

The young stellar content of the giant H II regions M 8, G333.6–0.2, and NGC 6357 with VLT/KMOS^{★,★★}

M. C. Ramírez-Tannus^{1,2}, J. Poorta², A. Bik³, L. Kaper², A. de Koter^{2,4}, J. De Ridder⁴, H. Beuther¹, W. Brandner¹, B. Davies⁵, M. Gennaro⁶, D. Guo², T. Henning¹, H. Linz¹, T. Naylor⁷, A. Pasquali⁸, O. H. Ramírez-Agudelo^{9,10}, and H. Sana⁴

¹ Max Planck Institute for Astronomy, Königstuhl 17, 69117 Heidelberg, Germany
e-mail: ramirez@mpia.de

² Anton Pannekoek Institute for Astronomy, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands
e-mail: j.poorta@uva.nl

³ Department of Astronomy, Stockholm University, Oskar Klein Center, 106 91 Stockholm, Sweden

⁴ Institute of Astronomy, KU Leuven, Celestijnenlaan 200 D, 3001 Leuven, Belgium

⁵ Astrophysics Research Institute, Liverpool John Moores University, Egerton Wharf, Birkenhead CH41 1LD, UK

⁶ Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

⁷ School of Physics, University of Exeter, Stocker Road, Exeter EX4 4QL, UK

⁸ Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, Mönchhofstrasse 12-14, 69120 Heidelberg, Germany

⁹ UK Astronomy Technology Centre, Royal Observatory Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK

¹⁰ German Aerospace Center (DLR), Institute for the Protection of Maritime Infrastructures, Fischkai 1, 27572 Bremerhaven, Germany

Received 23 May 2019 / Accepted 17 November 2019

ABSTRACT

Context. The identification and characterisation of populations of young massive stars in (giant) H II regions provides important constraints on (i) the formation process of massive stars and their early feedback on the environment, and (ii) the initial conditions for population synthesis models predicting the evolution of ensembles of stars.

Aims. We identify and characterise the stellar populations of the following young giant H II regions: M 8, G333.6–0.2, and NGC 6357.

Methods. We have acquired *H*- and *K*-band spectra of around 200 stars using the *K*-band Multi Object Spectrograph on the ESO Very Large Telescope. The targets for M 8 and NGC 6357 were selected from the Massive Young Star-Forming Complex Study in Infrared and X-ray (MYStIX), which combines X-ray observations with near-infrared (NIR) and mid-infrared data. For G333.6–0.2, the sample selection is based on the NIR colours combined with X-ray data. We introduce an automatic spectral classification method in order to obtain temperatures and luminosities for the observed stars. We analysed the stellar populations using their photometric, astrometric, and spectroscopic properties and compared the position of the stars in the Hertzsprung-Russell diagram with stellar evolution models to constrain their ages and mass ranges.

Results. We confirm the presence of candidate ionising sources in the three regions and report new ones, including the first spectroscopically identified O stars in G333.6–0.2. In M 8 and NGC 6357, two populations are identified: (i) OB main-sequence stars ($M > 5 M_{\odot}$) and (ii) pre-main sequence stars ($M \approx 0.5\text{--}5 M_{\odot}$). The ages of the clusters are $\sim 1\text{--}3$ Myr, < 3 Myr, and $\sim 0.5\text{--}3$ Myr for M 8, G333.6–0.2, and NGC 6357, respectively. We show that MYStIX selected targets have $> 90\%$ probability of being members of the H II region, whereas a selection based on NIR colours leads to a membership probability of only $\sim 70\%$.

Key words. H II regions – stars: pre-main sequence – stars: formation – X-rays: stars – infrared: stars

1. Introduction

The assembly of massive stars is a key problem in modern astronomy. From an observational point of view, two lines of research are being pursued to better understand their formation process (Zinnecker & Yorke 2007; Motte et al. 2018). One approach is aimed at the very early stages of star formation, observing the cold cores in which massive stars are being formed. The central objects are surrounded by extended disks

and produce molecular outflows that can be studied at sub-mm and radio wavelengths (Beltrán & de Wit 2016; Cesaroni et al. 2017; Beuther et al. 2017; Ilee et al. 2016). A second strategy is to study the final phase and outcome of the formation process, characterising the recently formed stars and looking for remnant signatures of their assembly.

Here, we adopt the second strategy. Finding and characterising young massive stars is important for two reasons. First, the properties of the newly formed massive stars, such as mass, rotational velocity, magnetic field, and multiplicity provide us with information about the formation process. Second, the statistical distributions of these properties – once a large enough sample is scrutinised – define the initial conditions for population synthesis models in which the evolution of entire populations of massive stars is traced.

* Full Table 4 and normalized spectra are only available at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.128.5) or via <http://cdsarc.u-strasbg.fr/viz-bin/cat/J/A+A/633/A155>

** Based on observations collected at the European Southern Observatory at Paranal, Chile (ESO program 095.C-0048).

As formation takes place in opaque natal clouds dispersed throughout the Galactic disk, severe extinction constitutes an observational challenge when studying these young massive stars at optical or shorter wavelengths. Therefore, the youngest stars are studied at near-infrared (NIR) wavelengths, where the extinction is much lower. The technical development of instrumentation and detectors in this wavelength regime is also an advantage when studying the youngest populations of stars. For example, by studying stellar populations via NIR spectroscopy, [Bik et al. \(2012, 2014\)](#) present evidence of an age spread (2–3 Myr) in the giant H II region W3 Main, where the most massive star is $35 M_{\odot}$ (O5.6–O7.5V). They conclude that the high UV flux from this and other massive stars affects the disk fraction of the lowest mass stars. [Wu et al. \(2016\)](#) analysed 14 O-type stars with masses between 20 and $120 M_{\odot}$ in the giant H II region W49, and concluded that the cluster age is around 1.5 Myr. [Ramírez-Tannus et al. \(2017\)](#) performed a study of eleven massive young stellar object (mYSO) candidates in the giant H II region M17, with masses ranging from 6 to $25 M_{\odot}$ and a likely age < 1 Myr, and confirmed the pre-main-sequence (PMS) nature of six objects. This constitutes a small but unique sample of massive stars observed just after their formation in a very young cluster. Interestingly, they observed a lack of double-lined spectroscopic binaries and measured a narrow range of measured radial velocities ($-10 < v_{\text{rad}} < 20 \text{ km s}^{-1}$) which is in stark contrast to the overall properties of small and large samples of fully formed, main-sequence massive stars (e.g. [Mason et al. 2009](#); [Sana et al. 2014](#); [Kiminki & Kobulnicky 2012](#); [Peter et al. 2012](#); [Kobulnicky et al. 2014](#); [Dunstall et al. 2015](#)). The samples studied in these works are characterised by large radial-velocity amplitudes ($\Delta v_{\text{rad}} > 100 \text{ km s}^{-1}$), from which one can conclude that 40% to 50% of OB-type binaries have periods of one month or less.

Spectroscopic studies aiming to classify (massive) stars in star-forming regions have the problem that cluster membership determination using NIR photometry suffer from pollution by fore- and background stars (typically $\sim 50\%$ [Bik et al. 2012](#); [Wu et al. 2016](#)). The Massive Young Star-Forming Complex Study in Infrared and X-ray (MYStIX, [Feigelson et al. 2013](#)) applied a novel approach in which X-ray, NIR, and mid-IR (MIR) observations are combined. Since X-rays are tracers of PMS low- and intermediate mass stars (through emission produced by coronal activity), and of massive stars (through emission produced by shocks in their stellar winds or colliding winds), this approach allows for a more efficient removal of fore- and background stars and hence more robust constraints on cluster membership. NIR multiplexed instruments like the *K*-band Multi Object Spectrograph (KMOS) on the ESO Very Large Telescope allow for the simultaneous acquisition of tens of spectra in complex environments. The combination of this observing strategy with input catalogues containing MYStIX probable complex members (MPCM; [Broos et al. 2013](#)) provides an efficient way of obtaining spectra of young massive stars in regions with high extinction. *Gaia* DR2 ([Gaia Collaboration 2016, 2018](#)) information on distances and proper motions, as well as spectroscopic followup, provides a reliable way to assess cluster membership. We follow this approach here as well, and cross correlate the MYStIX and NIR photometry methods to assess their reliability. The latter methodologies are the main ones to use at large distances, where *Gaia* becomes unreliable or unfeasible.

In this paper we present *H* and *K* band spectroscopic observations performed with VLT/KMOS surveying three giant H II regions (M8, G333.6–0.2, and NGC 6357) in order to characterise their stellar populations. The three star forming regions

Table 1. Properties of the observed H II regions.

H II region	α (J2000)	δ (J2000)	Diameter (')	Distance (pc)
M8	18 ^h 03 ^m 41 ^s .6	−24°22′43″	10	1336 ⁺⁷⁶ _{−68} ^(a)
G333.6–0.2	16 ^h 22 ^m 12 ^s .3	−50°05′56″	7	2540 ± 0.71 ^(b)
NGC 6357	17 ^h 24 ^m 14 ^s .8	−34°20′06″	27	1770 ⁺¹⁴⁰ _{−120} ^(a)

References. ^(a)[Kuhn et al. \(2019\)](#). ^(b)This work.

observed, presented in Sect. 2, were selected to sample a range in cluster morphologies. In Sect. 3 we provide an overview of the target selection procedure. In Sect. 3.3 we describe the KMOS observations and in Sect. 3.4 the data reduction procedure. In Sect. 4 we present the stellar populations of the three giant H II regions by placing them in NIR and MIR colour-magnitude (CMD) and colour-colour diagrams (CCD). In Sect. 5 we match our sources with the *Gaia* DR2 catalogues in order to use their astrometric properties to further assess cluster membership. In Sect. 6 we describe the mYSO detected in our sample. We present a detailed NIR spectral classification of the massive stars and the PMS population in Sect. 7. With these results we confirm or reject the massive star nature of the candidates presented by [Povich et al. \(2017\)](#). We combine all the results to place the stars in Hertzsprung–Russell diagrams (HRDs) and obtain an estimate of the age of the regions and the mass range of the observed stars in Sect. 8. We discuss our results in Sect. 9 and summarise our findings in Sect. 10.

2. Sample of giant H II regions

We observed three H II regions with different mass ranges and physical sizes: G333.6–0.2, M8, and NGC 6357 (Table 1). The location of the selected targets is shown in Figs. 1–3. We have colour-coded the targets according to their temperature derived from our spectral classification (see Sects. 7 and 8). The candidate mYSOs (Sect. 6) are labelled as magenta diamonds. Below, we describe each region in some detail.

2.1. M8

Also known as the Lagoon Nebula (Fig. 1), the M8 H II region spans a $\sim 50' \times 40'$ region in the plane of the sky and is located at a distance of about 1.3 kpc in the Sagittarius–Carina arm in front of the Galactic centre region. It is associated with the open cluster NGC 6530 containing several O-type stars and about 60 B stars ([Feigelson et al. 2013](#)). The principal source of ionising flux in the H II region is the O4 star 9 Sgr ([Tothill et al. 2008](#)). The western part of the nebula is concentrated into a compact H II region with much denser ionised gas than the main body of the Lagoon Nebula. This region is called the Hourglass Nebula and it is powered by the O7 star Herschel 36. Since M8 is close to the Galactic plane and projected on top of the Galactic bulge, it is difficult to distinguish between the cluster members and foreground and background stars. Many molecular clumps are scattered through the nebula, suggesting that star formation is triggered by ionisation fronts produced by 9 Sgr and Herschel 36 ([Tothill et al. 2002](#)). The MPCM catalogue for M8 contains 2056 sources ([Feigelson et al. 2013](#)). Figure 1 shows a 2MASS colour image (*J*, *H*, and *K*) of the region. The western compact H II region around Herschel 36 is visible as diffuse emission. Overplotted as coloured dots are the cluster members

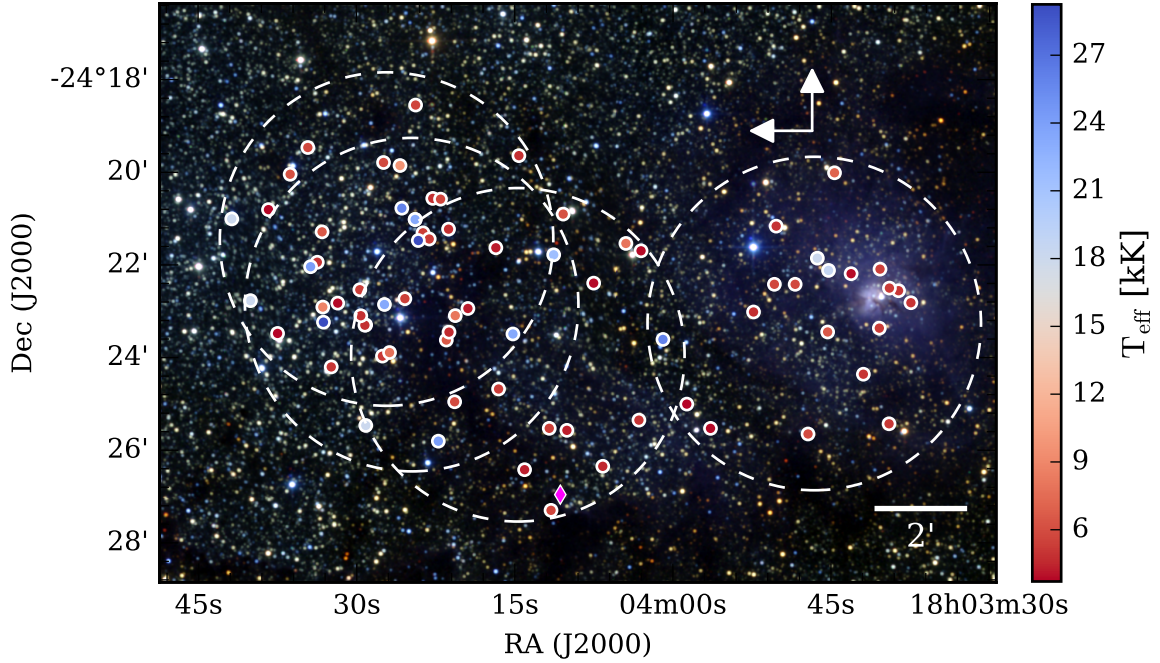


Fig. 1. 2MASS NIR colour image of the Lagoon Nebula (M8) with the 81 KMOS targets over-plotted. North is up, east is to the left. The white dashed circles show the $7.2'$ field-of-view of KMOS for each of the pointings. The colour of the targets represents the effective temperature according to their spectral type (see Sect. 8). The objects for which no spectral classification has been obtained are coloured grey. The magenta diamond corresponds to a massive YSO. The number of KMOS targets in this region is 81.

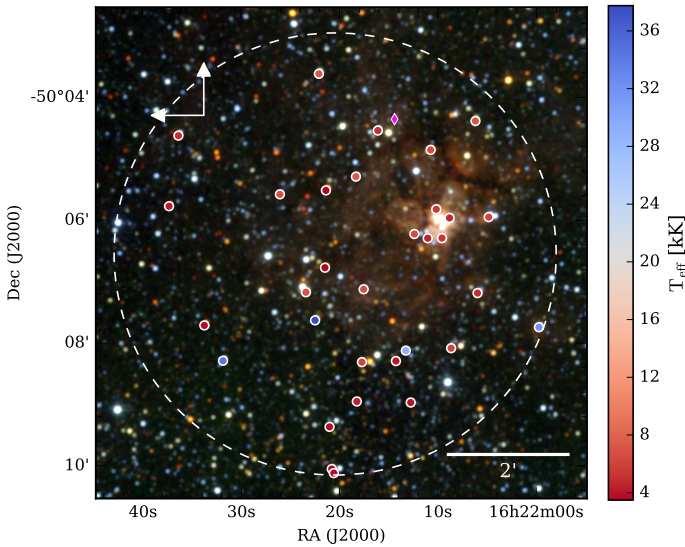


Fig. 2. Same as Fig. 1 for G333.6–0.2. The number of KMOS targets in this region is 39.

for which we obtained KMOS spectra; their colour represents the effective temperature according to their spectral classification.

Arias et al. (2007) performed intermediate resolution spectroscopy of PMS stars in M8. They identified 27 classical T Tauri stars, seven weak-lined T Tauri stars and three PMS emission objects with spectral type G. They identified the stars to be younger than 3 Myr and between 0.8 and $2.5 M_{\odot}$. Prisinzano et al. (2007, 2019) identified 237 PMS stars as members including 53 binaries. Kumar & Anandarao (2010) identified 64 class 0/I and 168 class II objects using *Spitzer* NIR photometry.

2.2. G333.6–0.2

G333.6–0.2 (Fig. 2) is an ultra-compact H II (UCHII) region embedded in the giant H II region RCW106. It is the most prominent massive star forming region in the G333 complex. Located at a distance of ~ 2.6 kpc (Figuerêdo et al. 2005), it is compact ($11''$ in diameter) and dusty, and hosts powerful ionising stars (Kumar 2013). G333.6–0.2 is embedded in a massive molecular clump of $15\,000 M_{\odot}$ (Mookerjea et al. 2004). Townsley et al. (2014, 2018) detected more than 100 X-ray sources clustered in the central region, and over 500 sources in the field of view. The central sources would be responsible for powering the H II region. Townsley et al. (2018) also detected diffuse X-ray emission that they associated with plasma flowing out of the embedded cluster. To our knowledge, there are no age estimates for G333.6–0.2 reported in the literature.

2.3. NGC 6357

NGC 6357 is a major star-forming region in the Sagittarius-Carina spiral arm (Fig. 3). Its brightest H II region is G353.1+0.9 located near the northern rim of the cloud and ionised by the massive stellar cluster Pis 24 (Feigelson et al. 2013). Maíz Apellániz et al. (2007) determined that Pis 24 contains two very massive stars of about $100 M_{\odot}$ each. The $H\alpha$ image of NGC 6357 shows a shell containing a Wolf-Rayet star; this could be a super-bubble produced by an earlier generation of massive stars. This hypothesis is supported by the discovery of an older X-ray population around Pis 24 (Wang et al. 2007). Massi et al. (2015) found that Pis 24 is very young (~ 1 – 3 Myr) and that it probably consists of a few sub-clusters. They found evidence that the most massive stars triggered the current star formation episode. The MPCM catalogue for this region contains 2235 sources (Feigelson et al. 2013).

