DEPARTAMENTO DE ASTROFISICA

Universidad de La Laguna

MASSIVE OBSCURED CLUSTERS IN THE MILKY WAY

Tesis doctoral que presenta

Don Klaus Simón Rübke Zúñiga

para optar a la titulación

de Doctor en Astrofísica.

Trabajo dirigido por el

Dr. Artemio Herrero Davó

INSTITUTO Đ ASTROFISICA Đ CANARIAS

La Laguna, junio de 2017

Examination date: JUN, 2017
Thesis supervisor: Artemio Herrero Davó

Space, the final frontier.

Mr. Spock.

Resumen

Esta thesis se concentra en dos principales bloques, por un lado en la búsqueda y caracterizacion de cúmulos masivos obscurecidos en el disco de la Vía lactea y en el análisis espectroscópico cuantitativo de estrellas O en el infrarrojo.

Cúmulos masivos son aquellos que poseen masa superior a $10^4 M_\odot$. El número de cúmulos masivo conocidos esta por muy debajo a la prediccion real, ya que por lo general estan obscurecidos por el gas y polvo que los rodea, lo que hace dificil encontrarlos en rango óptico, por ese motivo es necesario usar el infrarrojo cercano para encontrarlos y observarlos.

Bajo el proyecto **MASGOMAS** (**MA**ssive **S**tars in **G**alactic **O**bscured **MA**ssive cluster**S**), este estudio se dedica a la búsqueda de cúmulos massivos obscurecidos usando fotometria 2MASS en el disco norte de la galaxia , con el proposito de usar LIRIS@WHT (William Hershel Telescope) como principal telescopio para realizar las observaciones.

El método usado para encontrar candidatos a cúmulos masivos fue desarrollar un algoritmo automatico de deteccion, el que consta principalmente de cortes fotometricos (K , $(J-K)$ y Q_{IR}) que eliminan gran parte de la población vieja y no masiva de el disco y un algoritmo de sobredensidades sobre estas estrellas candidatas. De este estudio presentamos en este trabajo la asociacion MASGOMAS-6, la cual resulto finalmente en dos poblaciones a diferentes distancias. En la ambicion de obtener únicamente candidatos de alta probabilidad de ser masivos el estudio de selección de asociaciones nos llevo a desarrolar un estudio en detalle de los parametros de selecion para las sobre densidades. Como resultado de estas mejoras, presentamos la asociacion MASGOMAS-10. En ambas asociaciones la poblacion masiva esta confirmada lo que muestra que nuestro método de búsqueda sistematica es efectivo.

La segunda parte esta centrada al analisis espectroscopico quantitativo de 12 estrellas O con espectros en alta resolucion en el óptico ($R \sim 46000$), que tambien estan presentes en el un catálogo de alta resolucion en el infrarrojo cercano ($R \sim 12000$). Usando IACOB Grid-Based Automatic Tool (IACOB-GBAT), un preciso y rapido método que compara de manera automatica, usando el método de χ^2, el espectro observado con los perfiles de líneas sinteticas de H y He. Obteniendo los parametros estelares de el óptico hemos testeado la posibilidad de obtenerlos usando solamente el infrarrojo cercano. Este estudio nos llevo al estudio de el no bien conocido aglutinamiento en el viento de las estrellas, por lo que creamos dos redes completas de modelos sinteticos con aglutinamiento en el viento.

Finalmente el estudio del aglutinamiento en el viento nos llevo a mejorar considerablemente el ajuste de líneas que se forman en el viento, especialmente para estellas supergigantes.

Summary

This thesis comprises of a full search in our galaxy for the hidden population of massive stars using the near-infrared photometry with a novel method and a quantitative spectroscopy study of an O star sample in the optical and near-infrared.

Massive clusters are those with mass greater than $10^4 M_\odot$. The number of known massive clusters is significantly lower than the real prediction, as in general, due to the gas and dust that surround those clusters make difficult to find them in the optical range and due to it is necessary to use near infrared to find and observe them.

Under the **MASGOMAS** (**MA**ssive **S**tars in **G**alactic **O**bscured **MA**ssive cluster**S**) project, the main objective of this thesis is to find massive stars using 2MASS photometry on the north disc of the galaxy, using LIRIS@WHT (William Hershel Telescope) as the main telescope to make the follow-up.

The method that we used to find the candidates for the massive stars was to develop an automatic detection algorithm, which mainly using photometric cuts (K , $(J-K)$ y Q_{IR}), which eliminate a large part of the late not-massive population, and an overdensity algorithm over the stars which fulfill the photometric cuts. From this study is presented the association MASGOMAS-6, that result in two populations at two different distances. To find high probability massive cluster candidates, we have developed a study that critically improves the way to select the overdensities. As result of this improvement we present the association MASGOMAS-10. In both associations the massive star population is confirmed, supporting the automatic method.

The second part focused on the quantitative spectroscopy study of 12 O-type stars with high resolution optical spectra ($R \sim 46000$), which are also present in a survey of high resolution in the near-infrared ($R \sim 12000$). To estimate stellar parameters we use the IACOB Grid-Based Automatic Tool (IACOB-GBAT), an accurate and fast method that compares, by means of a χ^2 algorithm, observed and synthetic H and He line profiles in an automated way. Deriving the optical stellar parameters we test the viability of deriving parameters using only near-infrared. This study also lead us to study the not well know clumping condition for our sample, creating two full grids of clumping models

Finally the clumping study results in a drastic improvement for several wind lines, especially for supergiant stars.

Index

1

Introduction

Contents

1.1 Introduction to massive stars

Massive star are mentioned several times as the architects of the modern universe due the energy and mass returned to the ISM (Abbott 1982). They also generate most of the ultraviolet (UV) ionizing radiation in galaxies, and power the far-infrared luminosities through the heating of dust. And, massive stars serve as the primary source of carbon, nitrogen, and oxygen (CNO) enrichment of the interstellar medium (ISM, Maeder 1981). They have more than 8 Solar masses ($M_{\odot}$) as initial mass and end their lives with a supernova explosion (Zinnecker

& Yorke 2007). The most massive star have an initial mass estimated up to 300 $M_\odot$ (Crowther et al. 2010; mass obtained from evolutionary tracks) and are found in intense star forming region such as the Tarantula nebula, located in the Large Magellanic Cloud, a satellite galaxy around the Milky Way at 50 kpc of distance.

The study carried out in this thesis is concentrated on O-type stars. These objects follow the regular stellar classification from 0 to 9 which is related to the effective temperatures, which ranges from 30000-50000 K (Conti 1973). O stars are also classified by luminosity class and evolutionary state such as WR (not discussed in this thesis). The luminosity class which is related with the size and the surface stellar gravity factors affect the intensity of the winds, high wind are associated to class I (supergiant) and low winds to class V (dwarf).

Table 1.1 and 1.2 shows a reference of masses and radii for the different O type stars provided by observational T_{eff} scales by Martins et al. (2005).

Table 1.1: O-stars initial masses $M_\odot$ as a function of spectral types and luminosity class, obtained with the observationals T_{eff} scales. (Martins et al. 2005).

Mass ($M_\odot$)			
	Luminosity class		
Subtype	V	III	I
O3	57.9	55.95	67.53
O4	46.9	47.43	58.54
O5	38.0	40.43	50.72
O5.5	34.3	37.35	47.25
O6	30.9	34.53	44.10
O6.5	28.00	31.96	41.20
O7	25.29	29.59	38.44
O7.5	22.90	27.45	36.00
O8	20.76	25.49	33.72
O8.5	18.80	23.68	31.54
O9	17.08	22.04	29.63
O9.5	15.55	20.55	27.83

1.2 Introduction to massive clusters

Massive stars are born in dense regions called *molecular clouds* or H_2 regions. These regions have the necessary amount of material (gas and dust) to form massive stars $\rho \sim 10^{23} cm^{-2}$, Zinnecker & Yorke 2007). An example is the Tarantula nebula. Inside it we can find a very massive starburst cluster: R136. This cluster is probably the most massive young compact cluster in the Local Group with a total mass of $\sim 10^5$ $M_\odot$ and an age of ~ 3 Myr (Andersen et al. 2009).

Another example of massive star forming region found in our galaxy in Cygnus-OB2. Where

Table 1.2: O Stellar radius (in R⊙) as a function of spectral types and luminosity class, obtained with the observationals T_{eff} scales. Martins et al. (2005).

Radii ($R_\odot$)			
	Luminosity class		
Subtype O	V	III	I
O3	13.80	16.19	18.56
O4	12.42	15.61	18.99
O5	11.20	15.07	19.45
O5.5	10.64	14.82	19.70
O6	10.11	14.58	19.95
O6.5	9.61	14.36	20.22
O7	9.15	14.14	20.49
O7.5	8.70	13.93	20.79
O8	8.29	13.74	21.10
O8.5	7.90	13.55	21.41
O9	7.53	13.38	21.76
O9.5	7.18	13.22	22.11

there is evidence of increased oxygen abundances and age as a function of galactic latitude (Rodriguez-Berlanas et al in prep). Similar to a supernova explosion, but on a smaller scale, strong and ionized winds from massive stars pressure the surrounding gas and may trigger a chain reaction of star formation (Lee & Chen 2007).

In our galaxy a huge number of massive clusters ($> 10^4 M\odot$) may still remain undiscovered; (close to 100, according to Hanson & Popescu 2008; or 80, according to Ivanov et al. 2010). But several of these clusters are optically extincted by the same dense region. Extinction and crowding makes the discovery of massive star clusters a challenging quest.

Massive stars are rarely born alone, recent studies shows that 70 percent of massive stars posses a companion (Sana et al. 2012). They are born in dense cores of gas and dust, which also provide enough material to form more massive stars. According to Lada & Lada (2003) the mayority (70-90%) of the stars in the Milky Way (MW) are formed in clusters.

The search of massive clusters has exponentially increased during the last decade thanks to the development of infrared instrumentation, which resulted in surveys such as the Two Micron All Sky Survey (2MASS, Skrutskie et al. 2006) and Deep Near Infrared Survey of the Southern Sky (DENIS, Epchtein et al. 1997). Recently United Kingdom Infra-Red Telescope (UKIRT), Infrared Deep Sky Survey (UKIDSS; Lawrence et al. (2007)) and the ESO Public Survey Vista Variables in the Vía Láctea (VVV; Minniti et al. 2010) have scanned the Milky Way from the Northern and the Southern hemisphere, respectively. In this thesis we base the near-infrared study on 2MASS and as reference, the Figure 1.1 illustrates the range for 2MASS and VVV considering the extinction limits.

The mid-infrared data allow us to study the dust in our Galaxy. The survey GLIMPSE and

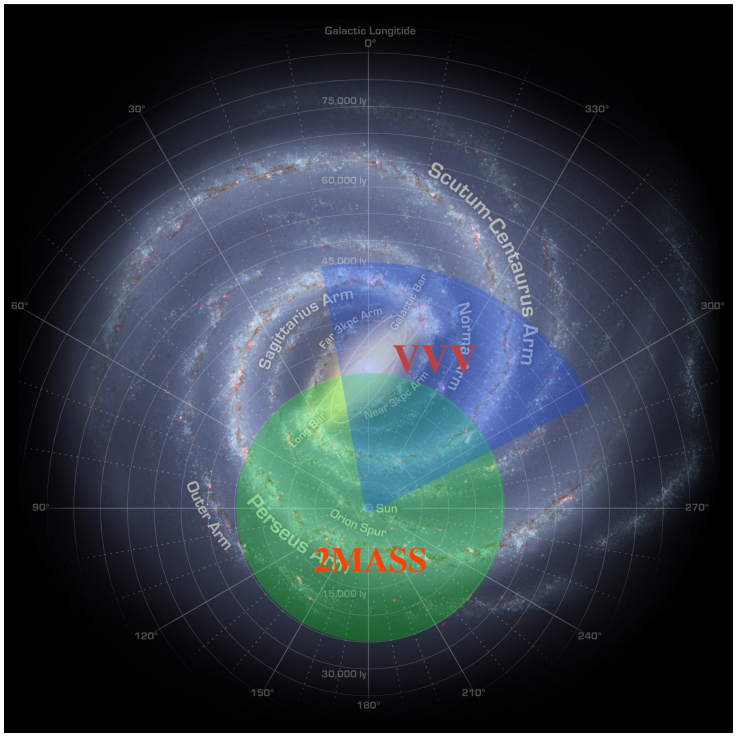

Figure 1.1: Survey areas of 2MASS full sky and VVV, which contains the center of the galaxy and the disk from galactic latitude of [10°,295°]. Milky Way image credits: NASA/Adler/U. Chicago/Wesleyan/JPL-Caltech.

GLIMPSE II (Benjamin et al. 2003; Churchwell et al. 2009), completed using the IRAC camera on Spitzer Space Telescope, gave us access to this kind of data. Considering that massive stars may push the gas and dust by their strong wind, they may create bubbles of material surrounding them (van Marle et al. 2015). The presence of bubbles inside clusters may be a product of the strong wind.

Table 1.3 summarizes all the present catalog of obscured clusters candidates. Only one of reported catalogues present a decontaminated photometry method to detect automatically clusters candidates.

Table 1.3: Summary of the principal near-infrared cluster candidate catalogs. The second column is the photometric database used. The third column includes the method used for candidate detection: VI, visual inspection; AOD, Automatic overdensities detection; Mpsr, Massive protostellar regions; SFR, Star Forming Regions; DZ, Decontamination Zones. The last column includes the number of reported candidates.

resizebox15cm!

Authors	Photometry	Method	Number of Candidates
Dutra & Bica (2000)	2MASS	VI	58
Dutra & Bica (2001)	2MASS	VI in SFR	42
Ivanov et al. (2002)	2MASS	AOD + VI	10
Dutra et al. (2003)	2MASS	VI in SFR	179
Bica et al. (2003)	2MASS	RFEs	167
Mercer et al. (2005)	GLIMPSE	AOD + VI	92
Kronberger et al. (2006)	2MASS	RFEs	66
Kumar et al. (2006)	2MASS	AOD in Mpsr	34
Froebrich et al. (2007)	2MASS	AOD	102
Glushkova et al. (2010)	2MASS	AOD	55
Lucas & Samuel (2010)	UKIDSS	AOD + VI	127
Borissova et al. (2011)	VVV	VI	96
Solin et al. (2012)	UKIDSS	AOD	137
Morales et al. (2013)	GLIMPSE	AOD + VI	75
Solin et al. (2014)	VVV	AOD	88
Borissova et al. (2014)	VVV	VI	58
Barbá et al. (2015)	VVV	AOD + VI	493
Ivanov et al. (2017)	VVV	AOD + Dec	9

1.3 Analysis of the light

The light from the stars or from any other body in our Universe is one of the main sources of information since humans rose their heads to look at the sky. In our era, using new telescopes, we are capable of looking deeper and obtaining information of what we are looking at.

Which range of the light is used and how we obtain the properties of the object that we are observing through the telescope (photometry and spectroscopy), are described in this section.

1.3.1 Wavelength range

Historically, massive stars have been studied using the optical range (0.4-0.7μm). In this range we find the Johnson-Morgan photometric system or UBV system (Johnson & Morgan 1953a). But owing to interstellar extinction several other massive stars may remain hidden to optical searches, embedded inside the gas and dust. Near-infrared data allows us to avoid the effects of the extinction in the optical range. Nevertheless atmospheric water molecules may absorbs photons (see Figure 1.2), making the presence of near-infrared windows essential for the observation of stars in the near-infrared.

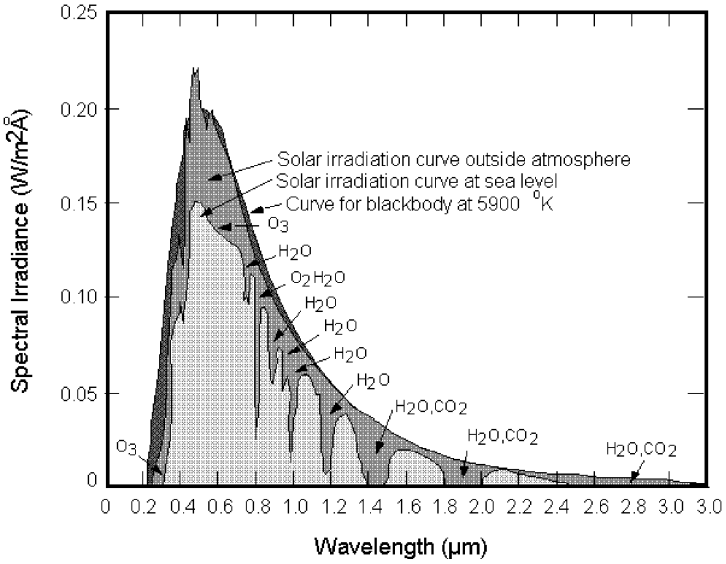

Figure 1.2: Example of the Solar spectrum in space and at sea level. The differences between the observations are mainly due to the water and carbon dioxide absorption (Hoffmann 2007)

Between 1μm and 3 μm the bands are covered mostly by the *J*, *H* and *K* filters (described in Table 1.4). Using these filters it is possible to observe hidden and obscured sources, including

populations of massive stars.

Table 1.4: Near-infrared filters and their wavelength coverage.

Band	Wavelength range [μm]
J	1.1 – 1.4
H	1.5 – 1.8
K	2.0 – 2.4

1.4 Photometry

Photometry is the procedure to measure the quantity of light. The brightness of the object is measured in magnitudes, and it is standardized using the star Vega as an object with magnitude zero. It mean that a star as bright as Vega also has a magnitude equal to 0.

Equation 1.1 represents how we interpret the light of the stars in terms of the apparent magnitude, m_λ; M_λ correspond to the absolute magnitude M (the apparent magnitude an object would have if it were located at 10 pc), d is the distance measured in parsec [pc] and A_λ corresponds to the reddening, that is basically how the light from the stars is absorbed or scattered by the dust and gas between the source and the Earth.

$$m_\lambda = M_\lambda + 5logd - 5 + A_\lambda \tag{1.1}$$

For the V band (optical), we can use the average value for A_V by Rieke & Lebofsky (1985).

$$A_v \approx 1.6\left(\frac{d}{kpc}\right) \tag{1.2}$$

They also provide an estimation of the reddening for different filters based on transformations between 1.2 and the filters. These approximations provide us a very good tool to interpret the position of the stars in the different diagrams.

1.4.1 Extinction and Reddening

The light from stars that reach the Earth suffer first, the extinction, which is the dimming of the light through absorption by the Inter-Stellar Medium (ISM) (see eq. 1.1) and second, the reddening which changes the color of the star (see 1.4.2) . Stars get redder to larger wavelengths due to there is more extinction for blue wavelength than red. One way to estimate the extinction is through the Red Clump (RC) stars. During this phase of evolution intermediate mass stars (the dominant star population) preserve their color and magnitude for a long time, making easy to distinguish in the diagram a specific cluster of stars. In the case of different clusters and extinction, this clump is spread through the color-magnitude diagram (CMD) and color-color diagram (CCD), and using the slope from the CCD it is possible to obtain the average extinction of the field.

Unfortunately RC stars are hard to distinguish from the other stellar populations because they share a position with the rest of the stars in the CMD diagram, and depending on the observed region we can find them spread along the diagram in different ways. On the other hand an, accurate identification of RC stars could be made using spectroscopy, but spectroscopy required a lot of time and resource. An example of the different spread of RC stars is given by Cabrera-Lavers et al. (2007), in two different galactic longitudes highlighted in the Figure 1.3, they can identify the red clump stars and show how they spread along the CMD.

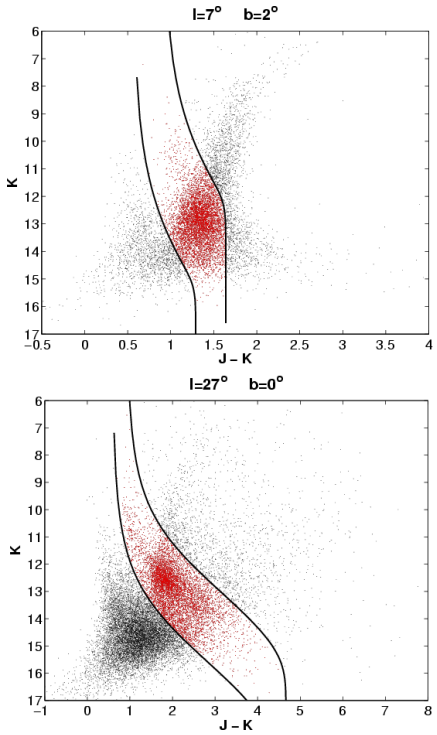

Figure 1.3: Different representations of the extinction for the red-clump, Cabrera-Lavers et al. (2007). Using theoretical traces for the giant population by means of the SKY model (Wainscoat et al. 1992), they isolate stars between both traces, which corresponding to the K0 and the M0 populations (probable red clump stars). Both regions corresponds to an area of 14.2 deg^2 centered; top : $l = 7°, b = 2°$, bottom : $l = 27°, b = 0°$

1.4.2 Color Magnitude diagram and Color Color Diagram

When we observe a star with a telescope, and after reducing the data, we can measure the photometry of the object. By subtracting the magnitudes in two filters, we can estimate the star color and derive from equation 1.1 the following distance independent equation:

$$m_{\lambda 1} - m_{\lambda 2} = M_{\lambda 1} + 5\log(d) + A_{\lambda 1} + 5 - M_{\lambda 2} - 5\log(d) - A_{\lambda 2} + 5 \tag{1.3}$$

$$m_{\lambda 1} - m_{\lambda 2} = M_{\lambda 1} - M_{\lambda 2} + A_{\lambda 1} - A_{\lambda 2} \tag{1.4}$$

This different provide us with specific information for the different type of stars. As an example, consider that O-type stars are located in the top of the Hertzsprung-Russell (HR) diagram by E. Hertzsprung in 1909 and independently by H.N. Russell in 1913. HR diagram, provide both one of the most stringent test of the theory of stellar evolution, and one of the most incisive tools for exploring the history of Galaxy as a whole (Binney J. 1998). It typically plots the absolute visual magnitude versus spectral type and O-type stars are found in the top left of the diagram. The color-magnitude diagram (CMD) is a variant of the HR diagram and provide us with also the type of population of the stars that we are observing. Figure 1.4 (left) shows different evolutionary tracks for four stars on the HR diagram, two of them non massive stars $0.8M_{\odot}$, $5M_{\odot}$, and two massive, $20M_{\odot}$, and $100M_{\odot}$, the figure on the right shows a CMD example from Perryman et al. (1995) for a population of 11,125 stars from the Hipparcos catalogue, the clear sequence for this star population is present when no reddening is present but extinction, and also it is possible to differs stars in different evolutionary states as the red-clump (RC) stars (the bump at $(B-V) \sim 1.2$ and absolute magnitude ~ 0).

Now if we plot two different colors for the same group of stars, is consider as the color-color diagram (CCD). Different spectral types stars also show different positions in the diagram (always depending on the distance and the extinction). In the CCD some combination of filters may provide us with different extra information.

In general, CMD and CCD diagram provides us with information about the star population, as an example, Figure 1.5 shows the position of different type of stars in this two diagram. We took the intrinsic colors from Tokunaga 2007 and calculate the position of the stars using using the average extinction obtained by Rieke (equation 1.2) on the equation 1.1 for two distances , 1[Kpc] and 5[Kpc]. In the Figure is possible to clear distinguish the position of the different type stars and how they are affected by the distance and reddening.

However, the complication occurs when we observe several population of different distances in the same field of view, were is not possible to clear distinguish which type of stars are. As an example, Figure 1.6 shows this two diagram with a real field of stars, using the photometry from 2MASS catalogue for a region of $6^{\circ} \times 6^{\circ}$ degrees^2 centered around galactic latitudes $l = 33^{\circ}, b = 0^{\circ}$ and how the stars positioned through the diagrams.

1.4.3 Reddening-free pseudo-color Q

One last diagram provide us with the possible correction of the colors for a group of stars, and to estimate the spectral type of the stars without using the spectroscopy. This method was developed by Johnson & Morgan (1953b) by defining the reddening-free parameter Q (see equation

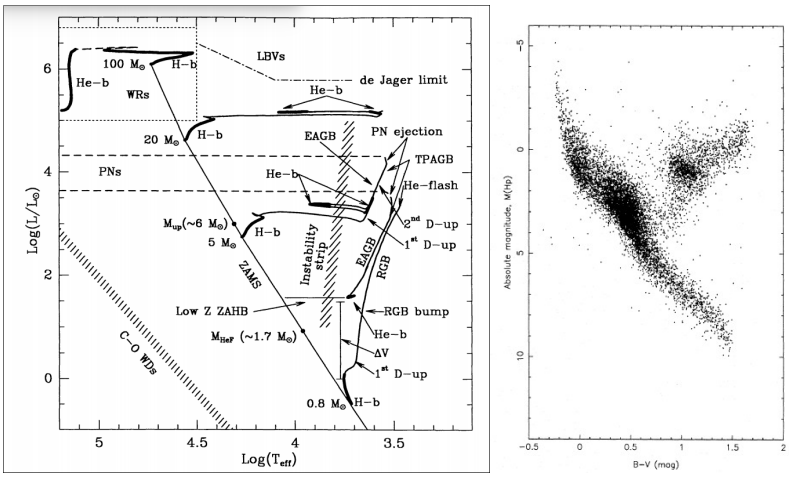

Figure 1.4: Left: HR Diagram taken from Chiosi et al. (1992). Evolutionary tracks of model stars $Y = 0.25$ and $Z = 0.008$ for $0.8M_{\odot}$, $5M_{\odot}$, $20M_{\odot}$, and $100M_{\odot}$. Thick lines represent zones where most of stars should be observed, corresponding to zones where the star is burning H or He in its core. Right: color-magnitude diagram for the Hipparcos catalogue H30 with 11,125 stars (Perryman et al. 1995).

1.5 for the Jonson photometric system), which depends only on the star intrinsic magnitude and is independent of the reddening.

$$Q \equiv (U-B) - \frac{E(U-B)}{E(B-V)}(B-V) \quad \simeq (U-B) - 0.72(B-V) \tag{1.5}$$

In the same way Comerón et al. (2005) uses these principles to define the Q_{IR} parameter for the JHK photometric system, providing a tool to be used in infrared catalogs

$$Q_{IR} \equiv (J-H) - \frac{E_{J-H}}{E_{H-K}}(H-K) \tag{1.6}$$

Figure 1.7 (left) shows the position of the same theoretical stars from Figure 1.5 (magnitudes and colors calculated with the same method than Figure 1.5), now on the Q_{IR} diagram using the extinction of $E(J-J)/E(H-K) = 1.78$ for the same two distances 1[Kpc] and 5[Kpc]. It is clear to notice that no shift is present in Q_{IR}, and only a difference in K magnitude is affected by using two distances (Reddening-free).

More important, in Figure 1.7 right (using a distance $d = 0$, to generalize the diagram), OB-type dwarf stars concentrate around $Q_{IR} = 0$, providing us with a very good tool to select only massive star candidates. The Q_{IR} parameter has been used to distinguish stellar population with great success in the past, (Negueruela & Schurch 2007; Ramírez Alegría et al. 2012, 2014).

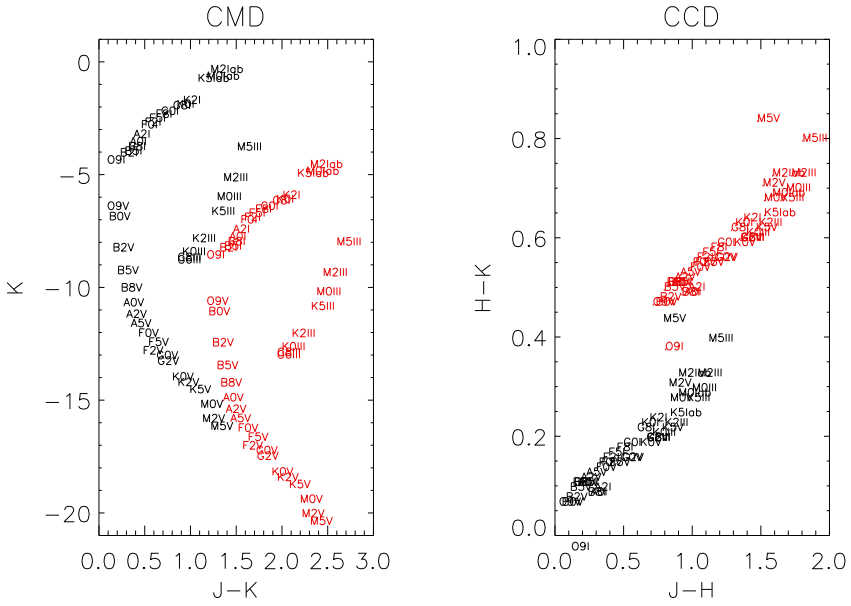

Figure 1.5: Right: Position of different type of stars with the calculated magnitudes (considering distance and reddening using equation 1.2) and equation 1.1) using intrinsic colors from Tokunaga 2007, for two distances, 1[Kpc](black) and 5[Kpc] Right . Left: Similar theorical stars over the CCD, same colors for the respective same distances.

