microturbulent velocity of 2 km s^{-1}. We direct interested readers to Knowles (2019) and Knowles et al. (2021) for further details. The star models in this work have effective temperatures of 3500, 3750, and 4000 K, for $\log g = 1.5$, [M/H] $=$ [α/M] $= 0.0$, and two different carbon abundances: scaled-solar ([C/Fe] $= 0.0$) and enhanced ([C/Fe] $=$ 0.25). [M/H] here is defined as a scaled metallicity in which all metals, apart from the α-elements and carbon if they are non-solar, are scaled by the same factor from the solar mixture (e.g. [M/H] $=$ 0.2 $=$ [Fe/H] $=$ [Li/H]). This definition results in [α/M] $=$ [α/Fe] and [C/M] $=$ [C/Fe]. The models are generated specifically for this work, with a wavelength range of 1675.1–24 002.1 Å and a resolution of $R \approx 100\,000$.

(iii) APOGEE (Majewski et al. 2017): APOGEE is an *H*-band (1.59–1.69 μm) spectral library of stars in the Milky Way from the SDSS with a resolving power of ∼22 500. The 16th data release provides reduced and pipeline-derived stellar parameters and elemental abundances (ASPCAP; García Pérez et al. 2016) for more than 430 000 stars. For our study, in the ASPCAP catalogue, we made a selection of a subset of stars.[4] We also removed stars with SNR $<$ 100 (per pixel) and those that have a radial velocity scatter greater than 1.5 km s^{-1}, and a radial velocity error greater than 3 km s^{-1}.

4 CO SPECTRAL INDICES

As described in Section 1, fitting SPS models to CO absorption features of galaxies in the *K* band has proven to be challenging. To gain further insights into the problem, we consider here not only the already studied CO feature at ∼2.30 μm, but also other CO features at 2.32 and 2.35 μm. These are the first-overtone CO bandheads at the red end of the *K* band; their depth increases with decreasing stellar temperature (Kleinmann & Hall 1986; Rayner et al. 2009) and luminosity (Origlia et al. 1993) and becomes progressively weaker with decreasing metallicity (Frogel et al. 1975; Aaronson et al. 1978; Doyon, Joseph & Wright 1994; Davidge 2018, 2020), being stronger in red giant and supergiant stars than in dwarf stars.[5] In addition, we also include in our analysis six second-overtone CO bandheads in the *H* band, namely at 1.56, 1.58, 1.60, 1.64, 1.66, and 1.68 μm. In the following text, we first show observed and model spectra around each CO absorption (Section 4.1), and then we discuss observed and model line-strengths in the *K* and *H* bands (see Sections 4.2 and 4.3), respectively.

4.1 CO indices: observed versus model spectra

In Fig. 1, we show the spectra of XSGs from LB19, around CO absorptions from *H* through *K* band, and compare them with model spectra. From light to dark, the shifted red spectra correspond to the individual XSGs, while the median-stacked spectrum is shown in black. The wavelength definitions of CO indices, from Eftekhari et al. (2021), are shown with shaded grey and orange areas corresponding

[2]Note that, although the IRTF library has been recently extended to a wider range in metallicity by Villaume et al. (2017), in this paper we used the original version of the IRTF library, as this library is actually used to build up our reference SSP models (E-MILES). Moreover, the extended library shows a significant improvement mostly in the low-metallicity regime (see fig. 1 of Villaume et al. 2017), which is not relevant for our samples of massive ETGs.

[3]http://research.iac.es/proyecto/ATLAS-APOGEE//

[4]We removed stars with bits STAR_BAD, TEFF_BAD, LOGG_BAD, and COLORTE_WARN set in the APOGEE_ASPCAPFLAG bitmask and also those with bits BRIGHT_NEIGHBOR, VERY_BRIGHT_NEIGHBOR, PERSIST_HIGH, PERSIST_MED, PERSIST_LOW, SUSPECT_RV_COMBINATION, and SUSPECT_BROAD_LINES set in the APOGEE_STARFLAG bit-mask.

[5]The first-overtone bands of ^{12}CO show bandheads at 22 929.03, 23 220.50, 23 518.16, and 23 822.97 Å (in air) and those of ^{13}CO at 23 441.84, 23 732.94, and 24 030.49 Å. The feature at 2.35 Å is dominated by the ^{12}CO bandhead at 23 518.16 Å but disturbed by the ^{13}CO line at 23 441.84 Å. The strong saturation of the ^{12}CO feature makes the ^{13}CO line clearly visible even for low ^{12}CO/^{13}CO abundance ratios (see the shoulder on the blue-ward of the central feature in panel CO2.35 of Fig. 1). Consequently, the first-overtone bands of CO can be used to estimate the ^{12}CO/^{13}CO isotope ratio providing, in principle, important clues about stellar evolution and nucleosynthesis.

Downloaded from https://academic.oup.com/mnras/article/512/1/378/6534922 by University of Central Lancashire user on 24 March 2022

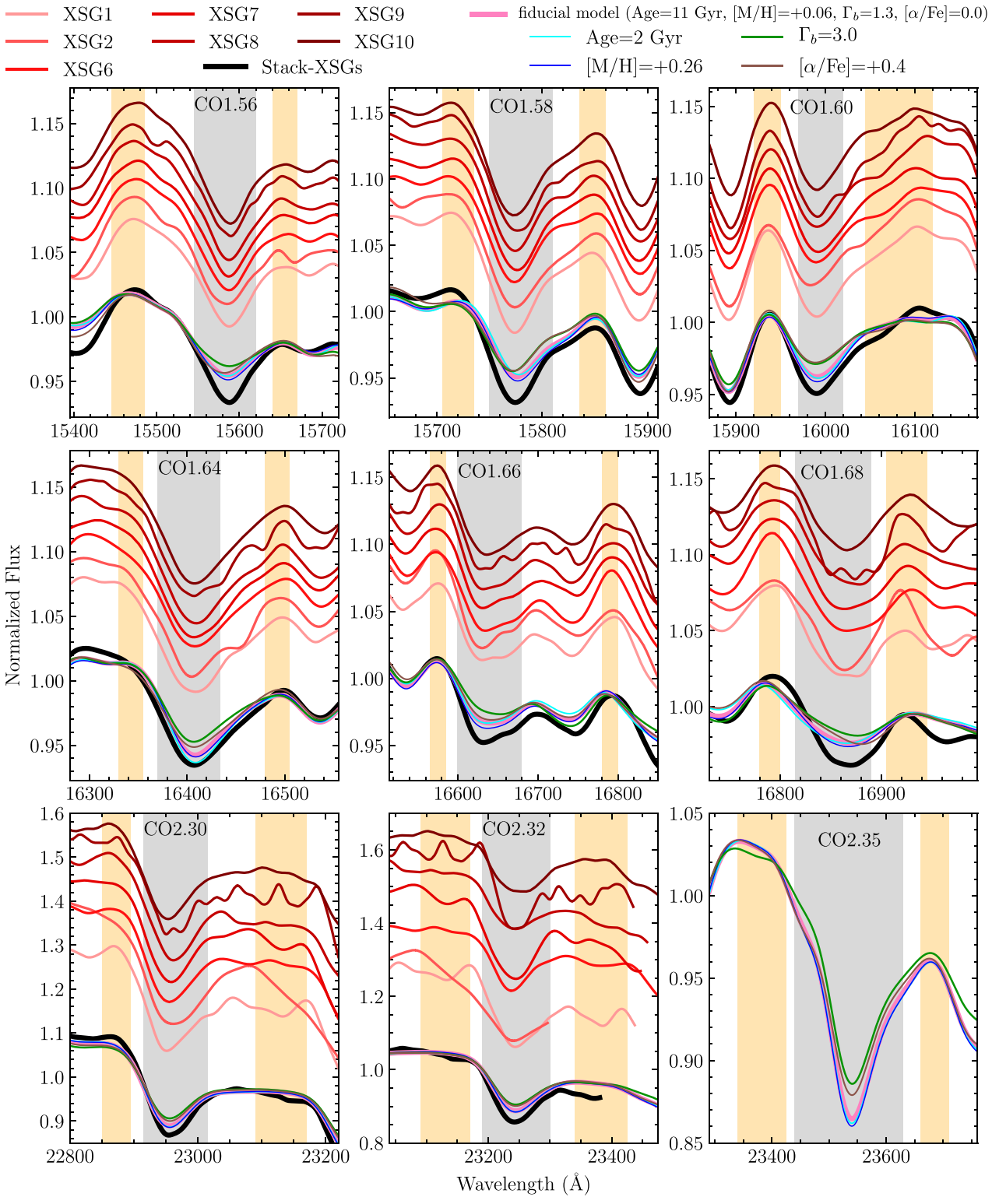

Figure 1. Spectra of XSGs and E-MILES models in the regions of CO features: individual XSG spectra (light to dark red); the stacked spectrum of XSGs (black); a 'fiducial model' with solar abundance and metallicity, an age of 11 Gyr, and MW-like IMF (pink); a model with Age = 2 Gyr (cyan); a model with [M/H] = +0.26 (violet), a model with a bottom-heavy IMF of $\Gamma_b = 3.0$ (green); and a model with [α/Fe] = +0.4 (brown). All these models have the same stellar population parameters as the fiducial model except for a given parameter (see labels on the top). The model and observed spectra have been convolved to a common resolution of $\sigma = 360\,\mathrm{km\,s^{-1}}$, as this is the highest velocity dispersion in the sample of LB19. The central bandpasses of CO indices, as well as the blue and red pseudo-continua, are from Eftekhari et al. (2021) and are shown as grey and orange areas, respectively. All spectra have been normalized to the mean flux within pseudo-continua bands. The XSG individual spectra have also been arbitrarily shifted to display galaxy-to-galaxy variations in the depth of the COs. Note that the spectral range of XSGs does not cover the CO2.35 index but this index is included in the figure as it is used for other galaxy samples. Remarkably, for all CO features, galaxies show stronger absorption than the models, regardless of the adopted model parameters.

Downloaded from https://academic.oup.com/mnras/article/512/1/378/6534922 by University of Central Lancashire user on 24 March 2022

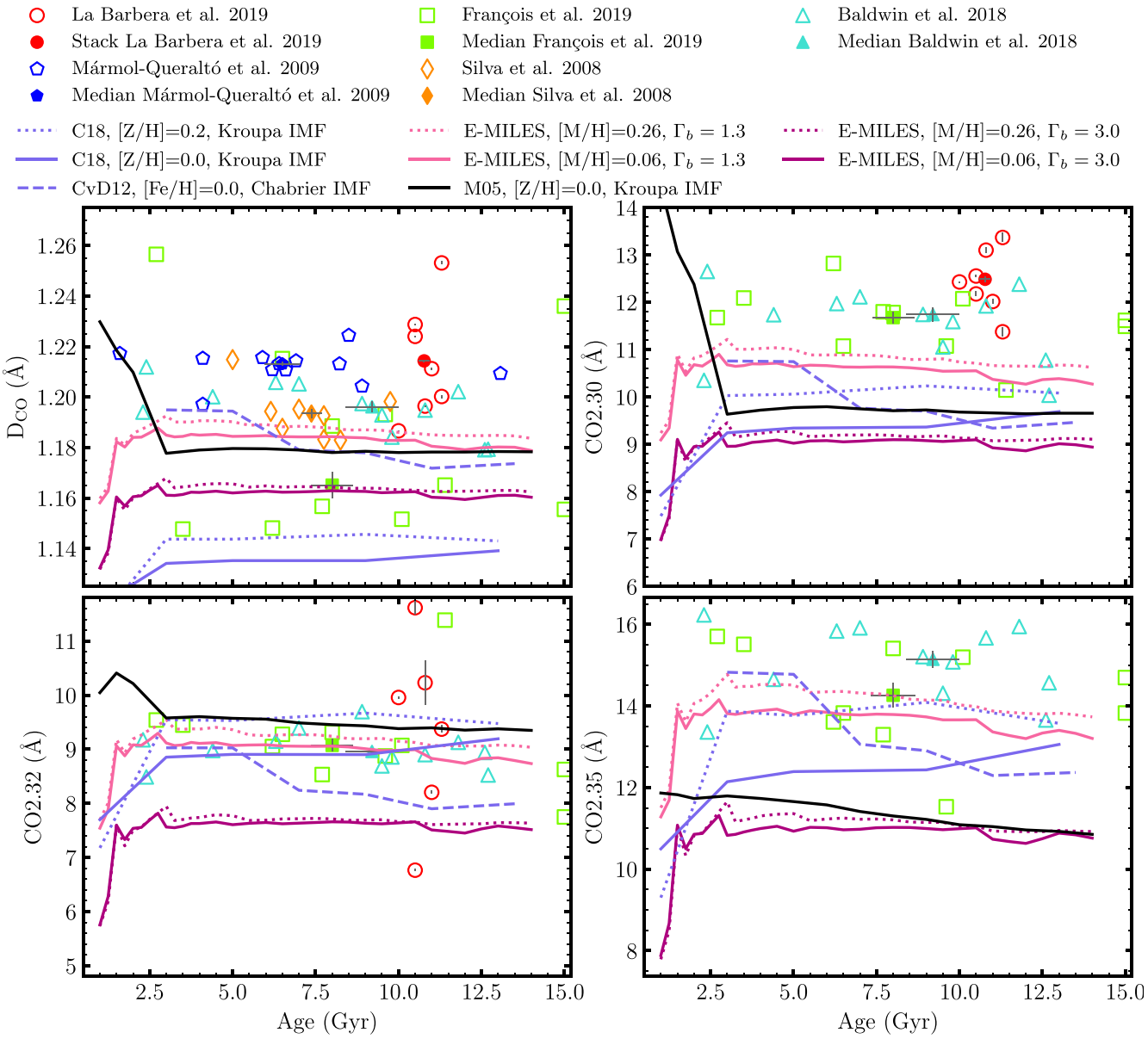

Figure 2. CO indices measured from different galaxy samples, with open (filled) lime squares, cyan triangles, blue pentagons, and orange diamonds corresponding to individual (median) values for galaxies from François et al. (2019), Baldwin et al. (2018), Mármol-Queraltó et al. (2008), and Silva et al. (2008), respectively (see labels on the top). The red open (filled) circles are line-strengths for individual spectra (median-stacked spectrum) of the XSGs (LB19). Observed line-strengths are compared to CO indices measured on different SPS spectra: two sets of E-MILES models with solar metallicity and bimodal IMF of slopes 1.3, corresponding to an MW-like IMF, and 3.0, representative of the bottom-heavy IMF of massive galaxies (solid pink and purple lines, respectively); two sets of E-MILES models with supersolar metallicity and bimodal IMF slopes of 1.3 and 3.0 (dotted pink and purple lines); one set of CvD12 models with [Fe/H] = 0 and Chabrier IMF (dashed violet line); two sets of C18 models with Kroupa IMF and [Z/H] = 0 and 0.2 (solid and dotted violet lines, respectively); and one set of M05 models with solar metallicity and Kroupa-like IMF (solid black line). The indices are measured on data and models all corrected to the same velocity dispersion of 360 km s^{-1}.

to indices bandpass and pseudo-continua bands, respectively. In the same figure, we also show a fiducial model (pink) corresponding to an E-MILES simple stellar population (SSP) with an age of 11 Gyr (mean age of the XSGs as derived from optical indices), solar chemical abundance pattern, and a bimodal IMF of logarithmic slope 1.3 (corresponding to an MW-like IMF). Note that although the spectra of XSGs do not cover the CO index at 2.35 μm, we also show this index as it is covered by other galaxy samples in our analysis (see Fig. 2). For clarity and ease of comparison, normalized spectra of XSGs are shifted upwards, while the stacked spectrum and model spectra are only normalized to the mean flux within pseudo-continua bands. Clearly, XSG9 is the one with the largest scatter among the spectra, likely because of the lower quality of the data for this galaxy. A clear mismatch between observations and models is seen in all panels. The CO indices of the XSGs are much stronger than those of the fiducial SSP model.