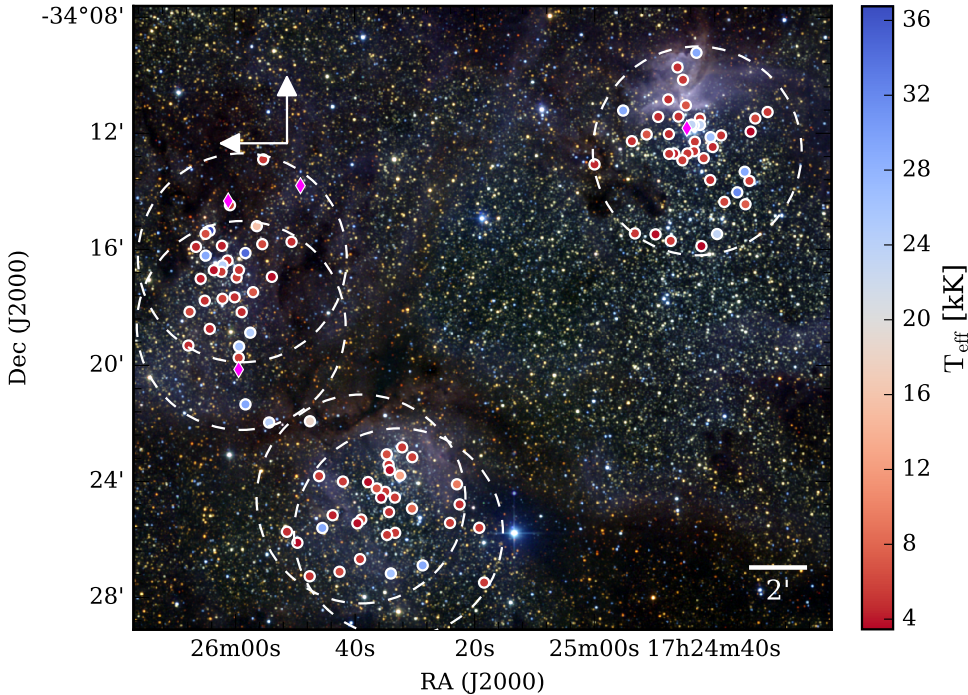


Fig. 3. Same as Fig. 1 for NGC 6357. The number of KMOS targets in this region is 123.

3. Observations and data reduction

The observations presented here were obtained between April and August 2015 using the KMOS (Sharples et al. 2013) mounted at the ESO VLT. For M8 and NGC 6357 the target selection was based on the results of the MYStIX team (Feigelson et al. 2013; Broos et al. 2013; Povich et al. 2017). G333.6–0.2 is not covered by MYStIX, therefore we adopted a different method in order to select the targets.

3.1. Target selection in M8 and NGC 6357

M8 and NGC 6357 are included in the MYStIX project (Feigelson et al. 2013; Broos et al. 2013) which provides a selection of MPCM based on the cross-matching of three samples: (i) X-ray sources with IR counterparts, variability, and/or spatial distribution consistent with being PMS stars; (ii) MIR excess sources consistent with having circumstellar material (i.e. possible YSOs with or without X-ray counterparts); and (iii) known OB stars from Broos et al. (2013), which do not necessarily have X-ray counterparts.

As the MPCM catalogue contains many more sources than we can possibly observe with KMOS, we applied several selection criteria to obtain a prioritised list of targets. We limited our sample to objects with $K_s < 13.0$ mag. Due to possible saturation of the KMOS detector we could not observe sources brighter than $K_s = 9.0$ mag. The main aim of the project is to classify the stellar spectra in order to derive the overall characteristics of the host regions. Therefore, our selection focuses on stars without infrared excess, as these offer the best chance of a direct NIR detection of the stellar photosphere. We identify three priority classes as follows.

As *Priority one* targets we selected all stars listed as known OB stars in the MPCM catalogue. We added a sample of MYStIX candidate OB stars from Povich et al. (2017). These candidate X-ray emitting OB stars are identified via two methods: (1) By fitting Castelli & Kurucz (2003) models to the

1–8 μm spectral energy distribution and then selecting all objects with $L > 10^4 L_\odot$, which approximately corresponds to main sequence B1 stars or earlier; (2) By inspecting the NIR CCD and CMD for the regions and identifying sources with X-ray detection that are reddened main-sequence stars according to their position in the CCD and that have dereddened J -band magnitudes $M_J \leq -2.4$ mag. By performing SED fitting, Povich et al. (2017) reported five new candidate OB stars (of which we observed two) in addition to the 28 previously published ones (of which we observed 12) in M8 (see Table 5). In NGC 6357 there were 16 previously published OB stars (of which we observed 4). Povich et al. (2017) increased this sample with 26 new candidate OB stars (of which we observed 17): 25 of those selected via SED fitting and one more selected via its NIR colours (see Table 7).

For the *Priority two* targets, we constructed colour-colour diagrams based on the MPCM JHK_s photometry, and selected the sources without infrared excess. This is similar to method two described above, but without a J -band magnitude limit. The targets already included under priority one and those with $9 < K_s < 13$ mag are removed from this subsample. In M8 we observed 55 priority two sources and in NGC 6357 82.

As low priority sources (*Priority three*) we added the infrared excess sources rejected in the selection of the priority two sources. We observed 12 sources in M8 and 20 in NGC 6357.

3.2. Target selection in G333.6–0.2

In the case of G333.6–0.2 we selected 23 stars by matching the X-ray observations of this region with the VISTA DR2 catalogue. We carried out the target selection using the *Chandra* X-ray catalogue of Townsley et al. (2014), but in Table 4 we report matches to the more recent catalogue of Townsley et al. (2018). This does not change any of the pairings but it does update the X-ray catalogue names.

Many of our objects are saturated in VISTA DR2, and so we replaced any J , H , or K magnitudes brighter than 13 mag with those of the nearest 2MASS star within $3''$. Both the VISTA

and 2MASS magnitudes should be viewed with some caution as they are often for objects which are crowded or superimposed on nebulosity. We did not replace the positions of saturated stars since VISTA is less susceptible to crowding than 2MASS, and VISTA positions remain relatively precise to quite bright magnitudes (we found that in the range $9 < K < 10$ mag 66% of 2MASS-VISTA matches are within $0.15''$ of each other in this region).

The matches were carried out using the Bayesian method described in Naylor et al. (2013), which assesses the likelihood of an X-ray/IR pairing using the uncertainty in position assigned to each X-ray source by the source extraction. It also allows for the fact that brighter sources in the IR are more likely to be the counterparts of a given X-ray source, and reduces the probability of a match if there are other possible counterparts. The mean probability of an infrared star being the counterpart is 97%, with the lowest being 82%.

We complemented this sample by adding 18 stars from the 2MASS catalogue with $H - K > 1$ mag in order to select some of the sources with high extinction. Also, these targets were split in three priority classes, where priority one contains 19 X-ray matched sources without NIR excess. Four X-ray matched sources show an infrared excess and are added as priority two sources. Finally, 18 2MASS sources without an X-ray match are the priority three sources.

3.3. VLT/KMOS observations

KMOS consists of 24 deployable integral field units (IFUs) and three spectrographs (Sharples et al. 2013). The IFUs are divided into three groups and the light is redirected into one of three spectrographs, each spectrograph disperses the light of eight IFUs. We divided the H II regions in several fields and for each field we took a long (300 s) and a short (30 s) exposure in order to cover the faint ($13 < K < 11$ mag) and bright ($K < 11$ mag) targets, respectively. For each exposure we applied a 9-point dither pattern both for the H - and the K -band grating. The fields observed per H II region are shown in Figs. 1–3. In the case of G333.6–0.2 and the upper right field in NGC 6357 we observed the same field twice, once for the bright and once for the faint stars. We used one IFU per spectrograph (three IFUs in total) to perform sky observations.

A set of calibration images was taken in addition to each of the observations. This set consists of dark frames, flat fields and an arc-lamp exposure; the dark and arc-lamp frames are taken at six different rotator angles. A telluric standard star was observed in one IFU per spectrograph.

3.4. Data reduction

The KMOS observations were reduced using the KMOS pipeline version 3.12.3 (Davies et al. 2013) in combination with the ESO Recipe Execution Tool (esorex). These routines were used to calibrate and reconstruct the cubes of the science and standard-star observations as described by Davies et al. (2013). The sky subtraction was performed using the standard KMOS routines, and the telluric correction was done using the default template for telluric correction: one telluric star is observed in one IFU for each of the three spectrographs. This telluric spectrum is then used to perform the telluric correction for all stars in that spectrograph. The data reduction performed by the pipeline can be divided in three basic steps: (i) flat field correction, (ii) extraction of the spectra, and (iii) telluric correction, instrumental effects correction, and flux calibration.

The first and third steps use procedures that are standard for NIR spectroscopy. Given that we are doing integral field spectroscopy, the second step is different from the standard procedure for other instruments. In the case of KMOS, the reconstruction of the 3D data cubes is performed by interpolating from look-up tables which provide the (x, y, λ) location in the final cube for each pixel on the detector. The interpolation is performed based on weighted averages of neighbouring points; therefore the weighted average cannot have a higher value than the neighbour pixel with highest flux. This procedure is problematic for observations of spatially compact, but spectrally extended objects, like our targets. The spatial axes are in that case under-sampled, resulting in (at times quasi-periodic) ripples which appear in the spectra after interpolation and reconstruction. This problem is described in detail in Sect. 4.2 of Davies et al. (2013). The ripple pattern is prominent and different for each star and is severe in some cases, as can be seen in the spectra displayed in Fig. 14 and at the CDS. For these reasons, we were not able to remove the ripples from the spectra. In the cases where the ripples are severe, they hamper our ability to extract physical information from the spectral lines.

The spectra have a resolving power R between 3570 and 4880 over the 1.46 – $2.40 \mu\text{m}$ range, corresponding to ~ 80 – 60 km s^{-1} respectively, per resolution element. Lapenna et al. (2015) performed a series of tests of the reliability of the KMOS pipeline wavelength calibration. They computed the radial velocity shifts from spaxel to spaxel in one IFU as well as the shifts for different IFUs. They conclude that, both within each IFU and among different IFUs, the wavelength calibration obtained from the KMOS pipeline is accurate at a level of a few km s^{-1} . They establish that the wavelength calibration by the KMOS pipeline is well suited for bulk kinematic studies, but if one would like to measure the radial velocity of an individual star, the wavelength calibration should be refined.

The normalisation procedure for the OBA stars was performed in two steps. We first fitted a second degree polynomial to all the spectra. Then we visually inspected all the spectra and selected the ones of sufficient quality (not too many ripples and detectable lines, which corresponds to signal-to-noise ratio, $S/N > 10$) to be classified spectroscopically. From this first step we could also distinguish the late-type objects and the YSOs from the OBA stars. We then re-normalised the OBA stars using the interactive spectrum normaliser `specnorm.py`¹ which allows to select continuum points of the spectrum interactively and fits a cubic spline through the selected points.

4. Photometric data

The photometric data were taken from the MYStIX catalogue (Feigelson et al. 2013; King et al. 2013; Kuhn et al. 2013) which includes the JHK -band magnitudes from the United Kingdom Infra-red Telescope (UKIRT) wide field camera (Lawrence et al. 2007; Lucas et al. 2008) or from the 2MASS point source catalogue (Skrutskie et al. 2006), plus the four *Spitzer* IRAC bands (GLIMPSE; Benjamin et al. 2003). In the CDS tables described in Table 4 we list the photometry for each of the observed objects.

The near- and mid-IR CMDs and CCDs for each of the regions are shown in Figs. 4–6. In these diagrams too, we have colour coded the sources according to the spectral classification described in Sect. 7. The diagrams show the zero-age main

¹ <https://github.com/ishivvers/astro/blob/master/specnorm.py>

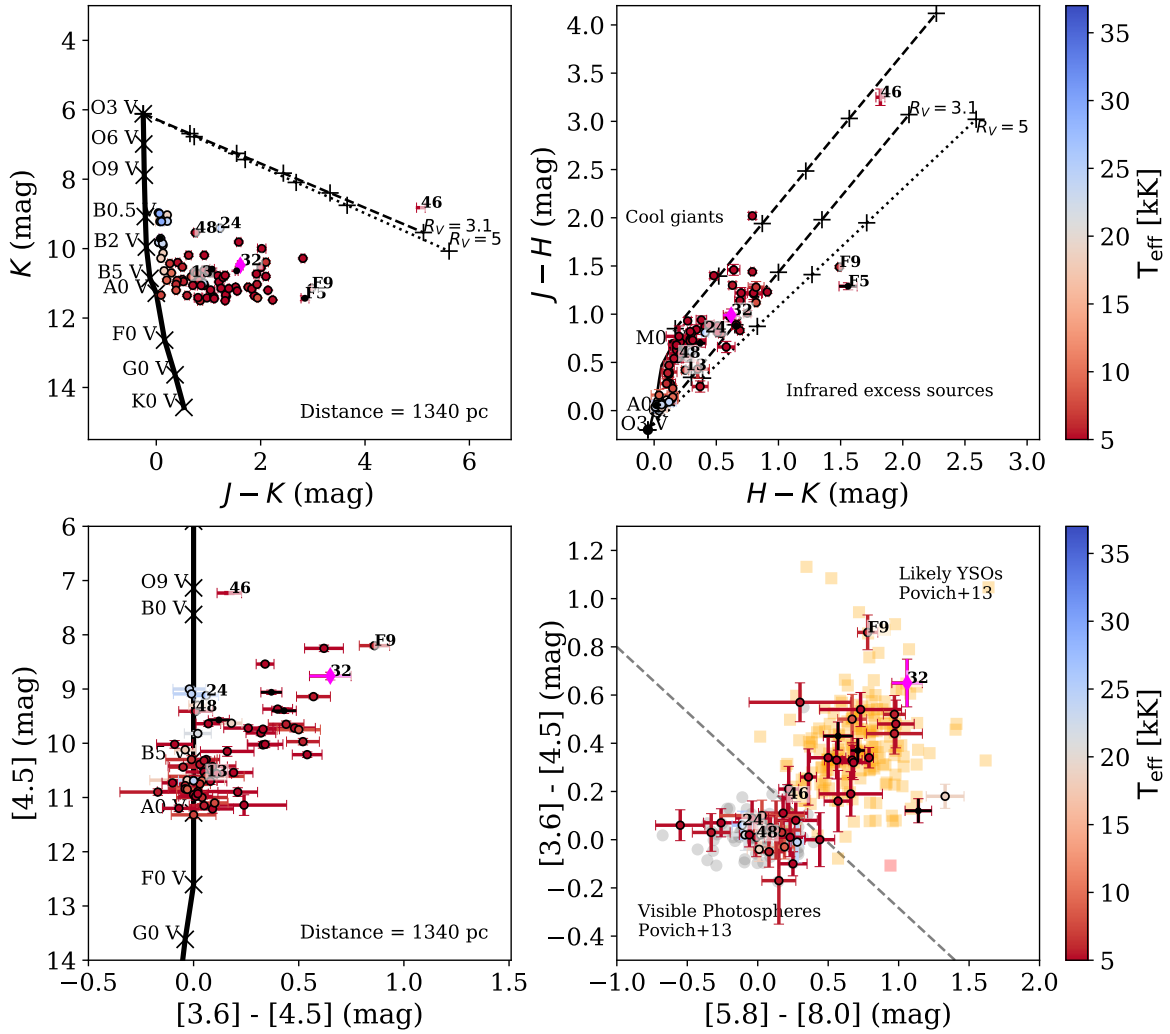


Fig. 4. NIR (*top*) and MIR (*bottom*) colour-magnitude (*left*) and colour-colour (*right*) diagrams for M8. The objects have been colour-coded based on their effective temperature as determined through spectral classification (Sect. 7). The magenta diamonds are the objects with CO bandheads in emission and the black dots the unclassified objects. The specific objects mentioned in the text are labelled with their respective names. The location of the zero-age main sequence at the distance of M8 is shown with a solid black line (Hanson et al. 1997; Cox 2000). The black-dashed line in the *top-left* panel represents the reddening line for an O3 V star for $R_V = 3.1$. *Top-right* panel: reddening lines of an O3 V (left) and a M0 V (right) star. In both cases the crosses on top of the reddening lines correspond to A_V values of 5, 10, 15, 20, and 30 mag. The grey-dotted lines show the reddening line of an O3 V star for $R_V = 5$. The dashed line in the *bottom-right* panel divides the CCD into the areas dominated by YSOs (grey dots) and visible photospheres (orange squares) in the MIREs catalogue (Povich et al. 2013).

sequence (solid line) and reddening lines (dashed and dotted lines). We adopted the extinction law by Cardelli et al. (1989) with the total-to-selective extinction value R_V set to 3.1 (dashed lines) and 5 (dotted lines). In the NIR diagram one can distinguish a reddened main sequence ($A_V \sim 2$ mag in M8 and $A_V \sim 5$ mag in NGC 6357) and a population of lower mass pre-main-sequence stars.

For G333.6–0.2 we could not rely on the MYStIX catalogue. We used the *JHK* CMD and CCDs of this region in order to check for membership of our selected targets. The spread of the objects on the CMD and the CCD is sizeable and shows that in G333.6–0.2 the extinction varies strongly from sight line to sight line.

In the MYStIX IR Excess Source catalogue (MIREs) catalogue compiled by Povich et al. (2013) the authors identified probable YSO members of the MYStIX regions by combining IR SED fitting, IR colour cuts, and spatial clustering analysis. In the MIR CCDs of M8 and NGC 6357 we plot the likely YSOs and IR excess sources from the MIREs catalogue (orange

squares) as well as the sources that Povich et al. (2013) identify as having visible photospheres (grey dots) in order to identify the region in the CCDs that is dominated by YSOs. G333.6–0.2 is not covered by the MIREs catalogue, so we used the cut found in M8 and NGC 6357 to show the YSO dominated region in the MIR CCD. In the bottom-right panels of Figs. 4–6 the cut used to separate YSOs and visible photospheres is shown with a dashed-grey line.

The CCD and CMD diagrams presented here are further discussed in Sect. 7.7, where they are also used to confirm or reject cluster membership.