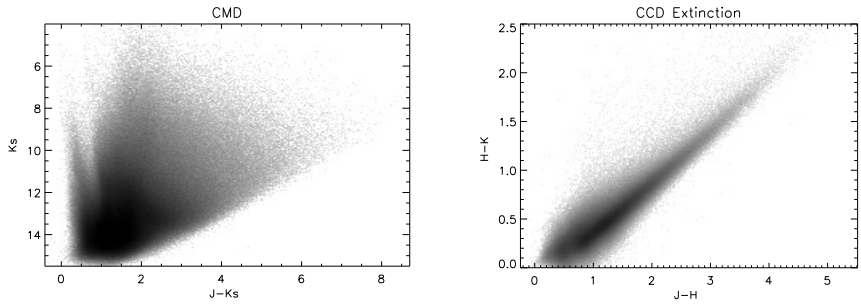

Figure 1.6: CMD (left) and CCD(right) for 2MASS catalogue for a region of $6° \times 6°$ degrees^2 centered around galactic latitudes $l = 33°, b = 0°$, as an example of the dispertion of the population in the diagrams product of the reddening plus the extinction.

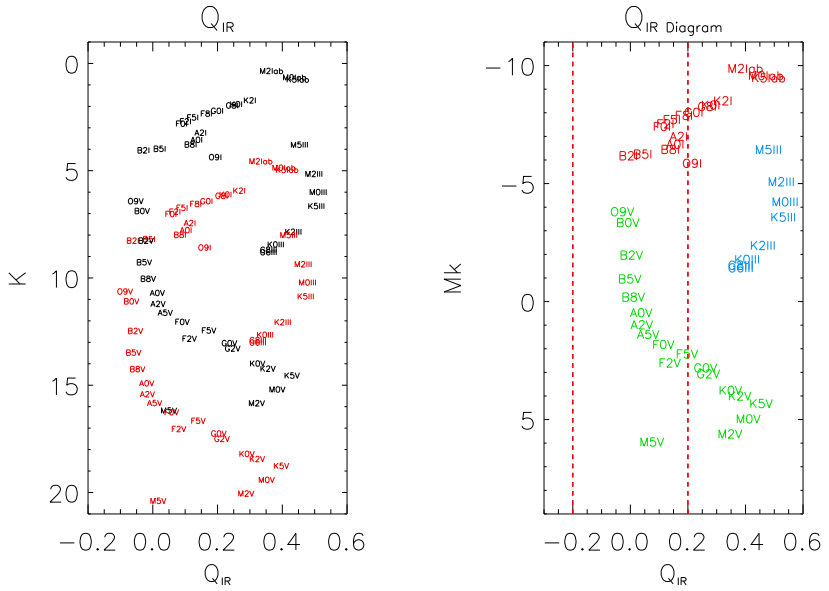

Figure 1.7: Left: Q_{IR} diagram using $E(J-J)/E(H-K)$ of 1.78 for two different distances of 1[Kpc] in black and 5[Kpc] in red, (magnitudes and colors calculated using same procedures than Figure 1.5). Right: Same Q_{IR} diagram, but distance $d = 0$. In green are shown the dwarf stars, blue for the giant and red for the supergiant stars using the intrinsic colors also from Tokunaga, Red dash lines delimited the limits used in Chapter 2.

However, the Q_{IR} parameter only depend on the near-infrared extinction used. During the last twenty years different estimations for the near-infrared extinction have been made, $E(J-H)/E(H-K) = 1.70$ was obtained from the extinction law of Rieke & Lebofsky (1985), $E(J-H)/E(H-K) \sim 1.9$ from Fitzpatrick 1999, $E(J-H)/E(H-K) \sim 1.8$ from Indebetouw et al. 2005. On this thesis we adopted a value of 1.78, similar to the estimation given by Wang & Jiang (2014)

1.5 Spectroscopy

Analysis of stellar spectra is one of the most powerful tools that we have in astrophysics to study the stars. Since we discovered that we could disperse the light and identify different chemical elements, we could *touch* the stars and take a sample of them. With the help of stellar evolutionary models it is possible to study properties such as the stellar age and mass and study the star formation history or the initial mass function of the clusters (Hillenbrand 2009; Bastian et al. 2010).

1.5.1 Stellar model atmospheres

Several codes have being developed to reproduce observed spectra and understand in the laboratory the process from stars. In the past, several codes intend to model hot stars such as PHOENIX (Hauschildt & Baron 1999), WMBASIC (Pauldrach et al. 2001), TLYSTY (Hubeny & Hubeny 1998, plane-parallel, no wind) and PoWR (Gräfener et al. 2002; Hamann & Gräfener 2003). In our time, the two principal codes to generate massive stars synthetic models are:

- CMFGEN: Created specifically for hot stars with atmospheres in expansion. It also solves the line formation calculation with spherical extensions for non-local thermodynamic equilibrium (NLTE, see section 1.5.5, for more detail) and line-blanketing. It is currently used to model O-type main sequence stars and supergiants, B supergiants, Population I & Population II Wolf-Rayet stars, and Luminous Blue Variables (LBVs). This code was designed to use H, He and metal atomic models (Hillier & Miller 1998).
- FASTWIND: A fast and easy-to-use NLTE line-formation code with spherical extension and stellar winds, developed for the spectroscopic analysis of luminous blue stars, covering the spectral range from "A" to "O". It posses a consistent photospheric stratification including continuum radiative acceleration and photospheric extension; "β-velocity law" for the wind; comoving frame or Sobolev plus continuum line transfer; fast solution algorithm for calculating line profiles, allowing for a consistent treatment of incoherent electron scattering (Santolaya-Rey et al. 1997; Puls et al. 2005).

The two codes differ primarily in their approach towards line blanketing, with CMFGEN treating all of the lines in the co-moving frame and FASTWIND taking an approximate approach which speeds up execution times considerably. While CMFGEN automatically computes the entire spectrum, including lines of C, N, and O (Massey et al. 2013), FASTWIND may only compute the H, He atomic models making it faster.

Because we expect to analyze spectra with H and He lines as the main spectral features, we used the lastest version of the FASTWIND code (Puls et al. 2005), for the spectrum characterization of the atmosphere of massive stars in the optical and near-infrared. With this code we are able to generate a full grid of models to test all the parameters for our sample of stars. However treating a huge amount of synthetic spectra and comparing it with the observational one, implies the usage of high efficiency computers programs.

1.5.2 Rotation

Rotation not only counteracts gravity such that stars behave as if they had less mass, it also changes the parameter determination $\log g$. This was studied by Repolust et al. (2005) , who introduced the equation:

$$g_{true} = g + \langle g_{cent} \rangle \tag{1.7}$$

$$\langle g_{cent} \rangle = \frac{(V_{rot} \sin(i))^2}{R_*} \tag{1.8}$$

Here $\langle g_{cent} \rangle$ is the centrifugal compensation of the gravity that must be added to the parameter obtained from the models (see Repolust et al. 2005, for more detail).

Rotational velocity also influences the broadening of the lines in a similar way to macro turbulence (each line must have a convolution with these two parameters to fit the observed model). These parameters, however, are slightly related considering that the latitudinal dependence of the centrifugal force results in the deformation of a star such that the equator feels the centrifugal force strongest, hence is pushed farthest away from the core and is cooler than the poles. Consequently, a thermal imbalance between the poles and the equator arises, thus mixing, that may be consider as macro turbulence, mixes the chemical layers of the star and may influence the massive stellar evolution. As a result, the atmospheric spectral lines are essentially the shape of the Doppler-shift distribution product of the rotation.

1.5.3 Stellar Wind: Mass-loss rate

The stellar winds of hot stars have an important influence on the evolution of massive stars and their environment by reducing the total mass of the star during their lifetimes. Due to this mass-loss the star may change its chemical profile, surface abundances, and luminosity. The environment gains mechanical, radiative momentum, energy and the injection of nuclear processed material.

During the massive star life, they expel mass in the form of stellar winds driven by momentum transfer from the photons to the atoms in the stellar atmosphere. The mass driven for massive stars by the wind varies between 10^{-8} and $10^{-4} M_{\odot}/yr$ (Najarro et al. 2006). The Figure 1.8 from Arnould et al. (1997) illustrates the mass-loss for a $60 M\odot$ star during its life.

The mass-loss is also related to the evolutionary state of the star. In general, stars with low mass-loss rate, correspond to the group close to the zero-age-main-sequence (ZAMS). As the star evolves, the burnt materials –which is responsible for absorbing photons– begins to be more abundant, increasing the wind. Together with the expansion of the star, which provide more area of burnt material and a low gravity on the outer layers plus rotational velocities produces higher winds. Metallicity also plays a roll in the mass-loss (Vink et al. 2001), but it is not considered for this work.

1.5.4 Velocity Field

The gas escapes from the outer layers of the photosphere as stellar wind, accelerated to supersonic velocities at large distances from the star ($120 R_*$ for FASTWIND code calculation). The acceleration of the wind may be approximated by a β law (see equation 1.10, with v_0 the velocity at radius $r_0 \sim R_*$).

$$v(r) = v_\infty \left(1 - \frac{b r_0}{r}\right)^\beta \tag{1.9}$$

$$b = 1 - \left(\frac{v_0}{v_\infty}\right)^\beta \tag{1.10}$$

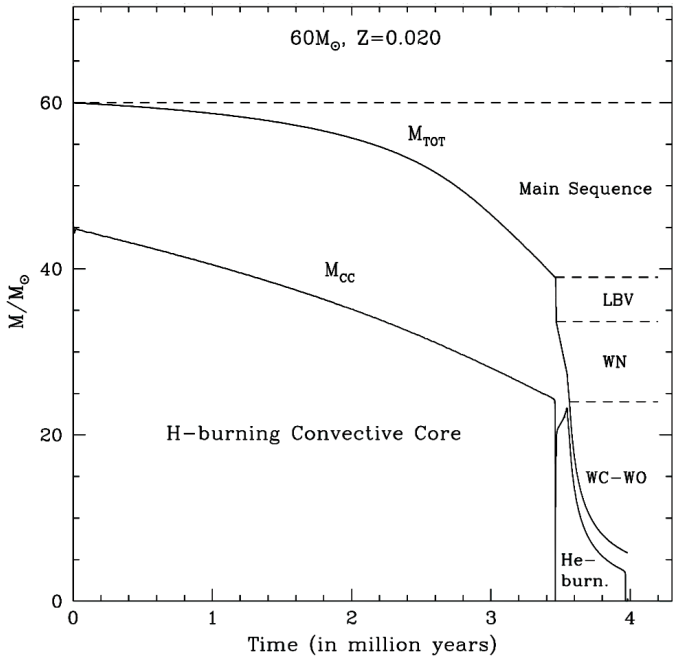

Figure 1.8: Evolution of the total mass M_{TOT} and of the convective core mass M_{cc} as a function of time for a 60 $M_{\odot}$ model with solar ($Z = 0.02$) metallicity (Meynet et al. 1994). The different stages from a O-type main-sequence star to an evolved WR star (Arnould et al. 1997).

Stars at different evolutionary points may present different velocity fields as result of the different mass-loss rate of the stars. The parameter β typically 0.8 for dwarfs may rise to 1.2 or even 1.5 for supergiant stars (Santolaya-Rey et al. 1997). But as we explain in Chapter 4. This parameter is dependent also on the mass-loss rate which in some cases its degenerate.

1.5.5 NLTE and Line Blanketing

Local thermodynamic equilibrium (LTE) corresponds to the assumption of a Maxwell distribution of particle velocities and a Saha-Boltzmann distribution of excitation and ionization states to be valid locally using local values of temperature (T) and electron density (n_e) which is simply represented by the Planck blackbody spectrum. For massive stars, as collisional or radiative excitation dominates over convection (non-LTE) effects dominate. NLTE process includes principally the *line blanketing*, which is the dimming of the blue part of the spectrum caused by the absorption of high UV energy photons by heavy elements (e.g. Fe) to then being emitted at longer wavelength. Figure 1.9 shows an example of the line models created considering line blanketing and not.

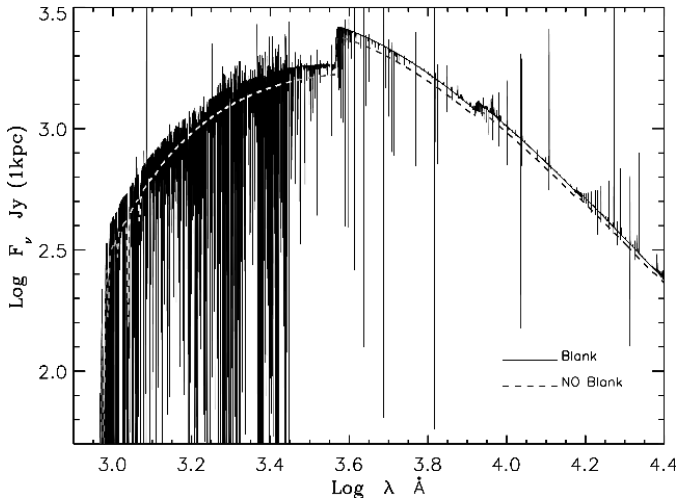

Figure 1.9: Effects of line blanketing (solid) vs. unblanketed model (dashed) on the flux distribution (Puls et al. 2008).

1.5.6 Clumping

Several studies have already indicated the presence of clumped winds on massive stars (Hillier et al. 2003; Bouret et al. 2005; Puls et al. 2008; Najarro et al. 2011; Sundqvist & Owocki 2013). Line-driven instability creates small scale density inhomogeneities distributed across the winds known as clumping. Including clumping leads to a reduction of inferred mass loss rates for O-type stars by about a factor of 23 (Smith 2014).

The presence of clumping in massive stars winds has two main effects: first, for a given atmospheric structure, it changes the shape of the lines affected by winds; second, due to the increased density in clumps, lines recombinations are stronger and thus the ionization structure is modified (Martins 2011).

If clumping is present, mass-loss-rate estimated by models are overestimated and also may change the velocity field of the wind. This change has an important impact on evolutionary models.

Under the known clumping condition, the mass-loss rate driven by the clumping models must be compensated to obtain similar continuum flux for the Hα line. Also the near-infrared and radio regimes may be more affected by clumping, as their lines are formed in regions closest to the wind, where the clumping is created. In Chapter 4, we discuss the clumping condition, the different laws, effects over the lines, and how the fit profile with clumping gets improved over the observed spectrum.

1.6 Aim and Outline of this Thesis

The study and discovery of **obscured massive clusters** has experienced a boost thanks to new instrumentation. Near-infrared instruments are key to observe these objects. Massive stars and clusters have been studied extensively in simulation and observations. But the current census reports a small fraction of the estimated total population of massive stars in our galaxy. The advent of new instruments and telescopes, will allows us to detect the still hidden population.

The studies driven on massive stars reveal several uncertainties on the accuracy of stellar parameters, particularly for stars with signs of clumping. Several studies have considered this effect, but for only one type of star in particular and not as a general sample of stars.

The first part of this thesis comprises a full search in our galaxy for the hidden population of massive stars using the near-infrared 2MASS survey using a novel method. Using properties of the already known massive stars, which are part of a massive cluster, and applying photometric cuts over the 2MASS database, we detect associations of stars via a friends-of-friends algorithm, determining which massive star candidates are probably associated. Most of the reported **masive cluster candidates** are based on overdensity searches or by visual inspection, which put our method into the top and most promising results methods to detect massive clusters candidates.

The second part of this thesis is also focused on the spectral analysis and physical characterization of massive stars. Considering that massive stars are mostly well observed in the optical, we took a sample of stars with high resolution optical spectra ($R \sim 46000$), which are also present in a survey of high resolution in the near-infrared ($R \sim 12000$). We generated a full model grids of models used to derive and compare stellar physical parameters. We aim to test the viability of deriving parameters using only near-infrared, which will give us independence from optical data, ideal for studying the new population to be discover using only near-infrared observations.

The study of massive stars in this second part also leads us to study the not well know clumping condition for our sample. We considered how the clumping condition is created and how this affects the lines in the optical and near-infrared regions. Using theoretical models with the clumping condition allows us to study **clumped stars**. This analysis helps us to understand fundamental processes of the line formation over the wind, where the clumping is created, and how the lines behaves. Our study allows for the first time the automatic determination of stellar parameters using a grid of clumped models.

The principal objectives of the thesis are:

- *To development an automatic algorithm to detect massive clusters candidates*

  To fulfill this we have to study the massive star population over the well know obscured clusters. It is necessary to characterize the photometric cuts to assure the selection of the massive star population. The second part aims to determine which are the best friends-of-friends parameters to assure the association detection. The description and result are included in Chapter 2.

- *To sweep the galactic disk using the automatic searching algorithm and to present a cat-*

alog of massive cluster candidates using 2MASS

To prepare for the arrival of new telescopes we need to have good candidate catalogue to observe. As most of the current candidates are identified using visual inspection and algorithm that identify density peaks, our method will play an important role in the selection of candidates undetected by the other methods, and include them in the future observations. Also presented in Chapter 2

- *To present new massive obscured clusters from the automatic method*

  We observed the candidate associations and cluster candidates from the preliminary automatic algorithm, and verified the method. Observations, analysis and result are included in Chapter 3

- *To study the viability of deriving the stellar parameters based on near-infrared observations*

  We studied the stellar parameters from the optical, and compared them with the ones obtained using only near-infrared data. We present our analysis and results in Chapter 4.

- *To understand the clumping condition and derived parameters for the clumped stars*

  We analysed the clumping condition to ensure the correct determination of stellar parameters of the stars, particularly when the automatic analysis present difficulties to provide good fits to all the spectral lines. The clumping and stellar analysis is included in Chapter 4.

2

MASGOMAS AUTOMATIC SEARCH OF CLUSTERS ALGORITHM *(MASCA)*

Contents

2.1 Initial steps of the MASGOMAS Project

The advent of IR surveys such as 2MASS Skrutskie et al. (2006), GLIMPSE Benjamin et al. (2003), UKIDSS Lawrence et al. (2007), or Vista-VVV Minniti et al. (2010), brought us a new vision of the Milky Way star formation, from a passive and quiet galaxy, with low formation rate to a galaxy full of regions of intense star formation.

During the last decade several new massive stellar clusters were discovered in our galaxy containing remarkable massive stellar populations (such as Quintuplet, Arches and RSGC1 clusters, Figer et al. 1999, 2002, 2006).

New massive stellar populations were also unveiled in already known stellar clusters (e.g. CygOB2, Knödlseder, 2000; Westerlund1, Clark & Negueruela, 2002). According to Hanson & Popescu (2008), hundreds of galactic massive stellar clusters would still remain unknown and hidden. This fact was the motivation to conceive the project **MASGOMAS** (**MA**ssive **S**tars in **G**alactic **O**bscured **MA**ssive cluster**S**), to discover and characterize massive cluster candidates, using near-infrared photometry and mid-resolution spectroscopy.

The initial phase of the project (hereafter phase A) was the study focused on the stellar population of the young known regions/clusters G61.48+0.09 (Marín-Franch et al. 2009), NGC 7538 (Puga et al. 2010) and Sh2-152 (Ramírez Alegría et al. 2011), in an attempt to discover and characterize new massive stellar population hidden within known clusters. The strategy was to spectrophotometrically observe the clusters, using LIRIS[1]. During this phase, we observed new massive stellar populations and their characteristic near-infrared photometric colors (all OB populations were found at magnitudes brighter than 12.5 Ks, which also correspond to the limit magnitude of LIRIS for spectroscopy, and with colors $(J - Ks)$ between 0.5 and 4). Table 2.1 summarizes the findings during this phase. Phase A gave us way to develop a preliminary tool to select OB-type star candidates stars from a catalogue for spectroscopic follow up. At this point of the project, we moved from characterizing known clusters to attempt to discover new obscured clusters based in the characteristic colors and magnitudes of a possible obscured population . This phase is called phase B.

In phase B, clusters candidates were initially selected using photometric cuts in color $J - Ks > 0.5$ mag and magnitude Ks < 12.5. To separate the massive young reddened stars from intrinsic red stars, for example giants stars, we use the Q_{IR} reddening-free parameter (Comerón et al. 2005 ;see Section 1.4.3). The 2MASS photometry (Skrutskie et al. 2006) was used as the input catalog for the search regions. We then inspected by eye the spatial distribution of the object fulfilling the photometric criteria. Apparent concentrations of stellar candidates were considered as a cluster candidate. Two new cluster candidates, MASGOMAS-1 and MASGOMAS-4, were found in this phase of the project. The spectroscopic follow-up was required to confirm the cluster candidates as real clusters and not asterisms

The analysis of these two new candidates supported the use of the Q_{IR} criterion as an ex-

[1] A near-infrared spectrograph (Manchado et al. 2004; Acosta Pulido et al. 2003), mounted at the Cassegrain focus of the 4.2 m William Herschel Telescope (WHT; Roque de Los Muchachos Observatory, La Palma)

cellent filter for selecting OB-type star candidates combined with the color and magnitude cuts and the spatial concentration. In MASGOMAS-1, from a total of 23 candidates, 16 were classified as late OB dwarfs. Three red supergiant stars and one A-supergiant also were discovered, which indicates a total cluster mass of $1.94 \times 10^4 M_\odot$ Ramírez Alegría et al. (2012). This include MASGOMAS-1 in the list of known Galactic clusters with more than $10^4 M_\odot$. Less than 20 clusters with more than $10^4 M_\odot$ are known.

The second cluster, MASGOMAS-4, has two cores. The analysis of their H+K spectra indicated a unique distance for both cores. The presence of Herbig Ae/Be star candidate in one of the cores of MASGOMAS-4 and surrounding HII region, implies that MASGOMAS-4 is a very young cluster with still on-going massive star formation. The cluster mass was estimated to be $2.19 \times 10^3 M_\odot$ (Ramírez Alegría et al. 2014). Table 2.1 summarize the result of phase B.

In phase C target concentration is automatically identified using a Friends-of-Frieds algorithm (FOFA) which is based on the code developed by Garcia et al. (2009). The code requires two parameters: The searching distance, D_S (the maximum distance to the nearest neighbor) and the minimum number of targets Nmin. Section 2.2.2 includes a detailed description of both parameters.

Base on the photometric cuts and magnitudes from MASGOMAS-1 (i.e. 1.5< (J-Ks)< 3.5, 6.5 < Ks< 11.5) and with Q_{IR} between −0.3 and 0.3, we started a search in an area of 3×3 degrees^2, centered around $l = 38°$ & $b = 0°$. The photometric survey used was also 2MASS. The search concluded with nine candidates. From them, we selected the most promising one by inspecting color-magnitude, color-color and spatial distribution diagrams of the candidates for follow-up spectroscopy. The cluster MASGOMAS-6 is described in the next Chapter as part of this thesis.

The aim of phase C is to develop an automatic search algorithm to construct a catalogue of automatically selected massive cluster candidates.The characteristic of the program is that it concentrates, through photometric cuts, in OB-type stars, instead of taking into account all the stars in the field of view as other programs use to do. The program structure is described in Section 2.2. Preliminary results are presented in Section 2.3. These result were used to improve the algorithm and this process is presented in Section 2.4. Finally Section 2.5 presents the catalogue of cluster candidates covering from $l = 12.5°$ to $l = 93.5°$ and with $-3° < b < 3°$

2.2 MASCA

MASCA (**M**SGOMAS **A**UTOMATIC **S**EARCH OF **C**LUSTERS **A**LGORITHM) is an IDL multi-routine with three main modules: a first module defining the set of photometric cuts for candidate selection (Section 2.1), a friends-of-friends algorithm (FOFA) module, and a third module to produce the graphics for inspection. In this Section, we describe the core of the tool and how the first version works.

Table 2.1: Summary of the results from the MASGOMAS phases.

resizebox15cm!

Cluster	Phase	Characteristic	Reference	Notes
G61.48+0.09	A	Known clusters	Marín-Franch et al. 2009	
NGC 7538			Puga et al. 2010	
Sh2-152			Ramírez Alegría et al. 2011	
MASGOMAS–1	B	Automatic star selection;	Ramírez Alegría et al. 2012	$1.94 \times 10^4 M_\odot$
MASGOMAS–4		cluster candidates selected by eye.	Ramírez Alegría et al. 2014	$2.19 \times 10^3 M_\odot$
MASGOMAS–6	C	Automatic cluster	Ramírez Alegría et al., in prep.	Presented in this work.
MASGOMAS–10		candidate selection.	Rübke et al., in prep.	Presented in this work.

2.2.1 Photometric cuts module

Using the photometric cuts from phase B of the MASGOMAS project, based on the properties of the OB-type stars in MASGOMAS-1, we build the photometric cuts module of MASCA for search the young reddening star population. Most of the stars with moderate colors (i.e., $1.5 < (J - K)$) are closest red intrinsic stars without larger extinction. The module uses the 2MASS catalog input for MASCA. Table 2.2 shows an example of the input format.

To set the limits in K we assumed a similar extinction than the one observed in MASGOMAS-1 i.e. 0.55 mag/kpc in the Ks-band. The magnitude of an O9V star at 1.5 kpc, the closest distance at which we expect to find an obscured cluster candidate, is Ks=8. To allow for a slightly lower extinction in clusters closer than MASGOMAS-1 and for the presence of RSGs, we adopt Ks=6.5 as our upper Ks magnitude limit.

The faint limit is set by LIRIS capabilities. Using LIRIS, we can obtain spectra with the signal-to-noise (S/N) needed for spectral classification in about 10 minutes for K=12 at 1.2 airmass and 1" seeing. An O9V star at 6.0 kpc will have K=11.5 (adopting again MASGOMAS-1 extinction per kpc). The final magnitude limits of K between 6.5 and 11.5 mag correspond to distances between 1.5 and 6 kpc for an O9V star.

For the cuts in $(J - K)$ color, we assumed an extinction given by Indebetouw et al. (2005) law. This can be expressed in the form :

$$(J - K) = (J - K)_0 + (A_J - A_K) = (J - K)_0 + A_K(\frac{A_J}{A_K} - 1) \tag{2.1}$$

With $(A_J/A_K) = 2.5$, we obtain $(J - K)$ colors between 1.2 and 4.8 for O-type stars. To exclude late-type stars we restricted the $(J - K)$ color between 1.5 and 3.5, which corresponds to distances between 2.0 and 4.5 kpc for an O9V star. Figure 2.1 illustrates the magnitude and color cuts. Table 2.3 presents the photometry obtained for the cluster MASGOMAS-1 as reference for the colors, magnitudes, extinction and distances.

With the previously mentioned magnitude and color cuts we can clean our sample of contaminats from the old Galactic population and concentrate on the massive and young stars when combining them with the pseudo-color Q_{IR} criteria, which depends on the ratio of colors excesses, E(J-H)/E(H-K). For this ratio we adopt the value of $E(J - H)/E(H - K) = 1.78$ deter-

Table 2.2: Sample of columns for a typical MASCA input file. The columns specify the object identification, right ascension and declination (in degrees), Galactic longitude and latitude (in degrees), 2MASS near-infrared magnitudes and their errors, and the 2MASS quality flag (stars which no posses errors in J, H or K). To assure a good quality photometric, we exclude from our file U-flagged stars. Moreover all the stars considered for the study must have errors below 0.1 mag in all three bands.

ID	RA	DEC	l	b	J	ϵ J	H	ϵ H	Ks	ϵ Ks	2MASS Flag
18395641-0142497	279.985063	-1.713826	30.0948	1.7749	9.947	0.022	8.419	0.047	7.697	0.015	AAA

mined by Indebetouw et al. (2005) for l=42°, close to our search direction.

Considering the possible distribution of stars in the Q_{IR} versus spectral type diagram we selected stars between $-0.15 < Q_{IR} < 0.2$(see Figure 2.2). These cuts may include A and F type stars (see Figure1.7 right), however with the combination of the color-magnitude and color-color criteria, the number of contaminants is considerably reduced.

As an example of this, from a total of 2,083,386 stars in Figure 2.1, 444,211 stars pass the

Table 2.3: MASGOMAS-1 Table from Ramírez Alegría et al. (2011), now including the color ($J - Ks$) and free-reddering parameter Q_{IR} . The stars observed in long-slit have an ID begginig with "s".

resizebox15cm!