As a first step, we investigate whether a young population, a non-solar metallicity, a dwarf-rich stellar population, and/or a non-solar chemical abundance pattern could significantly affect the CO lines, and explain the deep CO absorptions in the data:

Downloaded from https://academic.oup.com/mnras/article/512/1/378/6534922 by University of Central Lancashire user on 24 March 2022

(i) As mentioned in Section 1, the disagreement between *K*-band CO observations and models has been attributed to the presence of intermediate-age stellar components, dominated by stars in the AGB evolutionary phase. To test this scenario, in Fig. 1, we show an E-MILES SSP with the same parameters as the fiducial model but with an age of 2 Gyr (see cyan curves). Except for the CO1.64 index, a variation in age does not significantly change the depth of CO absorption features. We assess this issue, in more detail, in Section 5.2.

(ii) Since XSGs have metal-rich stellar populations, as shown by the analysis of optical spectral indices (see LB19), in the figure we also show the effect of increasing metallicity, with an SSP having the same parameters as the fiducial model but [M/H] = +0.26 (violet). According to E-MILES models, the increase in metallicity does not significantly affect the depth of (all) CO features.

(iii) The strong CO absorptions cannot be explained by IMF variations either; an SSP model with the same fiducial model parameters but with a steeper IMF slope (green) has shallower CO absorptions, hence worsening the fitting. This was first pointed out by Faber (1972), who showed that an increase in the number of dwarf stars drives the CO index at 22 800 Å to unacceptably low values. Fig. 1 shows that, indeed, all CO features exhibit a similar behaviour.

(iv) In Fig. 1, we also investigate non-solar [α/Fe] abundance effects. An SSP that differs from the fiducial model only in α-enhancement is shown with brown colour. As for the IMF, the enhancement in α weakens the strength of CO absorptions. Notice that a different result seems to hold for A-LIST SPS models (Ashok et al. 2021), which suggest an increase of CO strength with [α/Fe] (see their fig. 6a).

4.2 CO line-strengths in *K* band

Fig. 2 shows a quantitative comparison of line-strengths of CO indices in *K* band between data and different SPS models (see Section 3.1), i.e. E-MILES (solid/dotted pink and purple lines), CvD12 (dashed violet line), C18 (solid and dotted violet lines), and M05 (solid black line) models. We plot model line-strengths of the COs as a function of age, while for galaxies we plot observed line-strengths as a function of the age, as estimated from previous works (see Section 2). For each index, the measurements on the individual XSG spectra are plotted with open red circles, while the filled red circle corresponds to the measurement of their stacked spectrum. Individual galaxies from François et al. (2019), Baldwin et al. (2018), Mármol-Queraltó et al. (2009), and Silva et al. (2008) samples (hereafter F19, B18, M09, and S08, respectively) are shown with open lime squares, cyan triangles, blue pentagons, and orange diamonds, respectively. For each index, the median CO line-strength of each sample is shown with a filled symbol of the same colour. The indices were measured after smoothing all spectra to a common velocity dispersion of $\sigma = 360\,\mathrm{km\,s^{-1}}$.

4.2.1 D_{CO} *versus age*

In the upper left panel of Fig. 2, we consider the D_{CO} index, i.e. the definition of the first CO bandhead in *K* band from Mármol-Queraltó et al. (2008). This index is defined with two blue pseudo-continua and the absorption bandpass (see Mármol-Queraltó et al. 2008 for details). In the same panel, we also included measurements for the spectra of S08 and M09 (open/filled orange diamonds and blue pentagons). For ages greater than 3 Gyr, different models show similar trends, with D_{CO} showing no significant variation with age. Only for ages younger than 3 Gyr, there is a significant difference between E-MILES and M05 models since the contribution of AGB stars is more emphasized in the young populations of M05. We further discuss this issue in Section 5.2. All models underpredict the median value of D_{CO} for the samples except for that of F19. However, since the scatter of D_{CO} in this sample is far larger than that for the other samples, no firm conclusions can be drawn. In general, Fig. 2 shows that the models can barely reproduce only the galaxies with the smallest D_{CO} values. For instance, two galaxies of B18 at 13 Gyr are well matched with E-MILES and M05 models, and the same applies to two galaxies in the S08 sample with the weakest CO absorption, which are well matched with E-MILES models (see solid pink line and the orange diamonds for an age of ∼8 Gyr). Since the M05 model differs significantly from E-MILES model in the predictions for young populations, it can match the youngest galaxies of M09 and B18. Overall, for the D_{CO} index, the mismatch between observations and models applies to all models. C18 models predict the lowest values for D_{CO} and cannot match any data points.

4.2.2 D_{CO} *versus metallicity*

Another parameter that can be considered is the variation of metallicity among our galaxies, which span a wide range, from [M/H] = −0.4 to [M/H] = +0.5 dex. We investigate the effect of metallicity by showing the predictions of E-MILES models with an MW-like IMF ($\Gamma_b = 1.3$) and total metallicity of +0.26 dex (dotted pink line), and predictions of C18 models with a Kroupa IMF and [Z/H] = 0.2 dex (dotted violet line). Hence, the effect of metallicity can be seen by comparing dotted and solid lines in Fig. 2. The figure shows that the discrepancy between the dotted pink line and the filled blue pentagon and red circle (orange diamond) is almost eight (2) times larger than the increase in D_{CO} caused by variations in [M/H] from +0.06 to +0.26 dex. However, it is worth noticing that the IRTF stellar library, which is used to construct E-MILES models in the NIR, consists of stars in the solar neighbourhood, which are unavoidably biased towards solar metallicity. In fact, according to Röck et al. (2015), the quality of E-MILES models decreases at supersolar metallicities. Also, we cannot exclude a non-linear behaviour of CO absorptions in the very high-metallicity regime. Therefore, we looked at C18 models, which are based on the Extended IRTF Stellar Library with better coverage in metallicity (although this advantage mainly applies to the metal-poor regime). Comparing the difference between dotted and solid pink lines with the difference between dotted and solid violet lines shows that the effect of metallicity in C18 models is larger than the one in E-MILES models but it is not enough to match the high values of XSGs or M09 data. Also, notice that although C18 models are based on a stellar library with better coverage in metallicity, they predict the lowest values for D_{CO}, hampering the discrepancy to the observed data points. We conclude that the mismatch between data and models cannot be explained with metallicity variations alone.

4.2.3 D_{CO} *versus IMF*

We also investigated the effects of a bottom-heavy IMF ($\Gamma_b = 3.0$) on the CO features as shown by solid- and dotted purple lines for solar and metal-rich populations, respectively. The IMF slope of XSGs has been determined by LB19, using a combination of optical and NIR (Na) indices, finding that all galaxies have a bottom-heavy IMF in the centre. However, Fig. 2 shows that a dwarf-rich population does

Downloaded from https://academic.oup.com/mnras/article/512/1/378/6534922 by University of Central Lancashire user on 24 March 2022

significantly increase the discrepancy between models and observed CO indices. While this result seems to be consistent with Alton, Smith & Lucey (2018), who claimed an MW-like IMF for massive galaxies in their sample, based on *J*- and *K*-band spectral indices (including two CO bandheads in *K* band), it has to be noted that for an MW-like IMF the models do not match the observations either. In other words, any claim from CO lines on the IMF should be taken with caution.

4.2.4 D_{CO} versus environment

The samples of S08 and M09 allow us to assess the effect of galactic environment on the strength of D_{CO}, as galaxies in the former sample reside in a high-density environment (Fornax cluster) compared to the latter, which consists of field galaxies.

It is noteworthy to mention that within 20 Mpc, the Fornax cluster is the closest and second most massive galaxy cluster after Virgo. It has a virial mass of $10^{13} M_{\odot}$ (Drinkwater, Gregg & Colless 2001) and while most of its bright members are ETGs, mainly located in its core (Grillmair et al. 1994), its mass assembly is still ongoing (Scharf, Zurek & Bureau 2005), and therefore it is not fully virialized. Fornax cluster is an evolved, yet active environment, as well as a rich reservoir for studying the evolution of galaxies in a cluster environment, particularly within its virial radius.

The two samples of S08 and M09 have been observed with the same telescope and observational set-up (see Section 2), allowing for a direct comparison. Note that these two samples are also comparable with respect to velocity dispersion (see Section 2). As shown in the plot, M09 galaxies, located in the field, tend to have larger values of D_{CO} strength than S08 galaxies in the Fornax cluster. We see that the median value of these two samples cannot be matched with the current models. We can speculate that the origin of the difference between D_{CO} values of ETGs in low- and high-density environments might be due to a difference in the carbon abundance of field and cluster galaxies. Since star formation in dense environments takes place more rapidly than in isolated galaxies, carbon, which is expelled into the interstellar medium by dying stars of intermediate masses, cannot be incorporated in newer stellar generations. Therefore, the resulting stars in dense environments, like cluster galaxies, exhibit smaller carbon abundance with respect to their counterparts of similar mass in low-density environments. As the CO molecule has high binding energy, carbon mostly forms CO molecules. Thus, CO indices in the field galaxies are stronger than cluster galaxies as was suggested by M09. Moreover, Röck et al. (2017) found a dichotomy between the Na I2.21 values of the ETGs in the S08 and M09 samples. Indeed, they showed that one possible driver of NaI2.21 might be [C/Fe] abundance.

4.2.5 *Other* K*-band CO indices*

The upper right panel of Fig. 2 shows measurements for the same CO feature as for D_{CO}, but using the index definition, named CO2.30, from Eftekhari et al. (2021). While D_{CO} measures the absorption at ∼2.30 μm as a generic discontinuity, defined as the ratio between the average fluxes in the pseudo-continua and in the absorption bands, the CO2.30 index follows a Lick-style definition, with a blue and red pseudo-continua and the absorption bandpass (see fig. A3 in Eftekhari et al. 2021 for a comparison of the two definitions). Note that the red bandpass of CO2.30 is not covered by the S08 and M09 spectra and therefore these samples are not included in the upper right panel of Fig. 2. For ages of 3 and 5 Gyr, the predictions of CvD12 and E-MILES models for CO2.30 are very similar, while for older ages, CvD12 models are closer to the M05 predictions, leading to a larger mismatch. Also, C18 models with solar metallicity (solid purple line) have smaller difference with M05 models than E-MILES ones for ages older than 3 Gyr and their trend is very similar to E-MILES and M05 models, in contrast to CvD12. The difference between E-MILES, CvD12, C18, and M05 models for old ages (>5 Gyr) is quite significant, and comparable to the effect of varying the IMF slope (see pink and purple lines in the figure). However, similarly to D_{CO}, all models fail to match the observed strong CO2.30 line-strengths, and changing [M/H] of E-MILES models from +0.06 to +0.26 increases CO2.30 by only ∼0.2 Å, while the discrepancy between the pink line and the filled lime square and cyan triangle (red circle) is ∼1(2) Å. Moreover, although increasing the overall metallicity of C18 by 0.2 dex leads to an increase of ∼1 Å, the C18 models predict the lowest values for CO2.30 and cannot match the data. Even by considering only the relative changes and adding the effect of metallicity predicted by C18 to the E-MILES models, the models are not still able to match the high value of the stacked spectrum of XSGs. Hence, using the CO2.30 index, we end up with the same conclusions as for D_{CO}, i.e. an intrinsic offset exists between models and data, which is independent of the index definition.

In the lower panels of Fig. 2, we also show measurements for other two CO bandheads in *K* band, namely CO2.32 and CO2.35, which have been far less studied compared to the CO feature at ∼2.30 μm. Note that the XSGs spectra do not cover the wavelength limits of the CO2.35 index, and thus this sample is not included in the panel showing this index. Also, as can be seen in Fig. 1, the spectra of one XSG, and for the XSG stack, do not encompass the red bandpass of CO2.32. Hence, the corresponding measurements are not seen in the CO2.32 versus age panel. Unlike CO2.30, for CO2.32, the M05 models predict the highest line-strengths among all SPS models with solar metallicity (∼0.5 Å higher than E-MILES and C18 models with a Kroupa IMF). E-MILES models with solar metallicity and MW-like IMF match well the median CO2.32 value of the F19 and B18 samples, and C18 models with solar metallicity are very close to F19 and match well to B18, while the scatter of XSGs is large, likely because the feature is at the edge of the available spectral range for these galaxies, making all models compatible with them. However, it should be noted that galaxies in these samples span a wide range in velocity dispersion, and at the highest σ, galaxies should be better described by a bottom-heavy IMF (see e.g. La Barbera et al. 2013). The latter is particularly true for the XSGs, with $\sigma \gtrsim 300\,\mathrm{km\,s^{-1}}$. It is expected that MW-like IMF models (pink lines) describe low-σ galaxies, while bottom-heavy IMF models (purple lines) match high-σ galaxies, in particular the XSGs (but see Alton et al. 2018). However, predictions for a bottom-heavy IMF (purple lines) fall below the observed line-strengths for CO2.32, for all galaxies (except for one XSG). Thus, the mismatch between observations and models seems to be in place also for the CO2.32 index. Also, there is a closer similarity between the trend of C18 models and E-MILES and M05 models than CvD12 ones. Another interesting point is that C18 models with [Z/H] = 0.2 dex overpredict the mean CO2.32 line-strengths of F19 and B18 samples. Since the IMF of the most massive galaxies in F19 should be bottom-heavy, one might consider the effect of supersolar metallicity (from C18) and bottom-heavy IMF (from E-MILES) at the same time. In this case, the discrepancy between the C18 models and F19 sample gets even worse as the effect of IMF is larger than the effect of metallicity. The comparison for CO2.35 is similar to that for CO2.30, with solar and MW-like IMF models underpredicting the median values for F19 and B18 samples. However, the metal-rich E-MILES models (dotted pink line) can

Downloaded from https://academic.oup.com/mnras/article/512/1/378/6534922 by University of Central Lancashire user on 24 March 2022