5. Astrometric properties of our sample

We cross-matched our catalogue with the *Gaia* DR2 archive (Gaia Collaboration 2016, 2018; Lindegren et al. 2018). Because of the strong and highly variable extinction in the HII regions, we did not use the magnitude as a cross-match parameter, but relied on sky coordinates only, demanding a

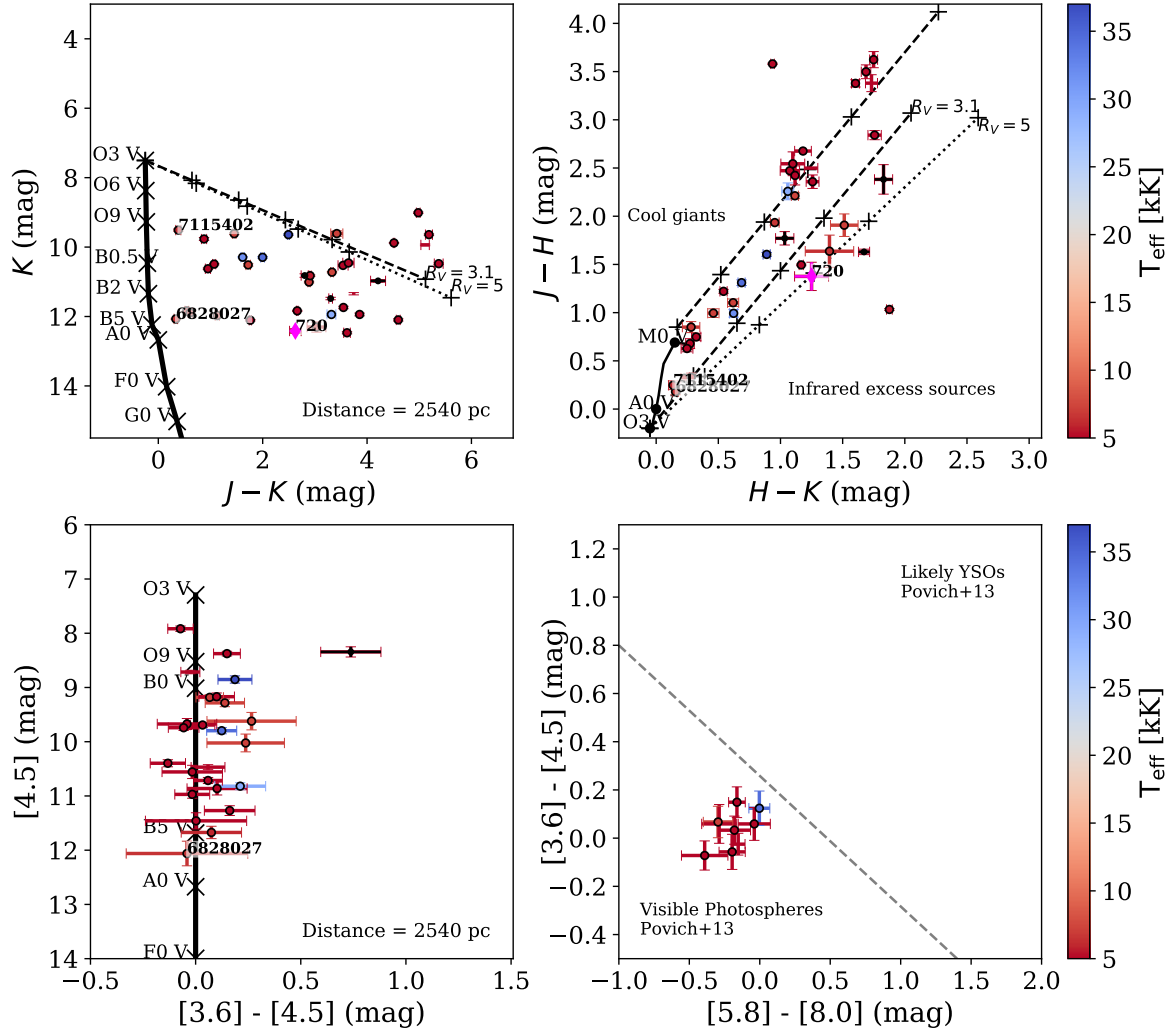


Fig. 5. Same as Fig. 4 for stars in G333.6–0.2.

separation no larger than $1''$. In addition we limited ourselves to $G \leq 19$ to reduce the risk of false matches with very faint stars. If more than one match was found within $1''$, the brighter one was chosen. In addition, we extracted all *Gaia* sources with $G \leq 19$ within a radius of 0.5° from the central coordinates given in Table 1. The astrometric properties for the stars in the three clusters are shown in Fig. 7. The sky positions of the *Gaia* sources are shown with blue, orange, and red dots in the first column of this figure. The grey contours represent the density of *Gaia* sources in the local neighbourhood and the black dots show all MYStIX sources (in M 8 and NGC 6357) and X-ray sources (in G333.6–0.2) that have *Gaia* counterparts within $1''$. We note that for G333.6–0.2 only 12 sources of our catalogue have *Gaia* counterparts that are brighter than $G = 19$.

The parallaxes (ϖ) and the proper motions ($\mu_{\alpha^*}, \mu_{\delta}$) of the *Gaia* sources are shown in columns two to four of Fig. 7. In blue we show our catalogue stars, and the red and orange dots correspond to the stars among our catalogue stars that are probable outliers based on their astrometric properties ($\varpi, \mu_{\alpha^*}, \mu_{\delta}$). We checked whether these properties are sufficiently reliable to reject their membership. We computed the renormalised unit weight error (RUWE) which quantifies the quality and the reliability of the astrometric fit (Lindgren 2018). Following the suggestion of the latter authors, we retain only those sources for which $\text{RUWE} < 1.4$ as astrometrically reliable.

For each of the clusters we find two outliers (plotted in orange in the Fig. 7) in $(\varpi, \mu_{\alpha^*}, \mu_{\delta})$ space among the cluster stars in our catalogues, which have nevertheless a reliable *Gaia* DR2 astrometric solution. For M 8 these are sources F13 (2MASS J18035006–2421079) and F24 (2MASS J18041025–2420524). For G333.6–0.2 the sources are 515756828027 (2MASS J16222585–5005334) and 515757115402 (2MASS J16220464–5005556). For NGC 6357 the astrometric solution of sources 91 and 8 (2MASS J17243100–3411143) indicate that the stars may not be cluster members.

The *Gaia* distances computed using the confirmed members (blue dots) are 1.34 ± 0.07 kpc for M 8 and 1.77 ± 0.12 kpc for NGC 6357, which are in agreement with the distances derived by Binder & Povich (2018) and Kuhn et al. (2019) using *Gaia* parallaxes of more complete samples. We derive a distance of 2.54 ± 0.71 kpc for G333.6–0.2 which is in agreement with the reported distance to this region by Figuerêdo et al. (2005).

6. Massive young stellar objects

In the sample selection we gave priority to objects with no IR excess. Nevertheless, our KMOS sample includes six objects that we classified as candidate mYSOs: object 32 in M 8, objects 33, F57, 114, and F68 in NGC 6357, and object 720 in G333.6–0.2. Their position in the CMD and CCD (Sect. 4) is

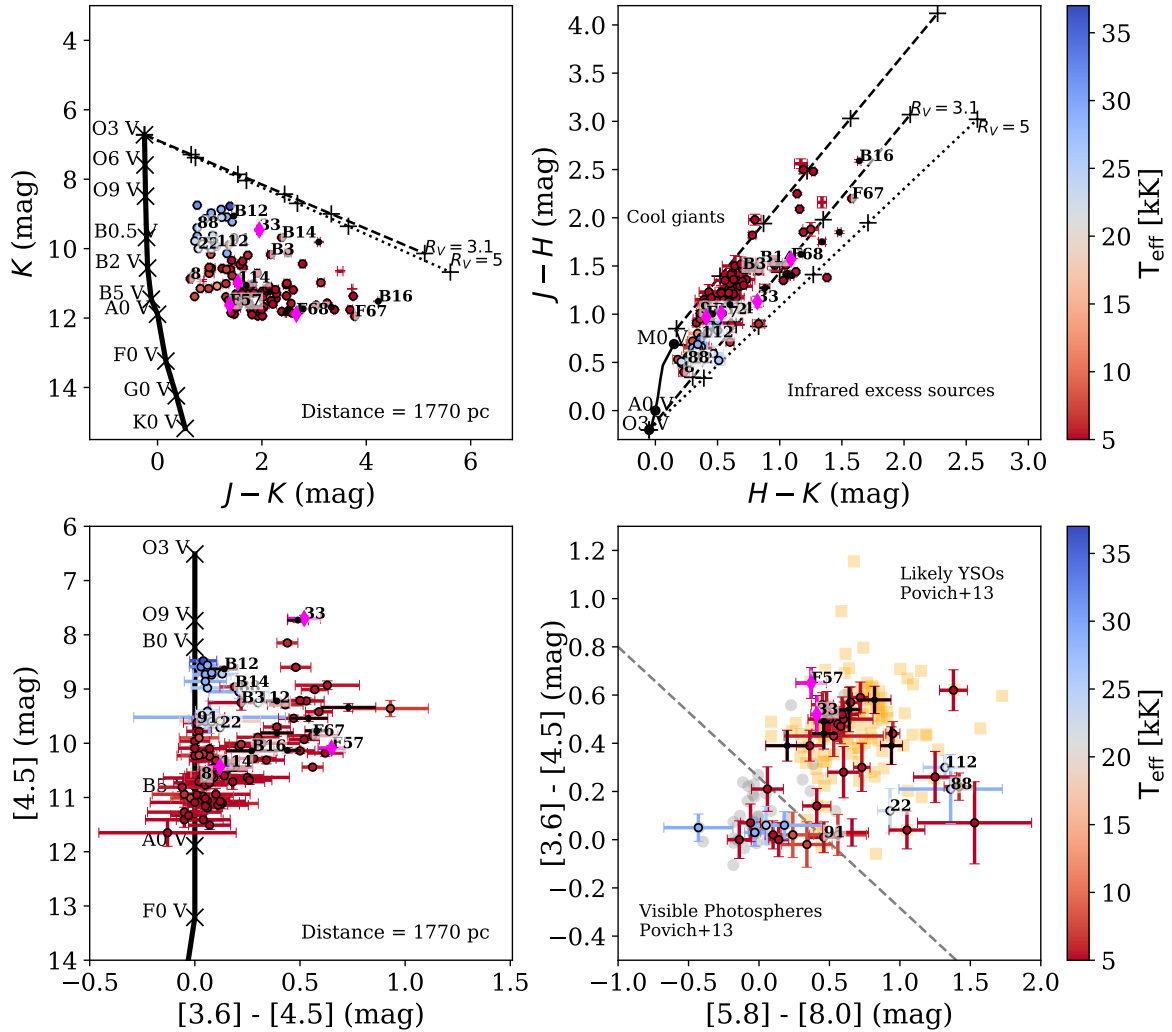


Fig. 6. Same as Fig. 4 for our stars in NGC 6357.

consistent with a mYSO and/or class II nature. We classify these objects based on the presence of double-peaked emission lines and/or CO overtone emission, indicative of a rotating circumstellar disk. CO overtone emission is produced in high-density (10^{10} – 10^{11} cm^{-3}) and high-temperature (2500–5000 K) environments (Charbonneau 1995). CO is easily dissociated, and therefore must be shielded from the strong UV radiation coming from the star. These conditions occur in the inner regions of accretion disks, which makes the CO bandheads an important tool to trace the inner disk structure around mYSOs (e.g. Bik et al. 2006; Ilee et al. 2013; Ramírez-Tannus et al. 2017). The shape of the CO lines probes the Keplerian rotation profile of the circumstellar disk (Bik & Thi 2004; Blum et al. 2004; Davies et al. 2010; Wheelwright et al. 2010; Ilee et al. 2013). The blue shoulder in the bandheads is a measure of the inclination of the disk: an extended blue shoulder indicates a high inclination angle (i.e. near “edge-on” view).

In Fig. 8 we plot the Brackett series profiles and the CO first overtone bandheads for the six candidate YSOs in our KMOS sample. Here we provide a description of the individual mYSO candidates:

NGC 6357-F68. We do not detect any CO bandheads in this object, but the Brackett (Br) lines are in emission. Bry is very strong in emission on top of a photospheric absorption line. Br-12, Br-11, and Br-10 appear to be contaminated by residuals

of the data reduction. This source is not part of the MIREs catalogue (Povich et al. 2013).

G333.6–0.2-720. CO bandhead emission is not present. The Brackett lines are in emission superposed on an absorption line; the emission is very narrow and likely has a nebular origin. In order to confirm the presence of a circumstellar disk, more and better signal-to-noise observations are needed.

NGC 6357-114. This object does not show CO bandheads in emission, but all the Brackett lines show clearly double-peaked emission profiles. This is the only object in our sample where the double-peaked structure is resolved. This source is not in the MIREs catalogue.

NGC 6357-F57. The CO bandheads are clearly present in emission and they show a pronounced blue shoulder. The Br lines seem to be filled in by disk emission. Bry shows a central emission, but the most pronounced part of this feature is very narrow and therefore is likely caused by residuals from the sky subtraction. This source is part of the MIREs catalogue and it is classified as a stage II YSO, dominated by optically thick circumstellar disks.

M8-32. The CO bandheads are clearly detected. The Br lines in this object are in emission; Br-11 and Br-10 seem to be contaminated by residuals of the data reduction process and Br-12

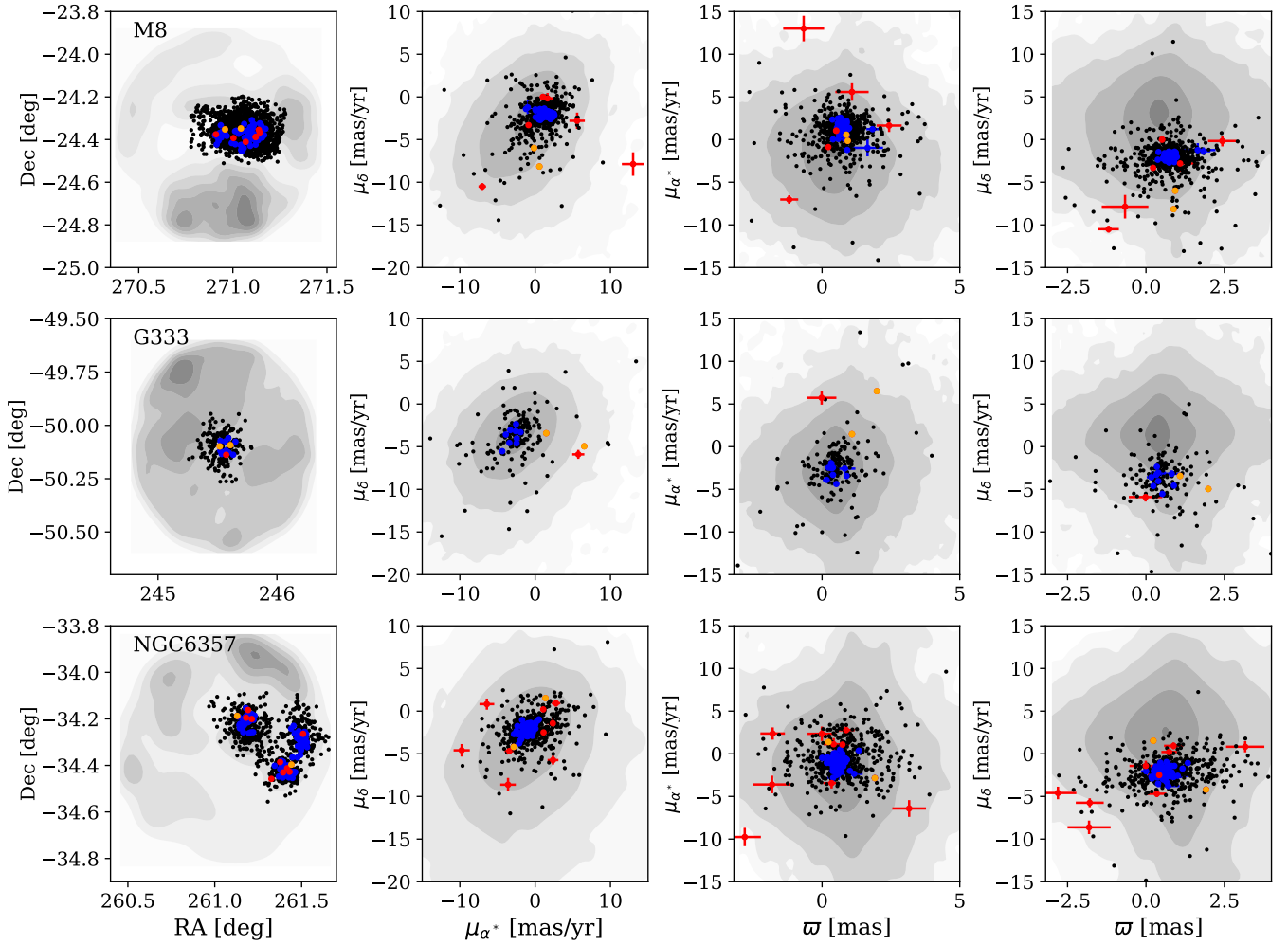


Fig. 7. Astrometric properties of the here studied cluster stars for M 8 (*up*), G333.6–0.2 (*centre*) and NGC 6357 (*bottom*). The blue, red and orange correspond to stars in our catalogues. The blue, red and orange dots correspond to the stars in our catalogue for which a reliable cross-match was found. The blue dots are confirmed cluster members, and the red and orange dots are candidate non-cluster members for which we verified the quality of the astrometric fit. The sources in orange have a reliable *Gaia* DR2 astrometric solution and therefore we confirm that they are not cluster members; for the sources in red we do not confirm nor reject cluster membership. The grey contours represent the density of *Gaia* DR2 sources brighter than $G = 19$ within a radius of 0.5° from the central coordinates listed in Table 1 and the black dots show all the MYStIX sources with *Gaia* counterparts within $1''$. The *first column* shows the equatorial *Gaia* sky positions. *Columns two to four* show the parallax ϖ vs the proper motion ($\mu_{\alpha^*}, \mu_{\delta}$).

and Br γ show emission. The Br lines are single-peaked and the CO bandheads lack a pronounced blue shoulder indicative of a high inclination angle of the disk. This source is part of the MIREs catalogue and it is classified as a stage II YSO, dominated by optically thick circumstellar disks.

NGC 6357-33. This object shows weak CO bandheads in emission. The Br lines seem absent, but this could be due to veiling. The mYSO nature of this object is unclear. This source is not in the MIREs catalogue.

7. Spectral classification

The spectral classification of the 243 observed stars was carried out on the basis of their KMOS H and K band spectra, spanning the wavelength range $1.5\text{--}2.4\ \mu\text{m}$. By a first visual inspection of the objects with no spectral signatures of them being mYSOs, we checked for the absence or presence of CO absorption bands; in this way all stars were categorised as either early (OBA) or late (FGKM) type, respectively (or both in case of doubt). For a

more precise classification two spectral libraries were employed: (1) 56 normalised, O3 to mid-B type, H and K band spectra from Hanson et al. (1996, 2005) and Bik et al. (2005) for the early types; and (2) 190 flux-calibrated, F0 to M9 type, H and K band spectra from Rayner et al. (2009)² for the late types. Both libraries span all luminosity classes. We note that for our three giant H II regions, given their young age, the late type stars probably correspond to low- and intermediate-mass PMS stars.

After initial separation in early and late types the spectral classification was done by two independent methods. On the one hand all spectra were classified by visual inspection. On the other hand, a computer code was developed to automatically classify spectra in each category (early or late), making use of the same spectral libraries as for the visual inspection. We now discuss each method separately and subsequently compare their results.