ID	RA (J2000) [h m s]	Dec (J2000) [° ′ ″]	J [mag]	H [mag]	K_S [mag]	$J - K_S$	Q_{IR}	Spectral type	A_{K_S} [mag]	Distance [kpc]
1	18 50 17.398	+00 22 04.77	14.843	13.101	12.204	2.639	0.1274	B1 V	1.83	3.69
2	18 50 21.558	+00 21 52.31	13.720	12.636	12.054	1.666	0.0364	A0 V	1.10	1.18
3	18 50 19.420	+00 21 41.15	14.089	12.051	10.925	3.164	0.0112	O9 V	2.21	2.95
4	18 50 16.434	+00 21 31.44	14.316	12.671	11.839	2.477	0.1474	B0 V	1.74	4.50
5	18 50 16.555	+00 21 21.69	13.817	11.987	11.033	2.784	0.1128	O9.5 V	1.95	3.14
6	18 50 16.500	+00 21 07.01	13.976	12.193	11.162	2.814	-0.0728	B0 V	1.96	2.98
7	18 50 17.481	+00 20 56.26	14.307	12.438	11.400	2.907	0.0006	O9.5 V	2.03	3.58
8	18 50 15.903	+00 20 44.31	13.821	12.261	11.317	2.504	-0.1392	O9.5–B0 V	1.76	3.70
9	18 50 16.185	+00 20 35.77	14.958	12.848	11.603	3.355	-0.131	O9.5 V	2.33	3.43
10	18 50 14.005	+00 20 06.94	11.281	9.966	9.472	1.809	0.4258	G9–K2 III	0.74	1.32
11	18 50 23.945	+00 19 57.52	13.762	11.880	10.995	2.767	0.289	O9 V	1.95	3.43
12	18 50 23.091	+00 19 47.64	13.535	11.496	10.456	3.079	0.167	O9 V	2.15	2.44
13	18 50 22.271	+00 19 28.18	15.505	13.357	12.205	3.3	0.0744	B0 V	2.29	4.15
14	18 50 20.709	+00 21 17.59	10.994	9.638	9.065	1.929	0.3246	G9–K2 III	0.82	1.06
15	18 50 17.303	+00 20 59.01	14.889	12.955	11.843	3.046	-0.0676	B0 V	2.12	3.73
16	18 50 16.837	+00 21 12.35	13.697	12.006	10.999	2.698	-0.1216	B0 V	1.90	3.51
17	18 50 17.036	+00 21 46.69	14.326	12.485	11.471	2.855	0.0158	B0 V	2.00	3.75
18	18 50 15.788	+00 21 46.06	10.752	9.775	9.391	1.361	0.2858	G0 III	0.61	0.81
19	18 50 13.554	+00 21 03.23	13.606	11.895	10.981	2.625	0.0658	B0 V	1.84	2.90
20	18 50 12.329	+00 20 46.68	14.829	13.25	11.96	2.869	-0.743	B0 V	2.00	4.23
21	18 50 10.573	+00 20 25.64	13.727	12.028	11.071	2.656	-0.0236	O9.5 V	1.87	3.32
22	18 50 10.445	+00 21 03.80	10.499	9.603	9.143	1.356	0.068	G9–K2 III	0.44	1.31
23	18 50 08.513	+00 22 11.42	11.019	9.528	8.94	2.079	0.4326	G6 III	1.01	0.67
s01	18 50 17.623	+00 21 30.65	10.812	8.893	7.839	2.973	0.0218	A2 I	1.89	3.58
s02	18 50 28.354	+00 21 13.44	11.348	8.935	7.818	3.53	0.4024	K–M I–III	···	···
s03	18 50 15.408	+00 20 58.07	9.232	6.625	5.299	3.933	0.2202	M2 I	1.99	4.00
s04	18 50 15.620	+00 21 07.46	9.563	6.984	5.649	3.914	0.176	M2 I	1.98	4.72

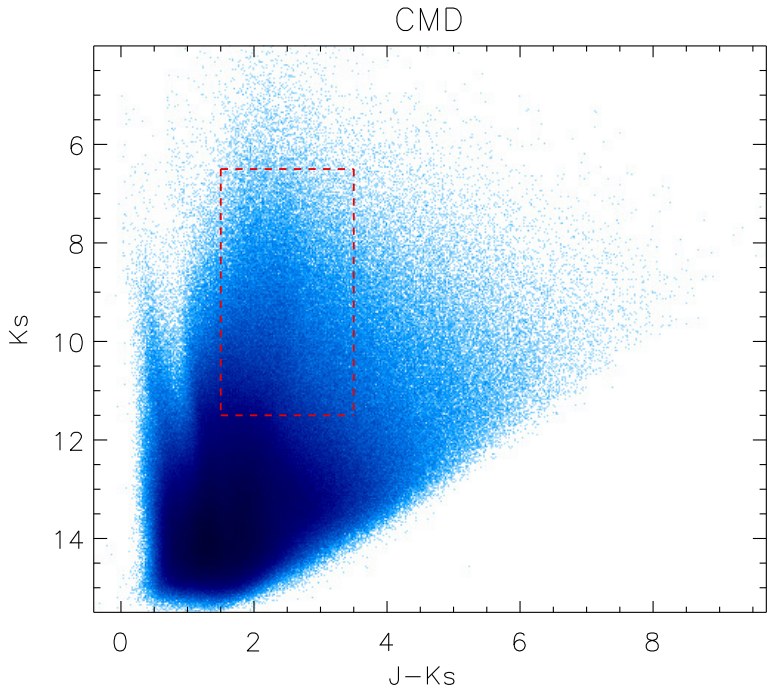

Figure 2.1: Photometric cuts used in the first automatic search over 2MASS photometry in the field of $6° \times 6°$ around Galactic latitudes $l = 33°$ and $b = 0°$. The red dashed lines limit the zone where the stars are selected.

error criteria (i.e < 0.1 mag in all bands), 180,122 stars also fulfill the photometric cuts in $(J-K)$ and K and finally only 7,683 also fulfill the criterium in Q_{IR}.

As result of the combination of all the photometrics cuts the program gave us a total of 7,683 stars in the $6° \times 6°$ region ($l = 33° \pm 3°, b = 0° \pm 3°$) . These candidate stars are used in the next step, where we look for the probable stellar associations.

2.2.2 Friends of Friends module (FOFA)

The automatic overdensity detection is based on an adapted version of the friends-of-friends algorithm AUTOPOP (Garcia et al. 2011), simply called **FOFA** (**F**riends **O**f **F**riends **A**lgorithm). The program is based on the friends-of-friends method, which declares that two stars may be part of an association if the distance between them is less than the searching distance D_S. The program searches the neighborhood around these stars, looking for a new companion within the

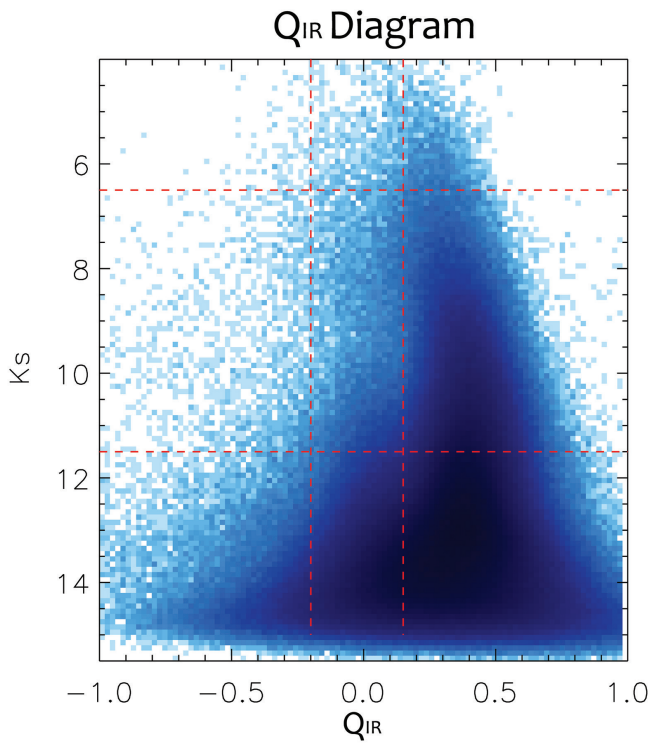

Figure 2.2: Illustration of the Q_{IR} cuts, the dashed red lines limit the search zone.

search distance. This process is repeated for all the targets until no star is found within D_S from any of the peripheral stars. A second input parameter is the minimum number of members (*Nmin*) for a group to be considered an association candidate. The groups found with equal or fewer members than *Nmin* are not considered cluster candidates. Figure 2.3 illustrates an example of a detection using this algorithm.

The program provides two main outputs. The first one is a distribution of the number of associations as function of *Ds* and *Nmin*. Figure 2.4 illustrates an example for 2500 stars in a field of $6^\circ \times 6^\circ$ centered around $l = 33^\circ$ and $b = 0^\circ$. The different symbols represent the different *Nmin*. In the Figure we can see that the number of association candidates depends on the D_S distance. The challenge is to find the optimal combination of *Nmin* and D_S (see Section 2.3.1 for detail).

The second output is a catalog of associated candidates, each of which is given in a separate archive with the candidate star data. A main file with all the associations and the number of

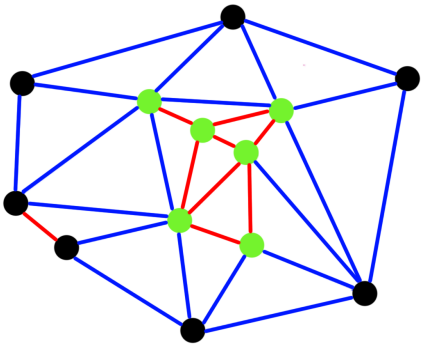

Figure 2.3: FOFA graphic illustration for a given D_S. All green points have at least one companion at a distance shorter than the adopted search distance, D_S. The segments shorter than the distance D_S have been drawn in red color. All black points except two are at a distance larger than Ds from its nearest companion. If the minimum number of individuals required (*Nmin*) is 6 or less, the green points will be selected as a cluster candidate.

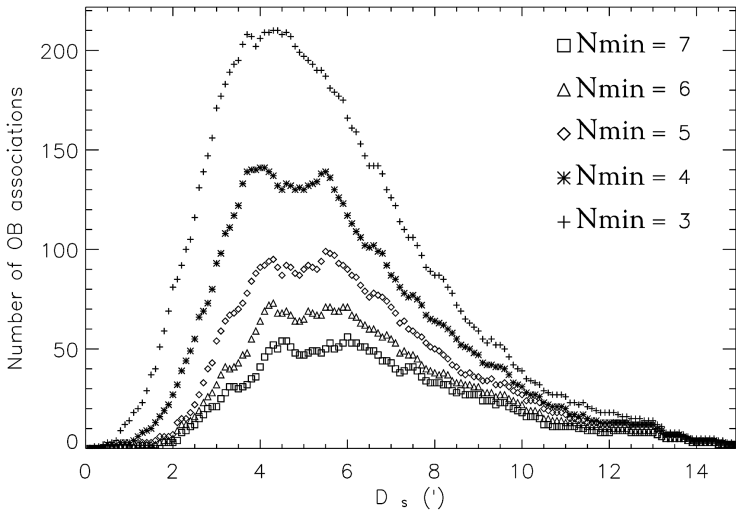

Figure 2.4: Example of the FOFA output, the abscissa axe corresponds to the distance D_S (in arcmin) and the ordinate the number of associations detected for different *Nmin* in the same field of $6^\circ \times 6^\circ$ field centered around $l = 33^\circ$ and $b = 0^\circ$, in this example 2500 stars fulfill the photometrics cuts.

components, radius, stellar density and eccentricity is also provided.

2.2.3 Output module and visual inspection

A new MASCA module was added to ease the interpretation of the results given by FOFA, (see Figure2.5). It consists in a series of graphs with the spatial distribution of the stars, parameters of the association (an identification number, radius & number of components) and the data of a control region around the association candidate. This control region has the same area than the association. Both areas are shown in the top-left panel, in Figure 2.5. Stars fulfilling the photometric cuts (OB candidates) are marked with blue and red symbols in the association candidate and control areas, respectively. This way we can realize whether we are in the presence of an interesting candidate. The diameter of the association candidate is arbitrarily defined as 1.05 times the maximum separation between OB-type star candidates.

The six sub-plots represent the photometric properties for the association candidate (central row) and the control zone (bottom row). We present the Ks v/s Q_{IR} diagram, the color-magnitude diagram (CMD) and color-color diagram (CCD).

Finally, after the automatic cluster candidate selection, it is useful to inspect the association candidates using ALADIN[2] or other equivalent tool. This allows one to display the gas and hot dust using the SPITZER false color images. In Figure 2.6 we show a 2MASS false color (red=K, green=H and blue=J filters), the two cluster used as reference in the program: MASGOMAS-1 and MERCER 30.

[2]http://aladin.u-strasbg.fr

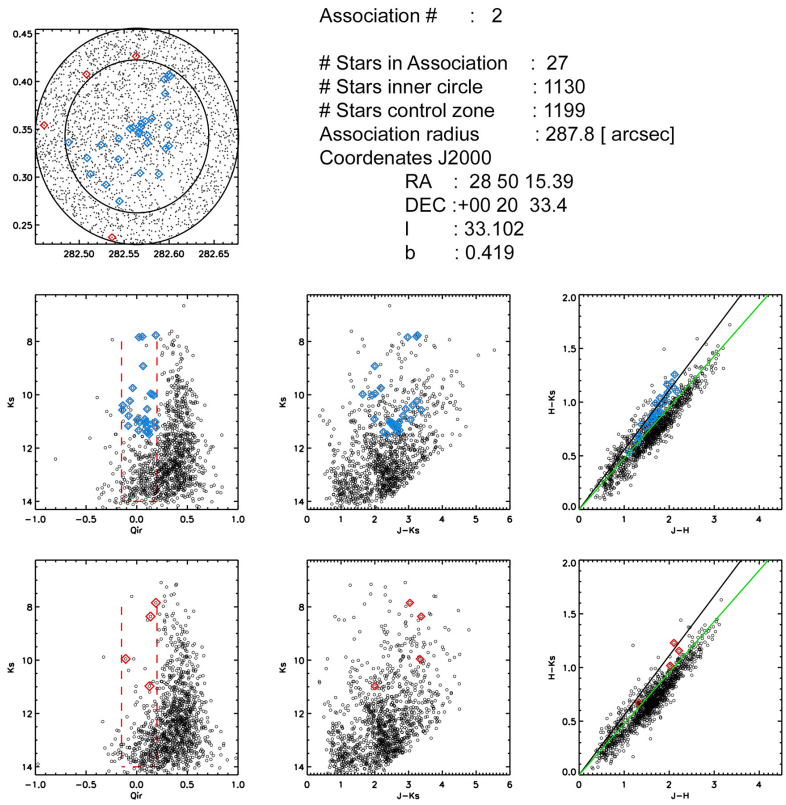

Figure 2.5: Diagnostic diagram produced by MASCA for MASGOMAS-1. At top left, we find the spatial distribution of stars selected as an association candidate. The blue symbols shows the position of the OB-type star candidates, probable members of the association, and the black dots show all the stars in the area from the 2MASS catalogue. The region between both circles is the control area and has the same area as the association candidate region. Stars fulfilling the photometric cuts in the control area are presented using red symbols. At top right we find the identification and some parameters of the association candidate. In the central and bottom rows, we include the photometric diagrams (K vs Q_{IR}, CMD and CCD) for the association candidate and the control zone.

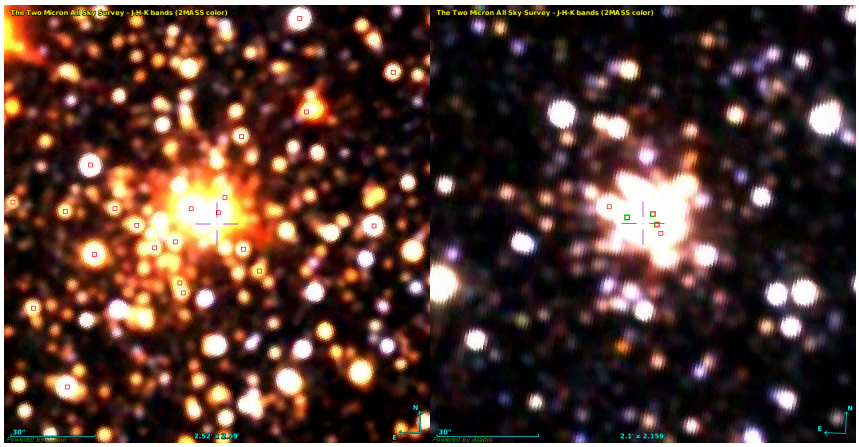

Figure 2.6: The Figure shows 2MASS false color images (red=K, green=H and blue=J filters) for MASGOMAS-1 left and MERCER30 right. The known massive stars in these cluster are highlighted.

2.3 First Results of MASCA

2.3.1 Search Area

W43 ($l = 29^\circ \pm 1.5^\circ$ and $b = 0^\circ \pm 1.5^\circ$) is a very active star-forming complex region at the closest intersection of Galactic bulge with the Scutum-Centaurus arm. It show an intense HII emission due to the gas crash of the solid rotating bulge (Izumiura et al. 1995) with the differentially rotating arm (Dame et al. 2001). MASGOMAS-1 is at the edge of this region, towards larger Galactic longitudes. We decided to focus our first search in this area, specifically in ($l = 33^\circ \pm 3^\circ$ and $b = 3^\circ \pm 3^\circ$). Figure 2.7 illustrates our search area, together with the nearby regions where MASGOMAS-4 and MASGOMAS-6 were found.

Our zone of interest is centered around W43 and the left part of the Sagittarius arm. In Figure 2.8 we show a 2MASS false color image of this region, with W43 to the right and arm to the left.

We have seen in section 2.2.2 that the number of association candidates depends on the combination of D_S and $Nmin$. To decide among the different combinations, we used MASGOMAS-1 as a reference. Because this is a massive and compact association, we decided to use $Nmin = 6$ (which corresponds to 1/3 of the MASGOMAS-1 confirmed OB stars) and D_S=95 arcsec (twice the average distance between OB stars in MASGOMAS-1). Figure 2.9 illustrates the distribution of the number of candidate associations as function of the distance D_S. This resulted in a total of 125 association candidates distributed across the $6^\circ \times 6^\circ$ searching area, as shown in Figure 2.10.

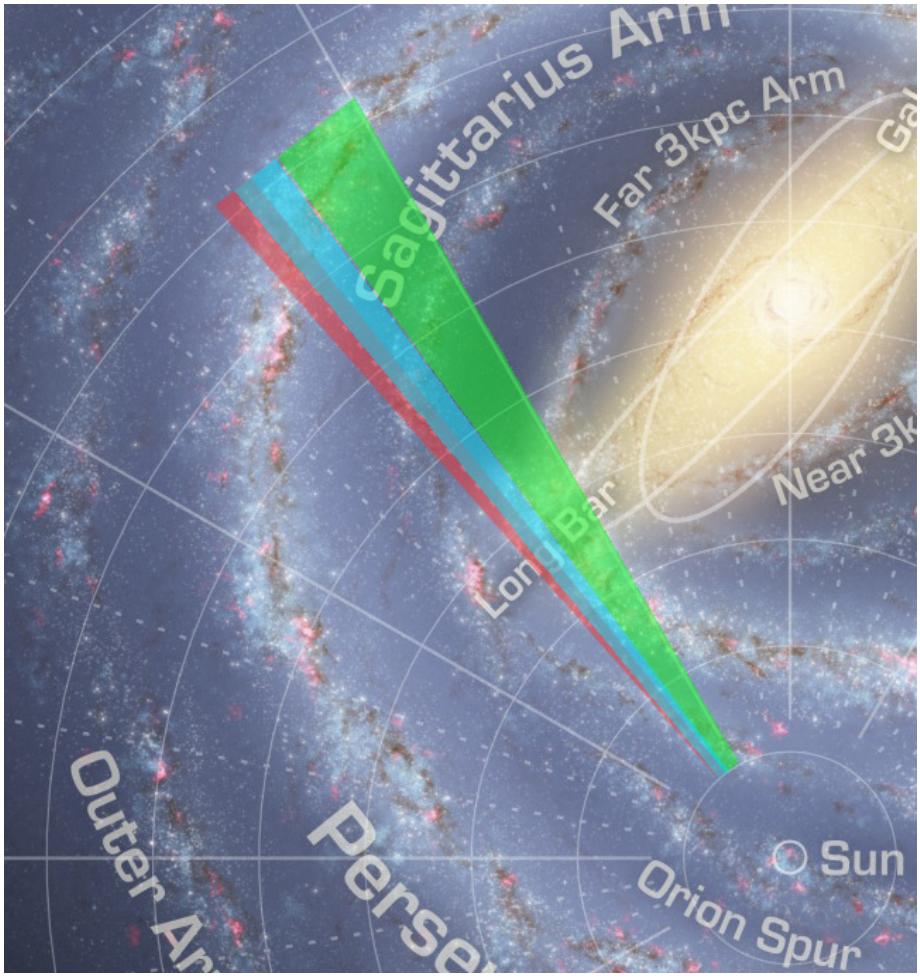

Figure 2.7: Searching areas of the MASGOMAS Project. Red correspond to the search on $l = 40°$ were MASGOMAS-4 was found. Cyan marks the search over $l = 38°$, resulting in the discovery of MASGOMAS-6. In green the new search over $l = 33°$, where MASGOMAS-1 was previusly found.

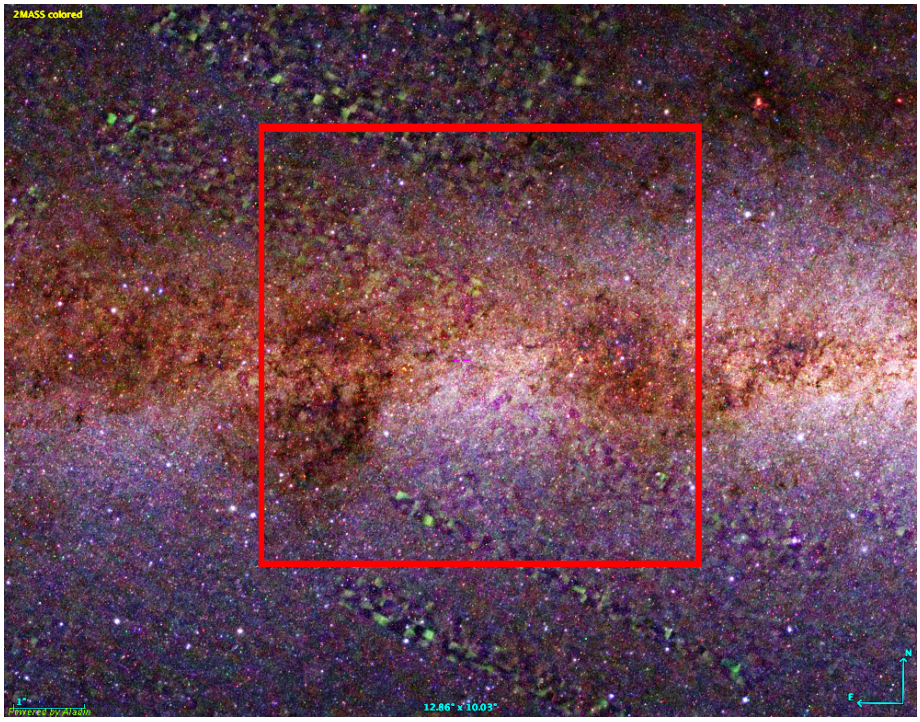

Figure 2.8: 2MASS false color centered around l=33°, highlighted the search area of $6° \times 6°$. (red=Ks, green=H and blue=J filters.)

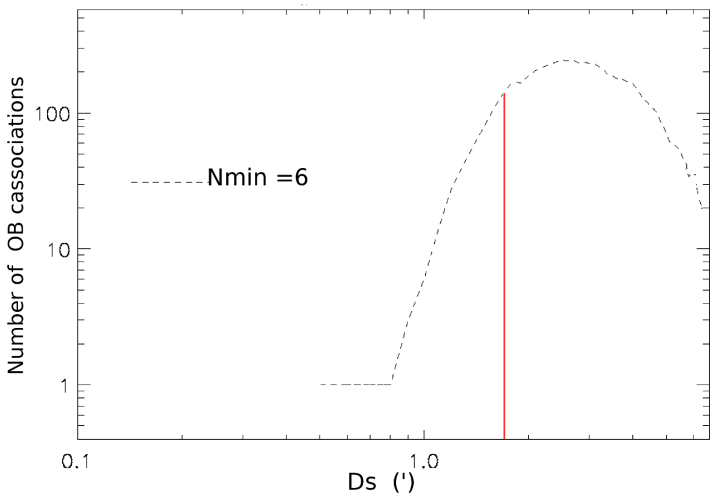

Figure 2.9: Distribution of the number of clusters as a function of the maximum distance for *Nmin* = 6. In red the position related to a distance of 95 [arcsec].

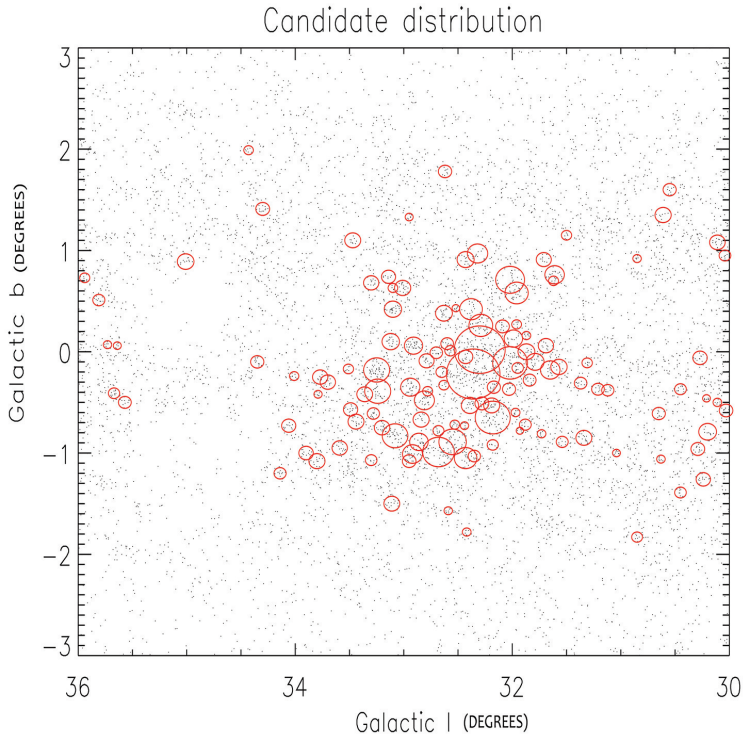

Figure 2.10: MASCA1.0 association candidates. The red circles are centered on each association and the radius of each circle represents the diameter of the association. Black dots represent the massive star candidates fulfilling the photometric cuts in the field $l = 33° \pm 3°, b = 3° \pm 3°$.

2.3.2 Candidate selection

We follow a series of selection criteria to pinpoint the most promising association candidate from the output of the first version of MASCA (MASCA1.0).

- **Spatial distribution of candidate stars**: We expect massive young stars to present a clear ovendensity as compared to their surroundings
- **Color-Magnitude Diagram**: Candidate stars should present similar colors and consistent magnitudes. Differential extinction will introduce some scatter.
- **Inspection of the Q_{IR} diagram**: We expect vertical sequence of candidates centered in $Q_{IR} = 0$.
- **Visual inspection**: We inspect the area surrounding the cluster candidate, looking for bubbles or ionized gas.

From the total 125 association candidates inspected following the previous criteria, only six of them were selected for spectroscopic follow up (associations candidates 61, 76, 88, 110, 113 and 118; see Figure2.11).

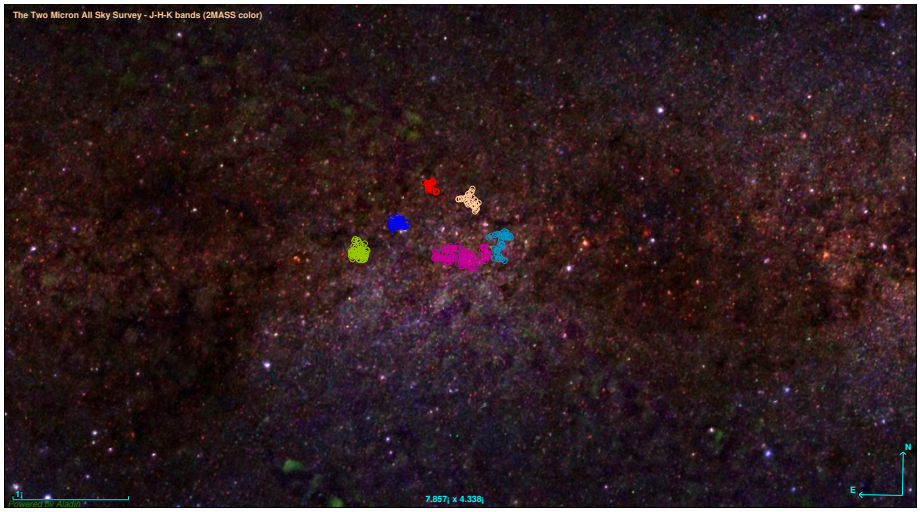

Figure 2.11: The six association candidates selected for the spectroscopic follow-up. All of them were detected close to the center of the searching region. Association candidates are presented in different colors: green (association 118), purple (association 113), red (association 110), pink (association 76) orange (association 88) and blue (association 61)

Table 2.4 list association candidates and their general properties, including the number of OB-type star candidates within the respective radius. Each association presents different characteristic. Association 61 shows a string form in right ascension. Association 76 also presents

Table 2.4: Selected associations for spectroscopic follow-up. In the table we included the candidate name, its central Galactic coordinates (in degrees), the number of OB star candidates (Ncand) and the association radius (in arcseconds)

Association candidate	l	b	Ncand	Radius["]
Assoc61	32.022	-0.111	35	577
Assoc76	32.361	-0.222	90	889
Assoc88	32.287	0.256	29	395
Assoc110	32.627	0.381	17	279
Assoc113	32.909	0.063	31	304
Assoc118	33.247	-0.183	44	430

a massive detection in a large region. It was selected considering the possible merger of three smaller association. Associations 88 and 110 present a smaller radius than the previews two with a more sparsely populated area. Association 113 presents a clear sequence in Q_{IR}. And finally association 118 shows a large population of candidate surrounding the central association. This last association candidate presents also an interesting near-infrared ionized region surrounding the central region plus a bubble of hot dust on the edges.