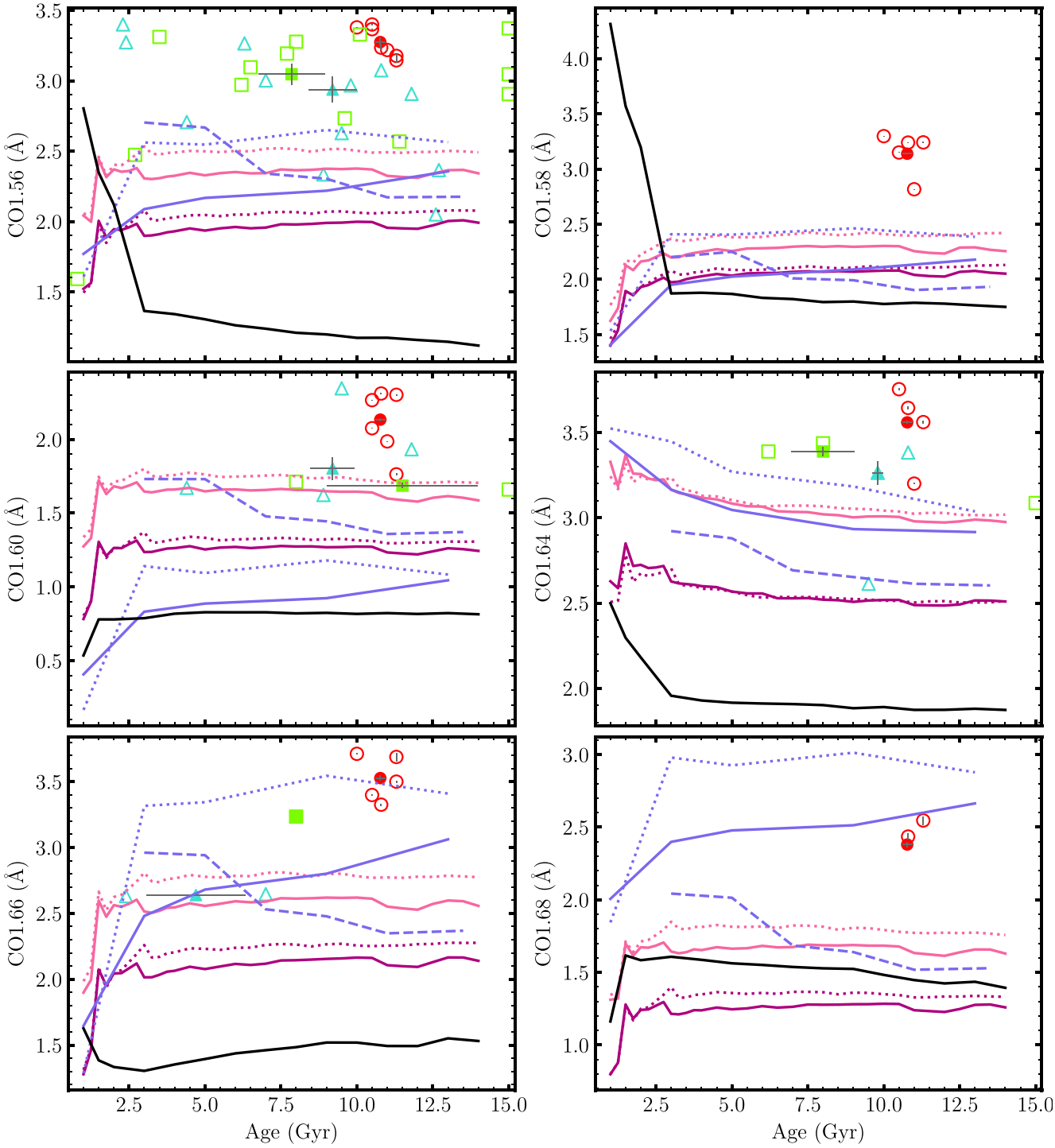

Figure 3. The same as in Fig. 2 but for *H*-band CO indices.

reproduce the median value of the F19 galaxies (filled lime square), but as already pointed out, the most massive galaxies in this sample are expected to have a steeper IMF, while bottom-heavy models fall below all the observed data-points for CO2.35 (the bottom-heavy E-MILES model predicts a CO2.35 of ∼11 Å, while the lowest value of CO2.35 in the F19 sample is ∼11.5 Å). We conclude that, in general, the disagreement between observations and models is present in all the *K*-band CO indices.

4.3 *H*-band CO indices

In order to assess whether the mismatch of observed and model CO lines is intrinsic to the *K* band, or whether it is a general issue in the NIR, we measured a whole battery of CO absorptions that populate the *H*-band spectral range (see Fig. 1). Fig. 3 shows the same comparison as in Fig. 2 but for the *H*-band lines. Note that for CO1.58 and CO1.68, the spectra of F19 and B18 are severely contaminated by sky, and thus we do not show the corresponding

Downloaded from https://academic.oup.com/mnras/article/512/1/378/6534922 by University of Central Lancashire user on 24 March 2022

line-strengths. For the same reason, only two XSGs are shown for CO1.68, and only a few galaxies are plotted in the panels for CO1.60, CO1.64, and CO1.66.

Remarkably, Fig. 3 shows that (i) LB19 galaxies have lower scatter in all plots compared to the CO indices in *K* band, most likely due to the high-quality of these data in *H* band, and (ii) these very massive ETGs show very high CO values with respect to the model predictions for all indices. The discrepancy between models and observations for *H*-band CO indices is similar to that found in *K* band. In particular, the median stacked spectrum of the XSGs shows *H*-band CO values ∼1.3 times larger than E-MILES models with MW-like IMF and solar metallicity.

For CO1.56, the offsets between the median values of F19 and B18 samples and the reference E-MILES model (pink line) are ∼0.7 and ∼0.6 Å. E-MILES models can reproduce the median value of F19 and B18 galaxies for CO1.60 and CO1.66, respectively. However, these indices have been computed only for two galaxies in either samples, and thus we are not able to draw any firm conclusion.

Although the updated version of CvD12 models with solar metallicity (solid violet lines) tend to increase slightly for ages older than 3 Gyr (except for CO1.64), the behaviour of supersolar C18 models (dotted violet lines) is more similar to the E-MILES models and in case of CO1.58 it even matches with the E-MILES (dotted pink and violet line). For CO1.56 and CO1.64, C18 models with [Z/H] = 0.2 dex predict the highest values among all models but still they cannot match the median values of galaxies. The mean value of CO1.66 index in the B18 sample is well fitted by a solar metallicity C18 SSP and interestingly, a C18 SSP with supersolar metallicity matches the CO1.66 index value of stacked XSGs well. This result is at first glance surprising, however, one should bear in mind that XSGs have bottom-heavy IMFs and according to E-MILES models this makes the predictions of the CO values lower. In the CO1.68 panel, surprisingly, C18 models overpredict the line-strengths of the XSGs. In general, E-MILES and CvD12 models are more self-consistent than C18 and M05 models as their deviation with respect to data is similar for all CO indices.

In all panels of Fig. 3, the M05 models predict the lowest CO strengths, compared to other models. For CO1.56, CO1.58, and CO1.64, M05 models predict a strong increase at ages younger than 3 Gyr, similar to what is found for CO2.30 in *K* band, while for other indices the opposite behaviour is seen (e.g. CO1.60 and CO1.68). On the contrary, for all CO indices, the XSGs data show, consistently, line-strengths significantly above the model predictions for old ages (except for CO1.66 in which the supersolar metallicity C18 model matches the stacked spectrum and for CO1.68 in which C18 models overpredict the line-strengths of the XSGs). Again, this points against a scenario whereby the CO line-strengths are accounted for by young stellar populations with an AGB-enhanced contribution such as in M05 models. As for the *K* band, CvD12 models show a trend for all CO indices to decrease with increasing age, while E-MILES models exhibit a nearly flat behaviour and C18 models a slightly increasing behaviour. For instance, in case of CO1.58, CO1.66, and CO1.68, for ages younger than 5 Gyr, CvD12 models predict ∼0.4 Å stronger line-strengths than E-MILES models (pink line), while the two models agree for populations with an age of ∼7 Gyr. For older ages, CvD12 models always predict lower CO index values than E-MILES models.

The effect of a bottom-heavy IMF in Fig. 3 is shown by the purple lines, corresponding to E-MILES models with a bimodal IMF slope of $\Gamma_b = 3.0$. Similarly to the *K* band, a bottom-heavy IMF leads to significantly shallower CO line-strengths, than those for a standard stellar distribution. Note also that for most indices, the discrepancy between MW-like IMF models and the XSG stack (filled red circle) is larger than the variations due to a change in IMF slope. In the case of CO1.56 and CO1.58, the discrepancy is about twice the difference between models with different IMF. As already quoted in Section 4.2.3, we emphasize that while a bottom-heavy IMF hampers the offset between model and observations, we are not able to match the CO line-strengths with a standard IMF (except for CO1.66 with supersolar metallicity C18 model). In other words, the CO-strong feature should not be interpreted as evidence against a varying IMF in (massive) galaxies.

From Figs 2 and 3, a very consistent picture emerges. CO lines throughout the *H* and *K* bands are stronger than the predictions of all the considered state-of-the-art SPS models with varying age, metallicity, and IMF.[6] Indeed, in order to reconcile models and observations, additional stellar population parameters should be taken into account, as we detail in the following sections.

5 EFFECTS OF VARYING OTHER STELLAR POPULATION PARAMETERS

To gain insight into the origin of the discrepancy between observed NIR CO features and model predictions, we scrutinize the effect of varying abundance ratios in the models (Section 5.1), as well as that of an enhanced contribution from AGB stars (Section 5.2). The main results of this analysis are shown in Figs 4 and 5, for *K*- and *H*-band indices, respectively. The figures are the same as Figs 2 and 3, but showing only median values of line-strengths for different samples, as well as CO indices for the XSG stack. To avoid crowding the figure, only E-MILES models are plotted, with different arrows showing the effect of varying different parameters in the models, as detailed below. Also note that in Fig. 4, we do not include the panel for D_{CO} (as in Fig. 2), as it does not add any further information with respect to the CO2.30 index, whose Lick-style definition is more similar to that of the other CO indices.

5.1 Abundance ratios

So far, we have considered only models constructed with stars following the chemical pattern of the solar neighbourhood. However, differences in the depth of CO features between models and data might also arise because of variations in [α/Fe], or other elemental abundances, with respect to a scaled-solar composition. According to Eftekhari et al. 2021 (see column 'd' of their figs 4 and 6), the maximum change in the strength of CO indices due to elemental abundance variations comes from carbon and α-enhancements. Therefore, in Figs 4 and 5, we show the effect of α-enhancement, based on α-enhanced E-MILES models with an age of 13 Gyr (brown arrows), and that of an enhancement in carbon abundance, based on CvD12 models with an age of 13.5 Gyr and a Chabrier IMF (see the violet arrows). The violet and brown arrows correspond to variations of δ[C/Fe] = +0.15 dex and δ[α/Fe] = +0.4 dex, respectively, which can be representative of massive ETGs, such as those of the XSG sample (see e.g. La Barbera et al. 2017). Note that we also have used the updated version of CvD12 models to see the effect of C-enhancement. Since the size of the violet arrow did not change with respect to that of CvD12, we do not include it in the plots to avoid crowding the figure. Moreover, since we are interested in the relative

[6]At this point, we remind the reader that it is yet unclear why some indices, e.g. CO2.30, increase significantly for young populations in M05 models, while they decrease in other models.

Downloaded from https://academic.oup.com/mnras/article/512/1/378/6534922 by University of Central Lancashire user on 24 March 2022

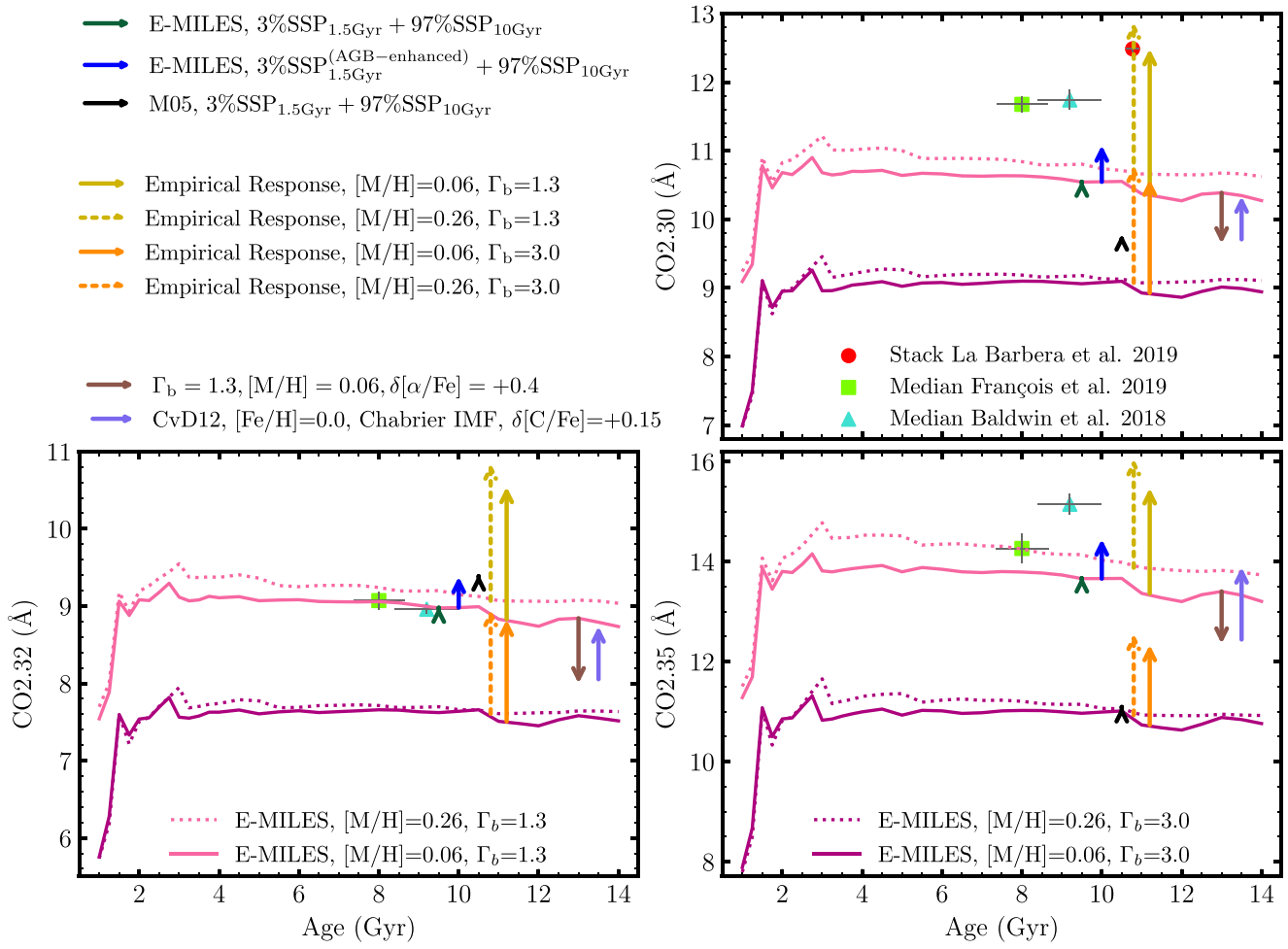

Figure 4. CO indices in *K* band, as measured from different galaxy samples. The filled lime squares and cyan triangles correspond to median values for galaxies from F19 and B18, respectively. The red circles are measurements for the XSG stack (see the text). Observed COs are compared to predictions of E-MILES SPS models: solid (dotted) pink and purple lines are models with solar (supersolar) metallicity and bimodal IMF of slopes 1.3 and 3.0, corresponding to an MW-like and bottom-heavy IMF, respectively. The (small) green arrows indicate the effect of adding a fraction of an intermediate-age E-MILES SSP on top of an old SSP (see the text); the blue arrow is the same but with an enhanced contribution of AGB stars for the intermediate-age component. The black arrow is the equivalent of the green arrow for the M05 models. The solid and dotted khaki (orange) arrows mark the effect of the 'empirical corrections' to E-MILES SSPs with solar and supersolar metallicities, respectively, for an MW-like (bottom-heavy) IMF. The brown and violet arrows show the effect of [α/Fe] and [C/Fe] abundance ratios on CO indices. The indices are measured after smoothing all data and models to a velocity dispersion of 360 km s^{-1}.

response of models to variations of [C/Fe] and [α/Fe], we shifted the starting point of the violet arrow to the end point of the brown arrow.