² These template spectra can be found in http://irtfweb.ifa.hawaii.edu/~spex/IRTF_Spectral_Library/index.html

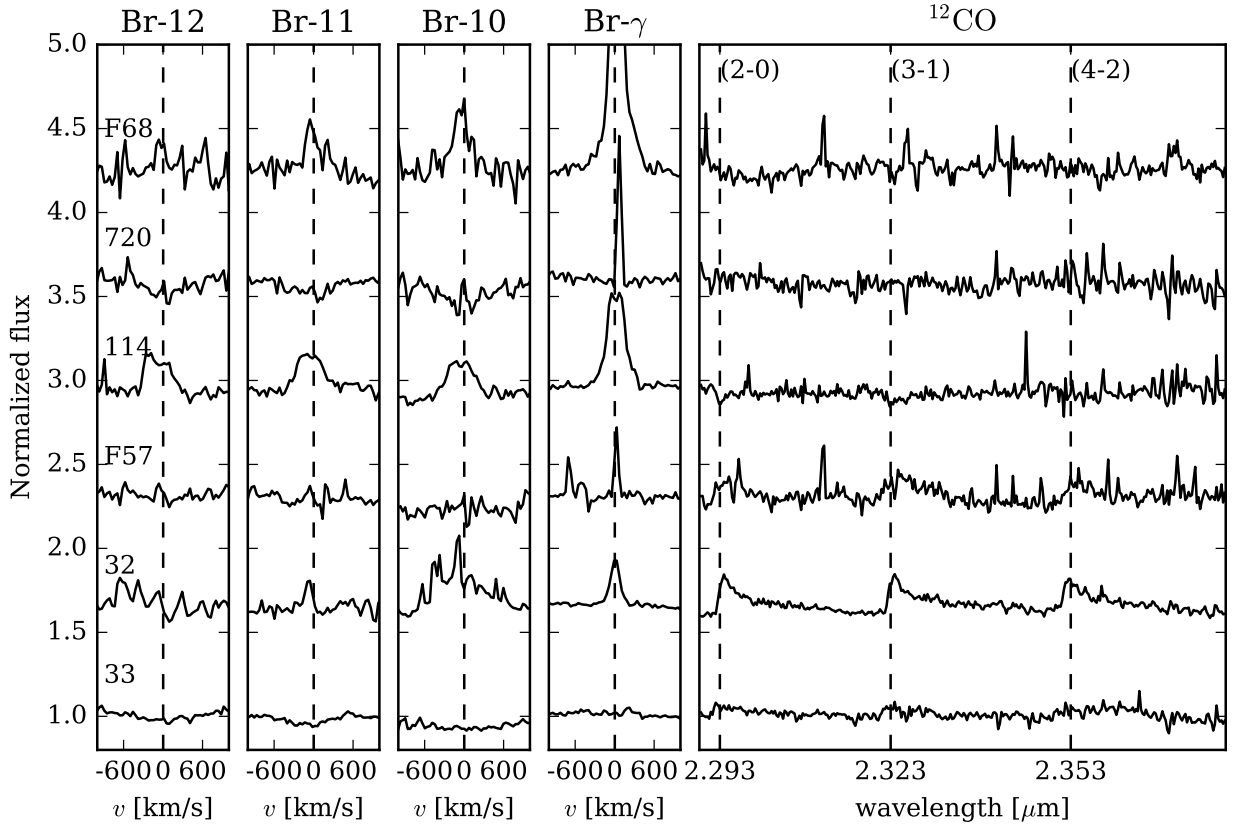


Fig. 8. *Left:* Brackett series line profiles for the six YSOs in our sample: object 32 in M8, objects 33, F57, 114, and F68 in NGC 6357, and object 720 in G333.6–0.2. *Right:* CO first overtone bandheads CO(2-0), CO(3-1), CO(4-2).

7.1. An automatic method to classify stars in the NIR

In the presentation of their template library Rayner et al. (2009) demonstrate that for a few spectral features (Ca II, Na I, Al I and Mg I; their Fig. 35) a relation exists between equivalent width (EW) and spectral type. They suggest that these relations can be used for spectral classification. Based on this idea and in order to have a reproducible and efficient method at hand, we wrote a Python code to do an automatic spectral classification using *H* and *K* band spectra³. To the set of features from Rayner et al. (2009) we add features from literature, adjusting the definitions where necessary and including a few of our own. Though the spectral features are often centred on a line, they can also involve multiple lines, a molecular bandhead or only a part of a line, depending on the wavelength range and spectral type considered. The features and continuum regions that were used are listed in Table B.1 for the late-type spectra and in Table B.2 for the early-type spectra.

The automatic method was tested on the spectra of the template libraries themselves. In each case the spectrum under examination was excluded from the fits mentioned in step two below. Taking the root-mean-square of the errors on the results we determined the mean error of the method for the template spectra to be two spectral subtypes for the late, and one spectral subtype for the early types.

The method treats the early- and late-type stars separately, but the basic steps of the procedure are the same for both cases, as follows.

First, the EWs, together with their variance σ_{EW}^2 , of the defined spectral features are calculated for all template spectra.

Following Cushing et al. (2005) they are given by:

$$EW = \sum \left[1 - \frac{f(\lambda)}{f_c(\lambda)} \right] \Delta\lambda \quad (1)$$

$$\sigma_{EW}^2 = \sum \Delta\lambda^2 \left[\frac{\sigma^2(\lambda)}{f_c^2(\lambda)} + \frac{f^2(\lambda)}{f_c^4(\lambda)} \sigma_c^2(\lambda) \right], \quad (2)$$

where f and f_c are the observed feature and estimated continuum flux, σ and σ_c are the errors on f and f_c , respectively, and the summations are over the number of wavelength intervals $\Delta\lambda$ across the feature. The $\Delta\lambda$ intervals were chosen such that the apparent Doppler shifts in the observed spectra do not affect the EW measurements. We follow the procedure as described in the appendix of Sembach & Savage (1992, Eqs. (A1)–(A20)). Wavelengths are converted to velocities with the centre of the feature at 0 km s⁻¹, and further normalised to the highest velocity so that all (velocity) values fall within the interval [−1,1]. The latter in order to simplify the error calculation (for more details see Sembach & Savage 1992); conversion back to Å is done at the end. We estimate f_c by fitting a first degree polynomial through the defined continuum regions and σ (the error on the flux f) is determined as the standard deviation of the observed values around these continuum fits, in the fitted regions. σ_c is calculated from the covariance matrix produced by the fit, as detailed in Sembach & Savage (1992).

Second, a relation between EW value and spectral type is established per luminosity class. As an example, in Fig. 9 we visualise this relation by plotting the EW values of all templates in the library for four different features for the late types and two features for the early types. The relation is quantified by fitting a 5th (for the late types) or 3rd (for the early types) degree

³ The Python scripts are available upon request.

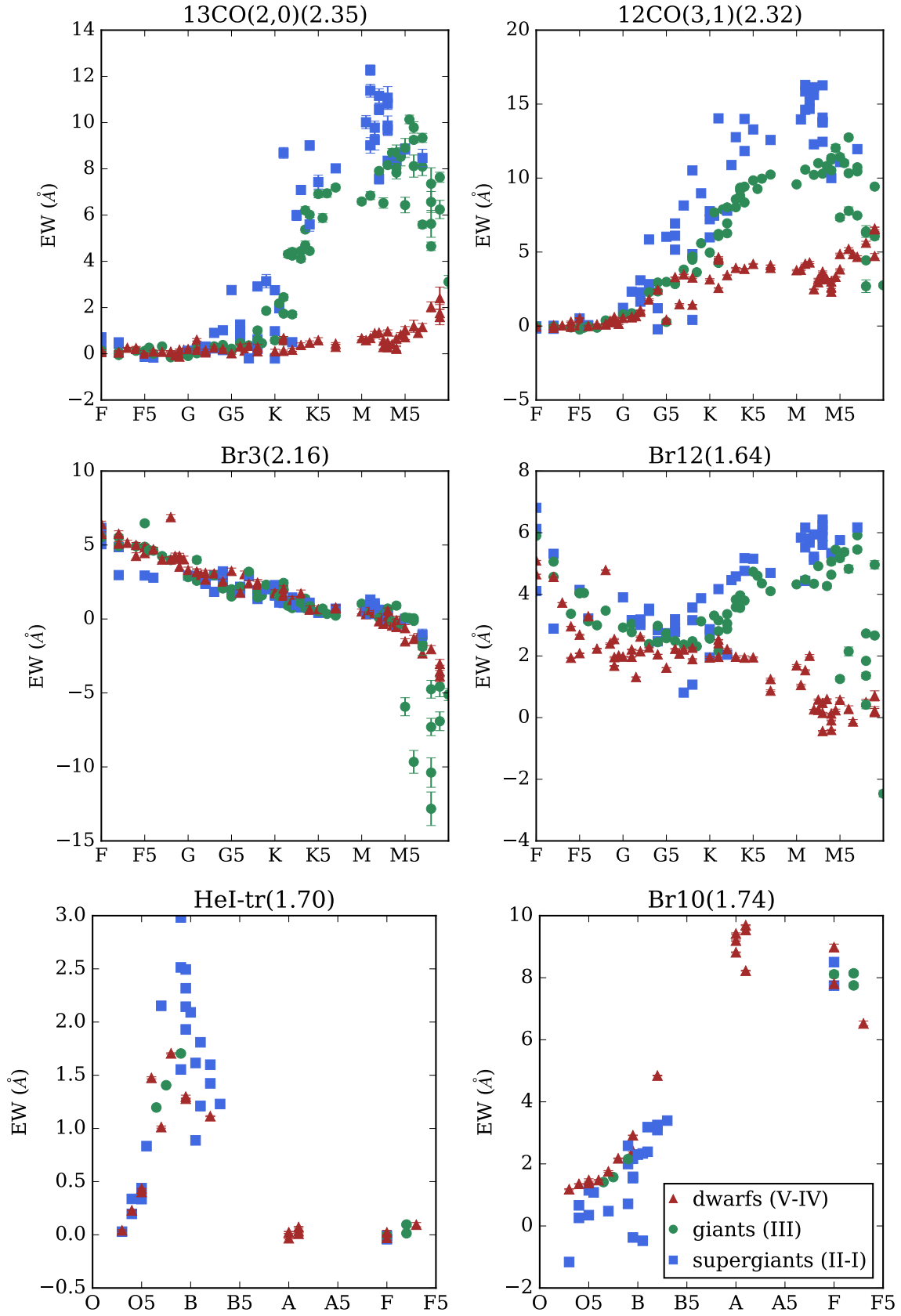


Fig. 9. Relation between EW-value and spectral type on the basis of late type (*four top panels*) and early type (*bottom two panels*) templates for four features. The symbols and colours represent different luminosity classes: red triangles for dwarfs (V-IV), green circles for giants (III) and blue squares for supergiants (II-I).

polynomial $p_{\text{fit}}(\text{SpT})$ through the EWs for each given luminosity class. Thus, for a certain EW value of a particular feature there is one or a few likely spectral types per luminosity class.

Third, the EWs are determined for a science spectrum. Taking into account the errors, each EW is mapped to a likelihood for spectral types, using the polynomial fits obtained in the previous step. In order to quantify how far the determined EW lies from the polynomial, the fitted polynomial $p_{\text{fit}}(\text{SpT})$ is mapped to a Gaussian $G(x; \mu, \sigma)$ with $\mu = \text{EW}$ and $\sigma = \sigma_{\text{EW}}$:

$$L(\text{SpT})_{\text{EW}} = G(p_{\text{fit}}(\text{SpT}); \text{EW}, \sigma_{\text{EW}}). \quad (3)$$

The resulting likelihood distribution $L(\text{SpT})_{\text{EW}}$ (one for each feature) peaks at the most likely spectral type(s) according to the EW for that feature. To avoid very sharp peaks a minimum σ of 0.2 \AA is applied. To combine the information contained in each distribution and obtain one final likelihood distribution for each luminosity class we tried several approaches among which: (1) taking the logarithm of the $L(\text{SpT})_{\text{EW}}$ before adding all distributions and (2) adding one to $L(\text{SpT})_{\text{EW}}$ to avoid numerical zero's before multiplying all distributions. In both cases all EWs were weighted equally. Testing the results on the template spectra as described before, the lowest errors were obtained when applying method (2):

$$L(\text{SpT})_{\text{tot}} = \prod_{\text{EW}} [L(\text{SpT})_{\text{EW}} + 1]. \quad (4)$$

A detailed explanation is provided in Appendix A. An example of a final distribution is shown in Fig. 10.

Finally, the method returns the value of the highest peak of the final distributions as the spectral type of the object and we assign an uncertainty of two spectral sub-types to the obtained value. The luminosity class is determined by the distribution with the highest peak.

For the late types, both EWs and ratios of EWs are taken into account. Steps two and three of the procedure are applied in the same way to the EW ratios. The resulting likelihood distributions are simply multiplied together with those from the EWs, so that one final distribution is obtained for each object: one that includes the information both from the EWs and the ratios. Ratios of which at least one of the EWs has a negative value are excluded in order to avoid negative ratios. Since no real new information is added, one could view the inclusion of ratios as a way of giving extra weight to certain lines or features, specifically to those that do not have negative EWs. We list the ratios used for the classification in Table B.3.

7.2. Spectral classification by eye

As an additional test on the method all science spectra were classified by eye. Each spectrum was individually overplotted with all available templates of the early or late type library. For each star the best matching type or (in the majority of cases) a range of types was selected, within the limitations of the quality of the KMOS spectrum and the completeness of the template spectra. Some of our stars were not classified in this way due to poor data quality, which was a result either of the ripple patterns caused by the data reduction procedure (see Sect. 3.4) or of poor S/N due to extinction, or both. The S/N in the H band is about 20 on average for the regions G333 and M 8; for NGC 6357 this is close to 30. In the K band a S/N of around 30 was reached in all regions, being somewhat lower for G333 and somewhat higher for NGC 6357. The spectra that were not classified by eye due to poor S/N typically had values less than ten in the H band.

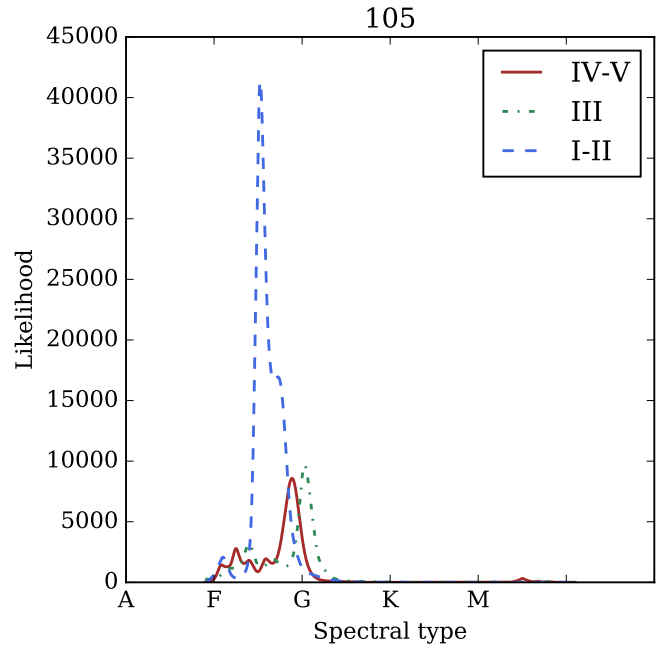


Fig. 10. Example of a final likelihood distribution resulting from the automatic classification method for star 105 in H II region NGC 6357. The different coloured lines are for different luminosity classes: red for dwarfs (IV,V), green for giants (III) and blue supergiants (I,II). The method returns the spectral type corresponding to the highest peak, in this case F5 I-II. We assign an uncertainty of two spectral sub types to this value for all the classified stars.

For the early types, hydrogen (e.g. H I $2.16 \mu\text{m}$) and helium (e.g. He I $2.11 \mu\text{m}$ and He II $2.19 \mu\text{m}$) lines were the only distinguishing features. For the late types both hydrogen and atomic metallic lines (the most prominent being Mg I, Si I, Ca I, Al I, and Fe I) were used, as well as CO ro-vibrational bandheads for the lower temperature types (mostly detected in the K band).

For F-type stars the spectrum is dominated by the hydrogen lines but some metallic lines start to appear. The hydrogen lines significantly weaken at early G spectral types so that by G0 the spectra are dominated by metals. The ratio of Mg I/Bry is a good temperature indicator for F-type stars. In order to distinguish the stars in luminosity class, the width of the hydrogen lines was considered: giants and supergiants have narrower lines than dwarf stars.

In the case of G-K stars the K band is best suited to do the spectral classification. The ratios between Bry and Na I at $2.26 \mu\text{m}$, and Bry and Ca I around $2.20 \mu\text{m}$ can be used, as these metallic lines are more sensitive to the decrease in temperature than others (Gray & Corbally 2009). For later types, the CO bandheads both in the H and in the K -band were mostly used as temperature and luminosity class indicators.

7.3. Comparison of the automatic method with the classification by eye

Before we use the spectral classification to determine the age of the studied regions, we first compare the results of the two methods. Stars for which the two spectral types did not match within error bars we carefully reinspected and a final decision was made (see columns three and four of Table 4). For stars that could not be classified by eye due to poor data quality, we kept the results of the automated classification, unless we judged any

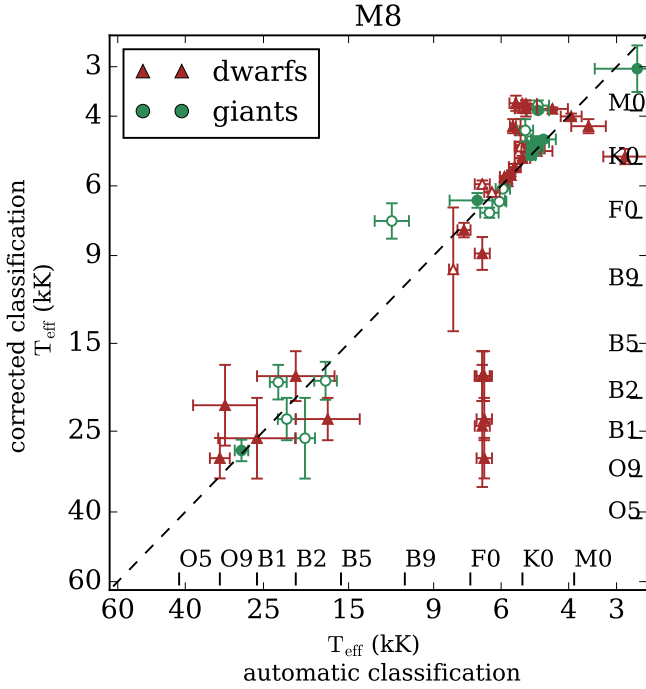


Fig. 11. Comparison of the automatic classification and the one after re-inspection of the spectra (corrected classification) for all stars classified by both methods for M 8. The corrected classification corresponds to either the automatic classification or the classification by eye depending on each specific case (see column four in Table 4). The red triangles represent dwarfs (IV–V), and the green circles (super)giants (I–III) according to the automatic classification. Full symbols signify agreement in luminosity class with the corrected classification. The outliers in the bottom-right of the plot correspond to the early B-type stars that are misclassified as F-type by the automatic method.

result to be untrustworthy based on the spectral quality. We label our final decision as “corrected classification”.

Figures 11–13 present the comparison of the corrected classification with the automatic method for each region. The comparison is shown in effective temperature rather than spectral type (see Sect. 8), to better reflect effects on the final results. There is an overall good agreement except for some B stars (that are misclassified as early F types by the automatic method) and some late M stars (that are misclassified as K types by the automatic method). These are outliers for different reasons: in case of the early types the miss-classification is due to a lack of late-B and A-type templates, requiring an interpolation from B3 to F0–F3 with only a few A0–A1 templates in between.

This interpolation helps to classify the late-A and early-F type stars correctly, but causes some B-types to be mistaken for F-types. In case of the late types it is the (relative) importance of the ^{13}CO feature that causes the confusion for M types: detecting this feature by eye assures one that the star must be M type, however, no special weight is assigned to this feature by the automatic method, which takes into account 23 features for the late types.

To quantify the accuracy of the automatic method further, we compare the number of stars it misclassifies with the number of stars misclassified by initial visual inspection (Sect. 7.2) taking the final corrected classification as reference. The results are shown in Table 2. We see that over the sample that we test, 74% is correctly classified by the automatic method and 73% by initial visual inspection.