2.3.3 Observations and data reduction

We carried out an exploratory near-infrared, long-slit spectroscopic program in the K band. All data were acquired with LIRIS. The candidates were observed during four half-nights on 2015 from 28 to 31 of July.

To obtain spectra with similar S/N, we paired stars with similar magnitudes using a 1.0 arcsec slit. For the observations, we used the K pseudogrism (Fragoso-López et al. 2008), which resulted in $\lambda/\delta\lambda \sim 1700$.

The star HIP90967 (spectral type A0 Vn) was used as telluric standard. Observations were performed following an ABBA telescope nodding pattern. This procedure places the star in two positions along the slit, separated by 12". With this mode, is possible to remove the sky contribution from the spectra.

We reduced the data using LIRISDR[3] under the IRAF[4] environment. The package the procedures of include first a correction of the pixel-mapping anomaly of LIRIS using *lcpixmap* and $fixpix$ tasks include in the package, then using the $LSPECT$ package inside LIRISDR, which include the subrutines *lspskynod* and *lspnodcomb*, we automatically subtract sky spectrum from ABBA sequences, combine sky subtracted frames, compute offsets, correct flat-field and compute wavelength calibration (with the required files using $identify$ $reidentify$ and $fitcoords$ task from ONEDSPECT package, using Argon and Xenon lamps, both lamps (continuum-subtracted)

[3]Package written by Jose Acosta Pulido to manage data produced by the LIRIS instrument www.iac.es/galeria/jap/lirisdr/LIRIS_DATA_REDUCTION.html

[4]Distributed by the National Optical Astronomy Observatory, which is operated by the Association of Universities for the Research in Astronomy, Inc., under cooperative agreement with the National Science Foundation (http://iraf.noao.edu/)

being used to calibrate the K-band spectra to finally extract the combined individual spectra for each star). We divided the extracted spectra by the corresponding A0Vn-type star to remove the telluric contamination using the XTELLCOR package Vacca et al. (2003).

2.3.4 Spectral classification

The classification for early type stars was based on the atlases by Hanson et al. (1996, 2005) and Ivanov (2004). Late type stars, the ones presenting the characteristic CO bands, were classificated using the Meyer et al. (1998) atlas. We assumed an error of ±3 subtypes and ±1 luminosity class for the assigned spectral type, considering that only the K- band was available.

Figure 2.12 shows the reduced spectra. We can easily separate them in three groups. The first one is formed by stars with HeI 2.11 μm, HeII 2.18 μm, and Brγ, the characteristic OB-type stars lines (or their equivalents in emission for WRs). The second group is formed by stars where the main feature present in the spectra are the CO-bands (^{12}CO (2, 0) at 2.29μm, ^{12}CO (3, 1) at 2.32 μm and ^{12}CO (4, 2) at 2.35 μm), characteristic of late type stars. And the third the spectrum with broad emission lines which correspond to a WR star. Table 2.5 summarizes the observed stars and their classification.

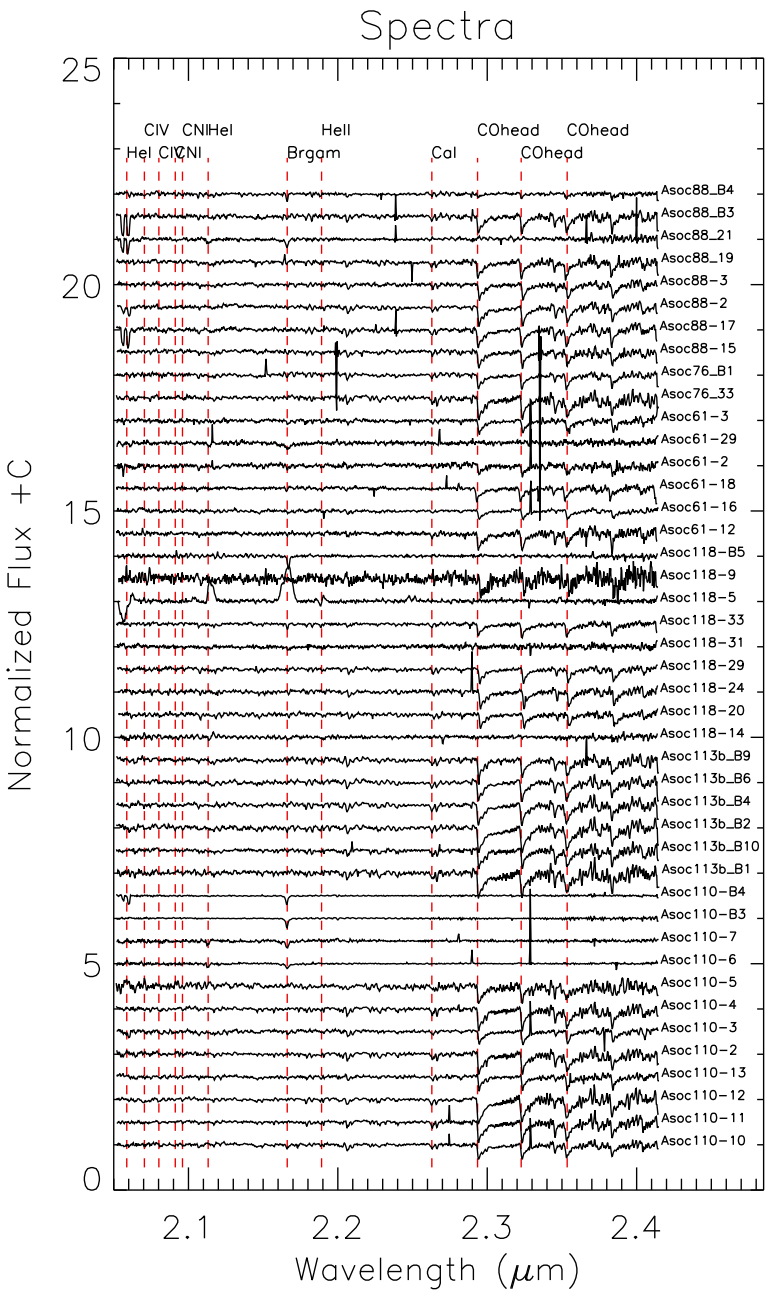

Figure 2.12: K-band spectra obtained with LIRIS@WHT during July 2015 observing run. The numbers to the right identify individual targets within each association candidate. Stars marked with a "B" (e.g, Assoc88_B3) were expected to be late type stars.

Table 2.5: Clasification of the spectra presented in Figure2.12.

Association 61												
ID	12	16	18	29	2	3						
Spectype	K4 III	G6 II	G6 I	B0.5 III	K0 III	K2.5 III						
Association 76												
ID	33	B1										
Spectype	M7 III	K2 III										
Association 88												
ID	31	19	3	15	24	17	2	21	B3	B4	33	
Spectype	O9 V	K4 III	K4 III	K4 III	K4 III	K5 III	K5 III	B3 I	M1 III	G3 III	K2 III	
Association 110												
ID	6	7	B3	B4	3	10	13	5	11	12	2	4
Spectype	B0.5 III	B0.5 III	A0 I	A0 I	K2III	K2.5 III	K3 III	M1 III	M3 II	M4 I	M6 III	M6 III
Association 113												
ID	B4	B6	B1	10	B9	B2						
Spectype	K1.5 I	K1.5 I	K5 I	M4 III	M5 I-II	M4 I						
Association 118												
ID	14	31	B5	5	20	24	29	33	9			
Spectype	O8 III	O9V	A0Ia	WN8	K2 III	K2 III	K2 III	K2 III	K2 III			

Table 2.6: Results from observation of association candidate 118.

ID OBS	RA	DEC	J	H	K	Spectral type
5	18:52:43.69	+00:08:41.5	12.490	10.684	9.614	WN8
14	18:52:46.82	+00:07:47.3	12.812	11.008	10.041	O8 III
31	18:52:49.10	+00:12:24.2	13.022	11.660	10.886	O9 V
B5	18:52:38.85	+00:12:09.9	14.110	11.824	10.629	A0Ia
9	18:52:37.21	+00:11:02.6	12.764	11.080	10.224	K2 III
24	18:52:31.93	+00:11:31.8	13.475	12.123	11.460	K2 III
20	18:52:36.94	+00:12:29.9	13.032	11.706	10.948	K2 III
29	18:52:49.65	+00:12:46.0	12.408	11.317	10.778	K2 III
33	18:52:48.24	+00:11:43.7	14.273	12.184	11.065	K2 III

Unfortunately only 7 early-type stars from a total of 32 OB-type candidates were found. On the other hand, from 11 supergiant-type star candidates, 9 were classified with luminosity class I, particularly in the association 113 where from 5 out of 5 stars were classified as supergiants.

Despite the huge amount of giant-stars, association 118, which includes an already known Wolf-Rayet (Mauerhan et al. 2011), presents some young OB stars. Unfortunately it was not possible to observed all the candidates. Table 2.6 shows all the properties of the stars from this association

2.3.5 Lessons Learned

After analyzing the results from MASCA1.0 we detected the need of two mayor improvements.

- We observed a large number of late-type stars due to a high stellar density in the field. There is still a large number of field population contaminants. These must be better discriminated by the color cuts.
- We need to define an optimization criterium for our choice of D_S and *Nmin*. that were arbitrarily set in our search. This seems to be the most relevant point because of (a) the large numbers of candidate associations found, and (b) the extended areas found for some of the candidates.

2.4 Improvements in MASCA 1.1 .

The first version of the automatic code did not consider the higher rate of spurious candidates in crowded fields. Because of the limited sensitivity of the Q_{IR} pseudo-color and the photometric errors, the identification of cluster candidate is dominated by random clustering. Thus improve MASCA we focused on the selection of OB-type star candidates and how to avoid the spurious detection of random clusters.

2.4.1 New color-cuts, extinction and Q_{IR} parameter

The photometric cuts used in the previous section defined a box in the CMD plus the Q_{IR} cuts. These limits included a large fraction of late-type stars with $(J - Ks) < 2$mag and $K > 10$ mag (See Figure2.1). To improve the selection and reduce the number of late type stars, we decided to use the *curve of extinction* of an O9V star across the CMD.

The total extinction in the K-band for a particular star is given by A_K. We estimate it from $A_K = a_k \cdot d$ where d is in [kpc], and a_k is the mean extinction per kpc.

To estimate a_k we choose an intermediate value between the average Galactic extinction given by Indebetouw et al. (2005; $a_k = 0.15 \pm 0.1$ [mag/kpc]) and the value provided by MASGOMAS-1. The first, presents a smooth behavior in the galaxy, and the second, a strong case of absorption. Thus we assumed $a_k = 0.34$ [mag/kpc] in our search. We added 0.5 mag in K to the curve generated from the extinction of an O9 V star to consider the effect of differential extinction. We also modify the Ks limits from $6 < K < 11.5$ to $9 < K < 12.5$. Figure 2.13 illustrates the new photometric cuts in the CMD, together with the position occupied by MASGOMAS-1 stars.

Changes in the K-magnitude versus Q_{IR} diagram do not present large differences with the previous photometric cuts. Q_{IR} depends on the value of $E(J - H)/E(H - K)$, which we adopted to be 1.78 (Wang & Jiang 2014).

We extended the limits for the Q_{IR} criteria to ±2 (see Figure 2.14). As a reference, the stars from MASGOMAS-1 and MERCER30 are shown in the plot. Table 2.7 summarizes the new photometric cuts.

Table 2.7: Old and new photometric cuts adopted in MASCA

Old	New	Comments
$6 < K < 11.5$	$9 < K < 12.5$	Apparent magnitudes optimized for MOS with LIRIS@WHT.
$1.5 < J - K < 3.5$	extinction curve $+0.5 < J - Ks < 3.5]$	Average extinction of an O9V.
$-0.15 < Q_{IR} < 0.2$	$-0.2 < Q_{IR} < 0.2$	Massive hot stars around $Q_{IR} \sim 0$

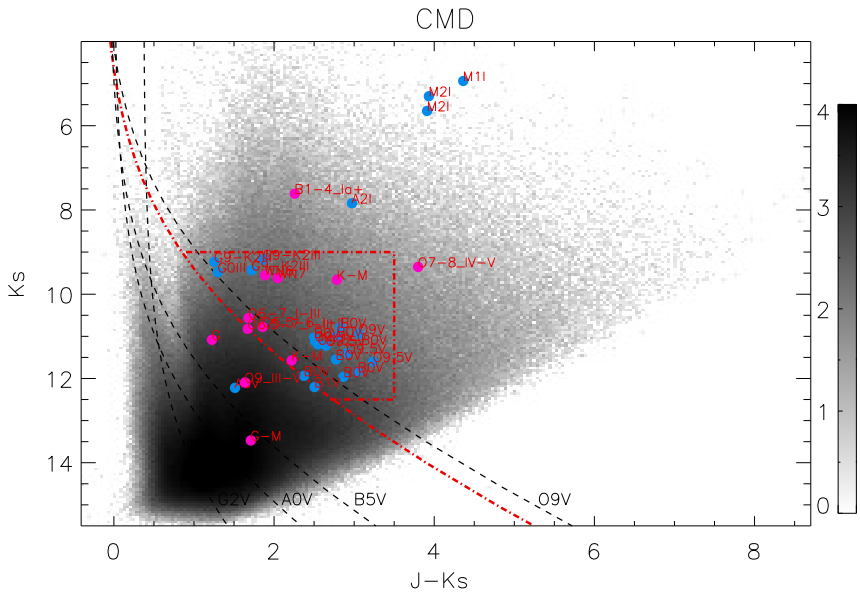

Figure 2.13: CMD of the searching zone centered around $l = 33°$ and $b = 0°$. Blue dots show the 2MASS data.The black curve represent the loci ocuppied by the main sequence stars at different distances, assuming an average extinction of $a_k = 0.34$mag/ kpc (see text). The red curve add 0.5 mag to take into account differential extinction The red dash-dotted curve corresponding to a O9V star and indicates the photometric cut applied in the program, complemented by the cuts in K-magnitude and color (indicated by the thick dash-dotted red line). Adopted intrinsic magnitudes and colors are from Tokunaga (2000). The bue circles show the position of MASGOMAS-1 stars and the pink circles show the position of MERCER30 (de la Fuente et al. 2016) stars as reference.

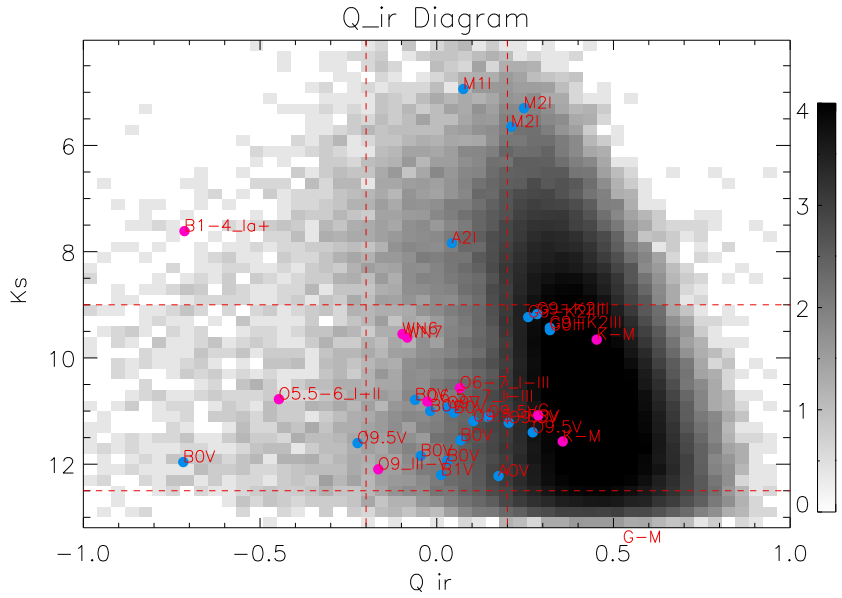

Figure 2.14: Same objects, colors and field as in Figure2.13, now in the Ks magnitude vs Q_{IR} diagram).

2.4.2 Improvements introduced in FOFA

In this section we explore how to obtain a trustworthy catalog of association candidates by studying the differences between combinations of D_S and *Nmin* in different search scenarios. This study attempts to improve the selection of candidate associations.

One of the factors not taken initially into account was the random distribution of field stars in the different areas, which create a Gaussian-like profile in the distribution plot obtained by FOFA. This could explain why during the previous search most of the candidates associations were not confirmed as clusters.

Another factor is the relation between the spurious detection and the stellar densities in the field. We took the density of candidates selected by the photometric cuts from the first section (i.e approximately 8000 targets in a $6^\circ \times 6^\circ$ area) and studied how FOFA behaves for different densities.

The two input parameters for FOFA, *Nmin* and *Ds* are critically affected by the target surface density, as described in this section. To establish the optimal values for these parameters, we explore the distribution of the number of associations as function of *Ds* for several values of *Nmin* using the first output from FOFA.

Optimal values of D_S and Nmin

The random distribution of contaminant stars (mainly field late giants) may be detected as over-densities. We developed a method to find the optimal combination of values for D_S and *Nmin* which minimizes the detection of false/random candidates, that may also explain the large radii found for some associations.

We created 50 mock fields of $6^\circ \times 6^\circ$ (same searching area used in 2.3.1), containing 3000, 5000, 8000, 10000 and 15000 stars located at random positions (10 mock field for each number of stars). These fields were generated using the *random* routine of IDL, and without any synthetic cluster added.

Figure 2.15 illustrates the behavior of FOFA within these random fields. We plot the number of detected spurious associations versus the search distance for different values of *Nmin* and stellar densities. If very small searching distances are used, no associations are found. When we increase the searching distance (allowing association members to be more and more separated) we increase the number of associations until we reach a maximum. For larger distances, associations begin to merge and their number decreases. For extremely large searching distances, every star in the field belongs to a single association. The maximum number of associations increases with stellar density and decreases with *Nmin*, whereas the corresponding distance *Dmax* follows the opposite trend (increases with *Nmin* and decreases with stellar density, see Table 2.8).

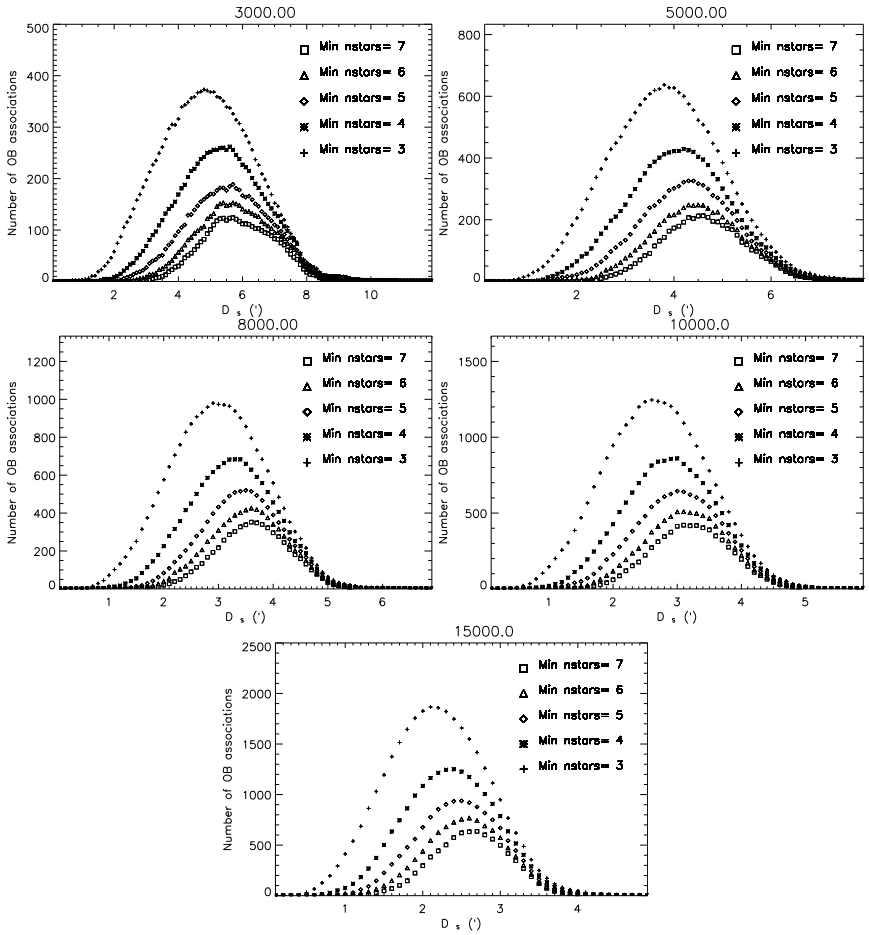

Figure 2.15: Distribution of association candidates detected in random fields for different values of *Nmin* (minimum number of association members) and stellar density. The total number of random stars are showed in the top of each distribution.

Table 2.8: Distances, in arcmin, where the maximum number of association candidates is found, obtained from Figure 2.15. Numbers in parenthesis give the total number of stars in the $6° \times 6°$ field

	DENSITY				
	$\rho(3000)$	$\rho(5000)$	$\rho(8000)$	$\rho(10000)$	$\rho(15000)$
$Nmin = 3$	4.8'	3.8'	2.9'	2.6'	2.1'
$Nmin = 4$	5.4'	4.2'	3.3'	3.0'	2.4'
$Nmin = 5$	5.7'	4.5'	3.5'	3.0'	2.5'
$Nmin = 6$	5.7'	4.4'	3.6'	3.0'	2.6'
$Nmin = 7$	5.7'	4.5'	3.6'	3.1'	2.7'

The following summarizes the new parameters and notations used in our study.

Ncum(D_S, $Nmin$): Number of associations candidates.

Density:(ρ_*): Stellar surface density of the field [stars/deg^2].

Dtyp(ρ_*): Typical mean distance between stars for a specific density in a regular grid.

Dopt(ρ_*): Maximum distance between stars for which no random associations are found in the field. This distance is obtained using the detection distribution output from FOFA for each Nmin.

D′opt(ρ_*): Approximate value for Dopt

Duse: Distance to be used as D_S, derived from the tests. It will be based on Dopt and is a function of $Nmin$ and ρ_*.

Consider an 2D homogeneous distribution of stars in a square region of side one degree, with total numbers of stars $\#stars$. First we determined the typical distance between stars, $Dtyp$. This is the mean distance for stars regularly distributed in space in a rectangular grid (aside from border effects). Figure2.16 illustrates the situation and eq. 2.2 gives the value of $Dtyp$.

$$Dtyp = \frac{\sum distances}{N} = \frac{4l + 4l\sqrt{2}}{8} = \frac{l(1+\sqrt{2})}{2} = 1.21 \times l \tag{2.2}$$

Now if we consider a column of the square with length $large = 1[deg]$ the total numbers of stars in this column is $\sqrt{\#stars}$. Dividing the square distance by the numbers of stars in this column we obtain the homogeneous distance between the stars in the column. Thus l can be well approximated by :

$$l = \frac{1}{\sqrt{(\#stars/degree^2)}} \tag{2.3}$$

Which is the inverse of the stellar surface density of the zone. Thus:

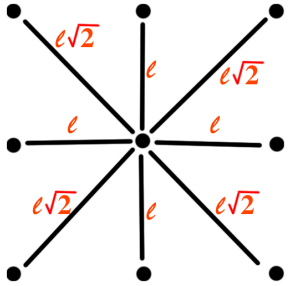

Figure 2.16: Determination of typical mean distance, *Dtyp*

$$Dtyp = \frac{1.21}{\sqrt{\rho_*}} \tag{2.4}$$

For example, take the case of 8000 stars on $6° \times 6°$ field. This corresponds to $\rho_* = 222.2$ and considering the equation 2.4 we obtain $Dtyp = 4.86$. This implies that equidistant stars on a regular grid are separated by $4.86'$. Table 2.9 shows the values of ρ_* and Dtyp for the different test fields.

Table 2.9: Stellar surface density and typical separation distances in the different fields.

#stars \field	ρ(3000)	ρ(5000)	ρ(8000)	ρ(10000)	ρ(15000)
ρ_* [$\#stars/degree^2$]	83.3	138.9	222.2	277.8	416.7
Dtyp [arcmin]	7.93	6.15	4.86	4.35	3.55

The Figure2.17 illustrates the relation between the number of detected candidates and D_S for the particular case of a field with 5000 randomly distributed targets. We see that there is a minimum searching distance below which no spurious candidates are found, We call this distance *Dopt*. It is the largest distance for which no association is found for a specific *Nmin*. We see that we find no cluster for *Dopt*=0.3 arcmin if $Nmin = 3$ and $Dopt = 2.1$ arcmin if $Nmin = 7$. As these values result from random points distributions, we know that these association are not real. For small values of D_S we only find apparent associations with a low number of members. As we increase the minimum number of members we have to increase the allowed separation between them to find an apparent association. This relation gives us an indication of the values that we have to use in our searches and allows us to obtain an estimation of the optimal search distance, *Dopt*.

The Figure2.18 shows distribution for the 10 tests for different densities of the distribution of cluster candidates found as a function of *Nmin*. Table 2.10 shows the average values obtained. It is possible to see the dependence of the values on the field surface density and *Nmin*.

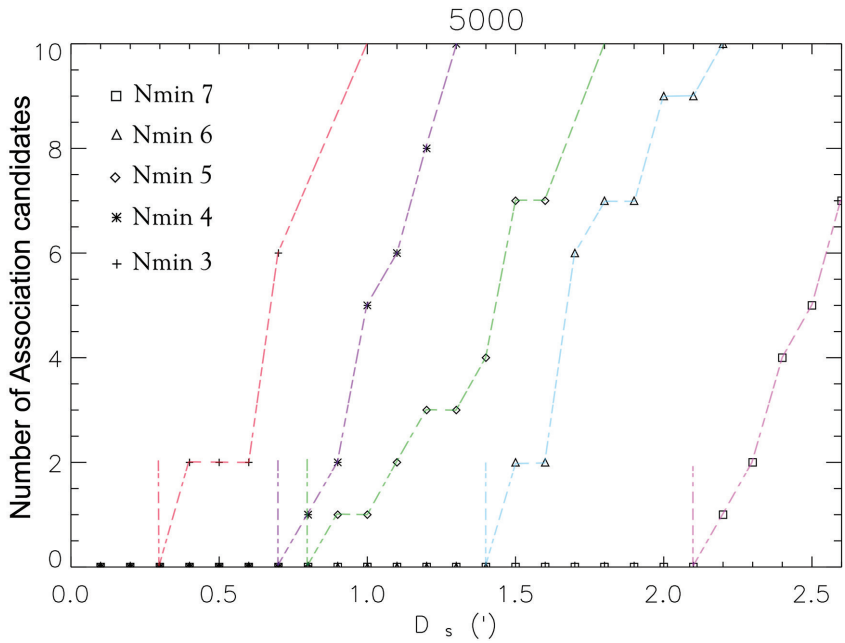

Figure 2.17: Distribution of number of associations founded by the module of FOFA in function of the D_S distance, for the field with 5000 randomly distributed targets in a $6^\circ \times 6^\circ$ for five different *Nmin* values, from where the *Dopt* values are obtained, in this case of 5000 stars, the value is marked by the short vertical lines.