As expected, increasing [C/Fe] results in stronger CO lines in all cases, pointing in the right direction to match the data. However, the effect is counteracted by that of α-enhancement (except for CO1.56, where the effect of α-enhancement is negligible), so that, for all the *K*-band CO indices, the brown and violet arrows tend to cancel each other for the adopted abundance variations in the CvD12 models. Hence, the picture emerging from Fig. 4 is similar to that of Fig. 2. For CO2.30 and CO2.35, even considering the effect of varying abundance ratios, the models are not able to match the observations, especially in the case of a bottom-heavy IMF. For CO2.32, the median index values for the F19 and B18 samples can be matched, but only with models having an MW-like IMF.

For the *H*-band CO line-strengths (Fig. 5), carbon abundance has a larger effect than [α/Fe], compared to *K*-band indices, i.e. the relative size of violet versus brown arrows in Fig. 5 is larger than in Fig. 4. This shows, once again, the importance of studying lines from the same chemical species (CO) at different wavelengths (*H* and *K*). Even so, summing up the violet and brown arrows in Fig. 5 does not allow us to reach the high CO values of massive XSGs. For instance, in the case of CO1.58 and CO1.60, summing up the effect of [C/Fe] and [α/Fe] would result in a (modest) increase of δ(CO1.58) ∼0.1 Å and δ(CO1.60) ∼ 0.2 Å, respectively, these variations being far smaller than the deviations between MW-like IMF E-MILES models and the XSGs stack, corresponding to ∼0.9 Å and ∼0.5 Å for CO1.58 and CO1.60, respectively. Note also that indices, such as CO1.56, for which the effect of abundance ratios is smaller (compared to e.g. the effect of varying the IMF) do not show smaller deviations of data compared to models. In other terms, also a qualitative comparison of data and model predictions seems to point against the effect of abundance ratios as the main culprit of the CO mismatch problem. However, one should bear in mind that the effect of abundance ratios on SSP models relies completely on theoretical stellar spectra, as well as on molecular/atomic line lists, that are notoriously affected by a number of uncertainties, particularly in scarcely explored spectral regions, such as the NIR. Hence, we cannot exclude that the effect of [α/Fe] and [C/Fe] is underestimated by current SPS models.

Downloaded from https://academic.oup.com/mnras/article/512/1/378/6534922 by University of Central Lancashire user on 24 March 2022

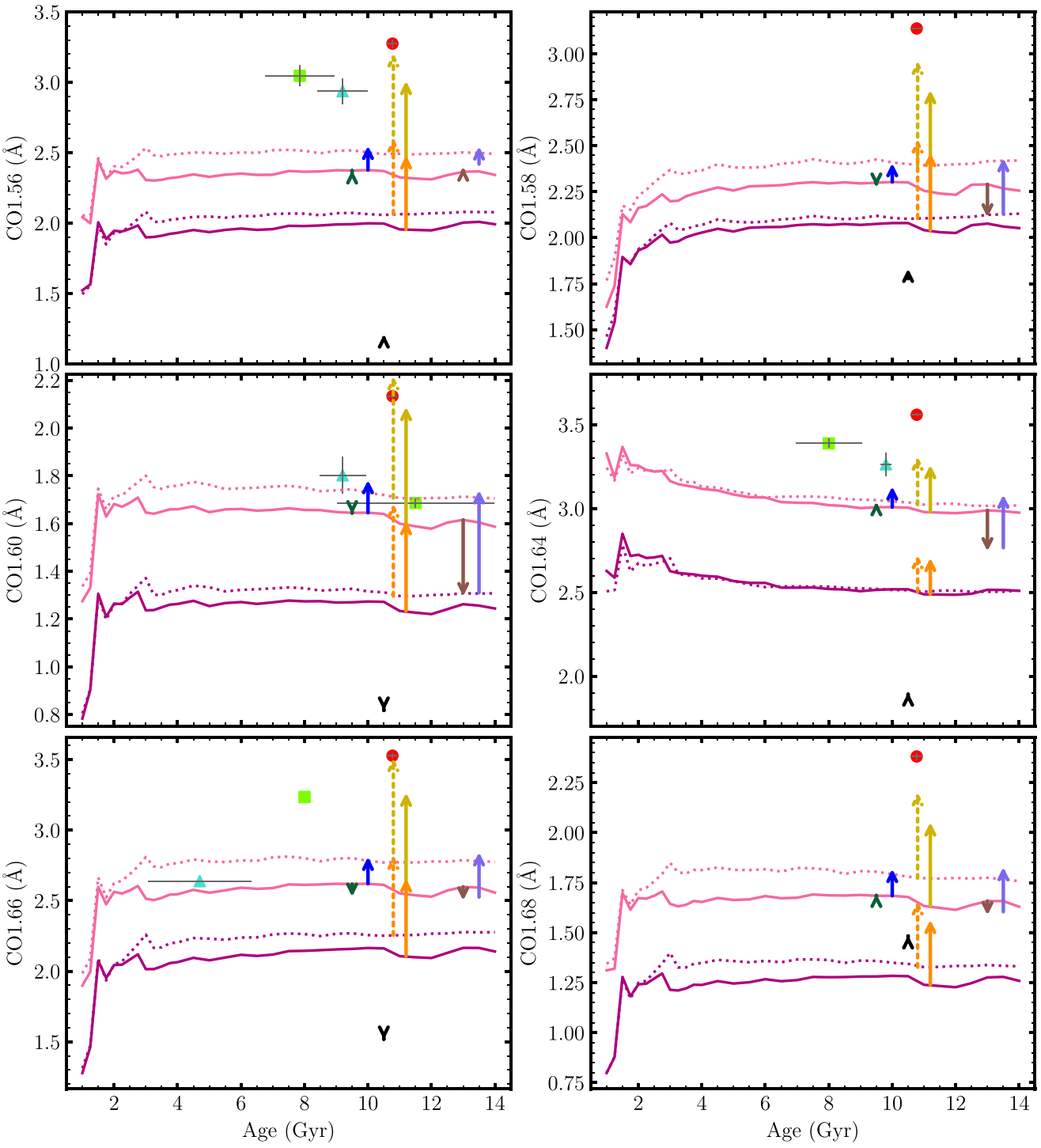

Figure 5. Same as Fig. 4 but for *H*-band indices.

Alternatively, one should seek for other possible explanations, as discussed in the following section.

5.2 Intermediate-age stellar populations

Since stars in AGB phase contribute most to the *K*-band luminosity of stellar populations with ages between 0.3 and 2 Gyr (∼30 per cent in E-MILES models, with M05 models having the largest contribution from AGB stars among other models, i.e. >70 per cent), it has been suggested that the deep CO band-heads of ETGs in the *K* band are due to the presence of AGB stars, from intermediate-age stellar populations (e.g. Mobasher & James 1996, 2000; James & Mobasher 1999; Davidge, Beck & McGregor 2008; M09). It is important to assess, in a quantitative manner, if this hypothesis can account for the observed high CO line-strength values. To this effect, we first fit the observed XSG stacked spectrum with different SSP models, assuming a non-parametric star formation history (SFH), and then simulate the effect of

Downloaded from https://academic.oup.com/mnras/article/512/1/378/6534922 by University of Central Lancashire user on 24 March 2022

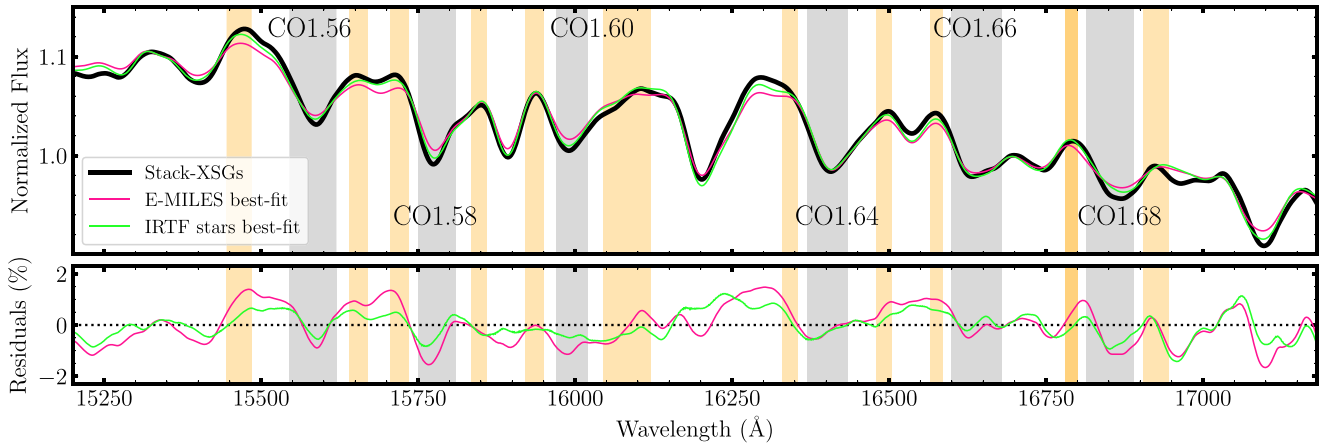

Figure 6. The upper panel shows the stacked spectrum of XSGs (black) in *H* band, together with the best-fitting spectra obtained with E-MILES SSPs (pink) and IRTF stars (lime), running the spectral fitting code STARLIGHT (see the text). CO features, with central absorptions and their pseudo-continua, are plotted as grey and orange shaded areas. All spectra have been normalized to their median flux. The lower panel shows the relative residuals of the stack with respect to each best-fitting spectrum. Notice the improvement in the matching of CO features when using IRTF stars, despite the fact that only four stars received a non-negligible weight in the STARLIGHT best-fitting mixture (see the text).

intermediate-age populations by constructing ad hoc two-component models.

To fit the XSG stack, we use the software STARLIGHT (Cid Fernandes et al. 2005), a full spectral fitting code that allows us to fit a galaxy spectrum with a generic linear combination of input model spectra, i.e. the so-called 'base' spectra. First, we use scaled-solar E-MILES SSPs as a base, including models with different ages and metallicities, and a Kroupa-like IMF.[7] Note that this approach does not make any assumption about the SFH, which is treated in a non-parametric way. Hence, the effect of young populations is taken into account in the most general manner, without any restriction from the optical range. The STARLIGHT fitting was carried out in the *H* band, as CO absorptions dominate this spectral range. Fig. 6 compares the stacked spectrum of XSGs (black line), with the best-fitting composite stellar population model of E-MILES SSPs (pink line). The best-fitting model shows deviations at a level of ∼2 per cent from the stacked spectrum in the region of the CO bandpasses. Note that this is similar to what was found when comparing individual E-MILES SSPs to the XSGs' stack (see the fiducial model, plotted as a pink line, in Fig. 1). Since in the STARLIGHT fits, there is no constraint on the age of the best-fitting SSPs, these results show that young populations do not help to resolve the tension between observations of CO lines and model predictions.

Based on near-ultraviolet (NUV) photometric data, Yi et al. (2005) found that roughly 15 per cent of bright ETGs at $z < 0.13$ show signs of young ($\lesssim$1 Gyr) populations at the level of 1 per cent–2 per cent mass fractions. Also, Schiavon (2007) generated two-component models, showing that a mass fraction of the young component of ∼0.5 per cent–1 per cent provides a reasonably good match to the blue indices of nearby ETGs. This result has been recently confirmed, based on a combination of NUV and optical absorption lines for the XSGs, by Salvador-Rusiñol et al. (2021), who found that the centre of massive ETGs are populated by a 0.7 per cent mass fraction of stars formed within the last 1 Gyr.

[7]Including bottom-heavy SSPs does not improve the STARLIGHT fits significantly, as expected by the fact that CO line-strengths get weaker for $\Gamma_b = 3$ (see Section 4).

To test this scenario, we contaminated the light of an old (10 Gyr) E-MILES SSP by a small fraction (3 per cent in mass) of an intermediate-age (1.5 Gyr) E-MILES SSP. The effect is shown for a solar metallicity and MW-like IMF population by the green arrows at ∼10 Gyr in Figs 4 and 5. Indeed, the arrows are small for all CO indices and, for CO1.58, CO1.60, and CO1.66, they also point in the 'wrong' direction, i.e. that of decreasing (rather than increasing) model line-strengths. However, when considering an E-MILES model with age of 1.5 Gyr, solar metallicity, and an MW-like IMF, AGB stars contribute by approximately one-third to its bolometric luminosity, while such fraction is larger for M05 models. Hence, one might attribute the small effect of the intermediate-age population to the less emphasized contribution of AGB stars to E-MILES, compared to M05, young SSP models. To address this issue, we used the AGB-enhanced version of E-MILES SSPs constructed by Röck et al. (2017). They calculated an AGB-enhanced E-MILES model of 1.5 Gyr, solar metallicity, and Kroupa-like IMF using 'partial SSPs', i.e. computing two SSPs, one by integrating stars along the isochrone without the AGB phase, and the other one by integrating only AGB stars, and combining the two models by assigning 70 per cent luminosity-weight to the model made up of AGB stars only. This synthesized AGB-enhanced stellar population is added on top of an old population of 10 Gyr assuming a 3 per cent mass fraction. The effect on CO line-strengths is shown by the blue arrows at 10 Gyr in Figs 4 and 5. Although the blue arrows are larger than the green ones, they are not large enough to fit the median values for the F19 and B18 samples, except for CO2.35, CO1.60, and CO1.66. In the case of CO2.35 (see Fig. 4), the median value of the F19 sample could be matched with a metal-rich and MW-like IMF E-MILES model, if one assumes that the effects of the blue, violet, and brown arrows (i.e. AGBs + [C/Fe] + [α/Fe]) sum up to ∼0. For CO1.60 and CO1.66, the comparison is limited by the small number of galaxies available for the F19 and B18 samples. As a further test, we added a 3 per cent mass fraction of a 1.5 Gyr M05 SSP with solar metallicity and Kroupa-like IMF to an old (10 Gyr) M05 SSP. The effect is shown by the black arrows at ∼10 Gyr in Figs 4 and 5. The effect of adding an emphasized-AGB intermediate-age population on top of an old one turns out to be negligible, and for the CO1.60 and CO1.66 indices, it goes also into the opposite direction compared to

Downloaded from https://academic.oup.com/mnras/article/512/1/378/6534922 by University of Central Lancashire user on 24 March 2022

Table 1. Properties of four IRTF stars that best fit the XSG stacked spectrum in *H* band. Columns 1 and 2 give the name of each star and its weight in light to the best-fitting STARLIGHT model (see the text). The effective temperature, surface gravity, and metallicity of the stars are given in Columns 3, 4, and 5, respectively.

Star	Weight (%)	T_{eff} (K)	log *g* (dex)	[Fe/H] (dex)
(1)	(2)	(3)	(4)	(5)
HD 219734	43	3730	0.9	0.27
HD 36003	18	4465	4.61	0.09
HD 187238	17	4487	0.8	0.177
HD 10465	14	3781	0.5	− 0.458

the data. Note also that there is no way the AGB-enhanced models can consistently match the strong CO line-strengths of the XSGs, in the *H* and *K* bands.

As a final remark, we emphasize that our analysis does not rule out the presence of intermediate-age populations in ETGs, but it points against a scenario where the observed strong CO absorptions are mainly due to the presence of intermediate-age populations (i.e. AGBs) in these galaxies.