We present the summary of the spectral classification for all the studied sources in Table 3. The tabulated spectral types can be

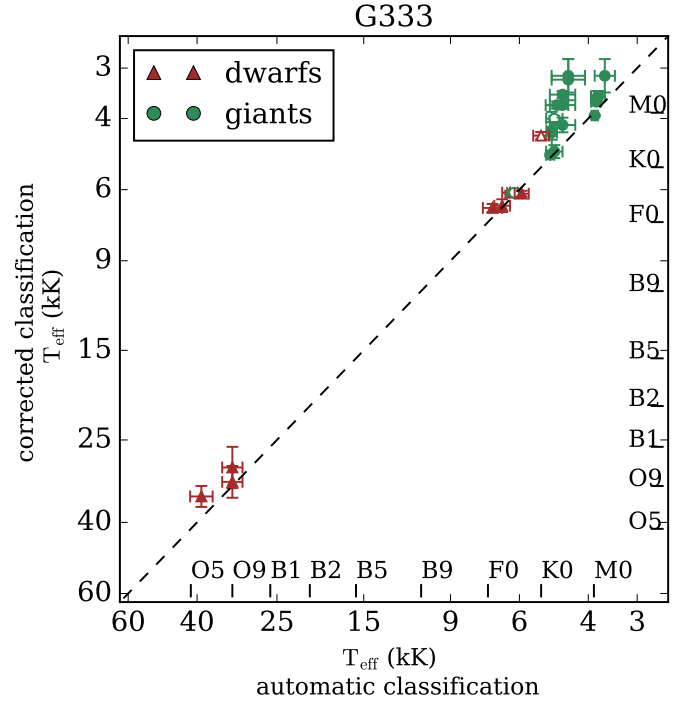


Fig. 12. Same as Fig. 11, but for G333.6–0.2.

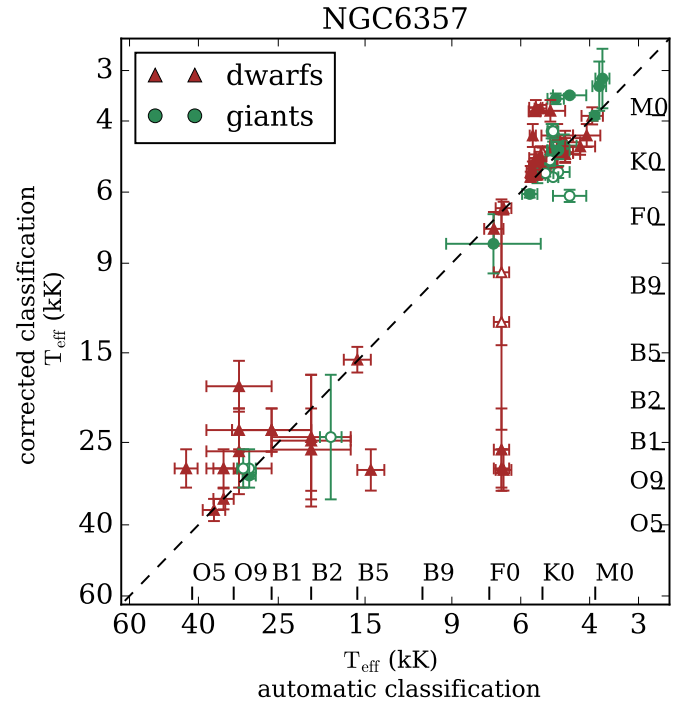


Fig. 13. Same as Fig. 11, but for NGC 6357.

found in the CDS Table 4 and the classified spectra are presented at the CDS. Some examples of spectra are shown in Fig. 14.

7.4. Spectral classification in M 8

We observed four fields in M 8 which resulted in spectra for 81 stars. From the K -band spectral classification we confirm that 21 are OBA stars (16 OB and 5 A). We have 1 candidate YSO (see Sect. 6 for details about the mYSO classification) and 55

Table 2. Comparison of the performance of the automatic classification (method one) and classification by eye (method two).

Region	# stars	Classified by both	Misclassified by method one	Misclassified by method two
M 8	81	60	16 (27%)	17 (28%)
G333.6–0.2	39	27	7 (26%)	5 (18%)
NGC 6357	123	80	20 (25%)	23 (29%)
Total	243	167	43 (26%)	45 (27%)

Table 3. Spectral classification.

H II region	OBA	YSOs	Late type	Unclassified	Total
M 8	21	1	55	4	81
G333.6–0.2	8	1	27	3	39
NGC 6357	31	4	81	7	123

late-type stars, identified by the presence of CO band-heads in absorption. We were not able to classify 4 stars due to severe ripples in their spectra or to insufficient signal to noise. The earliest type star in our sample is the source O10, which we classified as a O9-B1 type star. The full H and K -band spectra are displayed at the CDS. At the left side of the plots the name of the star is shown, at the right side we list the NIR-based spectral classification.

In the top section of Table 5 we present the spectral classification from this work and compare it with the literature classification (columns four and 5); our K -band classification is consistent with the previous classification by Wegner (2000) and Skiff (2009) for the overlapping sources. We also list the two candidate OB stars presented by Povich et al. (2017) and observed by us (second panel); we confirm that source 24 is effectively an OB star, while source 48 presents weak CO band-heads in absorption and no IR excess, and therefore it could be either a foreground low-mass PMS star or a background field giant. In the third section of Table 5 we list four new OB and four A stars classified by us that are not listed as massive stars in the literature. The last three columns of this table list the K -band extinction (A_K), effective temperature (T_{eff}), and luminosity ($\log L/L_{\odot}$); see Sect. 8.

7.5. Spectral classification in G333.6–0.2

7.5.1. Spectral classification

We observed two fields in G333.6–0.2 and obtained spectra of 39 stars. From the K -band spectral classification we confirm that 8 are OBA stars (4 OB and 4 A), 1 candidate YSO and 27 late-type stars, identified by the presence of CO band-heads in absorption. We were not able to classify 3 spectra. The most massive star in our sample is 6829462 which is an O6 star. The candidate YSO (source 720) does not present CO band-head emission, Bry seems to be in emission with a central nebular component, and the other lines of the Brackett series show a photospheric absorption line with a central emission component. The H and K -band spectra are available at the CDS. We present four new OB stars and six A stars in our observed sample. In Table 6 we list these stars together with their derived K -band extinction, effective temperature, and luminosity (see Sect. 8).

7.5.2. X-ray counterparts for sources in G333.6–0.2

Recently, Townsley et al. (2018) presented the second instalment of the Massive star-forming regions Omnibus X-ray Catalogue

Table 4. Properties of the KMOS targets.

Col.	Label	Description
(1)	Name	Source name in KMOS catalogue
(2)	X-ray ID	Source X-ray ID (CXOU J)
(3)	kSpT	K -band spectral type
(4)	SpT_source	Source of the spectral classification (b if both methods agree, e for the classification by eye, and a for the automatic classification)
(5)	α (J2000)	Right ascension (J2000, hh:mm:ss)
(6)	δ (J2000)	Declination (J2000, dd°:mm':ss'')
(7)	J	J -band magnitude (mag)
(7)	J_{err}	J -band magnitude error (mag)
(8)	H	H -band magnitude (mag)
(8)	H_{err}	H -band magnitude error (mag)
(9)	K	K -band magnitude (mag)
(9)	K_{err}	K -band magnitude error (mag)
(10)	[3.6]	3.6 μm magnitude (mag)
(11)	[3.6]_err	3.6 μm magnitude error (mag)
(12)	[4.5]	4.5 μm magnitude (mag)
(13)	[4.5]_err	4.5 μm magnitude error (mag)
(14)	[5.8]	5.8 μm magnitude (mag)
(15)	[5.8]_err	5.8 μm magnitude error (mag)
(16)	[8.0]	8.0 μm magnitude (mag)
(17)	[8.0]_err	8.0 μm magnitude error (mag)

Notes. This table is available in its entirety at the CDS.

(MOXC2) where they list X-ray point sources in several star forming regions including G333.6–0.2. By cross-matching their catalogue with the coordinates of our sources we find that 22 out of our initial 39 sources have X-ray counterparts. Those sources are included with *Chandra* IDs in the second column of Table 4. Source 378 is a late-O, early-B type star, but it does not have an X-ray counterpart. We assume that this massive star is part of G333.6–0.2. Given that we are not certain of the cluster membership of the other (low-mass) stars with no X-ray counterparts, in the remainder of this paper we treat those sources, except 378, as non members of G333.6–0.2 and do not take them into account when estimating the age of the region (Sect. 8).

7.6. Spectral classification in NGC 6357

We observed six fields in NGC 6357 resulting in 123 spectra. From the IR spectral classification we confirm that 31 are OBA stars (22 OB and 9 A), we have 4 candidate YSOs and 81 late-type stars, identified by the presence of CO band-heads in absorption. We were not able to classify 7 stars from the obtained spectra. The H and K -band spectra are available at the CDS, in the left side of the plots, the name of the stars is shown, and in the right part we show the NIR-based spectral classification.

In Table 7 we show the previously classified OB stars in NGC 6357 that overlap with our sample. The spectral classification obtained for these sources agrees with the previous studies. Out of the 26 new candidate OB stars presented by Povich et al. (2017) we observed 17 with KMOS. We confirm that most of these stars are OB stars with the exception of B3 which presents CO bandheads in absorption and is therefore classified as a K4-K5V type star, and B14 where the spectrum is more consistent with it being an F star. We were not able to classify B12 due to the very strong ripples affecting its spectrum nor B16 because the signal-to-noise is too low to recognise any spectral features. Finally, five newly classified OB stars and nine A stars are also listed in Table 7. As in Tables 5 and 6, the last three columns show the physical properties derived in Sect. 8.

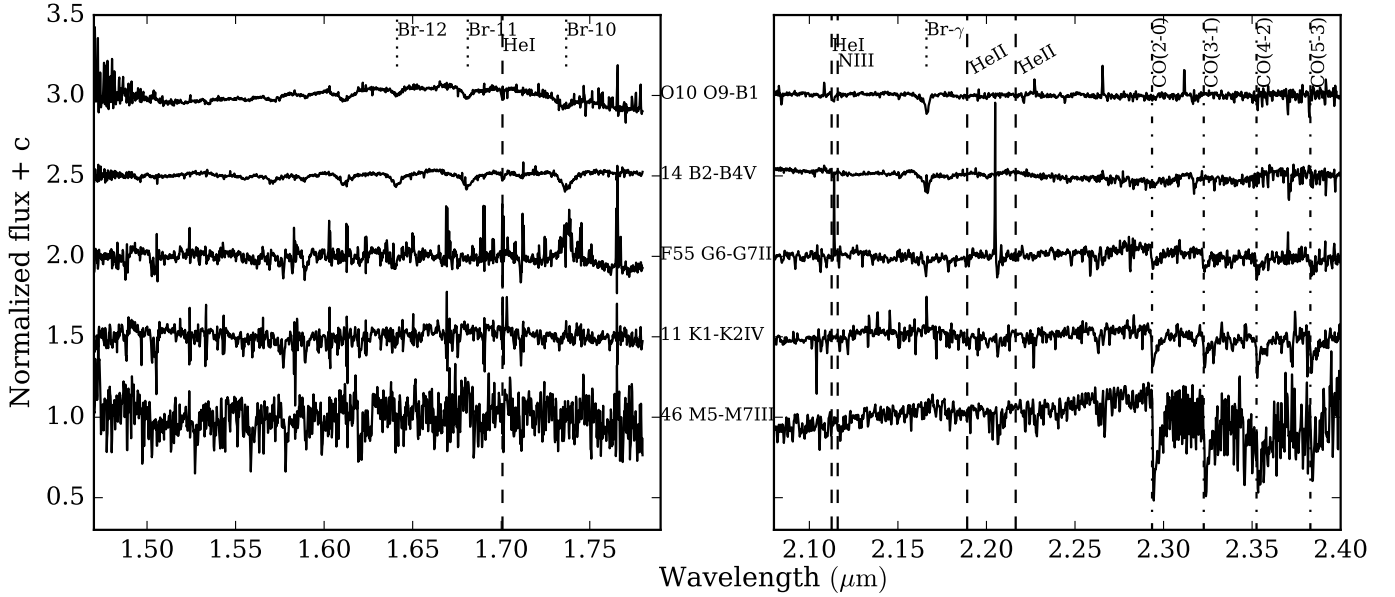


Fig. 14. Examples of the H - and K -band KMOS spectra across the spectral range. The figures containing the spectra of all sources are available at the CDS.

7.7. Relation between photometry, astrometry, and spectral classification

Here we discuss some individual objects whose position in the CMDs and CCDs (Figs. 4–6) or their astrometry from *Gaia* DR2 data (Sect. 5) suggest that they have peculiar properties.

Source 46 in M 8 seems to be too bright for its spectral type, one possibility is that the X-ray and IR sources were incorrectly matched. If the sources are correctly matched this source could be a cool giant foreground star where the X-ray emission is originated by binary interaction. Sources 24 and 48 seem to be highly reddened in comparison to the rest of the population in the NIR CMD, but a strong excess is not apparent in the NIR nor MIR CMDs. Source 24 is also amongst the candidate non-members according to its *Gaia* DR2 astrometric properties, but the quality of the astrometry is insufficient to firmly exclude membership of M 8. From the photometric properties of sources F13 and F24 it does not appear obvious that these stars are not members of M 8, but we discard them as cluster members based on our results from the *Gaia* DR2 data. We consider these four sources not to belong to M 8 and exclude them from the extinction calculation and age determination.

In the NIR CCD of NGC 6357 source 33 is below the reddening line for $R_V = 3.1$, and shows evidence for the presence of a circumstellar disk which indicates that it might be a massive YSO with a NIR excess (see Sect. 6). Sources 91 and 8 seem to be too bright for their spectral types, which is consistent with the astrometric solution asserting these stars to be non cluster members. We exclude these stars from the extinction and age determination.

In G333.6–0.2, the stars classified as M sources correspond to some of those which do not have X-ray counterparts and were selected only on the basis of their NIR colours. From the NIR diagrams we conclude that these stars are cool giants and are not part of the cluster. This shows the importance of the MYStIX project in order to identify cluster members. Stars 515756828027 and 515757115402 have X-ray counterparts but their position in the CMD and CCD seems inconsistent with their spectral type, this in agreement with the astrometric solution showing that they are not cluster members.

Our spectral classification agrees with the SED classification by Povich et al. (2013) in the MIR CCDs: the objects classified as OB type stars are located where the authors find the visible photospheres, and the sources that we identify as pre-main-sequence stars are likely YSOs. Objects 22, 88, and 112 in NGC 6357 are an exception because they are in the region dominated by YSOs, but do not show any gaseous disk signatures in the NIR diagram, nor in the H - and K -band spectra. This could be explained by a relatively large inner hole.

From the photometry, the spectral classification, and the X-ray counterparts (see also Sect. 4), it becomes clear that there are some stars that are not cluster members. These have been marked with grey symbols in the HRDs and have not been included in the extinction histograms nor taken into account for the age determination. Unfortunately, the quality of the astrometric fit for these stars is poor, and therefore we are not able to evaluate their membership with *Gaia* DR2 data. We do not discard the stars classified as giants because the fact that they have low surface gravities could mean that they are PMS swollen stars contracting towards the Zero age main-sequence (ZAMS).

8. Age of the giant H II regions

After the spectroscopic classification, we can place our KMOS sources in the HRD and match them with MESA isochrones and evolutionary tracks obtained from the MIST project (Paxton et al. 2011, 2013, 2015; Dotter 2016; Choi et al. 2016). The isochrones and tracks used have solar metallicity and initial rotational velocity $v_{\text{ini}} = 0.4v_{\text{crit}}$, where v_{crit} is the critical velocity, reached when the gravitational acceleration is exactly compensated by the centrifugal force.

The effective temperature, intrinsic colours $(V - K_s)_0$, $(J - H)_0$, and $(H - K)_0$, and bolometric correction of the luminosity class V stars are taken from Pecaut & Mamajek (2013); for the luminosity class III stars we used the calibrations from Alonso et al. (1999) for spectral types F0–F9, and for G4–M5 stars we used the calibration of Cox (2000). With the effective temperature and bolometric correction in the V -band, BC_V , we derive the absolute magnitude in the V -band. In order

Table 5. Spectral types and physical properties of the O, B, and A stars in the M 8 KMOS sample.

ID	X-ray ID (CXOU J)	Other ID	Previous ST	ST this work	A_K (mag)	T_{eff} (K)	$\log L/L_{\odot}$
Previously classified OB stars in M 8 included in our KMOS sample							
14	180345.10–242205.2	CD-24 13810	B... ^(a)	B2–B4V	0.23 ± 0.01	18 650 ± 1950	3.54 ± 0.14
33	180411.16–242145.1	NGC 6530 32	B3Ve ^(b)	B1–B3	0.16 ± 0.03	21 500 ± 4500	3.35 ± 0.3
36	180415.03–242327.8	HD 315032	B1V ^(b)	B1–B2Vz	0.16 ± 0.02	23 300 ± 2700	3.84 ± 0.17
O5	180423.99–242126.7	NGC 6530 60	B1Ve ^(b)	O9–B1–SB2	0.17 ± 0.01	29 250 ± 3250	4.03 ± 0.13
O6	180424.29–242059.5	NGC 6530 61	B2Ve ^(b)	B1–B2V	0.1 ± 0.02	23 300 ± 2700	3.47 ± 0.17
O7	180425.55–242044.9a	NGC 6530 66	B2Vep ^(b)	B0–B2V	0.13 ± 0.03	26 050 ± 5450	3.63 ± 0.28
O8	180427.19–242249.6	NGC 6530 70	B2Ve ^(b)	B1–B2V	0.19 ± 0.02	23 300 ± 2700	3.5 ± 0.17
60	180428.98–242526.2	NGC 6530 74	B2.5Ve ^(b)	B2–B5	0.1 ± 0.02	18 150 ± 2450	3.0 ± 0.19
O10	180432.98–242312.5a	NGC 6530 80	B1V ^(b)	O9–B1	0.17 ± 0.01	29 250 ± 3250	4.12 ± 0.13
73	180434.20–242200.6	NGC 6530 86	B2.5Ve ^(b)	B0–B3V	0.3 ± 0.04	24 250 ± 7250	3.78 ± 0.4
O11	180439.92–242244.4	NGC 6530 99	B2.5Vne ^(b)	B2–B5V	0.15 ± 0.02	18 150 ± 2450	2.97 ± 0.19
79	180441.68–242058.5	HD 164947A	B1/2Vne ^(b)	B2–B5V	0.11 ± 0.02	18 150 ± 2450	3.38 ± 0.19
Candidate OB stars in M 8 included in our KMOS sample							
24	180400.80–242334.6	–	–	B0–B2V	0.84 ± 0.03	26 050 ± 5450	4.03 ± 0.28
48	180423.55–242116.8	–	–	G8–K0IV-V	0.21 ± 0.01	5385 ± 105	1.73 ± 0.03
New OB and A stars in M 8							
16	180346.16–242149.4	–	–	B2–B5V	0.19 ± 0.02	18 150 ± 2450	2.84 ± 0.19
51	180424.29–242059.6	–	–	B2–B3V	0.07 ± 0.01	18 800 ± 1800	3.18 ± 0.12
55	180425.72–241949.6	–	–	A5–F2III-IV	0.17 ± 0.01	9752 ± 2948	2.29 ± 0.21
69	180433.03–242252.9	–	–	A0–A5V	0.01 ± 0.01	8890 ± 810	1.74 ± 0.12
F19	180404.31–242130.4	–	–	A5–F4V	0.02 ± 0.01	7940 ± 140	1.42 ± 0.03
F36	180420.48–242303.9	–	–	A8–F2III	1.38 ± 0.01	8055 ± 1251	2.4 ± 0.12
F44	180422.10–242546.6	–	–	B0–B3	0.63 ± 0.04	24 250 ± 7250	3.19 ± 0.4
F47	180426.73–242351.7	–	–	A5–A9V	0.2 ± 0.01	7760 ± 320	1.53 ± 0.05
Previously classified OB stars in M 8 not included in our KMOS sample							
–	180340.32–242242.8	HD 164740	O7.5V+O9V+B0.5V ^(b)	–	–	36 000	–
–	180352.45–242138.5	HD 164794	O4V((f)) ^(b)	–	–	42 000	5.82
–	180356.85–241845.0	HD 164816	O9.5V+B0V ^(b)	–	–	31 500	4.94
–	180414.50–241437.0	HD 315026	B1.5V ^(b)	–	–	24 000	–
–	180420.55–241354.7	NGC 6530 55	B2.5Ve ^(b)	–	–	19 000	3.4
–	180421.27–242118.2	NGC 6530 56	B0.5V ^(b)	–	–	28 000	4.25
–	180422.74–242209.8	NGC 6530 58	B2Ve ^(b)	–	–	21 000	–
–	180423.22–242616.7	HD 315033	B2Vp ^(b)	–	–	21 000	3.99
–	180425.83–242308.4	HD 164906	B0Ve ^(b)	–	–	29 500	–
–	180428.02–242142.9	HD 315031	B1V ^(b)	–	–	26 000	4.42
–	180432.91–241844.6	NGC 6530 83	B2.5Vne ^(b)	–	–	19 000	3.41
–	180435.99–241952.2	HD 315021	B1V ^(b)	–	–	26 000	–
–	180441.65–242055.1	HD 164947B	B2.5V ^(b)	–	–	19 000	3.42
–	180448.62–242635.5	NGC 6530 SJ 75	B2.5 ^(a)	–	–	19 000	3.5
–	180456.26–242403.4	HD 315034	B3 ^(a)	–	–	17 500	2.8
–	180459.03–242723.1	HD 315035	B3V ^(b)	–	–	17 500	–

References. ^(a)Wenger et al. (2000). ^(b)Skiff (2009).**Table 6.** Spectral types and physical properties of the newly classified O, B, and A stars in the G333.6–0.2 KMOS sample.