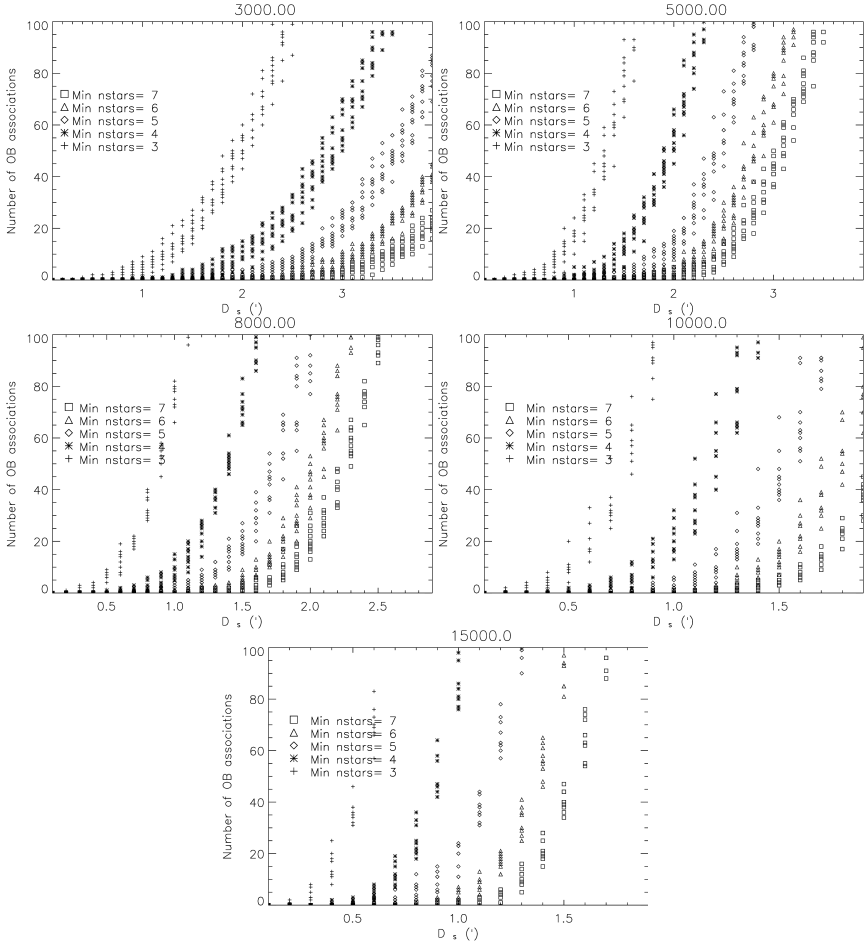

Figure 2.18: Distribution of number of associations founded by the module of FOFA in function of the D_S distance for the 5 different stellar surface densities. For each of the densities we display the result for the 10 random test fields generated for five different *Nmin* values.

Table 2.10: Dopt values in arcmin as function of *Nmin* and ρ_* for the differents fields.

	Nmin				
Density	*Nmin* = 3	*Nmin* = 4	*Nmin* = 5	*Nmin* = 6	*Nmin* = 7
ρ(3000)	0.55 ± 0.13	1.12 ± 0.25	1.63 ± 0.35	2.09 ± 0.41	2.55 ± 0.33
ρ(5000)	0.39 ± 0.12	0.92 ± 0.12	1.29 ± 0.20	1.56 ± 0.25	1.76 ± 0.27
ρ(8000)	0.23 ± 0.07	0.60 ± 0.13	0.84 ± 0.17	1.09 ± 0.23	1.31 ± 0.17
ρ(10000)	0.17 ± 0.08	0.45 ± 0.12	0.76 ± 0.16	0.99 ± 0.14	1.08 ± 0.18
ρ(15000)	0.15 ± 0.06	0.35 ± 0.07	0.55 ± 0.10	0.79 ± 0.17	0.96 ± 0.13

The associated errors correspond to the standard deviation of the mean value obtained for each of the 10 measurements. Figure 2.19 illustrates the values from table 2.10 with a characteristic decrease of *Dopt* as the density increases and *Nmin* decreases. But it also can be seen that it is not a linear function of the density particularly for high *Nmin*.

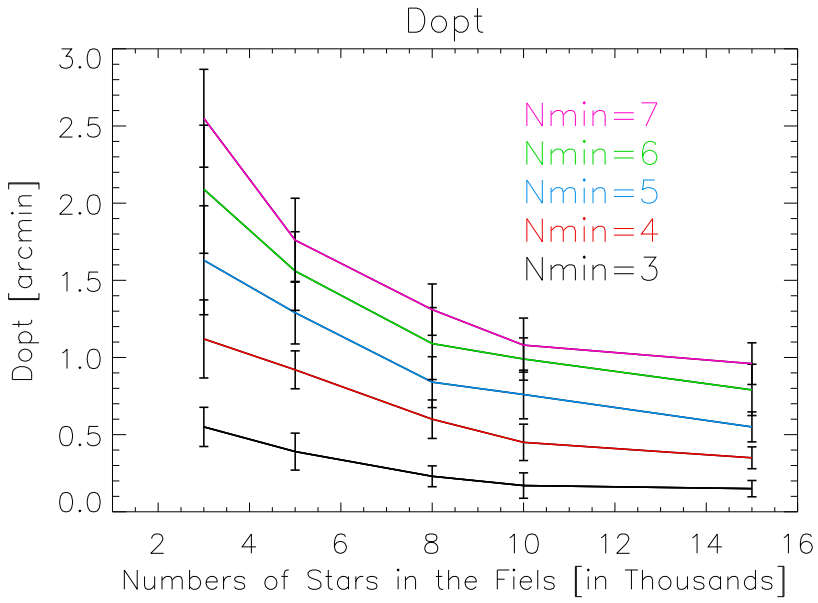

Figure 2.19: Dopt as a function of Nmin and ρ_*. Black Nmin=3; red Nmin=4; blue Nmin=5; green Nmin=6 ; purple Nmin=7

Instead of using a large table of Dopt values, we will try to derive a simple expression for *Dopt* as function of ρ_* and *Nmin*. The left-hand panel of Figure 2.20 shows the behavior of

Dopt with *Nmin*. It is close to linear, with a slope that varies with ρ_*. The right-hand panel of Figure 2.20 shows the behavior of the ratio *Dopt*/*Dtyp* (given in Table 2.11). We still see some dispersion, but the slope is very similar for all surface densities with the main difference being in the intercept (see Figure2.21). Using a second grade approximation we generate a single linear function for *Dopt*/*Dtyp* (see Figure2.21).

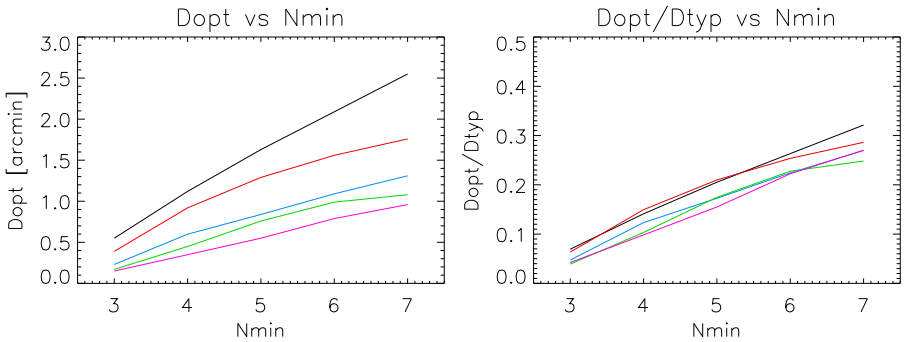

Figure 2.20: Left: Dopt as function of Nmin. Right: Dopt/Dtyp as function of Nmin. Different colors represents different densities. Black : ρ(3000); red : ρ(5000) ; blue : ρ(8000) ; green : ρ(10000) and pink : ρ(15000)

Table 2.11: Dopt/Dtyp values obtained for each field as function of *Nmin*.

	Nmin				
Density	*Nmin* = 3	*Nmin* = 4	*Nmin* = 5	*Nmin* = 6	*Nmin* = 7
ρ(3000)	0.069±0.016	0.141± 0.032	0.205± 0.044	0.263± 0.052	0.321± 0.040
ρ(5000)	0.064±0.019	0.150± 0.020	0.210± 0.033	0.254± 0.041	0.286± 0.044
ρ(8000)	0.047±0.014	0.123± 0.026	0.173± 0.034	0.224± 0.048	0.270± 0.034
ρ(10000)	0.039±0.019	0.103± 0.027	0.175± 0.036	0.228± 0.031	0.248± 0.040
ρ(15000)	0.042±0.015	0.098± 0.020	0.155± 0.027	0.222± 0.047	0.270± 0.038

Thus, let be

$$y = \frac{Dopt}{Dtyp} \tag{2.5}$$

Then the average relationship will be

$$y = m \cdot Nmin + n \tag{2.6}$$

When m and n are the average values obtained from the linear fitting of each surface densities.

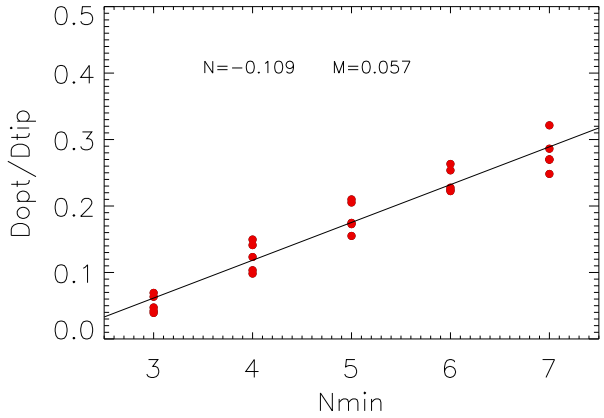

Figure 2.21: Average slope and the values of the y function (see eq 2.6)

$$m = 0.057 \pm 0.003 \tag{2.7}$$
$$n = -0.109 \pm 0.018 \tag{2.8}$$

However, for each surface density there is a slight dependence of intercept. So that.

$$y = m \cdot Nmin + n - c(\rho_*) \tag{2.9}$$

We fit $c(\rho_*)$ (the intercept) by a second order polynomial:

$$c(\rho_*) = a0 \cdot \rho_*^2 + b0 \cdot \rho_* + c0 \tag{2.10}$$

With :

$$a0 = -5.71510^{-7}$$
$$b0 = 4.25110^{-4}$$
$$c0 = -0.062$$

Thus we obtain :

$$y(\rho_*) = (0.057Nmin - 0.109) + 5.715 \times 10^{-7}\rho_*^2 + 4.251 \times 10^{-4}\rho_* - 0.062) \tag{2.11}$$

and multiplying this equation for $Dtyp$ (considering equation 2.2), we could obtain the $D'opt$

$$D'opt = \left(0.057Nmin + 5.715 \times 10^{-7}\rho_*^2 + 4.251 \times 10^{-4}\rho_* - 0.171\right)\frac{1.21}{\sqrt{\rho_*}} \tag{2.12}$$

We have written $D'opt$ and not $Dopt$ to emphasize that it is the result of the fitting process. For the final search distance we use $Duse = f\dot{D}'opt$, where f is a value close to 1.0 that can be used to modulate search, in this work, $f = 1.0$.

As *Duse* is mainly dependent on the *Nmin* value, the routine searches for different *Nmin*. This means that a candidate association can be found several times as long as it fulfills the requirement to form an association. This provides us with extra information about the association, (i.e an association candidate found in more than one combination of *Duse*, *Nmin* is more promising than another one found for only one combination.

Expression 2.12 was included in FOFA and is part of the determination of the distance *Duse* from the surface density of the zone for $Nmin = 3, 4, 5, 6, 7$.

Some extra tests

We finally carried out two additional tests. In the first one, we introduced five artificial clusters to simulate MASGOMAS-1 (21 stars within a circle with diameter 60", corresponding to a surface density of $\sim 10^5$ stars degree^{-2}). For the second test we added 5 simulated clusters with lower densities.

For each test we used the same fields as in Table 2.9. The results of the test will allow us to determine how effectively the program can identify simulated clusters.

Figures 2.22 to 2.26, top, illustrate the output distribution of candidate associations for each *Nmin* and bottom, the surface density obtained in the first test. It is clear that at least 5 candidates are found in each of the different density fields before the search distance reaches *Duse* (indicated by the blue dotted lines). The accompanying plots to the right correspond to the random field and the candidate found by FOFA. In a few cases random associations are also found but for only one *Nmin*. This supports our idea that an association found only once has a lower probability to be a real association.

A particular behavior in the distribution is a small bump to the left in each graph, approximately at 0.3 [arcmin]. This corresponds to multiple detection of the same simulated cluster, a behavior expected for small search distances when the stars form random subclusters separated by a distance slightly larger than *Duse*. This behavior is particularly seen at high stellar densities and low *Nmin* values, because foreground stars appear between the cluster stars helping to form small subclusters.

For the second test 5 different clusters were created with 7, 9, 12, 16 and 21 stars in an diameter of 60" (corresponding to densities between $3 \cdot 10^4$ and 10^5 stars degrees^{-2}). Figures 2.27 through 2.31 give the corresponding output in the same way as Figs 2.22 to 2.26, but now with simulated clusters of different densities. In all cases, the behavior is similar to that in previous Figures.

Our conclusion from our tests is that FOFA recovers the simulated clusters, but we have to pay attention to some artifacts:

- It is always possible that random clustering leads to false detections. This usually happens only for one value of *Nmin* (usually that with the largest *Duse*). Candidates detected only at the largest Nmin should have a smaller priority for follow-up, whereas candidates detected for several combinations of (*Nmin*, *Duse*) should have a higher priority.

- It is possible that we find subclusters, particularly at low values of *Duse*. These subclusters will probably be detected as a larger cluster at larger *Nmin*, *Duse* values
- Small clusters will only be detected with small values for Nmin (e.g. $Nmin = 3, 4$)

2.5 Application of the new method and first bona-field catalog for the inner disk.

In this section we run our improved MASCA (v1.1) over the same field studied in section 2.3 (i.e. centered around $l = 33° \pm 3°, b = 0° \pm 3°$).

Figure 2.32 shows the spatial distribution of OB-type stellar candidates. We notice that the surface density is not uniform (as it was in our tests). We highlight in red the region chosen to estimate the surface density in our field (the program allows to chose different regions). The selection of this area is somewhat arbitrary, and we decided to adopt the central region corresponding to $l = 33° \pm 3°, b = 0° \pm 3°$). The surface density obtained in the field was 283[star/deg^2] which correspond to a total of 10000 stars in a field of $6° \times 6°$. Nevertheless the total of stars in the field which pass the photometrics cuts is 5944 stars candidates.

Figure 2.33 shows the result of using 5 different *Nmin* (3, 4, 5, 6, 7) with the corresponding *Duse* (blue dashes lines) over the search region. For each of the *Duse* values, MASCA creates an independent detection that is represented with a different color in Figure 2.34

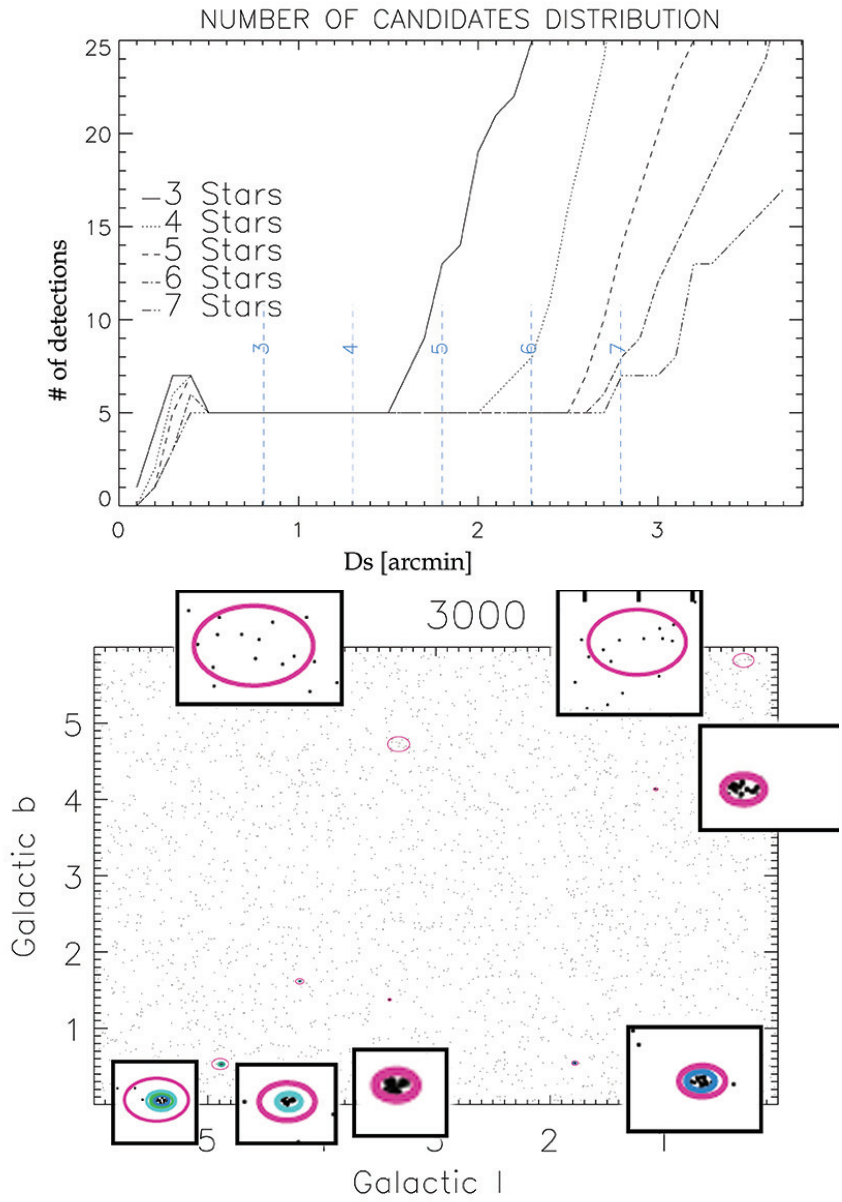

Figure 2.22. Top : Candidates distribution with 5 MASGOMAS-1 in a 3000 random stars field. The vertical lines mark the distance *Duse*. Note the bumps at very low search distances and the random candidate found for *Nmin* = 7 at distances close to *Duse*. Bottom: Spatial distribution of the candidates over the field. The inlets offer a zoom into the candidates found. The top ones are the spurius candidates found only at *Nmin* = 7. Each color represent an identification with a given (*Nmin*, *Duse*) combination.

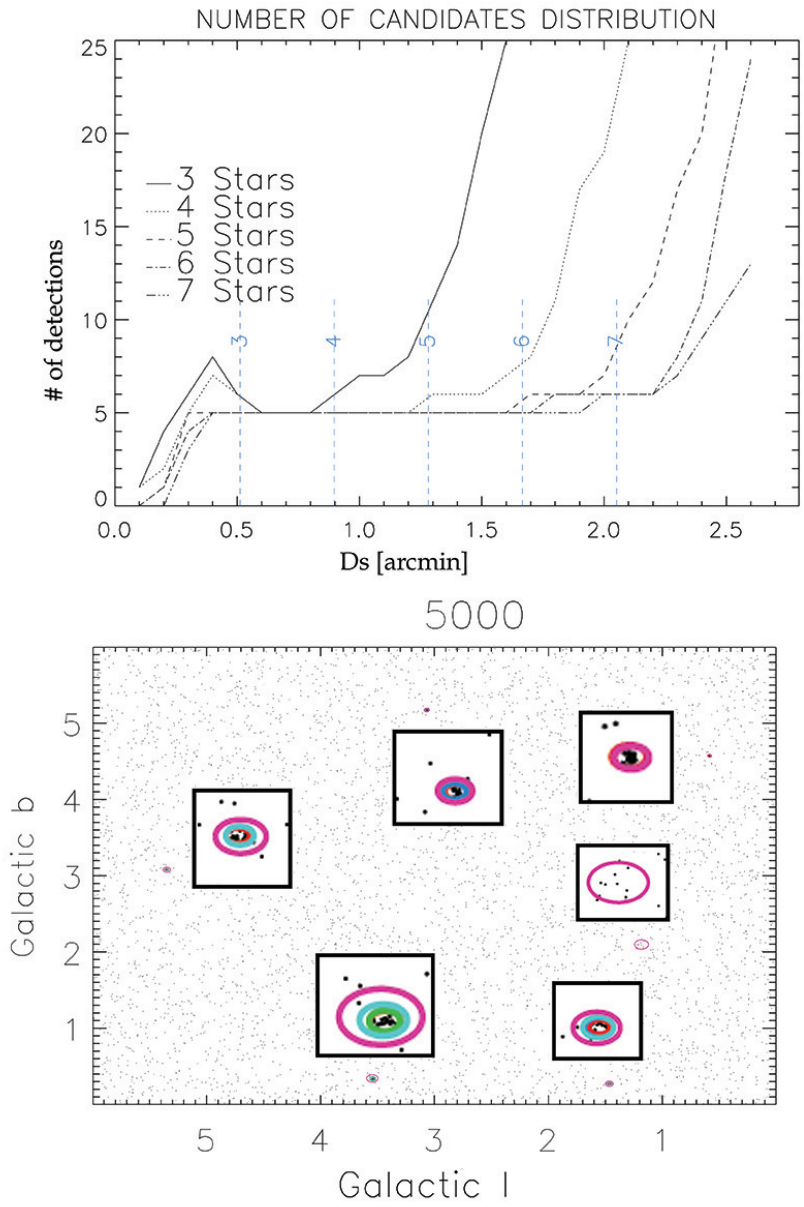

Figure 2.23: Same as Figure 2.22 for the 5000 random stars field

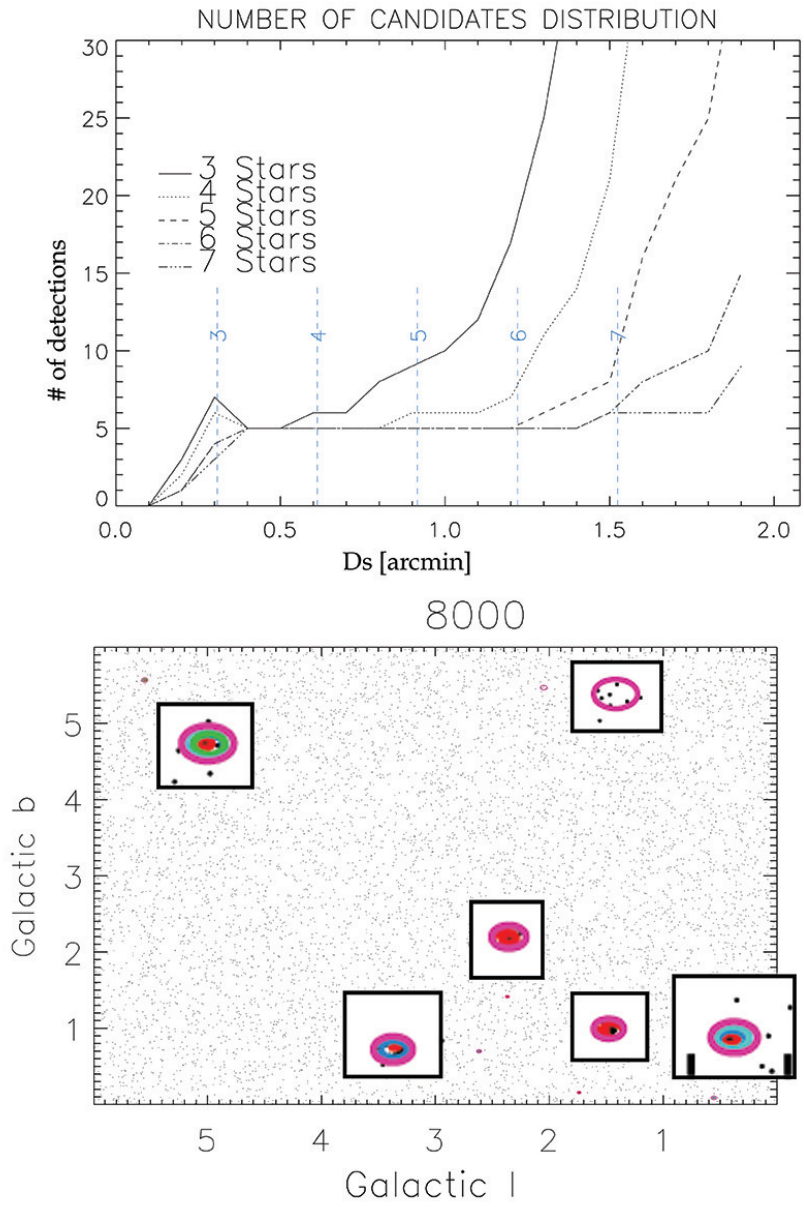

Figure 2.24: Same as Figure 2.22 for the 8000 random stars field

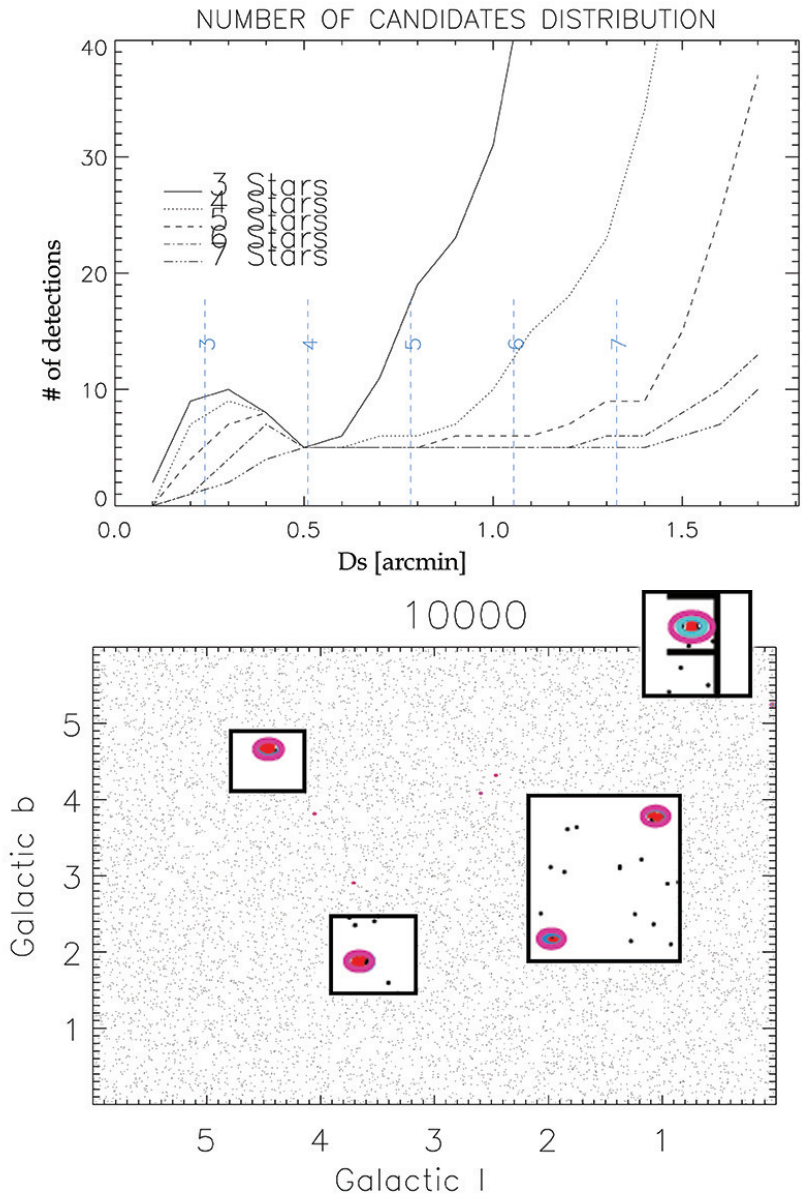

Figure 2.25: Same as Figure 2.22 for the 10000 random stars field

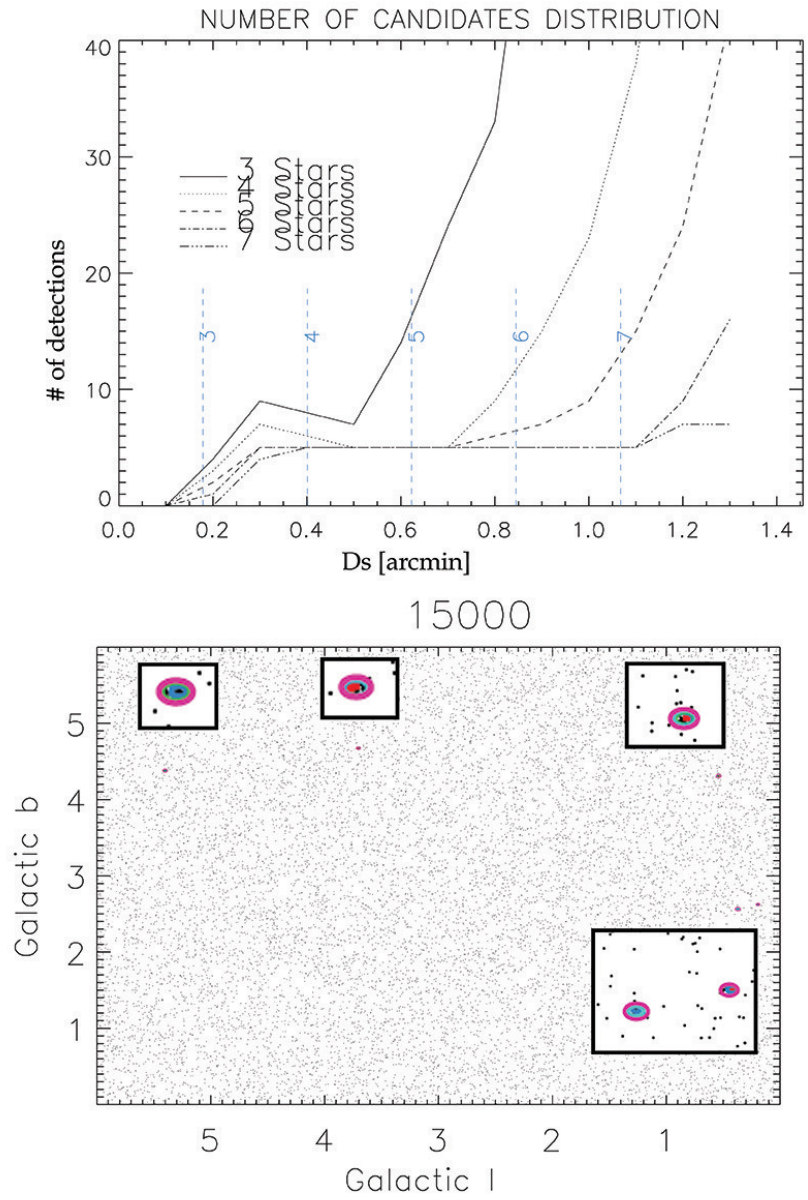

Figure 2.26: Same as Figure 2.22 for the 15000 random stars field

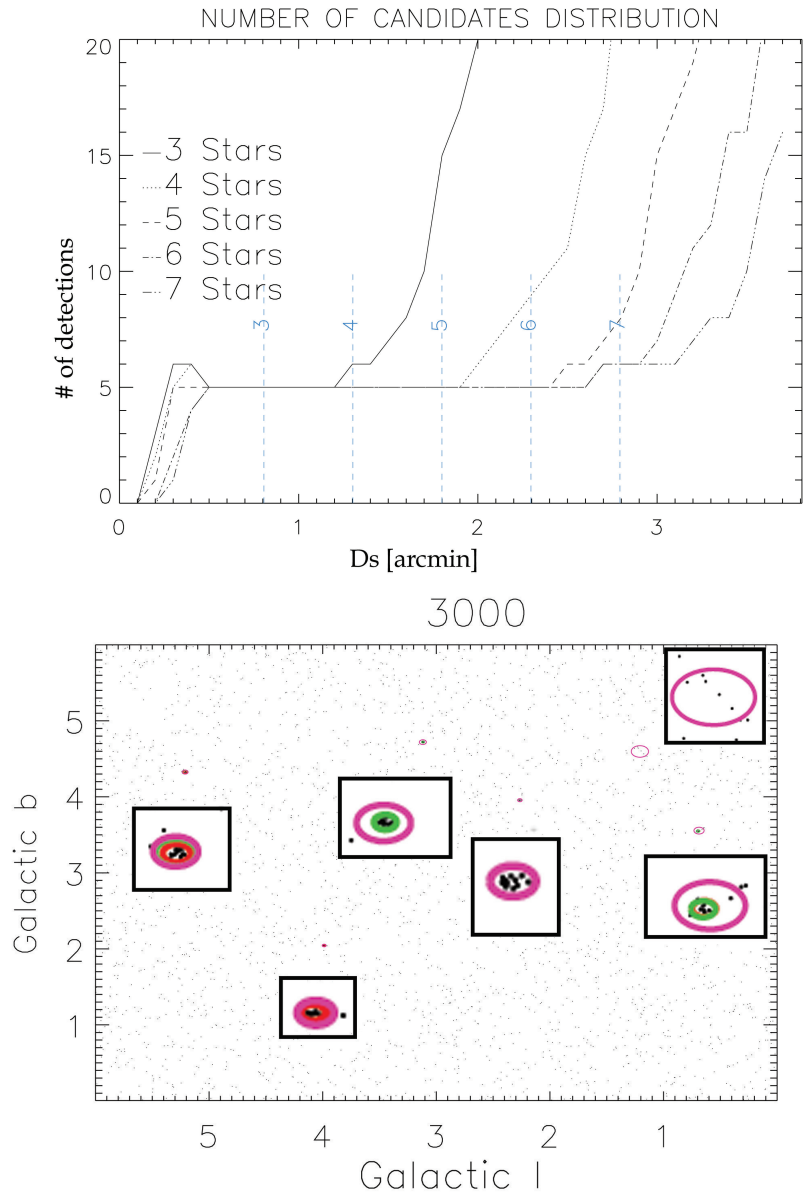

Figure 2.27: Same as Figure2.22, but with simulated clusters of different surface density in the 3000 stars random field.