6 AN EMPIRICAL MODELLING APPROACH

6.1 Searching for stars that match the strong CO lines

In order to identify the stars that might be responsible for the strong CO absorption observed in massive ETGs, we fitted the stacked spectrum of XSGs with STARLIGHT (see above), using as an input base all the 180 individual stars of the IRTF library that are used to construct E-MILES models in the NIR. As for the fitting with E-MILES SSPs (see Section 5.2), we fitted only the *H*-band region, where the signal from CO features is prominent with respect to that of other absorptions. The best-fitting mixture of IRTF stars is shown in Fig. 6, as a lime-coloured curve. The relative residuals between observed and model spectrum are shown in the bottom panel of the same figure. By comparing the residuals for the best fit of IRTF stars with that for E-MILES SSPs (pink curve), it can be seen that using the stars improves significantly the fit to the observed spectrum, with residuals in the CO lines at the level of ∼1 per cent, i.e. about half of those for the E-MILES best-fitting model. Although some improvement in the fitting may be actually expected when employing the IRTF stars, given that this band is populated with so many CO absorptions, it is still remarkable how much smaller the obtained residuals are. Indeed, no improvement at all would be achieved in the case where the input stellar library completely lacked those stars responsible for CO absorption. STARLIGHT also returns the weight (in light) of each star in the best-fitting mixture. Surprisingly, we found that only four (out of 180) stars received a significant weight (> 0.5 per cent) in the best-fitting spectrum, namely HD 219734, HD 10465, HD 36003, and HD 187238. The light-weighted contribution of these stars from STARLIGHT and their main stellar parameters from Röck et al. (2016) are summarized in Table 1. HD 36003 is a dwarf star (low-mass star), while the other three stars are giants (massive stars). Hereafter, we refer to these stars as the *H*-band best-fitting stars.

Since the XSG stack is best fit by only four stars, these stars have to be 'special' somehow, and their properties might help up to shed light on the nature of the CO absorptions. To address this point, we measured the line-strength of CO indices for the spectra of all IRTF stars, and marked the position of the *H*-band best-fitting stars in the CO versus effective temperature (T_{eff}) plots in Fig. 7. In this figure, different colours show different types of stars, according to the classification provided in table 2 of Röck et al. (2015), with blue and orange colours corresponding to AGB and M-dwarf stars, respectively. The five carbon stars in the IRTF library are shown in pink colour. Note that this classification is only available for stars cooler than 3900 K. The remaining IRTF stars are plotted with grey colour. Fig. 7 suggests that those stars with $T_{eff} < 5500$ K, which are not classified as carbon stars and M-dwarfs, seem to split into two sequences. Most of the stars trace a well-defined, narrow sequence, which we call the 'normal' CO sequence throughout the paper, where the star HD 219734 (one of the *H*-band best-fitting stars, see above) can be actually found, for *all* CO plots. Along this sequence, the CO line-strengths increase with decreasing T_{eff}. However, some stars do not fall on to this sequence, but form a sort of 'CO-strong' sequence (where two of the *H*-band best-fitting stars, namely HD 10465 and HD 187238, can be found in *all* CO plots).[8] To guide the eye, we performed a linear fit to the CO-strong sequence (see Appendix A for details), and marked such a sequence with black segments in Fig. 7. In all panels, we show the CO line-strengths for the XSG stack as horizontal red dashed lines. These lines intersect the locus of stars at an effective temperature of about 4000 K. This is the temperature where stars in the CO-strong sequence deviate the most from those in the normal sequence.

We also attempted to fit the spectrum of the *H*-band best-fitting stars with the MARCS (Gustafsson et al. 2008) library of very cool stellar spectra, and tried to extract additional information regarding their α abundance ratio. However, the best-fitting models do a poor job of predicting the shape of the spectra of such cool stars and CO indices; therefore the derived parameters are less reliable. Here, we only mention that the results point to a lower α abundance for stars in the CO-strong sequence compared to the one in the normal sequence (see Appendix B for details of this experiment).

As noted above, two stars that were assigned the highest weight in the STARLIGHT best-fitting spectrum (HD 10465 and HD 187238; see Table 1) occupy the CO-strong sequence, while the other two (HD 219734 and HD 36003) follow the normal sequence. This suggests that in order to match the observed spectrum of ETGs, a significant contribution from the CO-strong sequence might be required. However, as a general caveat, one should notice that massive ETGs might contain different stars than the few CO-strong-sequence stars in the IRTF library. An interesting point is that the two stars in the higher sequence have almost the same T_{eff} as the two other stars. This may explain why SPS models do actually fail to reproduce CO features. A nominal SSP model averages the available stellar spectra along the isochrones. Since stars in the normal sequence are in larger number compared to those in the CO-strong sequence, the contribution from the latter is diluted in the synthesized models. The STARLIGHT fitting results are suggesting instead that it should be the other way around, with a large weight from the CO-strong sequence.

[8]Note that for CO1.58 and CO1.64, a double-branch sequence is not so clear. However, this might result from some sky residuals in the wavelength range of the CO lines, or some contamination of the CO lines from different absorbers. According to panel (a) in fig. A2 of Eftekhari et al. (2021), the central bandpass of CO1.58 is severely contaminated by telluric absorption lines and its red bandpass is contaminated by a strong emission line at ∼15 844 Å. Moreover, a magnesium line contributes to this absorption feature. In the same figure, in panel (b), the presence of two strong emission lines can be seen in both blue and red bandpasses of the CO1.64 index. The central feature also has some contribution from atomic silicon lines.

Downloaded from https://academic.oup.com/mnras/article/512/1/378/6534922 by University of Central Lancashire user on 24 March 2022

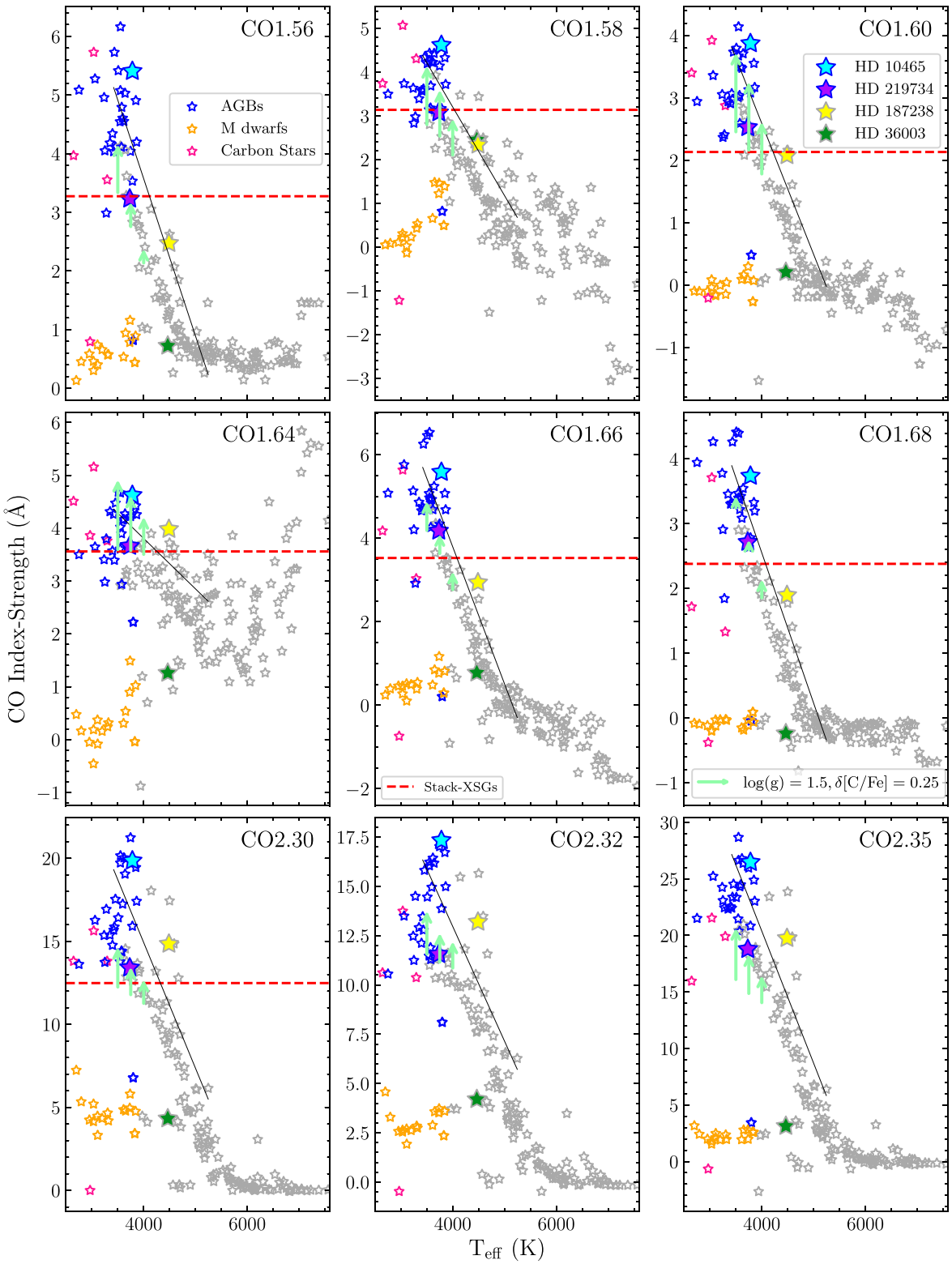

Figure 7. Index strength of CO indices for IRTF stars as a function of effective temperature. Indices have been measured on spectra smoothed at the common resolution of $\sigma = 360\,\mathrm{km\,s^{-1}}$. The blue, orange, and pink colours correspond to AGB, M-dwarf, and carbon stars at $T_{\rm eff} < 3900$ K, respectively. Remaining stars are shown with grey colour. The four stars that contribute most to the net flux of the STARLIGHT best-fitting model (see the text) are marked with filled star symbols of different colours (see labels in the top-right panel). The black lines are the linear fit to the CO-strong sequence of giant stars, with $T_{\rm eff} < 5500$ K (see the text for details). In each panel, the CO line-strength for the XSG stack is marked with a dashed red line. The light green arrows mark the increase caused by a [C/Fe] enhancement of 0.25 dex, from the theoretical stars of Knowles (2019), all having solar metallicity and $\log g = 1.5$.

Downloaded from https://academic.oup.com/mnras/article/512/1/378/6534922 by University of Central Lancashire user on 24 March 2022

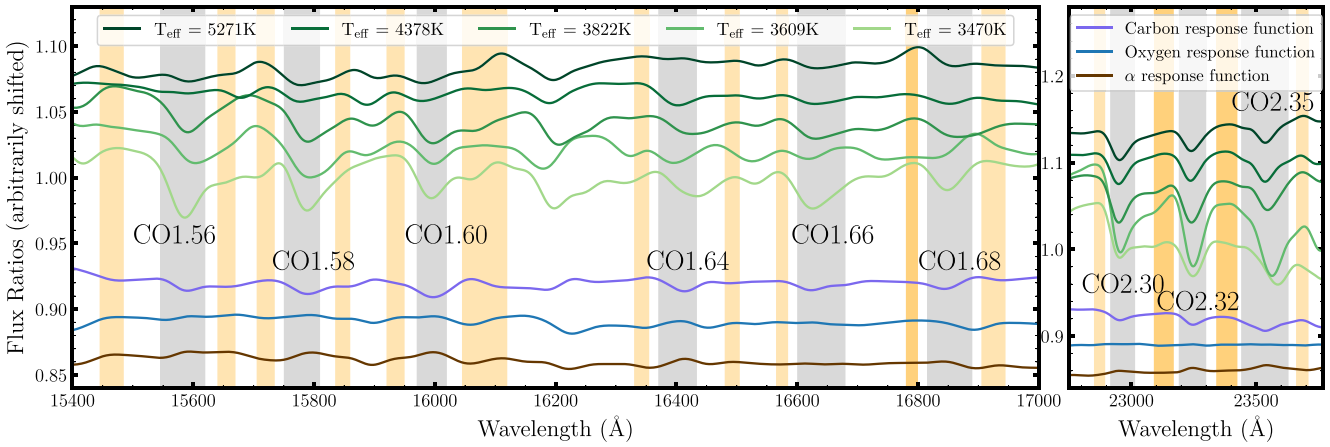

Figure 8. Comparison of empirical CO responses (light through dark green curves) with [C/Fe] (violet), [O/Fe] (blue), and [α/Fe] (brown) responses from CvD12 models. The light to dark green spectra are defined as the ratio between mean spectra of stars in the CO-strong and normal CO sequences, for stars with different temperature. The definition of indices are overplotted with grey (central bandpasses) and orange (pseudo-continua bandpasses) areas. The left- and right-hand panels refer to *H* and *K* bands, respectively. All spectra have been smoothed to $\sigma = 360\,\mathrm{km\,s^{-1}}$.

In order to remedy this situation, we constructed ad hoc SPS models, as detailed below.

6.2 Empirical corrections to E-MILES models

We modified E-MILES stellar population models by shifting stars in the normal CO sequence to those in the upper (CO-strong) one. The procedure is described in detail in Appendix A. In short, we systematically separated giant stars into the two sequences (according to all the available CO indices), and for stars that share similar stellar parameters, we divided the mean spectrum for the CO-strong sequence with that for the normal sequence, to obtain a (multiplicative) differential response of 'CO-enhancement', as illustrated in Fig. 8 (see light through dark green spectra). We point out that this procedure is possible because a number of stars are in the upper sequence for all the CO indices, i.e. there is actually a population of stars, with strong lines, that can be singled out from the normal sequence, which extend over a range of temperatures. The responses obtained in this way were interpolated at different temperatures, and applied to the giant star spectrum in the normal sequence. New SSP models were synthesized accordingly, using the 'empirically corrected' giant stars, for an age of 11 Gyr, [M/H] = +0.06 and +0.26, and Γ_b = 1.3 and 3.0, respectively, over a wavelength range 15 400 to 23 800 Å.

We measured the CO line-strengths on the empirically corrected models. In Figs 4 and 5, we show the variation of CO indices, compared to the reference E-MILES models, as khaki and orange arrows, for IMF slopes of Γ_b = 1.3 and 3.0, respectively. The solid and dotted arrows correspond to solar and supersolar metallicity models, respectively. The empirically corrected SSPs have significantly larger CO indices. In the case of CO2.30, CO1.60, and CO1.66, the khaki arrows would allow one to fit the stacked spectrum of XSGs. However, since the IMF has been shown to be bottom-heavy for these galaxies, it should be looked at the orange arrows, whose size is not enough to match the data. For some *H*-band indices, i.e. CO1.56, CO1.60, and CO1.66, the khaki arrows predict even larger CO values that the median line-strengths for the samples of B18 and F19. In the case of CO1.58, CO1.64, and CO1.68, although the empirically corrected models improve the predictions of CO indices, they cannot match those of massive ETGs (even for an MW-like IMF).

We note that the dotted arrows have approximately the same size as solid arrows, i.e. the effect of the empirical correction does not depend on metallicity. Perhaps, this is not surprising, as stars of the IRTF stellar library are biased towards solar metallicities (see Röck et al. 2017 and references therein). Also, khaki and orange arrows have approximately similar size, implying that the empirical CO response is not coupled to that of IMF, as it is the case for Na-enhanced E-MILES models (see La Barbera et al. 2017 for details). This stems from the fact that the CO correction is only performed on giant stars, while a bottom-heavy IMF increases the number of dwarf, relative to giant, stars.