ID	X-ray ID (CXOU J)	ST	A_K (mag)	T_{eff} (K)	$\log L/L_{\odot}$
6829201	162231.65–500815.6	O7–O9	1.37 ± 0.01	34 500 ± 2000	4.85 ± 0.04
6829462	162222.28–500736.7	O6–O7	1.72 ± 0.01	37 750 ± 1250	5.33 ± 0.03
7034854	162221.87–500335.6	A8–F2V	0.75 ± 0.01	7155 ± 345	2.88 ± 0.06
7117881	162159.55–500743.5	O8–B0.5Vz	1.25 ± 0.01	31 750 ± 2750	4.71 ± 0.08
7348675	162205.99–500421.5	A8–F2V	1.04 ± 0.01	7155 ± 345	2.64 ± 0.06
378	–	O9–B1	–	29 250 ± 3250	–
767	–	A8–F2V	–	7155 ± 345	–
789	162218.25–500514.8	A9–F3III	1.92 ± 0.02	7398 ± 777	3.44 ± 0.11

Table 7. Spectral types and physical properties of the O, B, and A stars in the NGC 6357 KMOS sample.

ID	X-ray ID (CXOU J)	Other ID	Previous ST	ST this work	A_K (mag)	T_{eff} (K)	$\log L/L_{\odot}$
Previously classified OB stars in NGC6357 included in our KMOS sample							
16	172436.03–341400.4	Pis 24 10	O9V ^(a)	O9–B0	0.62 ± 0.01	$32\,000 \pm 500$	4.73 ± 0.02
O0	172442.26–341141.1	Pis 24 12	B1V ^(a)	B1–B2	0.46 ± 0.02	$23\,300 \pm 2700$	3.87 ± 0.17
O1	172443.29–341142.1	Pis 24 18	B0.5V ^(a)	O9–B1	0.74 ± 0.01	$29\,250 \pm 3250$	4.59 ± 0.13
O2	172443.68–341140.7	Pis 24 19	B1V ^(a)	B1–B2	0.93 ± 0.02	$23\,300 \pm 2700$	4.12 ± 0.17
Candidate OB stars in NGC6357 included in our KMOS sample							
B0	172434.78–341318.0	–	–	O9.5–B1V	0.66 ± 0.01	$29\,000 \pm 3000$	4.22 ± 0.11
B1	172440.49–341206.3	–	–	O9.5–B2V	0.66 ± 0.03	$26\,300 \pm 5700$	4.31 ± 0.29
B2	172442.87–340911.6	–	–	O9.5–B1V	1.03 ± 0.01	$29\,000 \pm 3000$	4.59 ± 0.11
B3	172447.56–341048.6	–	–	K1–K5V	1.07 ± 0.03	4790 ± 380	1.93 ± 0.1
B4	172455.08–341111.5	–	–	O9.5–B1V	0.88 ± 0.01	$29\,000 \pm 3000$	4.67 ± 0.11
B6	172528.51–342652.5	–	–	O9.5–B1V	0.55 ± 0.01	$29\,000 \pm 3000$	4.23 ± 0.11
B7	172547.33–342154.1	–	–	B2–B5V	0.55 ± 0.02	$18\,150 \pm 2450$	3.52 ± 0.19
B8	172554.16–342156.4	–	–	B0–B3V	0.68 ± 0.04	$24\,250 \pm 7250$	3.88 ± 0.4
B9	172557.29–341850.8	–	–	O9–B3V	0.85 ± 0.05	$24\,750 \pm 7750$	4.3 ± 0.42
B10	172558.08–341606.1	–	–	O7–O9V	0.77 ± 0.01	$34\,500 \pm 2000$	4.4 ± 0.04
B11	172558.15–342119.2	–	–	O9–B1	0.92 ± 0.01	$29\,250 \pm 3250$	4.61 ± 0.13
B12	172602.07–341757.6	–	–	–	–	–	–
B13	172604.04–341520.0	–	–	O6–O8V	1.0 ± 0.01	$36\,750 \pm 2250$	4.99 ± 0.06
B16	172604.25–341557.0	–	–	–	–	–	–
B14	172604.71–341526.3	–	–	F2–F6V	1.43 ± 0.01	6575 ± 235	2.66 ± 0.05
B15	172604.72–341611.2	–	–	O9.5–B1V	1.03 ± 0.01	$29\,000 \pm 3000$	4.22 ± 0.11
118	172601.88–341631.1	–	–	B0–B3	0.75 ± 0.04	$24\,250 \pm 7250$	4.26 ± 0.4
New OB and A stars in NGC 6357							
14	172434.67–341425.2	–	–	A5–F3IV–V	0.25 ± 0.01	7400 ± 680	1.65 ± 0.12
22	172439.41–341526.8	–	–	B1–B2	0.61 ± 0.02	$23\,300 \pm 2700$	3.84 ± 0.17
34	172444.62–341100.6	–	–	A9–F3V	0.53 ± 0.01	7080 ± 360	1.98 ± 0.06
45	172451.14–341201.0	–	–	A9–F3V	0.51 ± 0.01	7080 ± 360	1.8 ± 0.06
65	172522.78–342404.6	–	–	A5–F5III–IV	0.39 ± 0.01	9477 ± 3222	2.34 ± 0.29
70	172530.30–342455.1	–	–	A8–F2III	0.44 ± 0.01	8055 ± 1251	2.76 ± 0.12
71	172532.29–342346.5	–	–	A0–F2III–IV	0.57 ± 0.05	$12\,579 \pm 5775$	2.68 ± 0.37
73	172533.82–342708.5	–	–	B0–B2V	1.02 ± 0.03	$26\,050 \pm 5450$	4.1 ± 0.28
88	172545.29–342535.2	–	–	O9.5–B1V	0.59 ± 0.01	$29\,000 \pm 3000$	4.34 ± 0.11
107	172556.17–341510.4	–	–	B4–B6V	0.4 ± 0.01	$15\,600 \pm 1100$	2.92 ± 0.1
108	172556.81–341727.0	–	–	A8–F2V	0.73 ± 0.01	7155 ± 345	2.11 ± 0.06
112	172559.18–341919.4	–	–	B0–B2V	0.76 ± 0.03	$26\,050 \pm 5450$	4.03 ± 0.28
Previously classified OB stars in NGC 6357 not included in our KMOS sample							
–	172428.95–341450.6	Pis 24 15	O7.5Vz ^(b)	–	–	$35\,000$	4.93
–	172442.30–341321.4	Pis 24 3	O8V ^(a)	–	–	$35\,000$	4.96
–	172443.28–341243.9	Pis 24 2	O5.5V((f)) ^(a)	–	–	$40\,000$	5.37
–	172443.49–341156.9	Pis 24 1 SW+NEab	O4III(f+)+O3.5If* ^(c)	–	–	$42\,000$	6.22
–	172444.48–341158.8	Pis 24 16	O7.5V ^(a)	–	–	$36\,000$	–
–	172444.72–341202.6	Pis 24 17	O3.5III(f*) ^(a)	–	–	$44\,500$	5.92
–	172445.77–340939.8	Pis 24 13	O6.5V((f)) ^(a)	–	–	$38\,000$	5.25
–	172508.85–341112.4	HD 157504	WC6+O7/9 ^(a)	–	–	$37\,000$	–
–	172529.16–342515.6	[N78] 51	O6Vn((f)) ^(b)	–	–	$41\,000$	5.35
–	172534.23–342311.7	[N78] 49	O5.5IV(f) ^(b)	–	–	$39\,000$	5.87
–	172538.60–340843.5	LS 4151	O6/7III ^(a)	–	–	$38\,000$	5.41
–	172626.45–341700.9	HD 319788	B2III: ^(a)	–	–	$21\,000$	4.22

References. ^(a)Skiff (2009). ^(b)Maíz Apellániz et al. (2016). ^(c)Maíz Apellániz et al. (2007).

to calculate the luminosity of the observed targets we used the absolute K -band magnitude assuming the Indebetouw et al. (2005) extinction law to derive A_K ,

$$A_K = \frac{(H - K) - (H - K)_0}{0.55}, \quad (5)$$

and calculated the absolute K -band magnitude by scaling it to

the region’s distance and correcting it by the extinction A_K . We obtained the bolometric correction in the K -band, BC_K , from BC_V and the absolute V and K -band magnitudes ($BC_K = BC_V + M_V - M_K$).

The K -band extinction, temperature, and luminosity derived for the OB and A stars in M 8, G333.6–0.2, and NGC 6357 are listed in the last three columns of Tables 5–7.

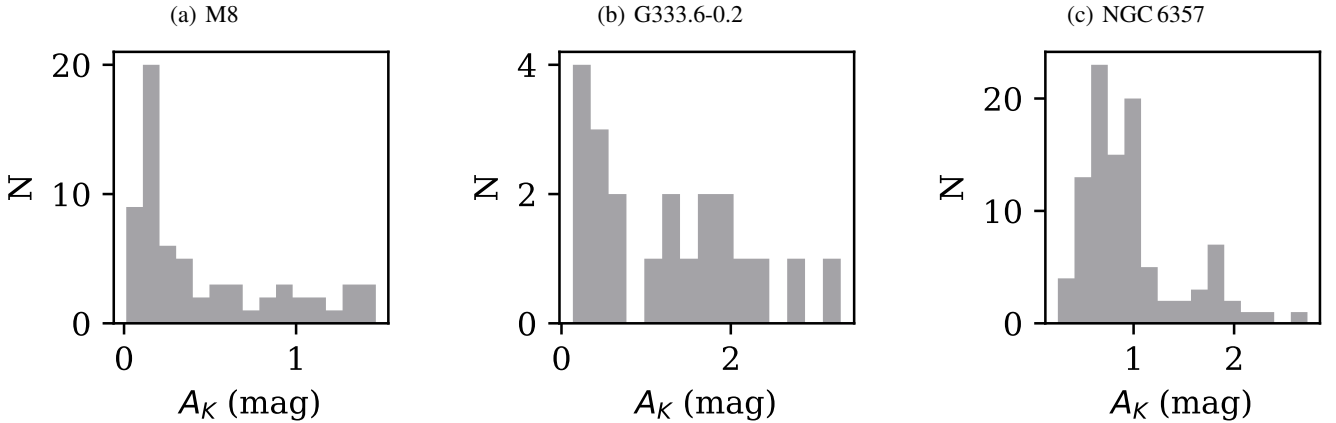


Fig. 15. Distribution of the extinction A_K for the three regions studied using the [Indebetouw et al. \(2005\)](#) extinction law.

8.1. Extinction A_K

To assess the sensitivity of our results to the choice of the extinction law, we also used the [Nishiyama et al. \(2009\)](#) extinction law ($A_H/A_K = 1.62$) to derive the luminosity as described in the previous section. With this extinction law the mean A_K values are slightly lower than the values derived with the [Indebetouw et al. \(2005\)](#) extinction law; the peak of the A_K distribution shifts at most 0.1 mag (Fig. 15). Given this small difference, adopting the [Indebetouw et al. \(2005\)](#) extinction law seems secure.

The extinction in M8 peaks at around $A_K = 0.1$ mag and has a tail extending to values of $A_K = 1.5$ mag. The object with the highest extinction is F5; its position in the CMD is consistent with having such a high extinction. With an extinction $A_K = 2.5$ mag the star F9 may be a non-member, thus we have excluded it from the histogram and the age classification. Although the trend is not very strong, the extinction seems to be higher in the western part of this region, close to the location of the massive multiple system Herschel 36.

In the case of G333.6–0.2 the A_K values have an almost flat distribution ranging from 0.2 to 3.5; this is consistent with what is observed in the CMD and CCD. Most of the OB stars have an extinction ranging from 1 to 2 mag. The objects with the highest extinction are found at the centre of the cluster while the lowest extinction is measured for the four objects to the north-east of the cluster centre.

The mean value of A_K for NGC 6357 is higher than for M8, which is also evident from the CMDs. In this region we can distinguish two peaks, one around $A_K = 0.7$ and a smaller one at about $A_K = 1.7$. The highest extinction corresponds to F67 which is also one of the reddest objects in the CMD and CCDs. The stars in the north-west region in NGC 6357 are on average less extinguished in the K -band than those in the south-east region, but other than that there is no obvious trend of extinction with location.

8.2. Age determination

Figures 16–18 show the HRDs for the three regions. We show the position of the observed targets as ellipses, where the height and width of the ellipses represent the errors in the temperature and luminosity resulting from the uncertainty in the spectral classification. The left panels show the MESA PMS (black) and post-main-sequence (dark grey) tracks for stars of 1, 2, 3, 5, 10, 20, 35, 70, and $100 M_\odot$, respectively; the thick black line shows the position of the zero-age main sequence. In the right

panels we show the isochrones for 0.1, 1, 3, and 5 Myr with dotted, dash-dotted, dashed, and solid lines, respectively. The thick dotted line shows the birth line for stars until $8 M_\odot$. The accuracy of the locations of the stars in the HRD is dominated by the uncertainty in the spectral classification.

To get a more complete picture of the studied regions, we add the previously known OB stars from the MPCM catalogue listed in [Povich et al. \(2017\)](#). The derived properties of these objects are listed in the last section of Tables 5 and 7 and shown as solid green and red dots corresponding to the giant and dwarf stars, respectively. We also plot 0.2 dex error circles to represent the scatter reported by [Povich et al. \(2017\)](#) around the ZAMS. The infrared spectral classification of the most massive stars (spectral types O3–O4) is normally degenerate (see e.g. [Wu et al. 2014, 2016; Bik et al. 2019](#)) so it could affect the lower age limit. To avoid this, we determined the age based on the whole population (including the low-mass PMS-stars) instead of focusing only on the massive stars.

In addition to the sources mentioned in Sect. 7.7 that were discarded as members of the regions, we marked the K -band excess sources (below the reddening line in the NIR CCDs) with empty ellipses filled with crosses. Their NIR excess emission could cause an overestimation of L_{bol} and affect the age estimation, therefore we did not take them into account when assessing the age of the regions.

The main sequence stars in M8 have masses between ~ 5 and $\sim 70 M_\odot$. The PMS stars will become ZAMS stars of ~ 0.5 – $3 M_\odot$. From comparing the observed population with MIST isochrones we conclude that the age of M8 is between 1 and 3 Myr.

For G333.6–0.2 the main sequence population ranges in mass between ~ 5 and $\sim 35 M_\odot$. As in this region the membership of the stars is more uncertain, we had to focus on the massive stars only to estimate the age. For this reason, we only provide an upper limit. By comparing the few confirmed OB stars with X-ray counterparts with MIST isochrones we estimate an age < 3 Myr for this region.

The massive stars in NGC 6357 have main sequence masses between ~ 10 and $\sim 100 M_\odot$. From comparing the position of the stars with MESA isochrones, we conclude that the mean age of the PMS and main-sequence populations of NGC 6357 is probably somewhat younger than that of M8. We determine the age of NGC 6357 to be 0.5–3 Myr.

We also tested the effect of the different extinction laws on the age determination. We conclude that the [Nishiyama et al. \(2009\)](#) and the [Indebetouw et al. \(2005\)](#) extinction laws result in similar age estimates for the studied regions.

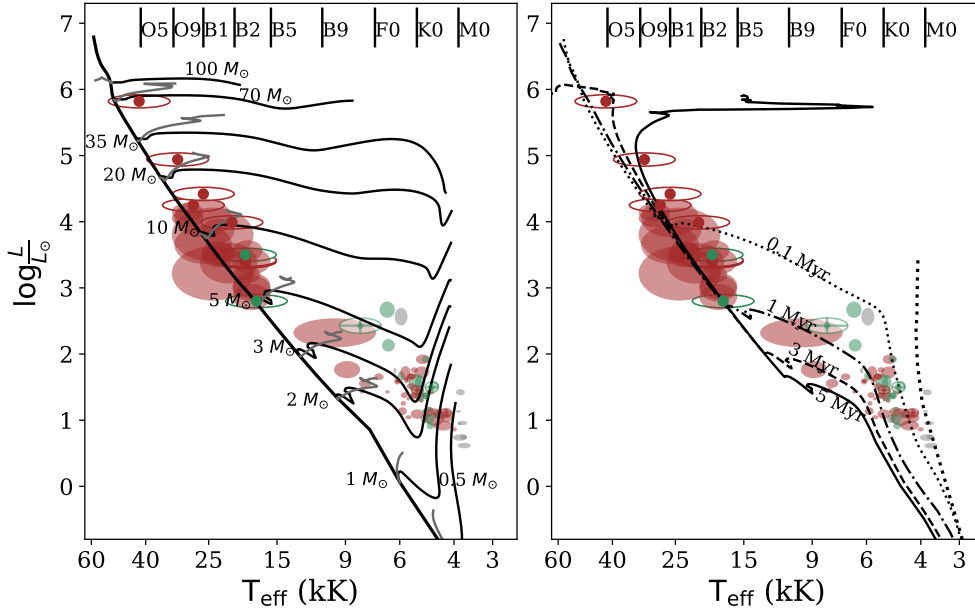


Fig. 16. Hertzsprung-Russell diagram (HRD) for the stars in M 8. *Left panel:* PMS (black) and post-main-sequence (grey) tracks for stars of masses of 1, 2, 3, 5, 10, 20, 35, 70, and $100 M_{\odot}$, respectively. *Right panel:* 0.1, 1, 3, and 5 Myr isochrones. The ellipses show the position of the observed stars. The dots show the position of previously classified OB stars reported in with 0.2 dex error ellipses representing the scatter about the ZAMS (Povich et al. 2017). The luminosity class IV and V stars are shown in red, the stars with luminosity class III are shown in green. In grey we show the late-K and M-type stars. The empty ellipses with crosses show the objects located below the reddening line in the NIR CCD, which have been excluded from the age determination. The age of M 8 is estimated to be between 1 and 3 Myr.