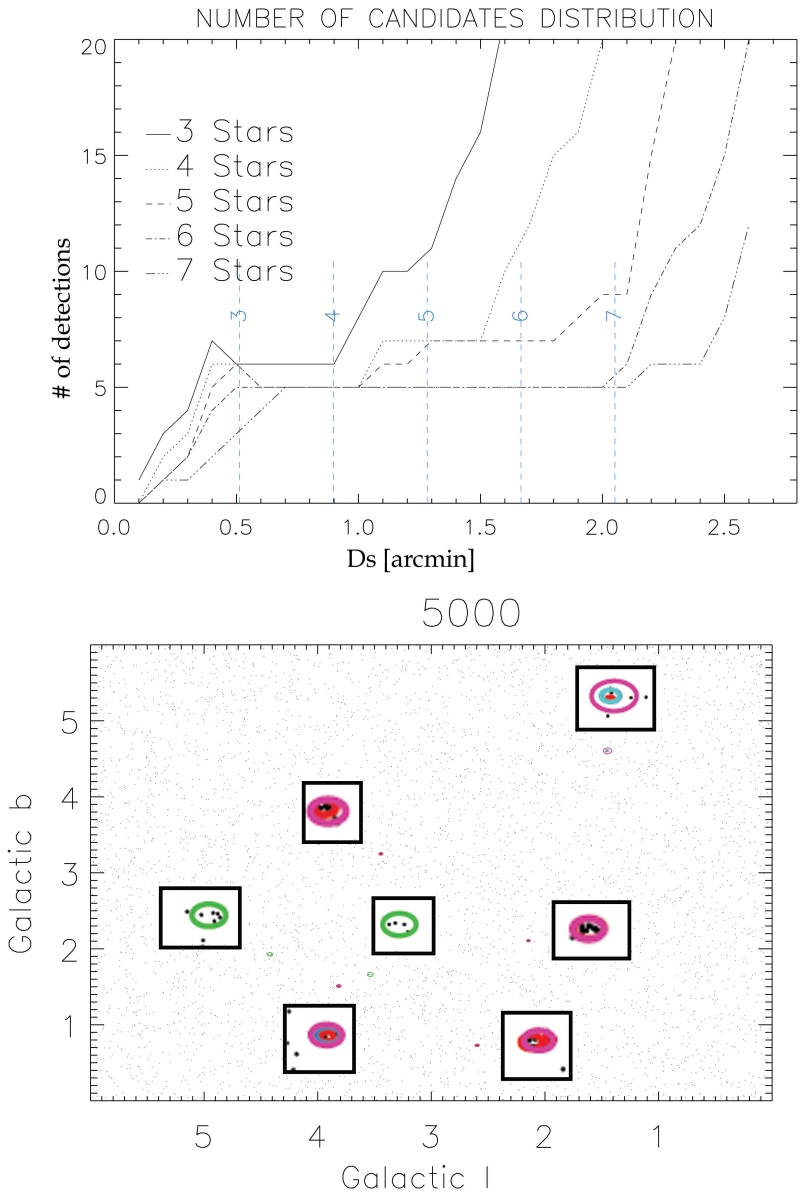

Figure 2.28: Same as Figure2.27 for the 5000 random stars field

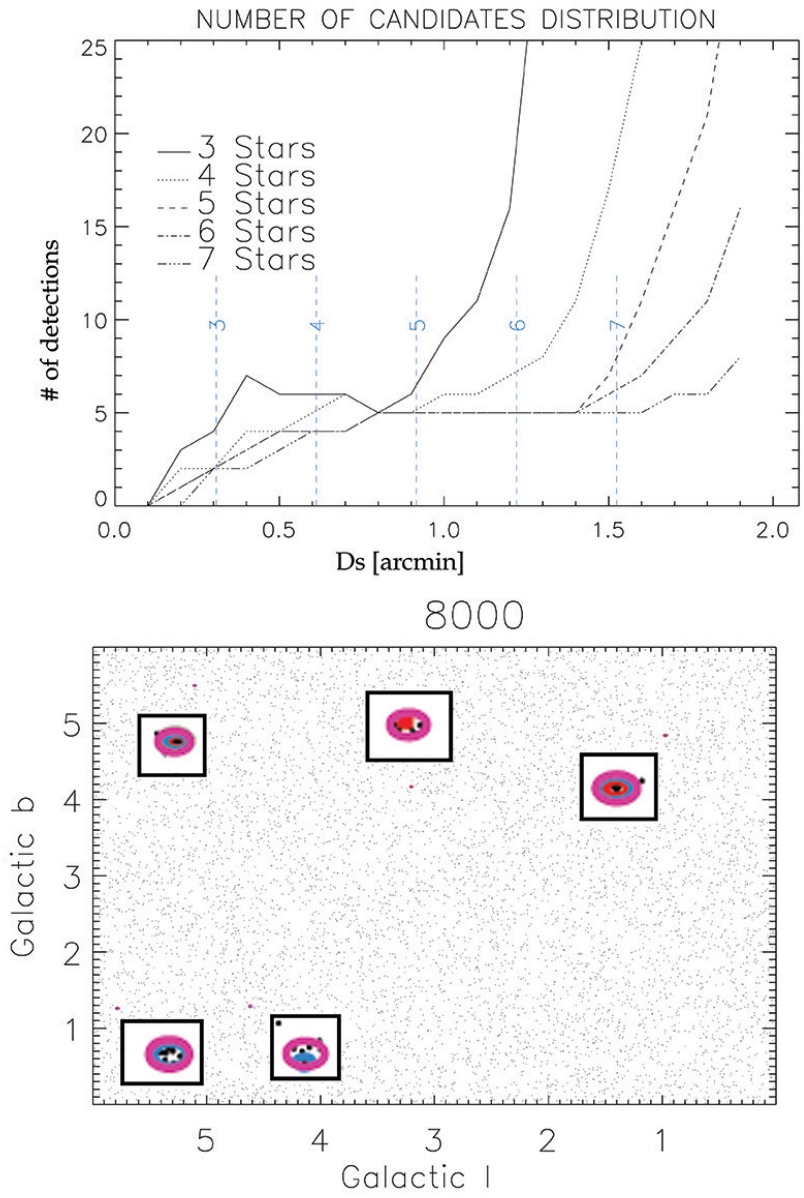

Figure 2.29: Same as Figure2.27 for the 8000 random stars field

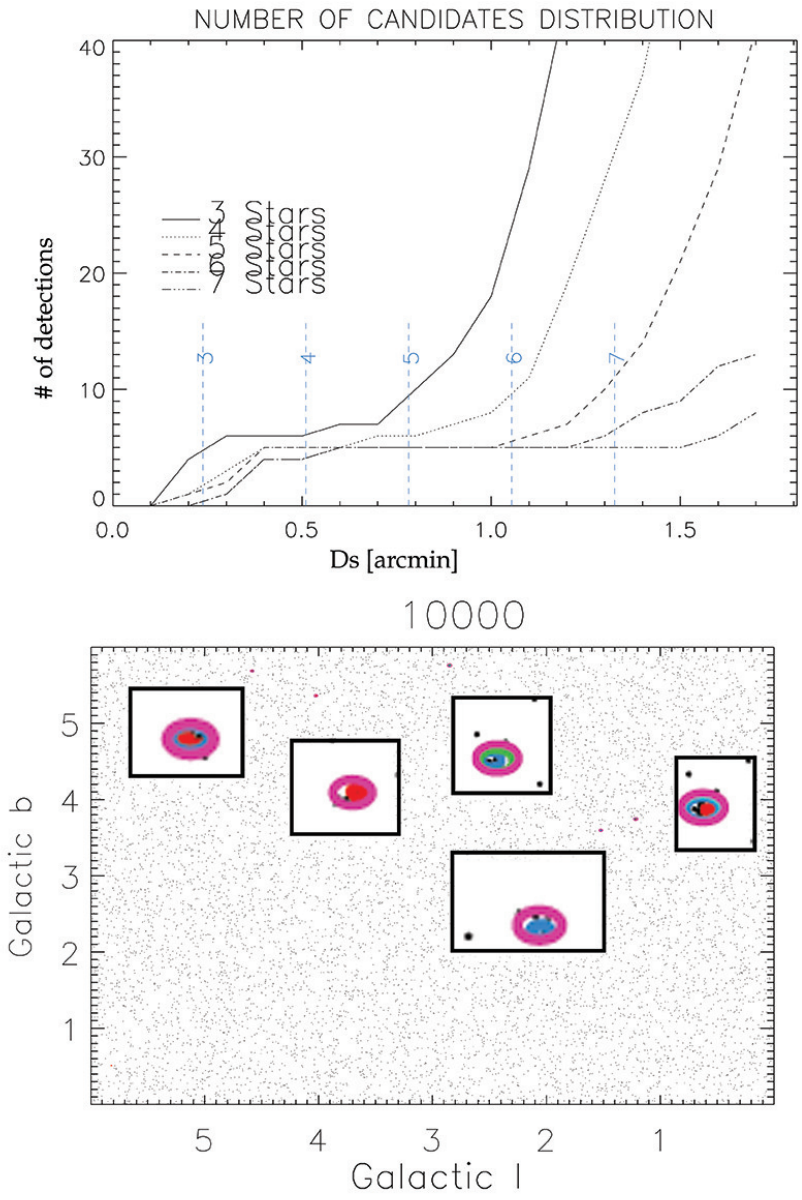

Figure 2.30: Same as Figure2.27 for the 10000 random stars field

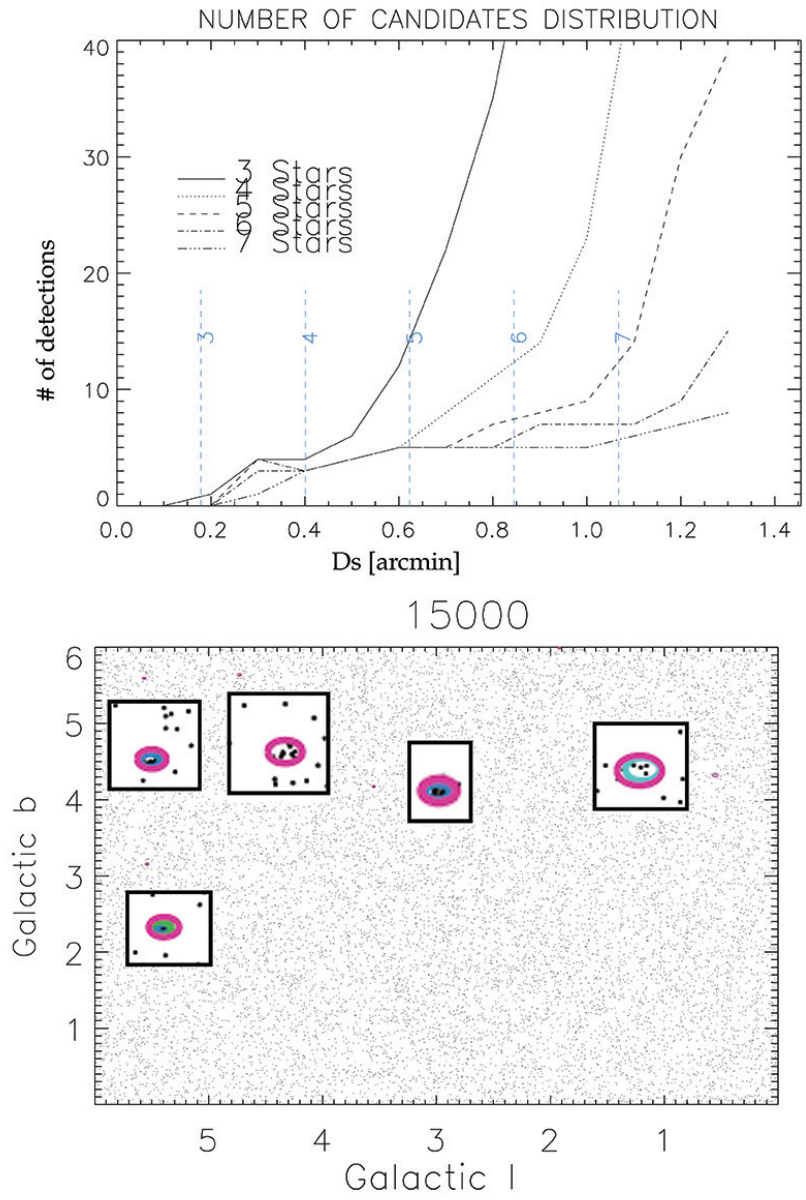

Figure 2.31: Same as Figure2.27 for the 15000 random stars field

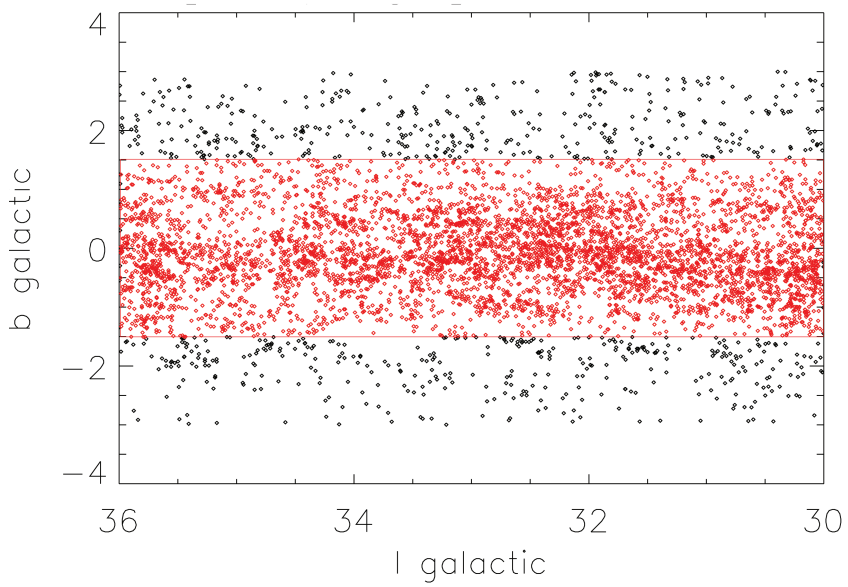

Figure 2.32: Spatial distribution of the stars selected in the $l = 33° \pm 3, b = 0° \pm 3°$ field with the new photometric cuts. The zone highlighted in red has been used to select the surface density, as an arbitrary density value for the field.

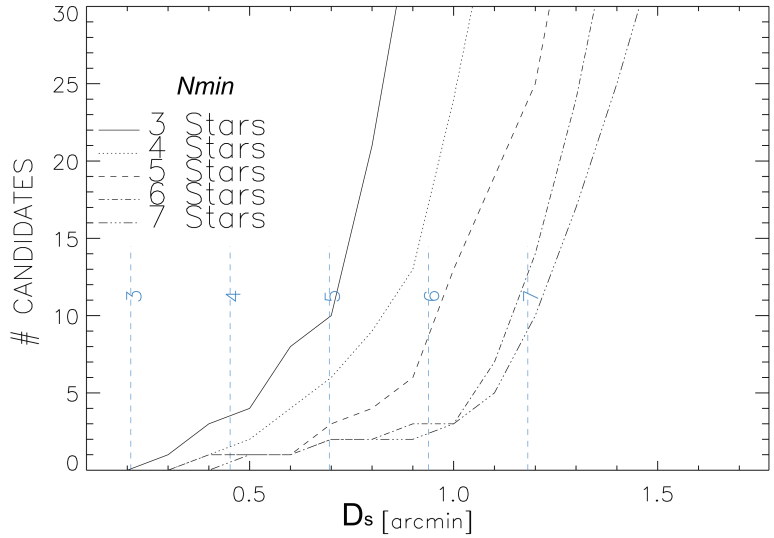

Figure 2.33: Distribution of number of associations founded by the module of FOFA in function of the D_S distance, for the new improve method in the field centered at $l = 33°, b = 0°$. The different lines represent the five *Nmin* used. In blue dash line is represented for each of the *Nmin* the $D'opt$ used.

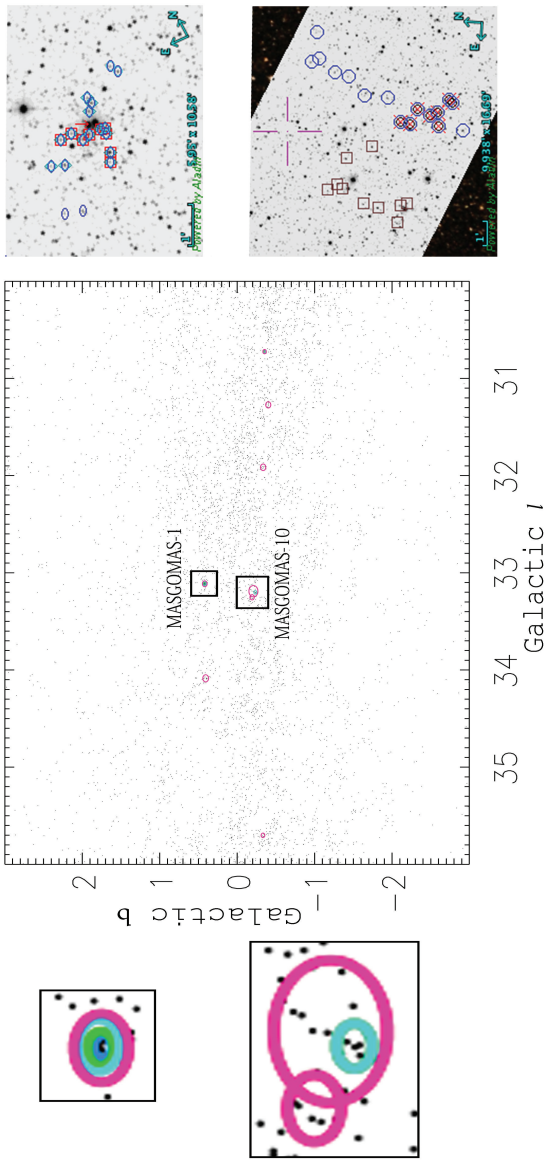

Figure 2.34: Spatial distribution of the association candidates. We highlighted MASGOMAS-1 and MASGOMAS-10. To the left, we zoom into the selection, with different colors representing detections with different (*Nmin*, *Dusar*). To the right the 2MASS Ks image with the candidate and different symbols representing detections with different (*Nmin*, *Dusar*).

Inspection of Figure2.34 shows that now MASCA retrives only 7 candidate (instead of the 125 found in Section 2.3.2). Table 2.12 summarizes the detections for each of the of *Duse* and *Nmin* combinations. The number of candidates has been considerably reduced but now the selected candidates have a good chance to be real associations.

In this Figure we highlight two promising candidates. The first one is again MASGOMAS-1, that we recover, as expected. It is detected for nearly all combinations of (*Nmin*, *Duse*). The second is a new cluster candidate that we call MASGOMAS-10. Actually, we find two nearby candidates. One of them is detected for more than one combination of (*Nmin*, *Duse*) and the second seems to overlap with it. In the Figure we also see a 2MASS image showing the candidate stars in the region. In next chapter we present the follow-up of MASGOMAS-10.

2.5.1 Output and visual inspection

Similarly to the output of MASCA1.0, Figure 2.35 shows an example of the graphic information distributed in seven plots: spatial distribution, CMD, CCD and K versus Q_{IR} for the association and the control region. Table 2.13 gives an example of the full output for the candidate. In this case the candidate 2.6 correspond to the core detection of MASGOMAS-10.

Table 2.12: Candidate detection for different *Nmin* with the corresponding *Duse*. For each combination of (*Nmin*, *Duse*) candidates are identified in the first column by an ordinal followed by the value of *Nmin* (thus, candidate 2.5 is the second candidate found in the search with *Nmin* = 5). The last column gives the number of OB-type stars candidates found in the cluster candidates. coordinates are give in degrees for the J2000.0 epoch. MASGOMAS-1 corresponds to candidates 1.4, 3.5, 3.6 and 6.7. MASGOMAS-10 corresponds to candidates 2.5, 2.6 4.7 and 5.7.

ID	RA	DEC	l	b	# stars
Candidate detection for *Nmin* = 3, *Duse* = 0.21					
–	–	–	–	–	–
Candidate detection for *Nmin* = 4, *Duse* = 0.46					
1.4	282.570	0.350	33.112	0.416	9
Candidate detection for *Nmin* = 5, *Duse* = 0.69					
1.5	282.160	-2.127	30.719	-0.348	6
2.5	283.190	0.132	33.201	-0.234	8
3.5	282.569	0.352	33.113	0.419	14
Candidate detection for *Nmin* = 6, *Duse* = 0.94					
1.6	282.163	-2.126	30.722	-0.350	7
2.6	283.190	0.132	33.201	-0.234	8
3.6	282.565	0.351	33.110	0.421	16
Candidate detection for *Nmin* = 7, *Duse* = 1.19					
1.7	282.167	-2.126	30.724	-0.353	8
2.7	282.459	-1.660	31.272	-0.401	8
3.7	282.693	-1.058	31.914	-0.335	8
4.7	283.163	0.133	33.189	-0.209	16
5.7	283.177	0.196	33.252	-0.193	10
6.7	282.569	0.352	33.113	0.418	18
7.7	283.023	1.212	34.085	0.406	10
8.7	284.419	2.312	35.701	-0.333	9

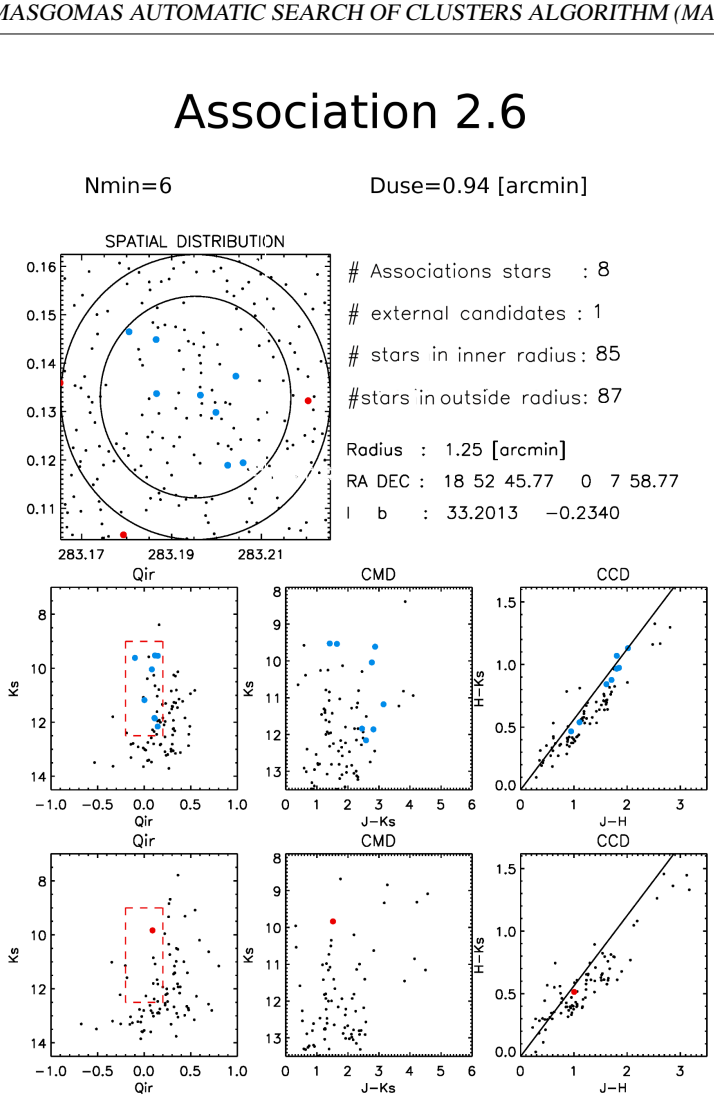

Figure 2.35: MASCA1.1 output information for the core of MASGOMAS-10. The central area shows a different kind of population than the control region. Similar to the previous version, the grafical output provide the spatial distribution of candidates which are part of the association in blue and in red the candidates which only fulfill the photometric cuts. The middle panel the photometric diagram of the association radius and the bottom pannel the control zone photometric diagrams.

Table 2.12 summarizes the candidates and their detection. Note that candidate 1.4 appears to be the same than 3.5, 3.6 and 6.7 but as the member number varies, the center is slightly different from each search. Candidate 2.5 also seems to be detected multiple times, 2.6, 4.7 and 5.7. This makes them particularly promising for follow up. After cross-checking the data, we identify candidate 1.4 with MASGOMAS-1, (found again as expected) and candidate 2.5 overlaps with association #118 from section 2.3.

Table 2.13: Example of the information provided by MASCA for one association candidate. Here we show a few rows from the total of 400 that correspond to the new cluster MASGOMAS-10. The collum "q" provide the extinction value used to calculate the Q_{IR} value. The last two columms give the flags (1 or 0) indicating if the star is an OB candidate and if it is part of the association (i.e., it is at a distance less than $Duse$ from its closest neighbour that belongs to the association).