6.3 What is driving the empirical corrections?

The empirically corrected models provide closer predictions to observed CO indices compared to E-MILES models, although still far from a perfect match. In order to make further progress, we tried to understand the physical drivers behind the empirical corrections, to possibly tune the models further.

Fig. 8 shows the CO-response functions of five selected stars, with different T$_{eff}$ but otherwise similar stellar parameters, that we used to construct the empirically corrected models (see Appendix A for details). The violet, blue, and brown lines in the figure plot responses corresponding to a carbon enhancement of 0.15 dex, an oxygen enhancement of 0.2 dex, and an α enhancement of 0.2 dex, based on CvD12 SSP models (with an age of 13.5 Gyr, [Fe/H] = 0.0, and Chabrier IMF), respectively. Indeed, the empirical responses look more similar to those for carbon enhancement, although the differences in the depth of CO indices are quite significant. On the contrary, the response to oxygen enhancement is very different from the empirical responses, being almost flat in the regions of CO absorptions. The effect of α enhancement is even more dissimilar, as it shows bumps in the regions of the CO central passbands, consistent with the fact that CO line-strengths anticorrelate with [α/Fe] (see Section 5.1).

Therefore, we speculate that the empirical corrections might be reflecting the effect of carbon enhancement on (cool) giant stars. To further test this hypothesis, the [C/Fe] of stars in the CO-strong and normal sequences should be compared. Unfortunately,

Downloaded from https://academic.oup.com/mnras/article/512/1/378/6534922 by University of Central Lancashire user on 24 March 2022

carbon abundances for giant stars in the IRTF library have not been measured yet. Hence, we relied on theoretical C-enhanced stars from Knowles (2019) to see if an enhancement in carbon abundance might explain the difference between CO line-strengths in the two CO sequences. In Fig. 7, we show the effect of a [C/Fe] enhancement of 0.25 dex on theoretical cool giant stars with light-green arrows. Interestingly, the size of the arrows increases with decreasing T_{eff}. However, one should bear in mind that theoretical stellar spectra are rather uncertain for very cool stars, and these models stop at 3500 K. Indeed, we may expect that this trend continues at lower T_{eff}, with CO absorptions getting even stronger. Focusing only on the difference due to enhancement and neglecting the starting point, we see that the arrows are comparable to the CO line-strengths difference between normal and CO-strong sequence stars. For CO1.58, CO1.60, and CO2.32 indices, the arrows can bring the stars from the normal to the CO-strong sequence, while for other indices, the arrows can explain only part of the difference between the two sequences, with a larger gap for cooler stars. For instance, the arrow at 3500 K for CO1.68 is too small, and it is unable to reach a group of three stars, with T_{eff} ∼3500 K and CO1.68 as high as ∼4.4 Å.

As shown in Figs 4 and 5, increasing [α/Fe] abundance causes the CO absorptions to weaken. However this prediction, which is qualitatively similar in both E-MILES α-enhanced and CvD12 models, is in contrast with predictions of A-LIST SSP models (Ashok et al. 2021). According to their fig. 6, CO absorptions get stronger by increasing [α/Fe]. A-LIST provides fully empirical SSP model predictions, based on the APOGEE stellar library, while α-enhanced E-MILES SSPs are semi-empirical models (i.e. the relative effect of α-enhancement is estimated through the aid of theoretical star spectra). Unfortunately, [α/Fe] abundance ratios have not been measured for all IRTF stars. However, we further assessed the effect of [α/Fe] on CO indices by looking at elemental abundance ratios for the APOGEE stellar library, as computed with ASPCAP. To this effect, we selected a set of APOGEE stars as described in Section 3.2. We attempted to single out the effect of surface gravity, metallicity, and carbon enhancement by only selecting stars within a narrow range of stellar parameters ($0.35 < \log g < 0.55$, $-0.1 <$ [M/H] < 0.1, and $0 <$ [C/Fe] < 0.05). Since the wavelength coverage of APOGEE spectra is divided across three chips with relatively narrow ranges (blue chip from 1.51 to 1.581 μm, green chip from 1.585 to 1.644 μm, and red chip from 1.647 to 1.7 μm), we were able to measure line-strengths for only three CO indices (i.e. CO1.56, CO1.60, and CO1.66, respectively). Fig. 9 shows CO line-strengths for APOGEE stars as a function of T_{eff}, log g, and [M/H] (left-, mid, and right-hand panels), respectively, with stars being colour-coded according to their [α/Fe]. According to the figure, stars with higher CO do not show a higher value of [α/Fe]. In many cases, stars with high CO seem to have lower (rather than higher) [α/Fe]. In the plot of CO1.60 versus T_{eff}, only one star (in red) has high α-enhancement and high CO1.60. For CO1.66, a correlation (anticorrelation) of the index with surface gravity (metallicity) is actually observed. Note that these results are in disagreement with predictions from the A-LIST models of Ashok et al. (2021), with the origin of such disagreement remaining unclear.

Overall, our analysis shows that it is very unlikely that α enhancement is the missing piece of the CO puzzle. On the other hand, the effect of carbon on low-temperature giant stars seems to be the most likely candidate to explain the strength of CO lines. However, the predictions from theoretical models should be improved, and extended to stars with lower temperatures ($\lesssim$3500 K) at high metallicity, in order to draw firm conclusions.

7 DISCUSSION

It is instructive to look at the mismatch between models and data using CO–CO diagrams, i.e. plotting one CO index against CO line-strengths for other features. In Fig. 10, we show two such diagrams, based on three CO indices (CO1.58, CO1.60, and CO2.30, respectively) as measured for IRTF stars (star symbols), the stacked spectrum of XSGs (red point), and E-MILES models with ages from 1 to 14 Gyr (see the pink and purple lines, corresponding to solar-metallicity models for MW-like and bottom-heavy IMF, respectively). In the CO versus CO plots, the locus of stars is well defined, forming a relatively narrow strip. Interestingly, the point of massive galaxies falls off the main strip: for values of CO1.60 ∼ 2 Å and CO2.30 ∼ 12 Å, stars have CO1.60 ∼ 2.4 Å, while the XSG stack has CO1.60 ∼ 3.2 Å. As expected, E-MILES model predictions follow the locus of stars, predicting lower CO values compared to the data. However, to be able to match the XSG stack, the models should not only increase the CO line-strengths but, also, move away from the main strip of IRTF stars.

Note that in both panels of Fig. 10, some dwarf stars (see the orange stars in the figure) do not share the same locus as the rest of the IRTF stars, but they are actually shifted to higher values of CO1.58 (∼1.4Å), at given values of CO1.60 (∼0.2Å) and CO2.30 (∼6Å), respectively. The result is that predictions of models with a bottom-heavy IMF (orange arrows in the figure) are slightly above the main star locus. However, at the same time, these models predict lower values of CO (ΔCO1.58 ≈−0.2 Å, ΔCO1.60 ≈−0.5 Å, and ΔCO2.30 ≈−2 Å, comparing the tips of the orange and khaki arrows in the figure). Therefore, while bottom-heavy models worsen the gap between observed and model line-strengths, on the other hand, the CO–CO plots suggest that IMF variations might help to reconcile models and data by producing greater CO1.58 line-strengths than CO2.30.

In Fig. 10, we also show, with blue arrows, the effect of an AGB-enhanced population (based on AGB-enhanced E-MILES SSPs; see Section 5.2), trying to mimic the presence of an intermediate-age population. The AGB-enhanced models increase CO1.58 by only 0.1 Å, while the discrepancy between the red point (i.e. the XSG stack) and the pink line (fiducial E-MILES model) is about eight times larger. Moreover, the change in CO1.60 (CO2.30) line-strength due to the blue arrow is ∼0.1 (∼0.5) Å, i.e. one-fifth (one-forth) of the offset between the models and data. We conclude, as already discussed in Section 5.2, that while AGB stars might have a relevant contribution to the NIR light of (massive) galaxies, they are likely not responsible for the strong CO line-strengths in the *H* and *K* bands.

The khaki and orange arrows in Fig. 10 are the same as in Figs 4 and 5, plotting the increase of CO line-strengths caused by the empirical correction on giant stars, for both MW-like and bottom-heavy IMF models, respectively. Both arrows increase the CO1.58, CO1.60, and CO2.30 line-strengths by ∼0.4, 0.4, and 1.5 Å, respectively. While the khaki arrow (compared to the orange one) brings the model indices closer to the XSG stack, the orange arrow is not able to reach the data, but it is slightly off the star sequence, similar to the XSG stack. This suggests that in order to match the CO line-strengths, one would need an effect similar to that of the empirical corrections, plus the slight offset due to a bottom-heavy IMF.

The effect of carbon enhancement from CvD12 models is also shown in Fig. 10 (violet arrows), to be compared with the empirical responses. In the CO1.58 versus CO1.60 diagram, the violet arrow, although increasing the CO indices (by ∼0.4 Å for CO1.58, and ∼0.4 Å for CO1.60), it is along the stellar locus, while in the CO1.58 versus CO2.30 plot, the arrow seems to point to the correct direction to match the XSG stack (though with a small overall variation of only ΔCO2.30 ≈ 0.5Å).

Downloaded from https://academic.oup.com/mnras/article/512/1/378/6534922 by University of Central Lancashire user on 24 March 2022

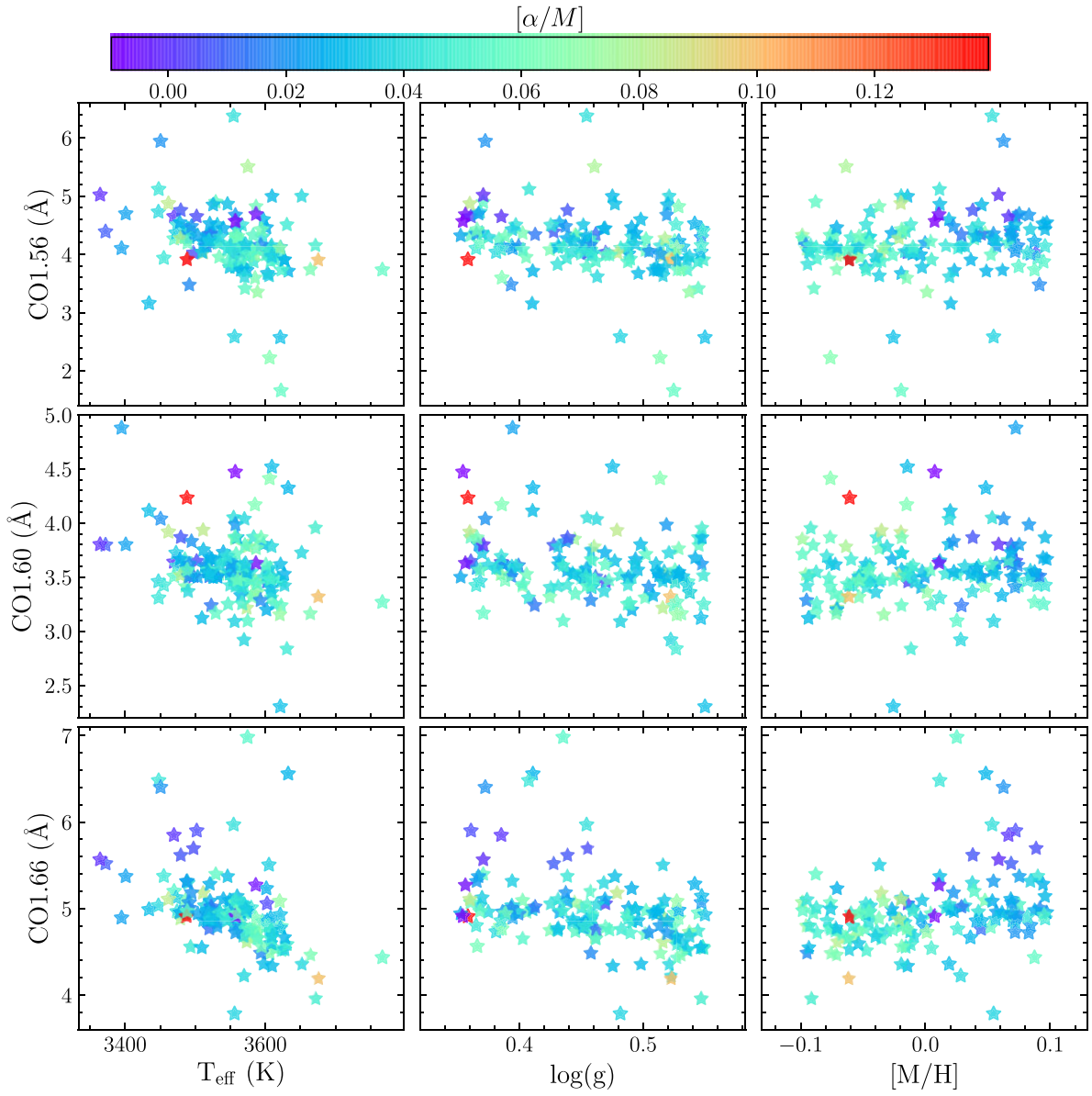

Figure 9. CO indices measured on the spectra of APOGEE stars as a function of stellar parameters, namely T_{eff} (left), log g (middle), and [M/H] (right). The stars are coloured according to [α/Fe], as shown from the colourbar on the top.

Similar to the effect of carbon enhancement, it seems that carbon stars (pink stars in Fig. 10) might also be able to bring the models out of the stellar locus in the CO1.58 versus CO2.30 diagram, while this is not the case for the CO1.58–CO1.60 plot, as in the latter case, pink stars are somehow aligned to the sequence of blue stars. However, according to Fig. 7, for most CO versus T_{eff} panels, carbon stars are not in the CO-strong sequence, i.e. they would not help in matching all the *H*- and *K*-band CO line-strengths. Again, this shows the importance of combining the largest available set of CO features, as we do in our analysis, and gives further support to our conclusion that adding an intermediate-age stellar population (that would also include carbon stars) to an underlying old component does not solve the issue with NIR CO spectral features.