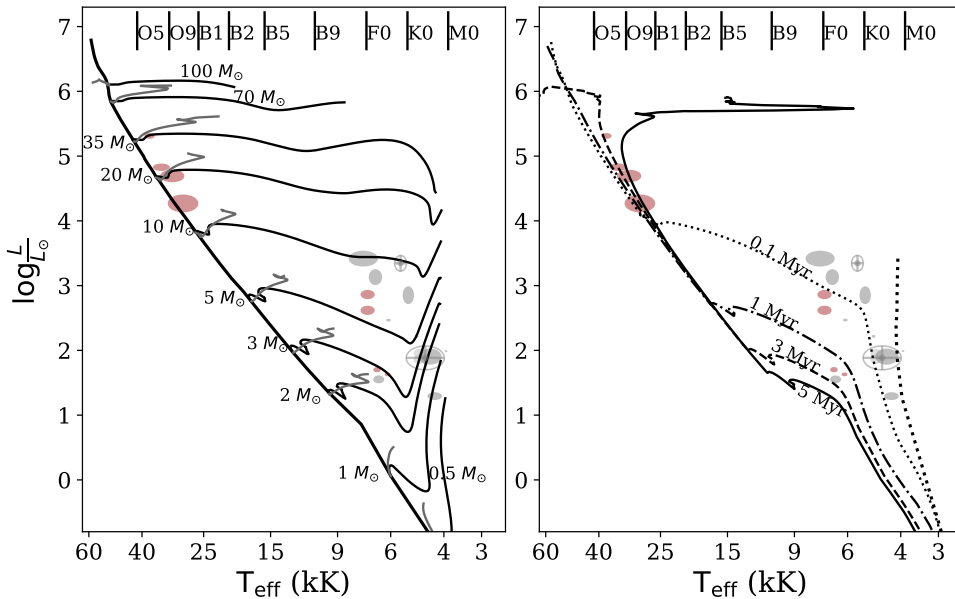


Fig. 17. Same as Fig. 16 but for the classified stars in G333.6–0.2. We estimate an age < 3 Myr.

9. Discussion

In this section we discuss the relation between the results obtained from photometry and spectroscopy. We consider some of the uncertainties that could affect the age determination and the impact that would have on our results.

9.1. The effectiveness of the MYStIX selection

When selecting cluster members based on their NIR colours, typically half of the sources are discarded after spectral classification because they are background or foreground stars. For example, Wu et al. (2016) selected 44 candidate OB stars

belonging to the cluster W49 to perform a spectroscopic follow up. They found that 22 of the observed spectra (50%) are dominated by CO in absorption, which led them to conclude that they are foreground or background stars. Gennaro et al. (2012) studied the CN15/16/17 molecular complex; they aimed at observing the massive stars in the complex and by classifying them spectroscopically, they constrained the distance and extinction towards the stars. The authors selected the spectroscopic sample based on the NIR colours (brightest stars in K and $H - K_s > 0.3$). Their sample consists of 14 stars out of which five (35%) are field giants polluting the sample of massive stars. Similarly, Maaskant et al. (2011) studied the massive star population of IRAS 06084–0611. They selected 39 stars to follow up

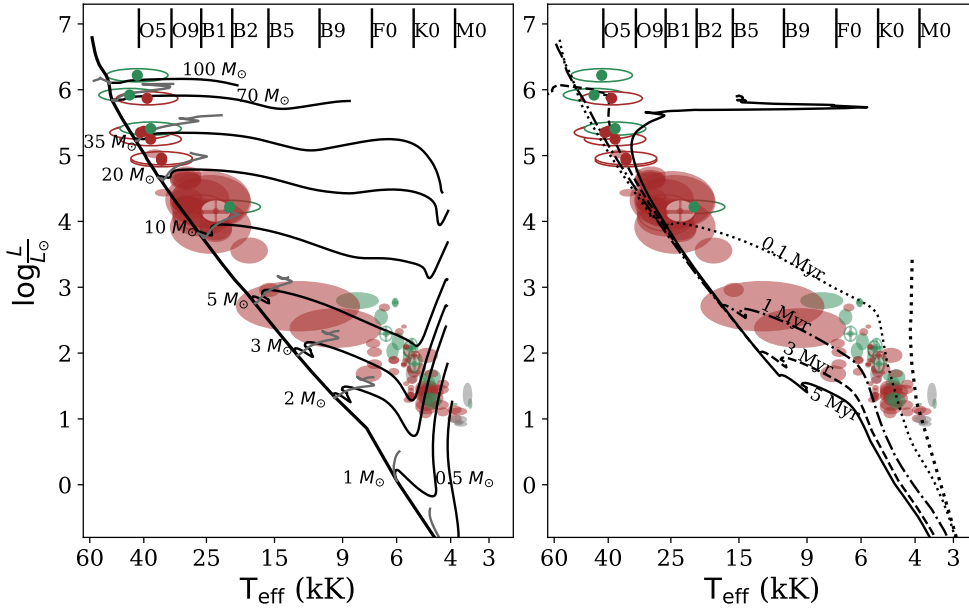


Fig. 18. Same as Fig. 16 but for the classified stars in NGC 6357. The age is estimated to be 0.5–3 Myr.

spectroscopically and after looking for highly reddened sources, IR excess in the IRAC bands, evidence of jets or outflows, and the presence of a radio counterpart, they identified 28 objects as cluster members (so 28% as foreground and background stars).

The importance of the MYStIX selection becomes very clear in this paper where two regions are included in the MPCM list (M 8 and NGC 6357) and one is not (G333.6–0.2). The MYStIX selection is especially important in the case of low-mass PMS stars. Given that they are more abundant than massive stars, when these stars are selected on the basis of their NIR colours only, the probability that these are foreground or background sources is high. Nevertheless, adding the X-ray information points towards these sources being PMS stars. In the case of our two MYStIX regions – M 8 and NGC 6357 – the fraction of stars whose membership is not confirmed based on their spectral type and astrometric properties is very low (<10%), while in G333.6–0.2 the fraction of stars with uncertain membership is around 30%.

In all HRDs a group of stars (grey symbols) is located at the right side of the 0.1 Myr isochrone near the birth line. These objects correspond to late-K and M-type stars and the most likely explanation for them to be at that location is that they are not members of the cluster, and so contaminate the MYStIX MPCM list.

9.2. Accuracy of the age determination

We constrained the age of each region in Sect. 8.2. Here we discuss some of the issues that might affect the age determination and compare our age estimates with those published in the literature. These issues are:

(i) Binarity, which could affect the luminosity determination by making stars look brighter than they actually are. No information about the multiplicity properties of our sources is however available. (ii) The use of different evolutionary tracks, which could lead to different age estimates, e.g. [Martins & Palacios \(2013\)](#) show that the main-sequence age for different models could vary by ~ 0.7 Myr. In our case, this uncertainty is likely smaller than that caused by the uncertainty on the spectral type. (iii) The adopted extinction law, which may impact the age determination. We tested the effect of two different extinction laws

on the age determination. By using the [Nishiyama et al. \(2009\)](#) extinction law, we obtain lower A_K values, and therefore slightly fainter objects. However, the implied difference in luminosity does not significantly affect the age determination.

[Prisinzano et al. \(2005, 2019\)](#) estimated the age of the PMS stars in M 8 between ~ 0.8 and 2.5 Myr, in agreement with our estimate – including the massive stars – of 1–3 Myr. [Massi et al. \(2015\)](#) determined the age of NGC 6357 at 1–3 Myr based on the visual and NIR CMD and CCDs. By comparing the whole population of stars with MIST isochrones, we estimate an age of 0.5 to 3 Myr for this cluster. For G333.6–0.2 we could not find a previous age determination.

The age obtained for the studied regions by considering only the massive stars is consistent with the age obtained only from the low-mass PMS stars. If the late-K and M-type stars near the Hayashi track were cluster members, we would obtain a younger age (~ 0.1 –1 Myr younger) for the PMS stars compared to that of the massive stars only.

9.3. Ionisation rates

[Binder & Povich \(2018\)](#) did a multi-wavelength study of Galactic massive star-forming regions where they studied their energy budgets and star formation rates. They obtained the total infrared luminosity processed by dust from the radio free-free continuum and the global infrared SEDs (3.6–500 μm). In the cases of M 8 and NGC 6357 they used a similar sample of ionising stars as we do in this paper, but in the case of G333 they had to rely on an incomplete stellar population. They concluded that in the first two cases there are more than enough stars to explain both the observed ionisation and the bolometric luminosity. For the entire G333 complex (of which G333.6–0.2 is the brightest component), they calculated that around a dozen O5V stars are necessary to explain the total luminosity of the complex, and therefore concluded that the principal ionising stars in the complex remain undiscovered. Due to the ambiguity in the stellar population of G333 [Binder & Povich \(2018\)](#) assumed the presence of five O5V stars in G333.6–0.2 based on a photometric study by [Fujiyoshi et al. \(2005\)](#). We only classified three new OB stars in G333.6–0.2, all of which have spectral types later than O5V (hence produce less ionising photons). Therefore, we

conclude that there are probably more OB stars to be discovered in this region.

10. Conclusions

We have presented a photometric and spectroscopic study of massive stars in three H II regions and characterised their stellar populations in terms of age and mass ranges.

- M 8, G333.6–0.2 and NGC 6357 contain young stellar populations: we detect O and B stars, several mYSOs as well as a low-mass PMS stellar population.
- We develop an automatic spectral classification method in the *H*- and *K*-band, that is accurate to within two spectral subtypes. We conclude that this method is at least as accurate as a classification by eye, while having the obvious merits of being fast and reproducible.
- The most massive stars in our KMOS sample are objects 73 (O9-B1V) in M 8, 6829462 (O6V) in G333.6–0.2, and 16 (O9-B0V) in NGC 6357. The most massive stars published in the literature and not included in our sample for these clusters are HD 164794 (O4V(f)) in M 8, and Pismis 24 17 (O3.5III(f*)) in NGC 6357.
- We cross matched our catalogues with the *Gaia* DR2 database in order to further constrain the cluster membership of our sources based on their astrometric properties. For each of the clusters we found two sources that are not cluster members. F13 and F24 in M 8; 515756828027 and 515757115402 in G333.6–0.2; 91 and 8 in NGC 6357.
- The sample selection for G333.6–0.2 was made based on the NIR-colours only. By classifying the stars spectroscopically and by using the information provided by *Gaia* (when available), we conclude that $\sim 30\%$ of the stars observed in G333.6–0.2 are not cluster members, but luminous background stars. For the other regions we used the MYStIX selection (Feigelson et al. 2013) in order to select cluster members; in these case the fraction of non-members is less than 10%. The latter demonstrates the importance and effectiveness of the MYStIX analysis when identifying cluster members.
- Based on the presence of CO bandheads in emission and/or (double-peaked) emission lines in the spectra, we find one mYSO candidate in G333.6–0.2 (720), one in M 8 (32), and four in NGC 6357 (33, F57, 114, and F68).
- We determined the *K*-band extinction of the observed stars by comparing their intrinsic NIR colours with the observed ones. In all regions A_K varies from ~ 0 to ~ 3 mag. The mean extinction of NGC 6357 is higher than that of M 8. We do not find strong evidence for a correlation between the extinction value and the location of the stars in the clusters suggesting that the extinction varies on small spatial scales, that is, is patchy.
- The observed stellar population of M 8 has an age of ~ 1 – 3 Myr, and that of NGC 6357 is 0.5 – 3 Myr old. From the OB stars in G333.6–0.2 we estimate an age < 3 Myr. Given that this region did not profit from the MYStIX analysis, we only use the main sequence stars to determine the age. A more accurate determination is at present not possible given the uncertainty of the spectral classification.

With present-day 8–10 m class ground-based telescopes we may resolve stellar populations in the Magellanic Clouds. With future extremely large telescopes, like the ELT and TMT, we are going to resolve young stellar populations in more distant galaxies. With the techniques used and introduced here we may charac-

terise these regions in terms of content, mass, age, and multiplicity properties even in highly obscured regions. This will allow us to probe massive star formation in galaxies spanning a wide range of types, masses, and metallicities.

Acknowledgements. Based on observations collected at the European Organisation for Astronomical Research in the Southern Hemisphere under ESO program 095.C-0048. The authors thank Leisa Townsley, Eric Feigelson, Matthew Povich and Patrick Broos for valuable discussions on the MYStIX catalogue and assistance with the construction of the final target lists for the KMOS observations. The authors thank Richard Davies, Ray Sharples, Nadine Neumayer and Michael Hilker for discussions on KMOS data reduction. JP and LK acknowledge support from their NWO/PAPESP grant for Advanced instrumentation (MOSAIC). HB acknowledges support from the European Research Council under the European Community’s Horizon 2020 frame-work program (2014–2020) via the ERC Consolidator grant “From Cloud to Star Formation (CSF)” (project number 648505). AP acknowledges support from the Sonderforschungsbereich SFB 881 “The Milky Way System” (subproject B5) of the German Research Foundation (DFG). O.H.R.A. acknowledges funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 665593 awarded to the Science and Technology Facilities Council. This research made use of Astropy, a community-developed core Python package for Astronomy (Astropy Collaboration 2018), APLpy, an open-source plotting package for Python (Robitaille & Bressert 2012), NASA’s Astrophysics Data System Bibliographic Services (ADS), the SIMBAD database, operated at CDS, Strasbourg, France (Wenger et al. 2000), and the cross-match service provided by CDS, Strasbourg. The VISTA Data Flow System pipeline processing and science archive are described in Irwin et al. (2004), Hambly et al. (2008), and Cross et al. (2012). We have used data from the 2nd data release of the VVV survey (Saito et al. 2012). This work has made use of data from the European Space Agency (ESA) mission *Gaia* (<https://www.cosmos.esa.int/gaia>), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, <https://www.cosmos.esa.int/web/gaia/dpac/consortium>). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.

References

- Alonso, A., Arribas, S., & Martínez-Roger, C. 1999, *A&AS*, **140**, 261
 Arias, J. I., Barbá, R. H., & Morrell, N. I. 2007, *MNRAS*, **374**, 1253
 Astropy Collaboration (Price-Whelan, A. M., et al.) 2018, *AJ*, **156**, 123
 Beltrán, M. T., & de Wit, W. J. 2016, *A&ARv*, **24**, 6
 Benjamin, R. A., Churchwell, E., Babler, B. L., et al. 2003, *PASP*, **115**, 953
 Beuther, H., Walsh, A. J., Johnston, K. G., et al. 2017, *A&A*, **603**, A10
 Bik, A., & Thi, W. F. 2004, *A&A*, **427**, L13
 Bik, A., Kaper, L., Hanson, M. M., & Smits, M. 2005, *A&A*, **440**, 121
 Bik, A., Kaper, L., & Waters, L. B. F. M. 2006, *A&A*, **455**, 561
 Bik, A., Henning, T., Stolte, A., et al. 2012, *ApJ*, **744**, 87
 Bik, A., Stolte, A., Gennaro, M., et al. 2014, *A&A*, **561**, A12
 Bik, A., Henning, T., Wu, S.-W., et al. 2019, *A&A*, **624**, A63
 Binder, B. A., & Povich, M. S. 2018, *ApJ*, **864**, 136
 Blum, R. D., Barbosa, C. L., Damineli, A., Conti, P. S., & Ridgway, S. 2004, *ApJ*, **617**, 1167
 Broos, P. S., Getman, K. V., Povich, M. S., et al. 2013, *ApJS*, **209**, 32
 Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *ApJ*, **345**, 245
 Castelli, F., & Kurucz, R. L. 2003, in *Proceedings of the 210th Symposium of the International Astronomical Union held at Uppsala University, Uppsala, Sweden, 17–21 June, 2002*, eds. N. Piskunov, W. W. Weiss, & D. F. Gray (Astronomical Society of the Pacific), A20
 Cesaroni, R., Sánchez-Monge, Á., Beltrán, M. T., et al. 2017, *A&A*, **602**, A59
 Charbonneau, P. 1995, *ApJS*, **101**, 309
 Choi, J., Dotter, A., Conroy, C., et al. 2016, *ApJ*, **823**, 102
 Cox, A. N. 2000, *Allen’s Astrophysical Quantities* (New York: AIP Press)
 Cross, N. J. G., Collins, R. S., Mann, R. G., et al. 2012, *A&A*, **548**, A119
 Cushing, M. C., Rayner, J. T., & Vacca, W. D. 2005, *ApJ*, **623**, 1115
 Davies, B., Lumsden, S. L., Hoare, M. G., Oudmaijer, R. D., & de Wit, W.-J. 2010, *MNRAS*, **402**, 1504
 Davies, R. I., Agudo Berbel, A., Wiezorrek, E., et al. 2013, *A&A*, **558**, A56
 Dotter, A. 2016, *ApJS*, **222**, 8
 Dunstall, P. R., Dufton, P. L., Sana, H., et al. 2015, *A&A*, **580**, A93
 Feigelson, E. D., Townsley, L. K., Broos, P. S., et al. 2013, *ApJS*, **209**, 26
 Figuerêdo, E., Blum, R. D., Damineli, A., & Conti, P. S. 2005, *AJ*, **129**, 1523
 Fujiyoshi, T., Smith, C. H., Moore, T. J. T., et al. 2005, *MNRAS*, **356**, 801
 Gaia Collaboration (Prusti, T., et al.) 2016, *A&A*, **595**, A1
 Gaia Collaboration (Brown, A. G. A., et al.) 2018, *A&A*, **616**, A1
 Gennaro, M., Bik, A., Brandner, W., et al. 2012, *A&A*, **542**, A74