[2MASS ID]	[ra]	[dec]	[l]	[b]	[j]	[errj]	[h]	[errh]	[ks]	[errks]	[j-ks]	[Q_IR]	[q]	[flag2MASS]	[flag Cand]	[flag Assoc]
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
18524227+0008473	283.17615	0.1464940	33.20670	-0.21500	10.940	0.025	9.993	0.034	9.526	0.026	1.414	0.11574	1.78	AAA	1	1
18525194+0007492	283.21643	0.1303460	33.21070	-0.25820	14.524	0.063	12.882	0.036	12.275	0.026	2.249	0.56154	1.78	AAA	0	0
18525167+0007360	283.21530	0.1322260	33.21180	-0.25640	11.354	0.028	10.349	0.035	9.835	0.026	1.519	0.09008	1.78	AAA	1	0
18524709+0009347	283.19623	0.1596540	33.22760	-0.22690	12.041	0.022	11.816	0.028	11.586	0.026	0.455	-0.18440	1.78	AAA	0	0
18525036+0008109	283.20987	0.1365780	33.21310	-0.24970	15.309	0.068	13.304	0.035	12.447	0.026	2.862	0.47954	1.78	AAA	0	0
18525072+0008386	283.21136	0.1440730	33.22060	-0.24750	13.532	0.023	10.671	0.040	9.310	0.026	4.222	0.43842	1.78	AAA	0	0
18524371+0008013	283.18216	0.1337050	33.19800	-0.22620	11.183	0.035	10.077	0.043	9.538	0.028	1.645	0.14658	1.78	EEA	1	1
18524790+0009045	283.19962	0.1512550	33.22160	-0.23370	15.043	0.060	12.237	0.042	10.940	0.028	4.103	0.49734	1.78	AAA	0	0
18525119+0008511	283.21332	0.1475420	33.22460	-0.24760	12.504	0.023	11.550	0.032	11.080	0.028	1.424	0.11740	1.78	AAA	0	0
18524412+0006506	283.18384	0.1140830	33.18140	-0.23670	12.893	0.032	12.020	0.043	11.607	0.029	1.286	0.13786	1.78	AAA	0	0
18524576+0006451	283.19070	0.1125370	33.18310	-0.24350	14.425	0.035	12.792	0.030	12.150	0.029	2.275	0.49024	1.78	AAA	0	0
18525067+0006209	283.21115	0.1058230	33.18650	-0.26470	12.600	0.029	12.136	0.033	11.937	0.029	0.663	0.10978	1.78	AAA	0	0
18525192+0009042	283.21634	0.1511870	33.22920	-0.24870	14.287	0.037	12.288	0.035	11.408	0.029	2.879	0.43260	1.78	AAA	0	0
18525277+0009034	283.21991	0.1509540	33.23060	-0.25190	15.166	0.046	13.617	0.021	12.955	0.029	2.211	0.37064	1.78	AAA	0	0
18524822+0009193	283.20096	0.1553880	33.22590	-0.23300	14.009	0.031	12.474	0.026	11.813	0.029	2.196	0.35842	1.78	AAA	0	0
18524950+0009117	283.20627	0.1532750	33.22650	-0.23870	12.293	0.026	11.330	0.032	10.982	0.029	1.311	0.34356	1.78	AAA	0	0
18524632+0006402	283.19302	0.1111820	33.18300	-0.24610	14.664	0.066	13.272	0.143	12.670	0.030	1.994	0.32044	1.78	ABA	0	0
18524743+0007080	283.19772	0.1188960	33.19200	-0.24680	14.325	0.045	12.309	0.050	11.178	0.030	3.147	0.00282	1.78	AAA	1	1
18524472+0007503	283.18637	0.1306480	33.19720	-0.23130	14.866	0.065	12.386	0.043	11.225	0.030	3.641	0.41342	1.78	AAA	0	0
18524672+0009323	283.19467	0.1589870	33.22630	-0.22580	13.929	0.041	12.999	0.049	12.603	0.030	1.326	0.22512	1.78	AAA	0	0
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

The following Figures 2.36 to 2.43 shows the output for the 8 candidates selected using $Nmin = 7$, starting with MASGOMAS-1, and with candidates 3.7 and 5.7 corresponding to MASGOMAS-10.

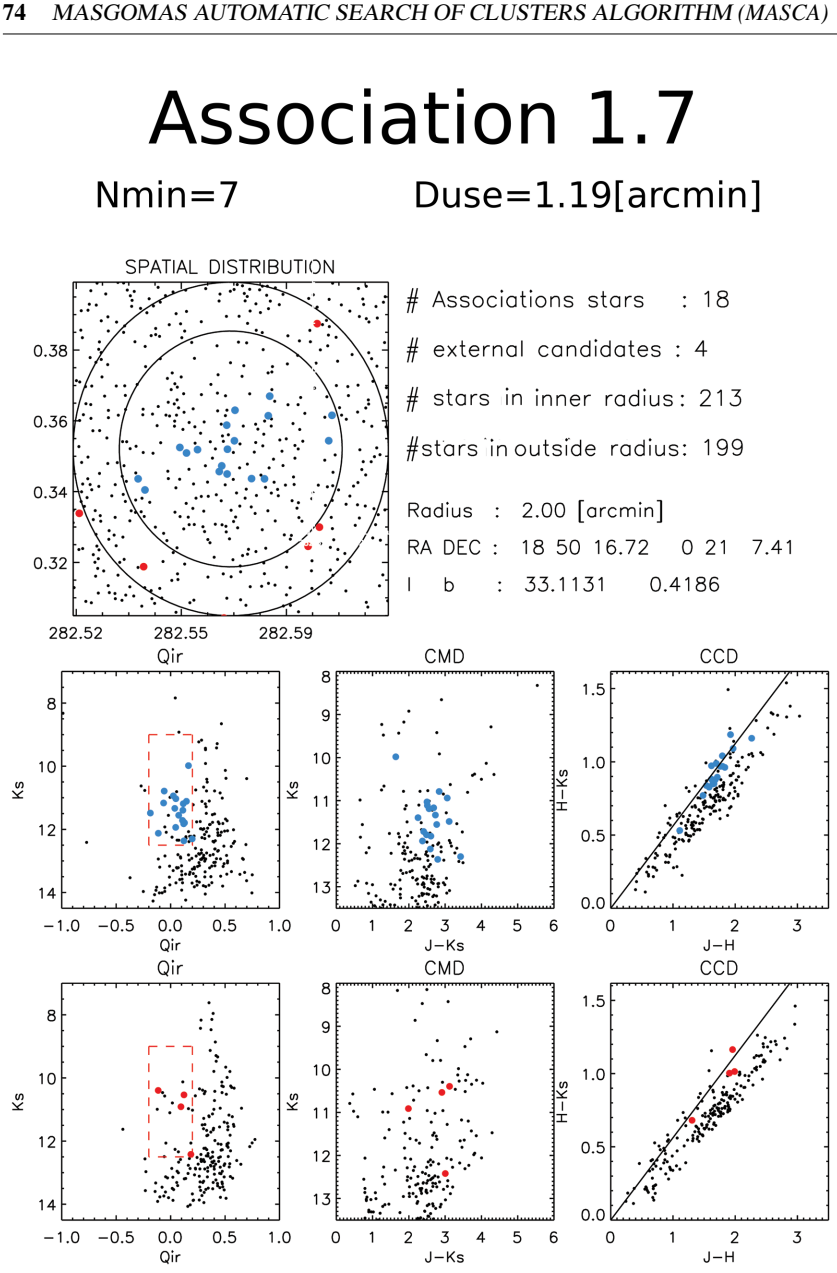

Figure 2.36: Same Figure as 2.35, but for Candidate 1.7

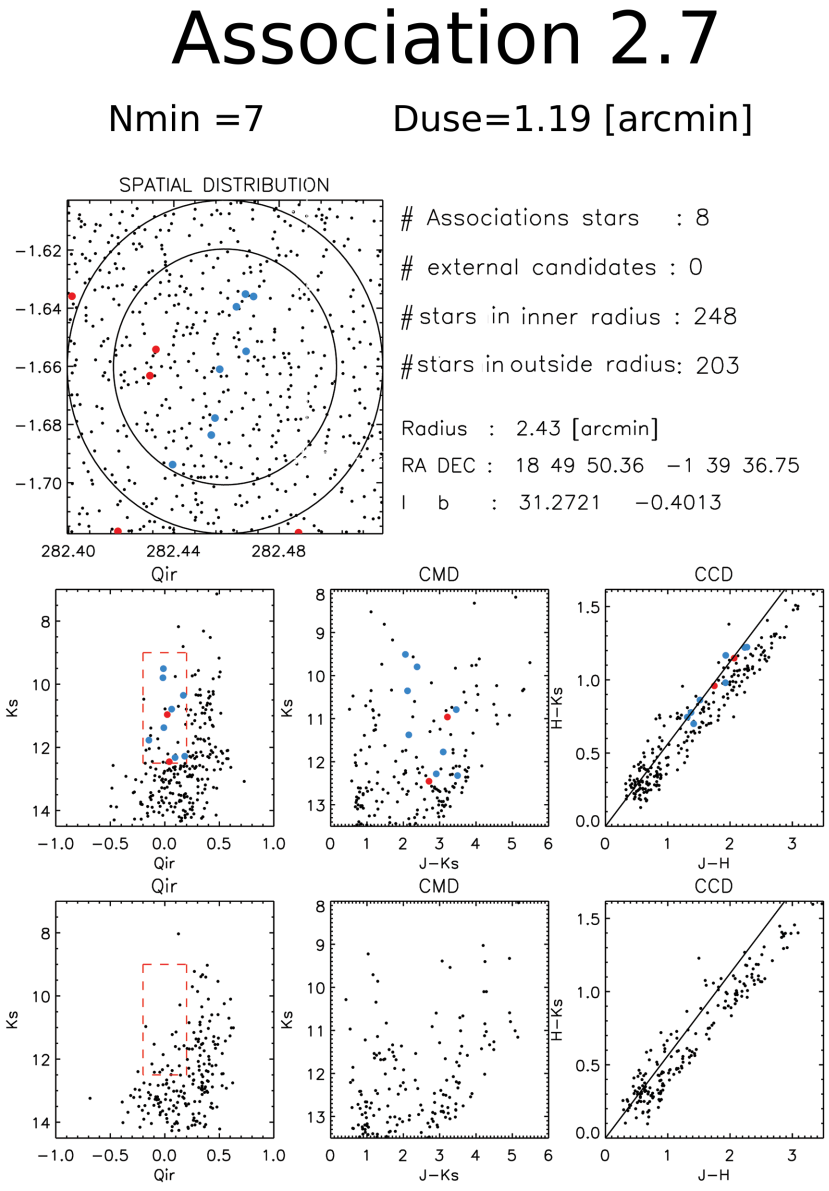

Figure 2.37: Same Figure as 2.36 but for Candidate 2.7 . In this case, two of the stars which fulfill the photometric cuts but not considered as part of the association (i.e. the distance from the closest association star is bigger than D_S) are inside the radius of the association.

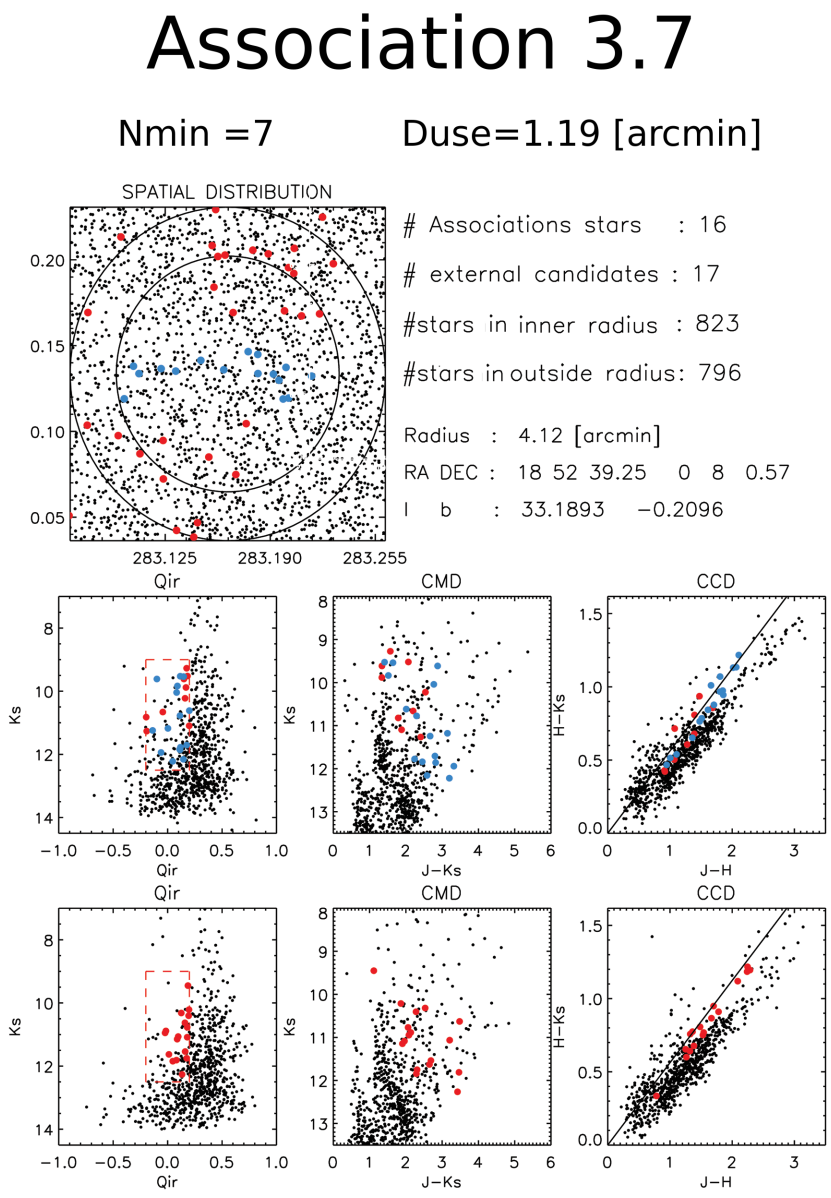

Figure 2.38: Same Figure as 2.36 but for Candidate 3.7

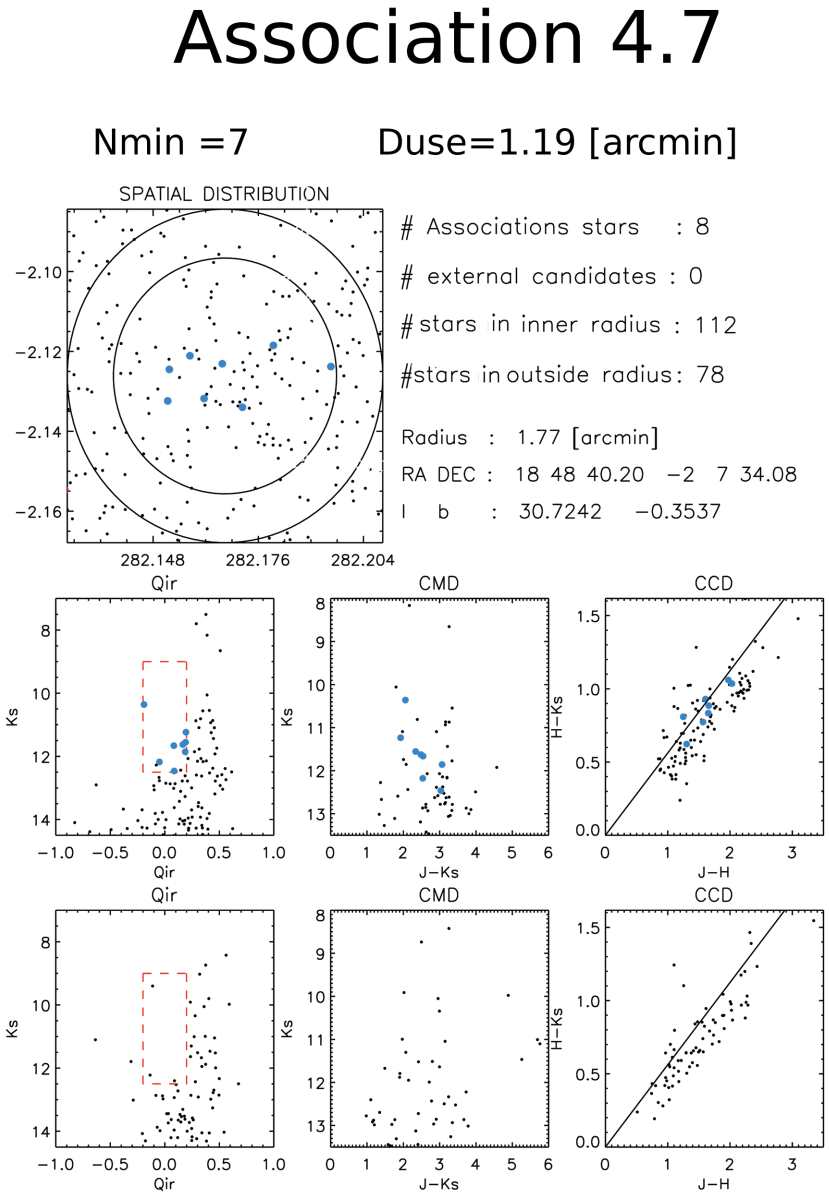

Figure 2.39: Same Figure as 2.36 but for Candidate 4.7

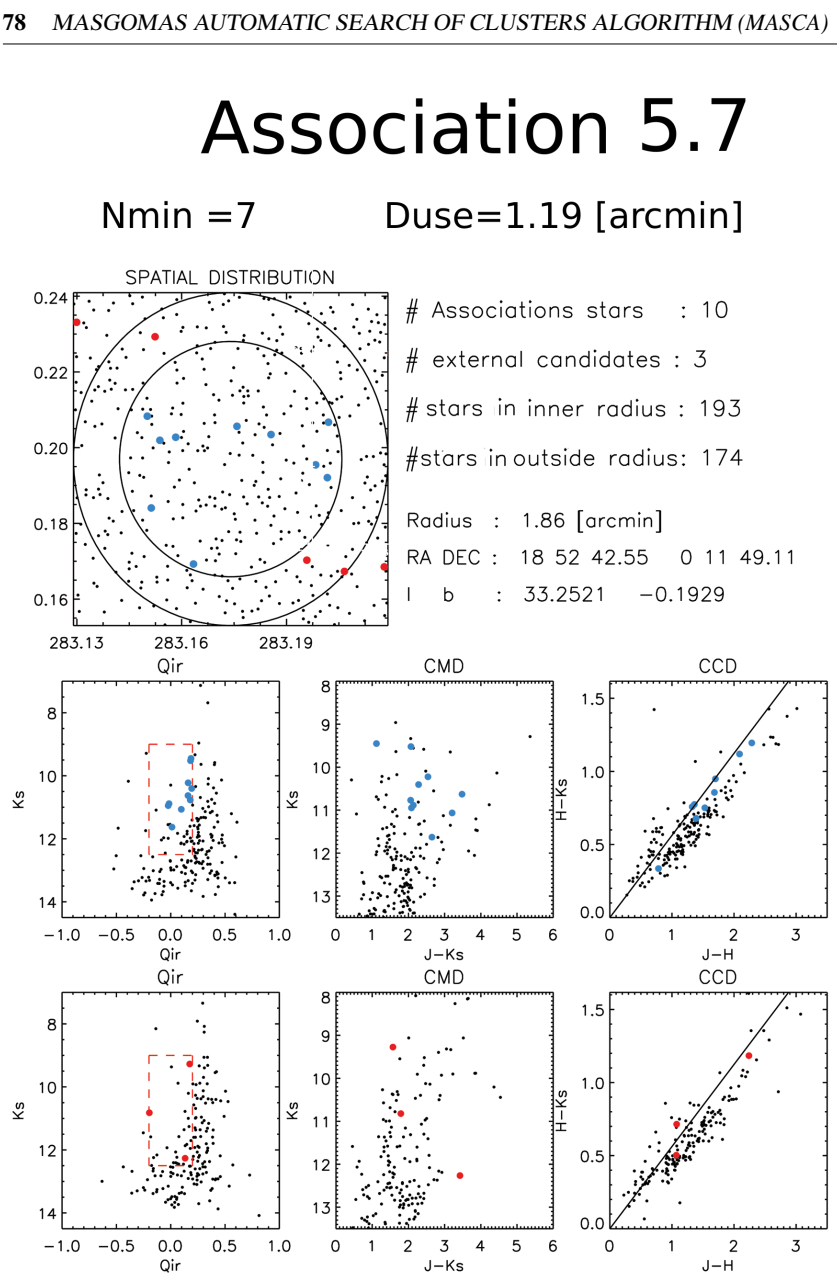

Figure 2.40: Same Figure as 2.36 but for Candidate 5.7

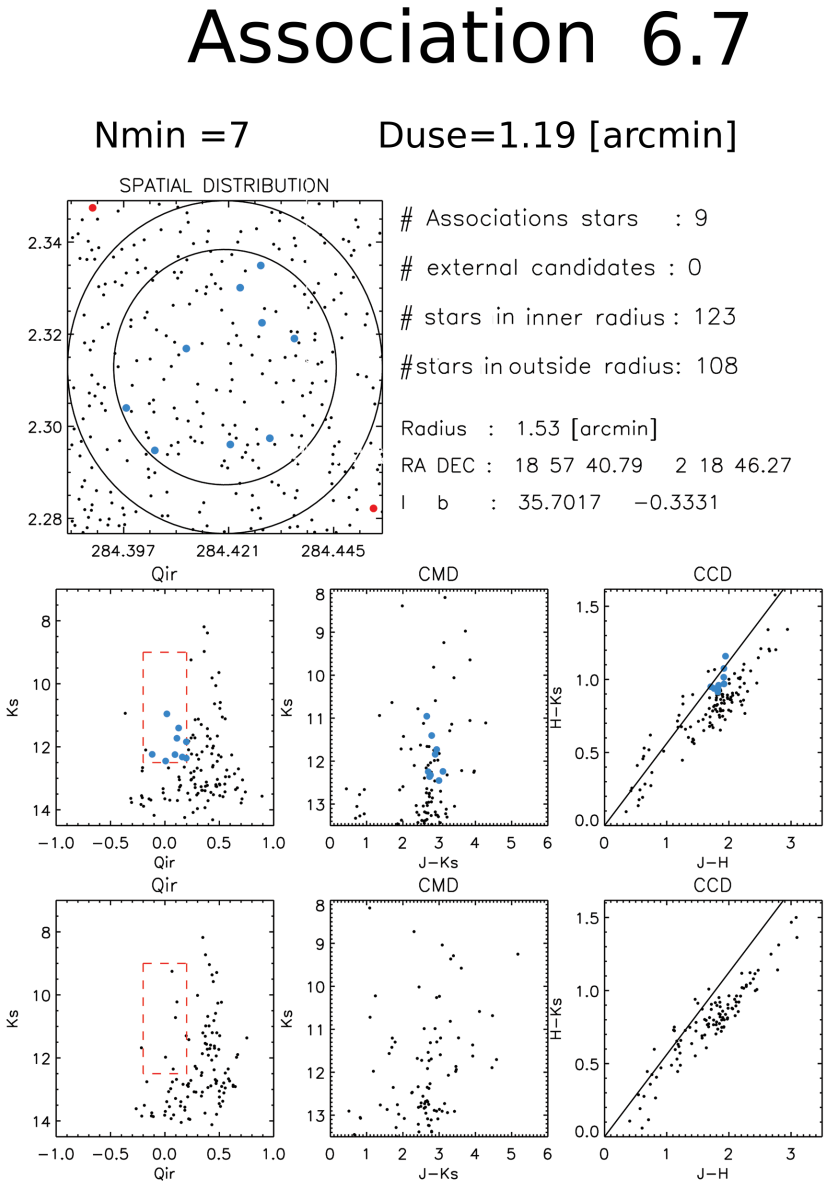

Figure 2.41: Same Figure as 2.36 but for Candidate 6.7

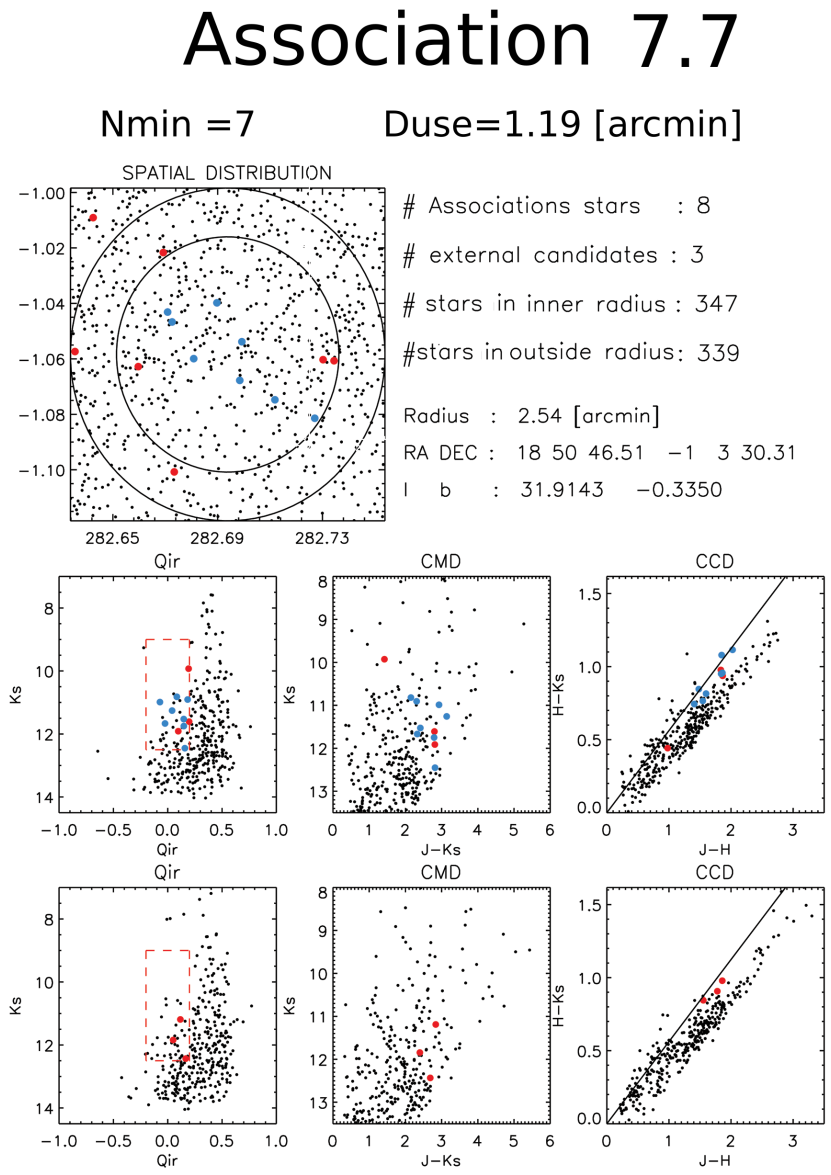

Figure 2.42: Same Figure as 2.36 but for Candidate 7.7

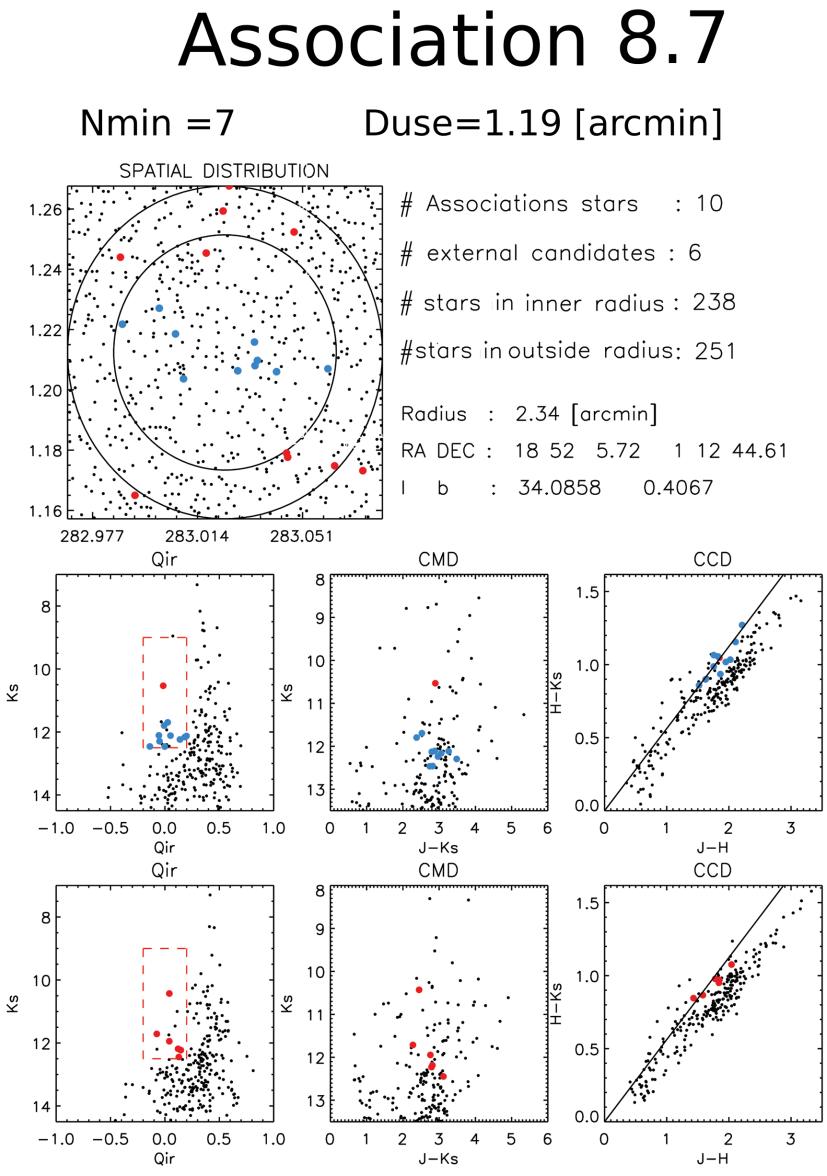

Figure 2.43: Same Figure as 2.36 but for Candidate 8.7

2.5.2 Catalog for $12.5^\circ < l < 93^\circ$

After the successful application to the small region around $l = 33^\circ, b = 0^\circ$ we decided to search in a more extended area, covering the inner Galactic region easily accessible from the Observatorio del Roque de los Muchcachos. We covered Galactic longitudes from $12.5^\circ < l < 93^\circ$ and between latitudes $-3^\circ < b < +3^\circ$. For the search we divided the total area in $6^\circ \times 6^\circ$ zones, similar to that explorer in the previous sections.