Using optical and *J*- and *K*-band absorption features, including the first two CO bandheads in *K* band, Alton, Smith & Lucey (2017) and Alton et al. (2018) studied stellar population gradients in the spectra of eight massive ETGs. They showed that models that do not account for the effect of [C/Fe] variations underpredict the CO bandheads in the *K* band. Moreover, they showed that to fit H_β, in the optical, a large enhancement in carbon abundance is also required. In other terms, an overabundance of carbon seems to have a prominent role in matching CO lines, in agreement with the suggestions from our analysis. However, Alton et al. (2017, 2018) also used CO features in the *K* band to conclude in favour of an MW-like IMF in the centre of (some) massive ETGs, in contrast to studies based on (optical) spectral features. Indeed, our analysis shows that current stellar population models in the NIR are still not accurate enough to allow for a quantitative matching of CO lines to be performed, and an even smaller effect to be constrained, such as that of a varying IMF. The results presented here demand a new generation of NIR stellar population models, after a significant effort to move beyond the current limitations of theoretical star spectra is made, particularly

Downloaded from https://academic.oup.com/mnras/article/512/1/378/6534922 by University of Central Lancashire user on 24 March 2022

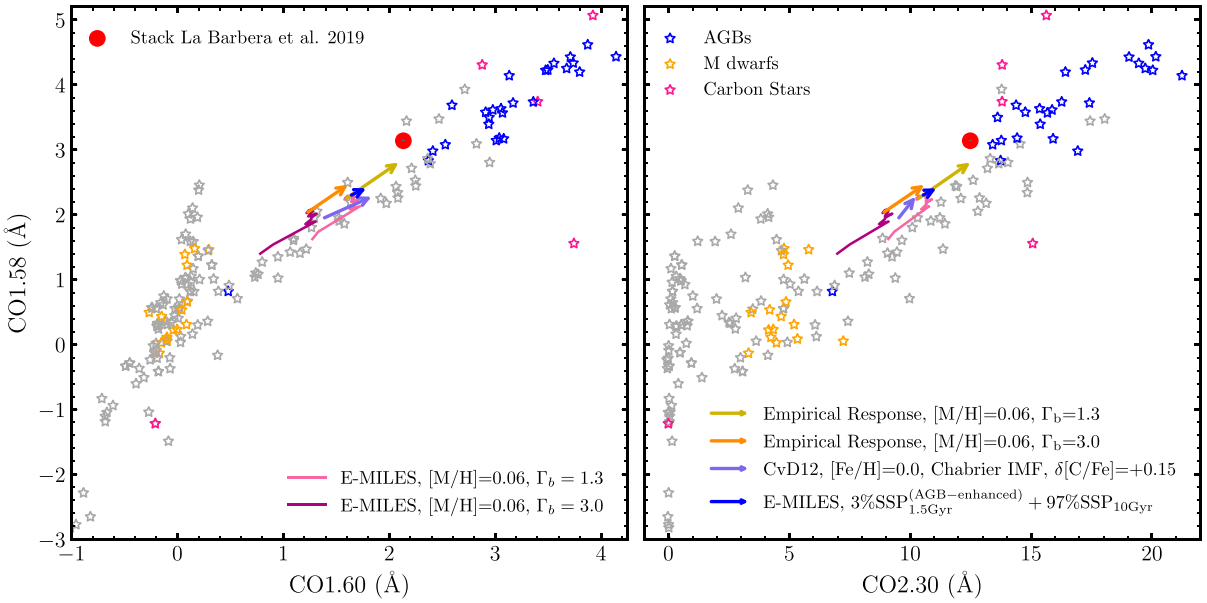

Figure 10. A selected pair of CO versus CO plots (CO1.58 versus CO1.60 on the left; CO1.58 versus CO2.30 on the right), showing IRTF stars (star symbols), the XSG stacked spectrum (filled red circle), and predictions for solar-metallicity E-MILES SSP models having ages from 1 to 14 Gyr, for an MW-like and bottom-heavy IMF (see pink and purple lines), respectively. The violet arrows show the effect of increasing carbon abundance for CvD12 models, while the (small) blue arrows plot the increase of CO indices when accounting for AGB-enhanced intermediate-age population on top of an old stellar component. The effect of our empirical corrections on SSP model predictions is shown with khaki and orange arrows, for an MW-like and a bottom-heavy IMF, respectively.

for the predictions of abundance ratio effects in low-temperature (giant) stars. Along the same lines, we point out that while our ad hoc empirically corrected SPS models do not match the observations yet, they tend to significantly reduce the discrepancy with respect to the observed CO strengths. Admittedly, the interpretation of our empirical corrections as an effect of [C/Fe] for low-temperature giants remains rather speculative, but urges for their study and opens up new avenues for improving SPS models in the NIR spectral range.

8 SUMMARY AND CONCLUSIONS

We have shown, in a comprehensive manner, that a whole set of CO lines in the spectra of massive ETGs, from *H* through to *K* band, lie significantly above the predictions of widely used state-of-the-art SPS models. We have explored different possible reasons for this 'CO-mismatch' problem, finding that an individual enhancement of carbon abundance might be the most likely explanation, compared to other scenarios (such as an enhanced contribution of AGB stars from intermediate-age populations). In general, our study highlights the importance of improving SPS models in the NIR, in the following aspects:

(i) *Non-solar chemical abundances:* we need substantial progress in the modelling of the response of stellar spectra to elemental abundance variations. In particular, the effect of varying abundance ratios for cool stars ($T_{\rm eff} < 4000$ K) is far from being well understood, and might be crucial to explain the current issues with NIR spectral lines. In addition, since current SPS modelling for varying abundance ratios is based on scaled-solar isochrones, we need a significant improvement on isochrones with non-solar abundances to create fully consistent SSP models with varying abundances. Moreover, the interplay between C, N, and O elements and their effects on CO indices are not yet fully understood. As some of the current models consider O as one of the α-elements, a different treatment of O from α-elements may be an interesting avenue for further investigation.

(ii) *Very cool stars:* current (theoretical) models struggle to reproduce atomic and molecular bands for stars with $T_{\rm eff} < 3500$ K. Moreover, SPS would benefit from an improved treatment of evolved stellar evolution phases, such as those of red giants and supergiants, and the AGB, which have a prominent contribution to the NIR light of a stellar population in various age regimes.

(iii) *High-metallicity stars:* empirical stellar libraries, used to construct SPS models, are based on MW stars, having an unavoidable bias towards solar metallicity. Based on current SPS model predictions, CO indices do not depend significantly on metallicity, but from the study of individual stars and clusters (Aaronson et al. 1978; Oliva et al. 1995), CO indices are found to increase with increasing metallicity. Therefore, models with a good coverage of stars in the supersolar metallicity regime might yield further important clues to understand the CO-mismatch problem.

As a final remark, we would like to emphasize that a revision of SPS models in the directions suggested by the NIR CO indices should take carefully into account the constraints provided by other spectral ranges, such as the optical and the UV. For example, fitting just a single CO line or a number of them could lead to misleading derivations of stellar population properties.

ACKNOWLEDGEMENTS

We are thankful to the reviewer, Dr. Russell Smith, for his careful reading of the manuscript and valuable comments. The authors acknowledge support from grant PID2019-107427GB-C32 from the Spanish Ministry of Science, Innovation and Universities (MCIU). This work has also been supported through the IAC project TRACES which is partially supported through the state budget and the regional budget of the Consejería de Economía, Industria, Comercio y Conocimiento of the Canary Islands Autonomous Community.

Downloaded from https://academic.oup.com/mnras/article/512/1/378/6534922 by University of Central Lancashire user on 24 March 2022

DATA AVAILABILITY

The E-MILES SSP models are publicly available at the MILES website (http://miles.iac.es). The updated version of Na-enhanced models of La Barbera et al. (2017) are also available from the same website (under 'Other predictions/data'). The Conroy & van Dokkum (2012) SSP models are available upon request to the authors (see https://scholar.harvard.edu/cconroy/projects). The Conroy et al. (2018) models are available for download at https://scholar.harvard.edu/cconroy/sps-models. The Maraston (2005) are downloaded from http://www.icg.port.ac.uk/~maraston/Claudia's_Stellar_Population_Model.html. Observations of La Barbera et al. (2019) sample made with ESO Telescope at the Paranal Observatory under programmes ID 092.B-0378, 094.B-0747, 097.B-0229 (PI: FLB). The central spectra and stacked central spectrum are available from FLB upon request. The spectra of Baldwin et al. (2018) sample were taken using the Gemini Near-Infrared Spectrograph on the Gemini North telescope in Hawaii through observing program GN-2012A-Q-22. The reduced spectra (FITS files) are available via https://github.com/cbaldwin1/Reduced-GNIRS-Spectra. Observations of François et al. (2019) sample made with ESO Telescopes at the La Silla Paranal Observatory under programme ID 086.B-0900(A). The reduced spectra (FITS files) are available via http://cdsarc.u-strasbg.fr/viz-bin/qcat?J/A+A/621/A60. Observations of Silva et al. (2008) sample performed at the European Southern Observatory, Cerro Paranal, Chile; ESO program 68.B-0674A and 70.B-0669A. Observations of Mármol-Queraltó et al. (2008) sample performed at the European Southern Observatory, Cerro Paranal, Chile, as well. Röck et al. (2017) corrected the data of these two samples to restframe and convolved them to a resolution of $\sigma = 360\,\mathrm{km\,s^{-1}}$and they are available from EE upon request. The IRTF Spectral Library is observed with the SpeX spectrograph, at the NASA Infrared Telescope Facility on Mauna Kea and the spectra are publicly available at http://irtfweb.ifa.hawaii.edu/~spex/IRTF_Spectral_Library/. Theoretical stars are computed by ATK specifically for this project and are available from ATK upon request. APOGEE stars can be downloaded through the https://www.sdss.org/dr16/ website.

REFERENCES

Aaronson M., Cohen J. G., Mould J., Malkan M., 1978, ApJ, 223, 824
Alton P. D., Smith R. J., Lucey J. R., 2017, MNRAS, 468, 1594
Alton P. D., Smith R. J., Lucey J. R., 2018, MNRAS, 478, 4464
Arnaud K. A., Gilmore G., Collier Cameron A., 1989, MNRAS, 237, 495
Ashok A. et al., 2021, AJ, 161, 167
Asplund M., Grevesse N., Sauval A. J., 2005, in Barnes T. G. III, Bash F. N., eds, ASP Conf. Ser.Vol. 336, Cosmic Abundances as Records of Stellar Evolution and Nucleosynthesis. Astron. Soc. Pac., San Francisco, p. 25
Baldwin C., McDermid R. M., Kuntschner H., Maraston C., Conroy C., 2018, MNRAS, 473, 4698 (F19)
Baldwin J. R., Danziger I. J., Frogel J. A., Persson S. E., 1973a, Astrophys. Lett., 14, 1
Baldwin J. R., Frogel J. A., Persson S. E., 1973b, ApJ, 184, 427
Baraffe I., Chabrier G., Allard F., Hauschildt P. H., 1998, A&A, 337, 403
Cappellari M. et al., 2011, MNRAS, 413, 813
Cenarro A., Cardiel N., Gorgas J., Peletier R., Vazdekis A., Prada F., 2001, MNRAS, 326, 959
Chabrier G., Baraffe I., 1997, A&A, 327, 1039
Choi J., Dotter A., Conroy C., Cantiello M., Paxton B., Johnson B. D., 2016, ApJ, 823, 102
Cid Fernandes R., Mateus A., Sodré L., Stasińska G., Gomes J. M., 2005, MNRAS, 358, 363
Conroy C., van Dokkum P., 2012, ApJ, 747, 69 (CvD12)
Conroy C., Villaume A., van Dokkum P. G., Lind K., 2018, ApJ, 854, 139 (C18)
Cushing M. C., Rayner J. T., Vacca W. D., 2005, ApJ, 623, 1115
Davidge T. J., 2018, AJ, 156, 233
Davidge T. J., 2020, AJ, 159, 186
Davidge T. J., Beck T. L., McGregor P. J., 2008, ApJ, 677, 238
Dotter A., Chaboyer B., Jevremović D., Kostov V., Baron E., Ferguson J. W., 2008, ApJS, 178, 89
Doyon R., Joseph R. D., Wright G. S., 1994, ApJ, 421, 101
Drinkwater M. J., Gregg M. D., Colless M., 2001, ApJ, 548, L139
Eftekhari E., Vazdekis A., La Barbera F., 2021, MNRAS, 504, 2190
Faber S. M., 1972, A&A, 20, 361
François P., Morelli L., Pizzella A., Ivanov V. D., Coccato L., Cesetti M., Corsini E. M., Dalla Bontà E., 2019, A&A, 621, A60 (B18)
Frogel J. A., Persson S. E., Aaronson M., Becklin E. E., Matthews K., Neugebauer G., 1975, ApJ, 195, L15
Frogel J. A., Persson S. E., Aaronson M., Matthews K., 1978, ApJ, 220, 75
Frogel J. A., Persson S. E., Cohen J. G., 1980, ApJ, 240, 785
García Pérez A. E. et al., 2016, AJ, 151, 144
Girardi L., Bressan A., Bertelli G., Chiosi C., 2000, A&AS, 141, 371
Gregg M. D. et al., 2006, in Koekemoer A. M., Goudfrooij P., Dressel L. L., eds, The 2005 *HST* Calibration Workshop: Hubble After the Transition to Two-Gyro Mode. p. 209
Grillmair C. J., Freeman K. C., Bicknell G. V., Carter D., Couch W. J., Sommer-Larsen J., Taylor K., 1994, ApJ, 422, L9
Gustafsson B., Edvardsson B., Eriksson K., Jørgensen U. G., Nordlund Å., Plez B., 2008, A&A, 486, 951
James P. A., Mobasher B., 1999, MNRAS, 306, 199
Kleinmann S. G., Hall D. N. B., 1986, ApJS, 62, 501
Knowles A. T., 2019, PhD thesis, Univ. Central Lancashire, available at http://clok.uclan.ac.uk/34353/
Knowles A. T., Sansom A. E., Allende Prieto C., Vazdekis A., 2021, MNRAS, 504, 2286
Koesterke L., 2009, in Hubeny I., Stone J. M., MacGregor K., Werner K., eds, AIP Conf. Proc. Vol. 1171, Recent Directions In Astrophysical Quantitative Spectroscopy and Radiation Hydrodynamics. Am. Inst. Phys., New York, p. 73
Kroupa P., Gilmore G. F., 1994, MNRAS, 269, 655
Kuntschner H., Davies R. L., 1998, MNRAS, 295, L29
Kurucz R., 1993, ATLAS9 Stellar Atmosphere Programs and 2 km/s grid. Kurucz CD-ROM No. 13. Cambridge, Mass.: Smithsonian Astrophysical Observatory, 1993., 13
La Barbera F., Ferreras I., Vazdekis A., de la Rosa I., de Carvalho R., Trevisan M., Falcón-Barroso J., Ricciardelli E., 2013, MNRAS, 433, 3017
La Barbera F., Vazdekis A., Ferreras I., Pasquali A., Allende Prieto C., Röck B., Aguado D. S., Peletier R. F., 2017, MNRAS, 464, 3597
La Barbera F. et al., 2019, MNRAS, 489, 4090 (LB19)
Lançon A., Wood P. R., 2000, A&AS, 146, 217
Lejeune T., Cuisinier F., Buser R., 1998, A&AS, 130, 65
Majewski S. R. et al., 2017, AJ, 154, 94
Maraston C., 2005, MNRAS, 362, 799 (M05)
Marigo P., Girardi L., Bressan A., Groenewegen M. A. T., Silva L., Granato G. L., 2008, A&A, 482, 883
Mármol-Queraltó E., Cardiel N., Cenarro A. J., Vazdekis A., Gorgas J., Pedraz S., Peletier R. F., Sánchez-Blázquez P., 2008, A&A, 489, 885
Mármol-Queraltó E. et al., 2009, ApJ, 705, L199 (M09)
Mészáros S. et al., 2012, AJ, 144, 120
Mobasher B., James P. A., 1996, MNRAS, 280, 895
Mobasher B., James P. A., 2000, MNRAS, 316, 507
Oliva E., Origlia L., Kotilainen J. K., Moorwood A. F. M., 1995, A&A, 301, 55
Origlia L., Moorwood A. F. M., Oliva E., 1993, A&A, 280, 536
Pietrinferni A., Cassisi S., Salaris M., Castelli F., 2004, ApJ, 612, 168
Rayner J. T., Cushing M. C., Vacca W. D., 2009, ApJS, 185, 289
Renzini A., Voli M., 1981, A&A, 500, 221
Riffel R. et al., 2019, MNRAS, 486, 3228
Röck B., Vazdekis A., Peletier R. F., Knapen J. H., Falcón-Barroso J., 2015, MNRAS, 449, 2853