- Gray, R. O., & Corbally, Jr., C. 2009, *Stellar Spectral Classification* (Princeton: Princeton University Press)
- Hambly, N. C., Collins, R. S., Cross, N. J. G., et al. 2008, *MNRAS*, **384**, 637
- Hanson, M. M., Conti, P. S., & Rieke, M. J. 1996, *ApJS*, **107**, 281
- Hanson, M. M., Howarth, I. D., & Conti, P. S. 1997, *ApJ*, **489**, 698
- Hanson, M. M., Kudritzki, R.-P., Kenworthy, M. A., Puls, J., & Tokunaga, A. T. 2005, *ApJS*, **161**, 154
- Ilee, J. D., Wheelwright, H. E., Oudmaijer, R. D., et al. 2013, *MNRAS*, **429**, 2960
- Ilee, J. D., Cyganowski, C. J., Nazari, P., et al. 2016, *MNRAS*, **462**, 4386
- Indebetouw, R., Mathis, J. S., Babler, B. L., et al. 2005, *ApJ*, **619**, 931
- Irwin, M. J., Lewis, J., Hodgkin, S., et al. 2004, in *Optimizing Scientific Return for Astronomy through Information Technologies*, eds. P. J. Quinn, & A. Bridger, *Proc. SPIE*, **5493**, 411
- Ivanov, V. D., Rieke, M. J., Engelbracht, C. W., et al. 2004, *VizieR Online Data Catalog*: 215
- Kiminki, D. C., & Kobulnicky, H. A. 2012, *ApJ*, **751**, 4
- King, R. R., Naylor, T., Broos, P. S., Getman, K. V., & Feigelson, E. D. 2013, *ApJS*, **209**, 28
- Kobulnicky, H. A., Kiminki, D. C., Lundquist, M. J., et al. 2014, *ApJS*, **213**, 34
- Kuhn, M. A., Povich, M. S., Luhman, K. L., et al. 2013, *ApJS*, **209**, 29
- Kuhn, M. A., Hillenbrand, L. A., Sills, A., Feigelson, E. D., & Getman, K. V. 2019, *ApJ*, **870**, 32
- Kumar, M. S. N. 2013, *A&A*, **558**, A119
- Kumar, D. L., & Anandarao, B. G. 2010, *MNRAS*, **407**, 1170
- Lapenna, E., Origlia, L., Mucciarelli, A., et al. 2015, *ApJ*, **798**, 23
- Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, *MNRAS*, **379**, 1599
- Lindgren, L. 2018, *Re-normalising the Astrometric Chi-square in Gaia DR2*, GAIA-C3-TN-LU-LL-124
- Lindgren, L., Hernández, J., Bombrun, A., et al. 2018, *A&A*, **616**, A2
- Lucas, P. W., Hoare, M. G., Longmore, A., et al. 2008, *MNRAS*, **391**, 136
- Maaskant, K. M., Bik, A., Waters, L. B. F. M., et al. 2011, *A&A*, **531**, A27
- Maíz Apellániz, J., Walborn, N. R., Morrell, N. I., Niemela, V. S., & Nelan, E. P. 2007, *ApJ*, **660**, 1480
- Maíz Apellániz, J., Sota, A., Arias, J. I., et al. 2016, *ApJS*, **224**, 4
- Martins, F., & Palacios, A. 2013, *A&A*, **560**, A16
- Mason, B. D., Hartkopf, W. I., Gies, D. R., Henry, T. J., & Helsel, J. W. 2009, *AJ*, **137**, 3358
- Massi, F., Giannetti, A., Di Carlo, E., et al. 2015, *A&A*, **573**, A95
- Mookerjee, B., Kramer, C., Nielbock, M., & Nyman, L.-Å. 2004, *A&A*, **426**, 119
- Motte, F., Bontemps, S., & Louvet, F. 2018, *ARA&A*, **56**, 41
- Naylor, T., Broos, P. S., & Feigelson, E. D. 2013, *ApJS*, **209**, 30
- Nishiyama, S., Tamura, M., Hatano, H., et al. 2009, *ApJ*, **696**, 1407
- Paxton, B., Bildsten, L., Dotter, A., et al. 2011, *ApJS*, **192**, 3
- Paxton, B., Cantiello, M., Arras, P., et al. 2013, *ApJS*, **208**, 4
- Paxton, B., Marchant, P., Schwab, J., et al. 2015, *ApJS*, **220**, 15
- Pecaut, M. J., & Mamajek, E. E. 2013, *ApJS*, **208**, 9
- Peter, D., Feldt, M., Henning, T., & Hormuth, F. 2012, *A&A*, **538**, A74
- Povich, M. S., Kuhn, M. A., Getman, K. V., et al. 2013, *ApJS*, **209**, 31
- Povich, M. S., Busk, H. A., Feigelson, E. D., Townsley, L. K., & Kuhn, M. A. 2017, *ApJ*, **838**, 61
- Prisinzano, L., Damiani, F., Micela, G., & Sciortino, S. 2005, *A&A*, **430**, 941
- Prisinzano, L., Damiani, F., Micela, G., & Pillitteri, I. 2007, *A&A*, **462**, 123
- Prisinzano, L., Damiani, F., Kalari, V., et al. 2019, *A&A*, **623**, A159
- Ramírez-Tannus, M. C., Kaper, L., de Koter, A., et al. 2017, *A&A*, **604**, A78
- Rayner, J. T., Cushing, M. C., & Vacca, W. D. 2009, *ApJS*, **185**, 289
- Robitaille, T., & Bressert, E. 2012, *Astrophysics Source Code Library* [record ascl:1208.017]
- Saito, R. K., Hempel, M., Minniti, D., et al. 2012, *A&A*, **537**, A107
- Sana, H., Le Bouquin, J.-B., Lacour, S., et al. 2014, *ApJS*, **215**, 15
- Sembach, K. R., & Savage, B. D. 1992, *ApJS*, **83**, 147
- Sharples, R., Bender, R., Agudo Berbel, A., et al. 2013, *The Messenger*, **151**, 21
- Skiff, B. A. 2009, *VizieR Online Data Catalog*: 1
- Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, *AJ*, **131**, 1163
- Tothill, N. F. H., White, G. J., Matthews, H. E., et al. 2002, *ApJ*, **580**, 285
- Tothill, N. F. H., Gagné, M., Stecklum, B., & Kenworthy, M. A. 2008, *The Lagoon Nebula and its Vicinity* (Astronomical Society of the Pacific), 533
- Townsley, L. K., Broos, P. S., Garmire, G. P., et al. 2014, *ApJS*, **213**, 1
- Townsley, L. K., Broos, P. S., Garmire, G. P., et al. 2018, *ApJS*, **235**, 43
- Wang, J., Townsley, L. K., Feigelson, E. D., et al. 2007, *ApJS*, **168**, 100
- Wegner, W. 2000, *MNRAS*, **319**, 771
- Wenger, M., Ochsenbein, F., Egret, D., et al. 2000, *A&AS*, **143**, 9
- Wheelwright, H. E., Oudmaijer, R. D., de Wit, W. J., et al. 2010, *MNRAS*, **408**, 1840
- Wu, S.-W., Bik, A., Henning, T., et al. 2014, *A&A*, **568**, L13
- Wu, S.-W., Bik, A., Bestenlehner, J. M., et al. 2016, *A&A*, **589**, A16
- Zinnecker, H., & Yorke, H. W. 2007, *ARA&A*, **45**, 481

Appendix A: “Plus one” method to combine the likelihood distributions

To combine the likelihoods of spectral type prediction by each feature’s EW and obtain the final likelihood distribution for each luminosity class, one needs to multiply all the likelihood functions for all possible EW available:

$$L(\text{SpT})_{\text{tot}} = \prod_{\text{EW}} L(\text{SpT})_{\text{EW}}. \quad (\text{A.1})$$

(To avoid numerical zeros, the multiplication was in fact carried out as a sum of logarithmic values

$\sum_{\text{EW}} \log(L_{\text{EW}})$ and then converted back to linear values). The resulting distribution function displays very sharp peaks, as the multiplication amplifies the overlapping peaks of different $L(\text{SpT})_{\text{EW}}$ by orders of magnitude while oppressing other peaks to nearly zero. Therefore, the final distribution obtained by this method does not provide a good uncertainty estimate for the spectral type predictions.

Alternatively, we combine the “plus one” quantity ($L(\text{SpT})_{\text{EW}} + 1$) by multiplication (Eq. (4)). The outcome of this operation gives a better likelihood distribution overall, with peaks and reasonable spreads around them to be used as the uncertainty estimation for the spectral type prediction.

This is not only a numerical trick to broaden the widths of the distribution functions and avoid numerical zeros, but also has physical meanings. To shorten the notation, the formula above is written as $L = \prod_i (G_i + 1)$ where i is the index of a list of n EWs. Expand the product as a binomial expansion we can obtain

$$\prod_{i=1}^n (G_i + 1) = 1 + \sum_{i=1}^n G_i + \sum_{i,j} G_i G_j + \sum_{i,j,k} G_i G_j G_k + \dots + \prod_{i=1}^n G_i \quad (\text{A.2})$$

where each collected term is the collection of terms with the same number of G -functions multiplied together, runs from zero to n . For example, $\sum_{i,j} G_i G_j$ means the sum of all terms with two G -functions like $G_i G_j$ taking all possible choices of indices (i, j) for $i \neq j \in \{1, 2, \dots, n\}$ and the different sequence of the indices with the same numbers are not counted as extra terms.

We can immediately find out the last term $\prod_{i=1}^n G_i$ is exactly the same as the result of direct multiplication: the combined likelihood function considering the predictions of all available EWs. Then, the terms with just one G -function missing represent all the possible cases where we have one of the EWs information missing, and use the rest $n - 1$ EWs to determine the combined likelihood. After that, we come to the terms with all the cases where two G -functions replaced by 1 and multiply the rest $n - 2$ G s together. . . all the way down to $\sum_{i=1}^n G_i$ in which cases only one EW is available and the combined distributions are the same as the G -functions. Summing up all the terms, we obtain a likelihood function which not only considers the combination of all available EWs, but all the possible cases where one or more of the EW information are missing. This method introduces robustness when there are a few unknown “bad” measurements mixed in the EW but the majority of the data are properly produced, as the G -function of each EW appears in only half of the terms in the binomial expansion, the contribution of a low quality EW measurement is reduced to half, at the cost of having a larger uncertainty range than the direct combination method.

Appendix B: Spectral features

Table B.1. Spectral features used for equivalent width (EW) calculations for the F-M type stars. All wavelengths are in μm .

Line (λ_c in μm)	Feature	Continuum 1	Continuum 2	Continuum 3
Mg I(1.485) ^(a)	1.475–1.497	1.477–1.485	1.491–1.497	
Mg I (1.50) ^(b)	1.5020–1.5060	1.4990–1.5020	1.5060–1.5090	
Fe I (1.58) ^(b)	1.5810–1.5850	1.5790–1.5810	1.5850–1.5865	
Si I (1.59) ^(b)	1.5870–1.5910	1.5830–1.5870	1.5910–1.5950	
CO (1.62) ^(b)	1.6175–1.6220	1.6145–1.6175	1.6255–1.6285	
Br 12 (1.64)	1.6385–1.6439	1.6281–1.6294	1.6503–1.6510	
Br 11 (1.68)	1.6781–1.6849	1.6700–1.6704	1.702 –1.7053	
Mg I a (1.71) ^(b)	1.7100–1.7130	1.7085–1.7100	1.7130–1.7160	
Mg I b (1.711) ^(a)	1.695 –1.726	1.702–1.708	1.715–1.720	
Na I (2.206) ^(a)	2.185–2.230	2.192–2.198	2.213–2.220	
Mg I (2.11) ^(b)	2.1040–2.1110	2.1000–2.1040	2.1110–2.1150	
Bry (2.16) ^(b)	2.1639–2.1686	2.0320–2.034	2.0907–2.0951	2.2873–2.2925
Na I a (2.21) ^(b)	2.2040–2.2110	2.2140–2.2200	2.2330–2.2370	
Na I b (2.21) ^(b)	2.2053–2.2101	2.0907–2.0951	2.2873–2.2925	
Ca I a (2.26) ^(b)	2.2580–2.2690	2.2490–2.2530	2.2700–2.2740	
Ca I b (2.26) ^(b)	2.2611–2.2662	2.0907–2.0951	2.2873–2.2925	
Mg I a (2.28) ^(b)	2.2780–2.2832	2.2720–2.2780	2.2832–2.2882	
Mg I b (2.28) ^(b)	2.2790–2.2850	2.2700–2.2740	2.2871–2.2901	
CO (2.29) ^(b)	2.2924–2.2977	2.2873–2.2919	2.305 –2.32	
¹² CO (2,0) (2.29) ^(b)	2.2930–2.3030	2.2873–2.2919	2.305 –2.32	
¹² CO (3,1) (2.32) ^(b)	2.3218–2.3272	2.2873–2.2919	2.340 –2.341	
¹³ CO (2,0) (2.35) ^(b)	2.3436–2.3491	2.2873–2.2925	2.399 –2.3997	
CO (2.35) ^(b)	2.3100–2.4000	2.2873–2.2925	2.399 –2.3997	

References. ^(a)Rayner et al. (2009). ^(b)Ivanov et al. (2004).

Table B.2. Spectral features used for equivalent width (EW) calculations for the O-A type stars, based on normalized spectra.

Line (λ_c in μm)	Feature	Continuum 1	Continuum 2
Br 12 (1.64)	1.6371–1.6451	1.705–1.710	
Br 11 (1.68)	1.6772–1.6850	1.705–1.710	
He II (1.69)	1.6895–1.6951	1.705–1.710	
He I triplet (1.70)	1.6993–1.7021	1.705–1.710	
Br 10 (1.74)	1.7327–1.7407	1.705–1.710	
C IV (2.08)	2.0773–2.0809	2.1813–2.1839	2.1987–2.1991
He I triplet (2.11)	2.1113–2.1132	2.1813–2.1839	2.1987–2.1991
He I (2.11)	2.1132–2.1146	2.1813–2.1839	2.1987–2.1991
He I (2.16)	2.1596–2.1639	2.1813–2.1839	2.1987–2.1991
Br 3 (2.16)	2.1639–2.1686	2.1813–2.1839	2.1987–2.1991
He II (2.19)	2.1861–2.1921	2.1813–2.1839	2.1987–2.1991

Notes. All wavelengths are in μm . Lines taken from [Hanson et al. \(2005\)](#).

Table B.3. Line ratios used for spectral typing of late-type stars.

$\frac{\text{CO1}(2.29)}{\text{Br3}(2.16)}$	$\frac{\text{CO2}(2.29)}{\text{Br12}(1.64)}$	$\frac{\text{CO}(2.35)}{\text{MgI}(1.50)}$	$\frac{\text{CO}(1.62)}{\text{Br3}(2.16)}$	$\frac{\text{Br3}(2.16)}{\text{NaI}(2.206)}$	$\frac{\text{Br3}(2.16)}{\text{MgI}(2.11)}$	$\frac{\text{Br12}(1.64)}{\text{MgI}(2.11)}$	$\frac{13\text{CO}(2,0)(2.35)}{\text{MgI}(2.11)}$
$\frac{\text{CO1}(2.29)}{\text{Br11}(1.68)}$	$\frac{\text{CO2}(2.29)}{\text{SiI}(1.59)}$	$\frac{\text{CO}(2.35)}{\text{NaI}(2.206)}$	$\frac{\text{CO}(1.62)}{\text{Br11}(1.68)}$	$\frac{\text{Br11}(1.68)}{\text{SiI}(1.59)}$	$\frac{\text{Br3}(2.16)}{\text{CaI1}(2.26)}$	$\frac{\text{Br12}(1.64)}{\text{CaI1}(2.26)}$	$\frac{13\text{CO}(2,0)(2.35)}{\text{CaI1}(2.26)}$
$\frac{\text{CO1}(2.29)}{\text{Br12}(1.64)}$	$\frac{\text{CO2}(2.29)}{\text{MgI}(1.50)}$	$\frac{13\text{CO}(2,0)(2.35)}{\text{Br3}(2.16)}$	$\frac{\text{CO}(1.62)}{\text{Br12}(1.64)}$	$\frac{\text{Br11}(1.68)}{\text{MgI}(1.50)}$	$\frac{\text{Br3}(2.16)}{\text{CaI2}(2.26)}$	$\frac{\text{Br12}(1.64)}{\text{CaI2}(2.26)}$	$\frac{13\text{CO}(2,0)(2.35)}{\text{CaI2}(2.26)}$
$\frac{\text{CO1}(2.29)}{\text{SiI}(1.59)}$	$\frac{\text{CO2}(2.29)}{\text{NaI}(2.206)}$	$\frac{13\text{CO}(2,0)(2.35)}{\text{Br11}(1.68)}$	$\frac{\text{CO}(1.62)}{\text{SiI}(1.59)}$	$\frac{\text{Br11}(1.68)}{\text{NaI}(2.206)}$	$\frac{\text{Br11}(1.68)}{\text{NaI1}(2.21)}$	$\frac{\text{CO}(2.35)}{\text{NaI1}(2.21)}$	$\frac{\text{CO}(1.62)}{\text{NaI1}(2.21)}$
$\frac{\text{CO1}(2.29)}{\text{MgI}(1.50)}$	$\frac{\text{CO}(2.35)}{\text{Br3}(2.16)}$	$\frac{13\text{CO}(2,0)(2.35)}{\text{Br12}(1.64)}$	$\frac{\text{CO}(1.62)}{\text{MgI}(1.50)}$	$\frac{\text{Br12}(1.64)}{\text{SiI}(1.59)}$	$\frac{\text{Br11}(1.68)}{\text{MgI}(2.11)}$	$\frac{\text{CO}(2.35)}{\text{MgI}(2.11)}$	$\frac{\text{CO}(1.62)}{\text{MgI}(2.11)}$
$\frac{\text{CO1}(2.29)}{\text{NaI}(2.206)}$	$\frac{\text{CO}(2.35)}{\text{Br11}(1.68)}$	$\frac{13\text{CO}(2,0)(2.35)}{\text{SiI}(1.59)}$	$\frac{\text{CO}(1.62)}{\text{NaI}(2.206)}$	$\frac{\text{Br12}(1.64)}{\text{MgI}(1.50)}$	$\frac{\text{Br11}(1.68)}{\text{CaI1}(2.26)}$	$\frac{\text{CO}(2.35)}{\text{CaI1}(2.26)}$	$\frac{\text{CO}(1.62)}{\text{CaI1}(2.26)}$
$\frac{\text{CO2}(2.29)}{\text{Br3}(2.16)}$	$\frac{\text{CO}(2.35)}{\text{Br12}(1.64)}$	$\frac{13\text{CO}(2,0)(2.35)}{\text{MgI}(1.50)}$	$\frac{\text{Br3}(2.16)}{\text{SiI}(1.59)}$	$\frac{\text{Br12}(1.64)}{\text{NaI}(2.206)}$	$\frac{\text{Br11}(1.68)}{\text{CaI2}(2.26)}$	$\frac{\text{CO}(2.35)}{\text{CaI2}(2.26)}$	$\frac{\text{CO}(1.62)}{\text{CaI2}(2.26)}$
$\frac{\text{CO2}(2.29)}{\text{Br11}(1.68)}$	$\frac{\text{CO}(2.35)}{\text{SiI}(1.59)}$	$\frac{13\text{CO}(2,0)(2.35)}{\text{NaI}(2.206)}$	$\frac{\text{Br3}(2.16)}{\text{MgI}(1.50)}$	$\frac{\text{Br3}(2.16)}{\text{NaI1}(2.21)}$	$\frac{\text{Br12}(1.64)}{\text{NaI1}(2.21)}$	$\frac{13\text{CO}(2,0)(2.35)}{\text{NaI1}(2.21)}$	