Downloaded from https://academic.oup.com/mnras/article/512/1/378/6534922 by University of Central Lancashire user on 24 March 2022

Röck B., Vazdekis A., Ricciardelli E., Peletier R. F., Knapen J. H., Falcón-Barroso J., 2016, A&A, 589, A73
Röck B., Vazdekis A., La Barbera F., Peletier R. F., Knapen J. H., Allende-Prieto C., Aguado D. S., 2017, MNRAS, 472, 361
Salvador-Rusiñol N., Beasley M. A., Vazdekis A., Barbera F. L., 2021, MNRAS, 500, 3368
Sánchez-Blázquez P. et al., 2006a, MNRAS, 371, 703
Sánchez-Blázquez P., Gorgas J., Cardiel N., González J. J., 2006b, A&A, 457, 809
Scharf C. A., Zurek D. R., Bureau M., 2005, ApJ, 633, 154
Schiavon R. P., 2007, ApJS, 171, 146
Silva D. R., Kuntschner H., Lyubenova M., 2008, ApJ, 674, 194 (S08)
Smith R. J., Lucey J. R., 2013, MNRAS, 434, 1964
Smith R. J., Alton P., Lucey J. R., Conroy C., Carter D., 2015, MNRAS, 454, L71
Valdes F., Gupta R., Rose J. A., Singh H. P., Bell D. J., 2004, ApJS, 152, 251
Vazdekis A., Casuso E., Peletier R. F., Beckman J. E., 1996, ApJS, 106, 307
Vazdekis A., Peletier R. F., Beckman J. E., Casuso E., 1997, ApJS, 111, 203
Vazdekis A., Sánchez-Blázquez P., Falcón-Barroso J., Cenarro A., Beasley M., Cardiel N., Gorgas J., Peletier R., 2010, MNRAS, 404, 1639
Vazdekis A., Koleva M., Ricciardelli E., Röck B., Falcón-Barroso J., 2016, MNRAS, 463, 3409
Vernet J. et al., 2011, A&A, 536, A105
Villaume A., Conroy C., Johnson B., Rayner J., Mann A. W., van Dokkum P., 2017, ApJS, 230, 23
Worthey G., Faber S., Gonzalez J. J., Burstein D., 1994, ApJS, 94, 687
Yi S. K. et al., 2005, ApJ, 619, L111

APPENDIX A: CONSTRUCTING EMPIRICALLY CORRECTED MODELS

We explain here the procedure of identifying stars in the CO-strong sequence and the derivation of spectral response functions to bring giant stars in the normal sequence to this CO-strong sequence. We identified for each index the CO-strong giants through a sigma clipping procedure within temperature intervals of 200 K. Those stars that were flagged in the CO-strong sequence for at least five CO indices were definitely classified as CO-strong, which are shown as filled cyan star symbols in Fig. A1. Then we selected five sets of stars, sharing similar T_{eff}, log g, and [Fe/H], within these two sequences, i.e. normal and CO-strong. These sets of stars are indicated with lime open circles in Fig. A1, and their parameters are given in Table A1.

These sets of stars allowed us to compute a 'response function' by dividing the mean spectrum of stars in the CO-strong sequence by the mean spectrum of stars in the normal sequence. These responses were interpolated to cover the temperature range of the giant stars in the normal sequence. Finally, we applied these response functions to all the giants in the normal sequence to correct them into the CO-strong sequence.

Table A1. Different sets of stars with similar stellar parameters selected for deriving empirical response functions as a function of temperature. Columns 1 and 2 indicate the group and sequence of stars to which they belong. Column 3 presents the ID of stars, and columns 4, 5, and 6 give the effective temperature, surface gravity, and metallicity, respectively.

Group	Sequence	Star	T_{eff} (K)	log g (dex)	[Fe/H] (dex)
(1)	(2)	(3)	(4)	(5)	(6)
Group 1	CO-strong	HD 14488	3509	− 0.16	− 0.77
		HD 14469	3551	− 0.16	− 0.69
	normal	HD 40239	3349	0.47	0.03
Group 2	CO-strong	HD 35601	3617	0.06	− 0.20
		HD 14404	3640	0.32	− 0.07
	normal	HD 4408	3571	0.99	− 0.51
Group 3	CO-strong	HD 10465	3781	0.50	− 0.46
	normal	HD 23475	3863	0.78	− 0.63
Group 4	CO-strong	HD 16068	4296	1.80	− 0.03
		HD 164349	4446	1.50	0.39
		HD 35620	4367	1.75	− 0.03
	normal	HD 161664	4405	1.63	− 0.03
	CO-strong	HD 179870	5246	1.96	0.11
		HD 42454	5238	1.10	0.03
Group 5		HD 3421	5383	2.45	− 0.17
	normal	HD 74395	5250	1.30	− 0.05
		HD 39949	5240	1.23	− 0.10

Downloaded from https://academic.oup.com/mnras/article/512/1/378/6534922 by University of Central Lancashire user on 24 March 2022

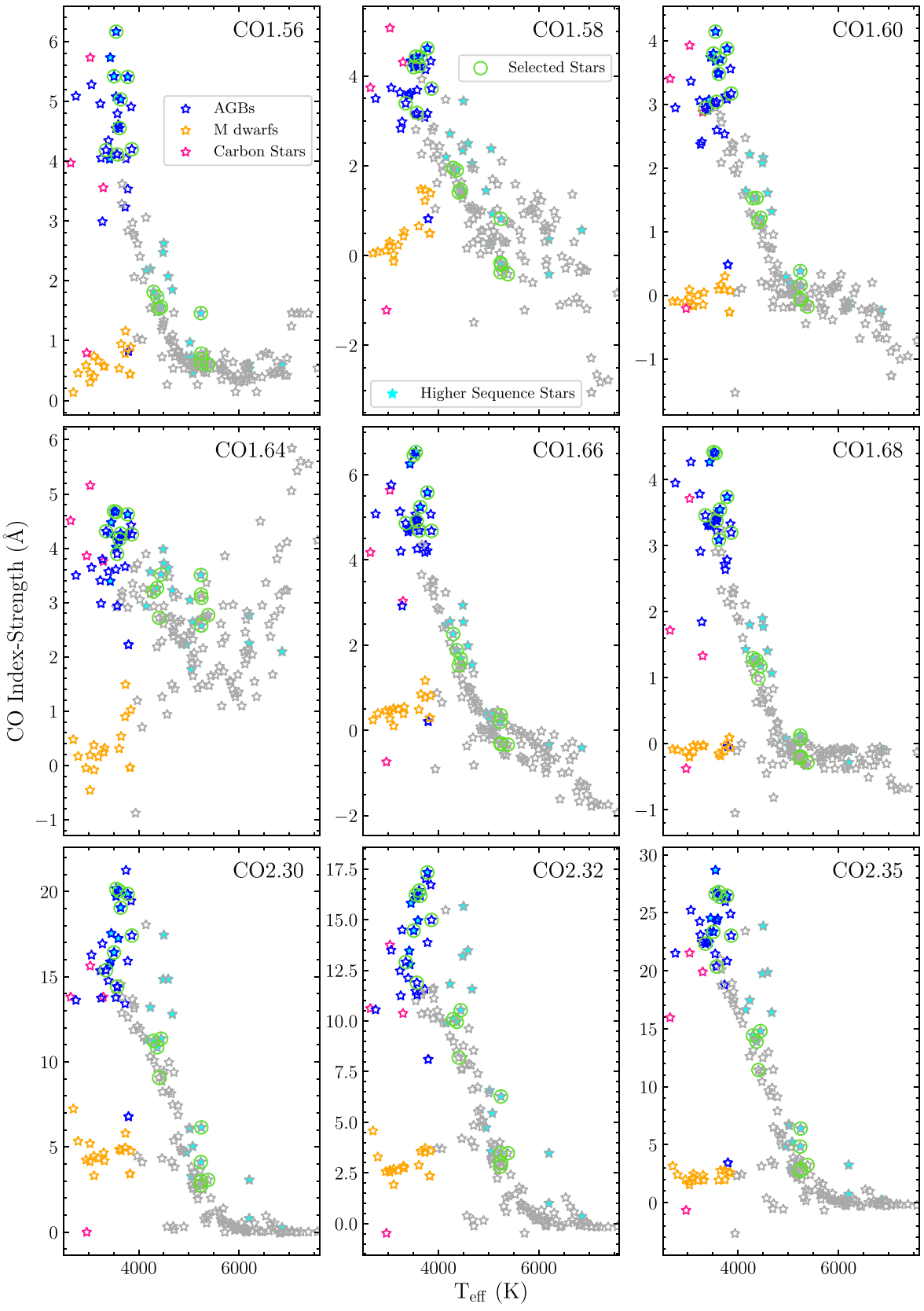

Figure A1. CO index-strengths for the IRTF stars plotted as a function of temperature, all measurements performed at resolution of $\sigma = 360\,\mathrm{km\,s^{-1}}$. The blue, orange, and pink colours represent AGB, M-dwarf, and carbon stars for $T_{\rm eff} < 3900$ K, respectively. Stars with higher temperatures are shown in grey colour. Giant stars belonging to the CO-strong sequence are shown as filled star cyan symbols. The open lime circles show different sets of stars sharing similar parameters in the normal and CO-strong sequences, which were used to derive the empirical responses (see the text).

Downloaded from https://academic.oup.com/mnras/article/512/1/378/6534922 by University of Central Lancashire user on 24 March 2022

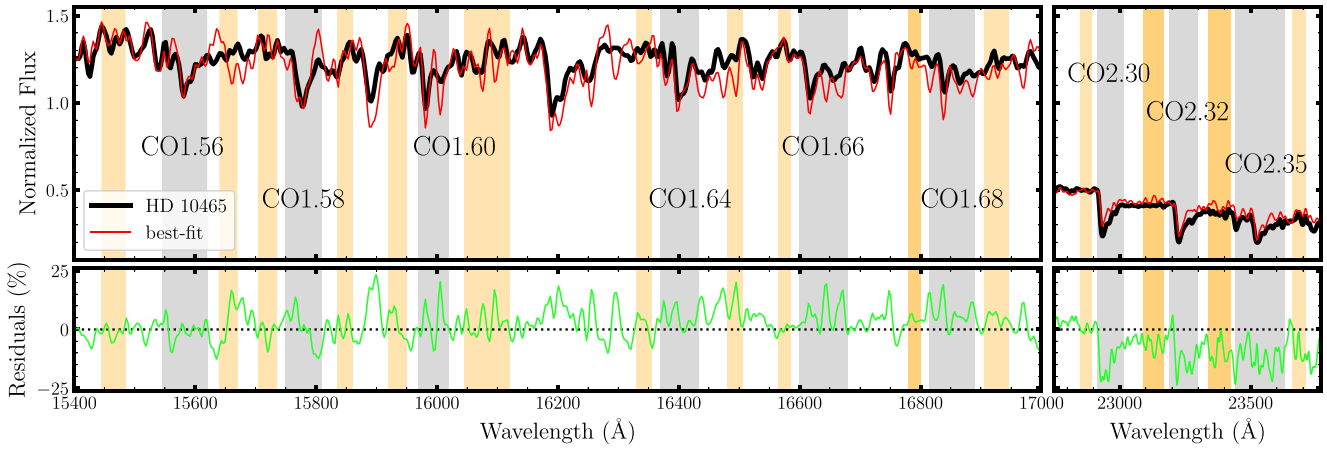

Figure B1. The upper panel shows the spectrum of one of the *H*-band best-fitting stars, HD 10465, (black) in *H* and *K* bands, together with the best-fitting spectrum from MARCS library of very cool stars (red). CO features, with central absorptions and their pseudo-continua, are plotted as grey and orange shaded areas. The lower panel shows the relative residuals of the observed spectrum with respect to the best-fitting spectrum.

APPENDIX B: CURRENT LIMITATIONS OF STELLAR MODELS IN THE LOW TEMPERATURE REGIME

One way to get information on the *H*-band best-fitting stars is by analysing their individual abundance ratios. For this, we need to derive the parameters of these stars using spectral template fitting with a synthetic spectral grid. However, at the very coolest temperatures, constructing stellar atmosphere models in the NIR is not a trivial task, because of the multiplicity of atomic lines and molecules. The MARCS (Gustafsson et al. 2008) models provide a good representation of the atmospheres of late-type stars down to effective temperatures of about 3000 K. These models have been chosen for the APOGEE survey (Majewski et al. 2017), and provide good agreement with the IR data, particularly for giant stars. We used the low-resolution spectra included with this collection of models to create a spectral library with a resolving power of 2000, matching that of the IRTF library.

We built two regular libraries, with equidistant steps in the atmospheric parameters. The first library spans 2500 K $\leq T_{\rm eff} \leq$ 4000 K in steps of 100 K, and $-0.5 \leq \log g \leq 5.5$ in steps of 0.5 dex. The second library spans 3500 K $\leq T_{\rm eff} \leq$ 6000 K in steps of 250 K, and $0.0 \leq \log g \leq 5.5$ in steps of 0.5 dex. In addition to $T_{\rm eff}$ and $\log g$ both libraries include two more parameters and [Fe/H], which spans $-2.5 \leq$ [Fe/H] ≤ 1.0 in steps of 0.5 dex. Both libraries span the full wavelength range of the IRTF spectra (8000–50 000 Å).

We fit the spectra of the stars in Table 1 using FERRE,[9] deriving the optimal set of parameters that best represented the observations. We preserved the SED in the observations and simply normalized the fluxes to have an average flux of 1, and fitted nearly the spectra between 8150 and 27 187 Å, where there is most information. An attempt to include longer wavelengths returned significantly poorer fits.

Table B1. Derived parameters of the *H*-band best-fitting stars from MARCS grid for cool stars. Column 1 gives the name of each star. The derived α abundance, effective temperature, surface gravity, and metallicity of the stars are given in Columns 2, 3, 4, and 5, respectively.

Star	[α/Fe] (dex)	$T_{\rm eff}$ (K)	$\log g$ (dex)	[Fe/H] (dex)
(1)	(2)	(3)	(4)	(5)
HD 219734	0.26	3720	0.15	−0.35
HD 36003	0.00	4578	5.02	−0.57
HD 187238	−0.05	3604	1.77	0.77
HD 10465	−0.01	3590	0.50	0.60

Fig. B1 shows the best-fitting spectrum (red line) to one of the stars (black line), in CO-dominated regions, as an example. It is clear that the best fit is relatively poor, with residuals at a level of ∼25 per cent in the regions of CO absorption. The derived parameters of the stars are listed in the Table B1. The two stars in CO-strong sequence (HD 10465 and HD 187238) have lower α abundances than their counterparts (with the same $T_{\rm eff}$) in the normal sequence (HD 36003, HD 219734); this is consistent with the behaviour of CO line-strengths in the α-enhanced E-MILES SSP models that decrease with increasing [α/Fe]. However, the derived parameters of these stars with MARCS models are significantly different than the ones from the literature (Table 1), ($\Delta T_{\rm eff} \sim -243$ K, $\Delta \log g \sim 0.2$, Δ[Fe/H] ∼0.1). This experiment clearly shows the current limitations of stellar atmospheric models in the low-temperature regime.

[9] http://github.com/callendeprieto/ferre

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

Downloaded from https://academic.oup.com/mnras/article/512/1/378/6534922 by University of Central Lancashire user on 24 March 2022