We found a total of 113 candidates in the whole area. Figure 2.44 shows the position of all candidates in Galactic coordinates. The areas on increasing stellar density (like the innermost Galactic region of the Cygnus arm direction around $l = 80^\circ$) are easily identified. The whole list of cluster candidates and their main characteristics is given in Table 2.14.

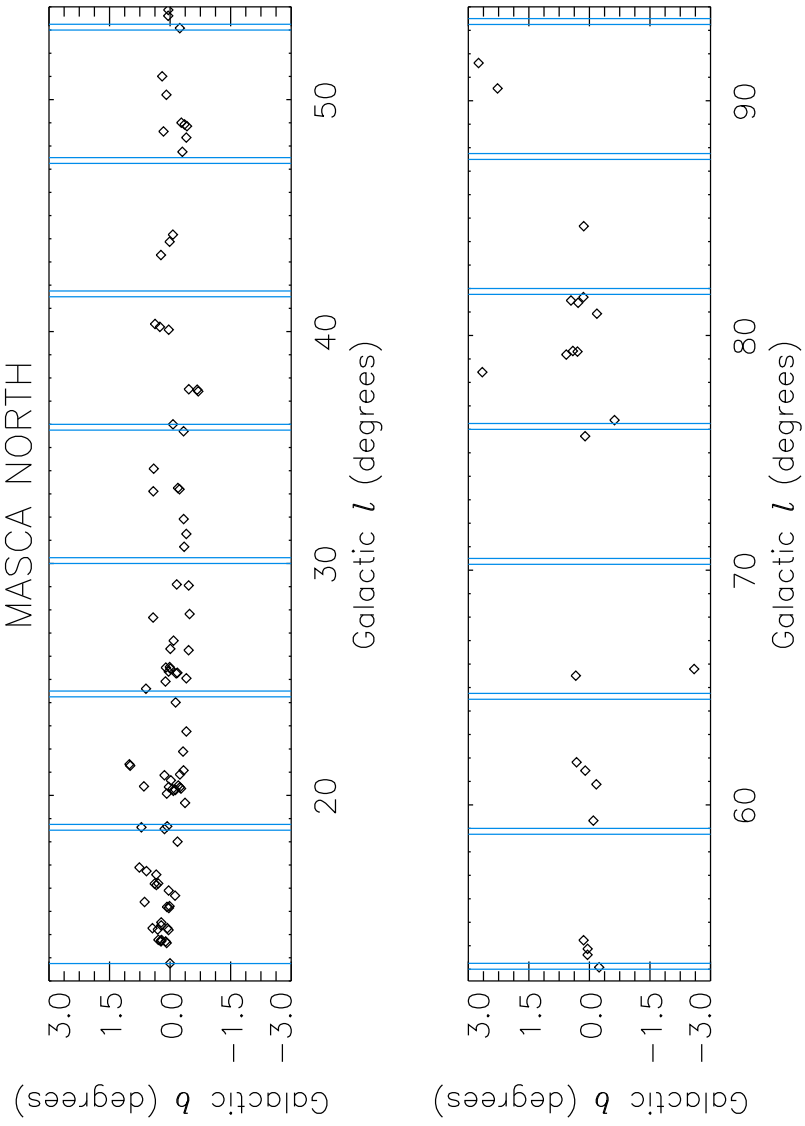

Figure 2.44: Cluster candidates from MASCA, sweeping the region from $12.5° < l < 93°$. The blue lines represent the limits of the different $6° \times 6°$ search zones, each region overlap $0.25°$ to consider possible association in the edges of the zones.

The table includes the ID number, RA and DEC (in degrees), galactic latitude, Galactic longitude and latitude, number of stars in the association candidate, number of detections for the association (an association may be detected a maximum of five times, one for each different combination of D_S, *Nmin*), and the average radius of the association. The next column gives a flag to highlight the associations that are detected in the intersection of the $6^\circ \times 6^\circ$ fields (i.e. two of the detections may be triggered by the same combination of D_S and *Nmin* when fields overlap). The last column list the relevant objects found (relative to massive stars forming region) in the radius of the association using SIMBAD[5] database.

The cross-match with the literature shows several already know clusters (22) plus a large number of uncatalogued cluster candidates. A proposal including $l > 35^\circ$ cluster candidates has already been submitted for follow-up with EMIR@GTC.

[5]http://simbad.u-strasbg.fr/simbad/

Table 2.14: List of candidates from MASCA, sweeping the regrion from $12.5° < l < 93°$. Including ID number, RA and DEC (in degrees), galactic latitude, Galactic longitude and latitude, number of stars in the association candidate, number of detections for the association, average radius of the association, a flag to highlight the associations that are detected in the intersection of the $6° \times 6°$ fields and list the relevant objects found in the radius of the association using SIMBAD database.

ID	RA	DEC	l	b	# STARS	# detections	Average Radius	flag	SIMBAD
1	273.5899048	-16.8424816	13.7741804	0.29178	12	4	0.74025	0	YSO
2	273.3478088	-17.8722248	12.7594299	0.00188	10	3	0.8830001	0	"Cl* 1813-178"
3	273.6177368	-16.9341412	13.7063999	0.22455	9	2	0.9	0	X-RAY
4	274.7150574	-14.7392817	16.1392097	0.33935	21	3	1.3446666	0	Bubble
5	274.8740234	-15.0943327	15.8989716	0.0361286	14	3	1.0653334	0	"YSO , DC"
6	273.6195374	-16.9062328	13.7317333	0.2364	9	2	0.7935	0	X-RAY
7	273.7056885	-16.3330059	14.2747831	0.43765	6	1	0.562	0	""
8	274.087616	-15.2469378	15.4047718	0.6321	11	2	1.2005	0	" IRS"
9	274.911377	-15.3641338	15.6781893	-0.1230778	9	2	1.023	0	"Several YSO and RS"
10	276.1804199	-12.634304	18.6664429	0.0727286	7	2	0.864	yes	"RS + Maser"
11	276.0650024	-12.7057934	18.5505657	0.1388417	20	4	2.2650001	yes	"Several IRS, RS, mm , YSO"
12	273.7218018	-16.9911785	13.7038441	0.1098556	11	2	1.0155001	0	""
13	274.5088196	-15.7599106	15.1460505	0.031325	14	2	1.0664999	0	"RS"
14	273.7115173	-17.060936	13.6378288	0.0851857	8	2	0.874	0	"IRS"
15	273.6596985	-16.8900528	13.7643003	0.210375	9	3	1.1003333	0	"HII,RS, IRS"
16	274.7875061	-14.7007341	16.2063427	0.2958286	16	3	1.6426667	0	"IRS"
17	274.0449219	-16.5069008	14.2770557	0.0687	9	1	1.21	0	"YSO and mm "
18	274.782074	-14.09692	16.7359409	0.58595	10	1	1.698	0	" Cl* NGC 6611 BKP 27561"
19	274.0319824	-16.5911865	14.1969891	0.0395	9	1	1.115	0	"HII, IR and YSO"
20	274.0291138	-16.21772	14.5241308	0.2197	13	1	1.6670001	0	"RS"
21	274.489563	-15.6980839	15.1917	0.0769727	11	1	1.346	0	"IRS, Bubble"
22	274.9278564	-14.3452148	16.5838795	0.34407	10	1	1.73	0	"Possible Open Cl*"
23	275.5825806	-12.3716908	18.6249218	0.7114334	9	2	1.678	yes	YSO
24	273.9632263	-16.3326607	14.3929625	0.2205625	8	1	1.186	0	YSO
25	276.0945435	-13.3381004	18.0052242	-0.18245	8	1	1.095	0	""
26	273.7969666	-16.4445667	14.2185297	0.30749	10	1	1.636	0	"RS, DN"
27	274.5625305	-15.6994629	15.2237797	0.01445	10	1	1.89	0	"RS"
28	274.6994629	-13.8764362	16.8922997	0.7608125	8	1	1.38	0	"Cl* NGC 6611 "
29	274.7007446	-14.6703587	16.1933994	0.3841375	8	1	0.778	0	"DN and YSO"
30	277.205719	-10.921773	20.6493244	-0.017425	7	4	0.34	0	"Open Gal Cluster (no spectroscopic follow up)"
31	278.5421143	-9.2314186	22.7565994	-0.4021333	6	2	0.405	0	" [MFD2010] B and MI stars"
32	277.2643127	-11.3542805	20.2929573	-0.2691	9	3	0.8710001	0	""
33	277.692749	-10.6876192	21.0785999	-0.3327333	7	2	0.897	0	""
34	277.0483704	-11.3507948	20.1977158	-0.0800167	33	4	1.4472499	0	""
35	277.0248108	-11.1509094	20.363945	0.0333556	22	3	1.6916666	0	"DC "

36	277.1147156	-11.339467	20.2379742	-0.13235	20	2	1.6315	0	"X-RAY "
37	277.0651855	-11.272501	20.2747002	-0.0582429	18	2	1.5165	0	"Several Maser, mm, X-RAY and DN"
38	276.5967102	-9.8958893	21.2788754	0.9900375	8	1	1.41	0	"Planetary nebula"
39	276.4822083	-10.8410168	20.3901749	0.649025	8	1	1.388	0	""
40	276.6092529	-9.8318777	21.3412743	1.0089124	8	1	1.939	0	""
41	277.5291138	-10.8019037	20.9028759	-0.243175	8	1	1.124	0	""
42	278.0643311	-9.9621029	21.8909492	-0.321325	8	1	1.35	0	"Bubble, IR, DN, X-RAY and DC"
43	276.8444214	-11.381793	20.0772629	0.0826	8	1	1.331	0	""
44	277.2683716	-11.2895994	20.3521004	-0.2426625	8	1	1.145	0	""
45	277.0639038	-11.9466314	19.6771851	-0.3704286	14	1	1.9680001	0	"X-RAY and several IRS"
46	278.8868713	-7.9959388	24.0102749	-0.1358125	8	1	1.13	0	""
47	277.1676331	-10.6541071	20.8690376	0.1400125	8	1	1.135	0	"YSO ?, X-RAY"
48	277.2662964	-11.2081079	20.4233131	-0.2030125	8	1	1.098	0	""
49	279.0797729	-7.080431	24.9111404	0.11548	6	2	0.2985	0	"CI* Quartet"
50	281.4987793	-3.6624582	29.0523605	-0.4605	13	4	0.74	0	"IRS"
51	280.2166443	-6.1444697	26.2608509	-0.4586667	6	1	0.754	0	""
52	279.5170593	-6.8836684	25.2852993	-0.1797143	16	2	1.4655	0	"CI* [FMR2006] RSG Cluster"
53	279.3492737	-6.7326188	25.3430004	0.0376714	21	2	2.0645001	0	"X-RAY"
54	281.255127	-3.4878144	29.0967293	-0.1643571	10	2	1.5190001	0	"HII, RS and DC "
55	279.6069031	-7.1961899	25.048542	-0.4023143	7	1	1.169	0	"WR star "
56	279.3656311	-6.5524435	25.5105381	0.1059375	11	2	1.374	0	"DC, mm and X-RAY"
57	279.480072	-6.8905578	25.2623215	-0.1502556	16	2	1.6340001	0	"CI* [FMR2006] RSG Cluster"
58	279.4616089	-6.577096	25.5324001	0.0099429	10	2	1.3585	0	"DN"
59	278.5033569	-7.1354976	24.5989666	0.5979111	9	1	2.0569999	0	"IR, SNR"
60	280.0723267	-5.6084614	26.6716499	-0.0856	12	1	1.798	0	"DNe, Rad,X,IR"
61	280.0791321	-4.4904532	27.6686115	0.4204222	9	1	0.962	0	"Emission Object, Masier"
62	279.8372498	-5.8858652	26.3179054	-0.0048062	16	1	2.6719999	0	"Several HII, RS and mm, star M3.5I present."
63	280.9539795	-4.7675219	27.8211117	-0.481725	8	1	1.273	0	""
64	279.444458	-6.6435208	25.4655533	-0.0054077	13	1	1.971	0	"X-RAY "
65	282.5707092	0.3503717	33.1120796	0.4167667	18	4	1.187	0	MASGOMAS-1
66	283.1907043	0.1329911	33.201313	-0.234125	16	3	1.8993334	0	MASGOMAS-10
67	282.1607056	-2.1278005	30.7196503	-0.34855	8	3	0.9683333	0	"mm , bubble and YSO"
68	282.4598389	-1.6602094	31.2721252	-0.401425	8	1	1.946	0	"IRS and mm "
69	283.1772766	0.1969744	33.2521286	-0.19301	10	1	1.6339999	0	"Variable star"
70	284.4199829	2.3128519	35.7017784	-0.3331778	9	1	1.408	0	""
71	282.6937866	-1.0584205	31.9143753	-0.33515	8	1	2.0380001	0	"Bubble, IR, RS and PN"
72	283.0238342	1.2123905	34.0858917	0.40654	10	1	2.207	0	"MC, HII, YSO"
73	285.3644104	3.8640621	37.5125732	-0.462975	4	1	0.231	0	"2 WR stars, Cluster of Stars, no spectroscopic follow up"
74	285.5316162	3.6774478	37.4229851	-0.6967143	20	3	2.4436667	0	"Dark nebula"
75	284.3233337	2.6962359	35.9987984	-0.0722286	10	2	1.9235001	yes	"RS, YSO, Gamma emission and DN"

76	286.1018677	6.3733716	40.0795441	0.0338444	9	1	3.1270001	0	"Several IRS"
77	285.961731	6.5827851	40.2017632	0.2536375	8	1	1.748	0	""
78	285.9127502	6.7485123	40.3267517	0.37285	8	1	3.3870001	0	"IRS"
79	285.5413208	3.7515996	37.4933548	-0.6713889	9	1	3.013	0	"Bubble and YSO"
80	288.1020508	9.9628372	44.1773796	-0.06902	30	5	1.9209999	0	"Cl* Canis Majoris stars in clusters"
81	287.4230652	9.3203592	43.2981987	0.22755	9	3	1.2183334	0	""
82	287.8855286	9.7301931	43.8723984	0.0123875	8	1	2.1070001	0	"Several RS"
83	290.6515503	13.9365396	48.8558197	-0.41764	6	2	0.897	0	"Several RS, YSO and DN"
84	290.8463135	15.3690748	50.207634	0.0924833	6	1	1.323	0	""
85	290.0117493	14.0088673	48.6275291	0.1642714	9	2	2.2065001	0	"DN, DC and YSO"
86	290.3981934	13.5099936	48.3635864	-0.4015	7	1	1.732	0	""
87	290.6328735	14.0295553	48.929287	-0.3577572	7	1	1.6210001	0	"Several DM, DC and YSO"
88	291.1424561	16.1220303	51.0058441	0.1975556	9	1	3.131	0	"Several DN"
89	290.0123596	13.0138283	47.7491493	-0.3032	10	1	5.3369999	0	"Several IRS and DN , YSO"
90	290.5976562	14.1334143	49.0047798	-0.27868	10	1	3.53	0	" CL* [BDS2003] 139cluster, no spectroscopic follow up"
91	292.5821228	18.3426304	53.6133766	0.0477875	14	4	3.7780001	0	""
92	292.8071899	18.9338284	54.2342987	0.1449333	6	1	3.0350001	0	""
93	292.7119751	18.5640182	53.8667107	0.0461	10	1	5.8049998	0	"Several YSO"
94	292.5791931	17.7375565	53.0812302	-0.24029	10	1	4.8790002	yes	"Several YSO and Bubble"
95	296.6743469	25.6226692	61.8167992	0.31908	5	1	2.5788	0	"Several YSO, DN and a Pulsar"
96	296.681366	25.2063293	61.4601517	0.104167	6	1	3.2509	0	"Cl* SH 2-88B IR Cluster"
97	296.6195984	24.5647831	60.8774719	-0.170057	12	2	7.2051497	0	"Cl* [BDS2003] 158, few spectroscopic follow-up"
98	295.7034912	23.2610798	59.3297157	-0.096414	7	1	3.4354	0	"Cl* NGC 6823, few spectroscopic follow-up"
99	301.7663574	27.4986305	65.7891998	-2.5942199	6	3	2.9666665	0	"Cl* [BDB2003] G065.79-02.6 NO SPECT"
100	298.7716675	28.8043919	65.504715	0.3375	8	3	7.6019998	0	""
101	305.6244202	37.248703	75.7085342	0.106737	8	1	10.7445002	0	"RS"
102	309.9096375	41.2900085	80.927475	-0.1847	12	5	2.2574	0	"Cl* DR 22 , only one spectra"
103	308.1183472	40.2879372	79.3111801	0.29596	13	5	1.7462	0	"Cl* DR 15 IR Cluster, few spectroscopic follow-up"
104	306.8679199	37.3869858	76.3918991	-0.62136	6	3	0.9226667	0	"Crowd of YSO and IRS"
105	304.9186401	40.9404907	78.4370499	2.6542001	4	1	0.331	0	"Crowd HII and IRS, a B0.5 star present"
106	308.0185852	40.3774223	79.3381653	0.4102778	12	2	4.0900002	0	"Several Emission lines, Herbig objects, YSO and snm "
107	310.1319275	42.0526428	81.6327744	0.1496143	13	2	3.4505	0	"Several RS, snm and YSO"
108	307.7270813	40.3506699	79.1853104	0.5735818	11	1	6.2090001	0	"Several IRS, YSO, mm and Emission line objects"
109	309.7944336	41.9414253	81.3915863	0.28045	8	1	4.0619998	0	"Cl* [BKH2010] G81.42+0.30"

110	309.6858826	42.1269684	81.4898682	0.45731	10	1	4.9099998	0	"Crowd of YSO, IRS and mm "
111	312.7286072	44.4095612	84.6547394	0.14118	10	3	2.1886666	0	"Cl* [BDB2003] G084.65+00.16"
112	16.5671692 5	1.3739243	91.6072998	2.7465999	5	2	0.9548	0	"IR and DN "
113	15.976593 5	0.2556305	90.5252533	2.2763669	8	5	3.17958	0	"Several Herbig star and crowd of YSO and Emission lines objects"

SNR: Super Nova Remanent
YSO: Young Stellar Object
Cl*: Cluster
X-RAY: X-ray source
Bubble: Bubble of gas or dust
DC: Dark nebula
IRS: Infra-Red Source
RS: Radio Source
mm: Milimetric source
smm: Sub-milimetric source
HII: HII region
MC: Molecular Cloud

2.6 Conclusions

In this chapter we have explained the motivations and construction of the MASGOMAS automatic search of clusters algorithm (MASCA). We described the several stages of MASCA and we can summarize the particular points of the chapter:

- When we search for massive star cluster candidates we find a large number of field population contaminants. These objects were efficiently excluded by improvements in the color cuts.
- We have investigated the detectability of overdensities in a random distribution of stars, which results in an optimization criteria for our choice of D_S and *Nmin*
- We have developed the bona-fide algorithm MASCA which allows us to automatically locate hidden clusters/associations candidates in the Milky Way.
- This algorithm was tested using 2MASS sky survey catalog, but can be used with any other near-infrared catalog.
- We have created the first bona-fide catalog of massive cluster candidates in the region of $12.5^\circ < l < 93^\circ$.

3

Two new MASGOMAS clusters

Contents

In the previous chapter we saw how to search for massive cluster candidates from a photometric survey using photometric cuts to exclude the late-old stellar population. We also described how we looked for overdensities ensuring that the candidates are related in an association.

In this chapter, we present spectroscopic follow-up, using LIRISWHT for two massive cluster candidates found by our algorithm MASCA. The first one, MASGOMAS-6 was found using a preliminary version of MASCA, described in section 2.3. The second one, MASGOMAS-10 was found using the lastest version of MASCA in the direction of MASGOMAS-1 [$l = 33° \pm 3°, b = 0° \pm 3°$].

3.1 MASGOMAS-6

3.1.1 Candidate selection

Following the project that led to the discovery of MASGOMAS-4 (Ramírez Alegría et al. 2014) found in the region around galactic coordinates $l = 40^\circ$ & $b = 0^\circ$, we expanded the search to a larger region of sky. This region is centered in the direction of the intersection of the Sectum arm and the bulge, between MASGOMAS-4 and MASGOMAS-1 ($l = 38^\circ \pm 3^\circ, b = 0^\circ \pm 3^\circ$).

Using the first version of MASCA, with square photometric cuts $K \epsilon [6, 12.7]$, $(J-K_S) \epsilon [1, 6]$ and $Q_{IR} \epsilon [-0.3, 0.3]$, with a ratio of color excesses of 1.78 ± 0.3 over the 2MASS catalogue survey, we discovered nine cluster candidates. The most prominent of these candidates, was MASGOMAS-6, which was selected for spectrophotometric follow up. Table 3.1 summarises the parameters obtained by MASCA.

Table 3.1: Candidate association MASGOMAS-6 and the parameters obtained by the preliminar version of MASCA.

ID	RA	DEC	l	b	# stars	Radius [']	Nmin	D_S
MASGOMAS–6	19:00:11.44	+03:46:16.7	37.28	-0.23	8	1.05	10	1.00

Figure 3.1 shows the output provided by the preliminary version of MASCA. We can observe the spatial distribution of the OB-type candidates and the photometric properties of the cluster candidate compared with all the remaining stars in the field. In this version, the program includes the surrounding stars which fulfill the photometrics cuts, in two radius around the association candidate (blue dots).

To prepare the spectroscopic follow-up we used UKIDSS photometry (Lawrence et al. 2007), which provides better spatial resolution than 2MASS, allowing us to improve the selection with photometric cuts. For OB-type candidates selection, we used $Q_{IR} \epsilon [-0.2, 0.2]$ plus the previous photometric cuts. The new selected candidates were divided in three groups: the blue group, with $(J - K) < 3$ (labeled as B); candidates with $(J - K) > 3$ formed the red group (labeled as R); and the supergiant candidates with $K_S < 10$ (labeled as LS; as they were observed with long-slit spectroscopic mode) formed the third group. Figure 3.2 shows all the stellar candidates which fulfill the photometric cuts using UKIDSS photometry, over GLIMPSE (Galactic Legacy Infrared Mid-Plane Survey Extraordinaire) false-color.

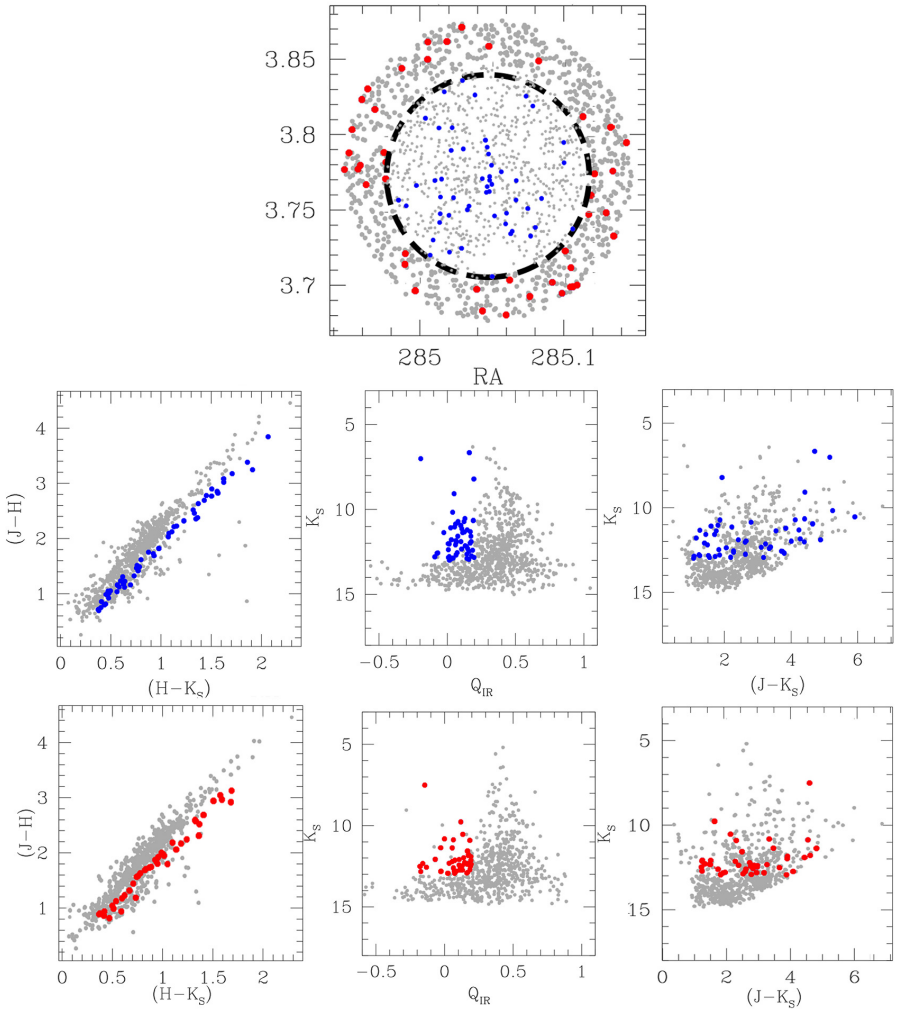

Figure 3.1: Spatial distribution, color-color, pseudocolor-magnitude and color-magnitude diagrams for the cluster candidate MASGOMAS–6, and its control field. The stars selected for follow-up are shown with blue (candidate field) and red (control field) symbols. The associated CCD, Q_{IR} vs K_S and CMD are displayed for the association and control field respectively.

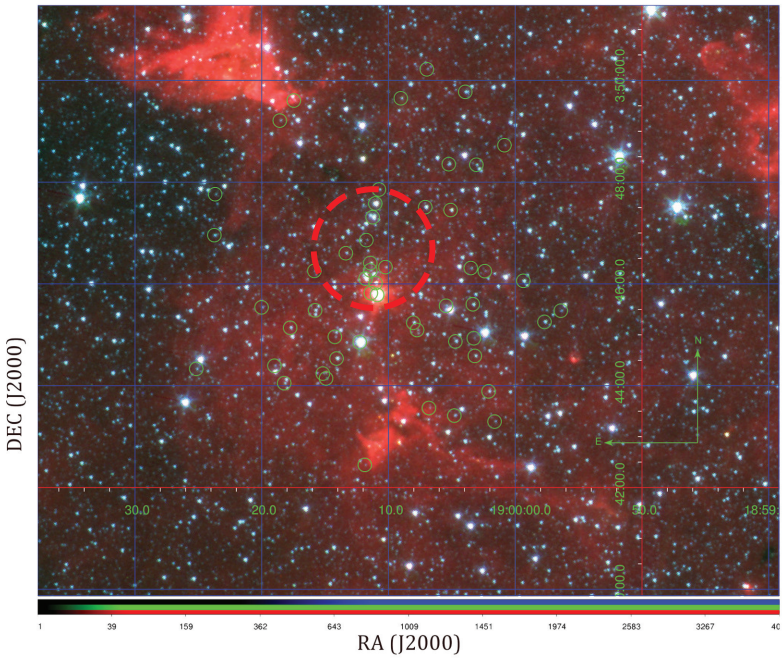

Figure 3.2: GLIMPSE false color (blue=IRAC I1, green=IRAC I2, red=IRAC I4) image for MASGOMAS-6. Using green circles we mark the stars which fulfill the photometric cuts using UKIDSS photometry. From MASCA we obtain a candidate radius of 1 arcmin (red dashed circle). The size of the image is 14.5×11.5 arcmin^2.

3.1.2 Mask design

For the spectroscopic follow-up we decided to observe the OB-type candidates using two multi-object masks, and to observe the brighter supergiant candidates stars using long-slit.

Due to the large number of candidate stars and the limited time, we focused on the B-group stars which have higher chance to be part of the obscured massive stellar population. Given the LIRIS field of view of $4.27' \times 4.27'$, we decided to concentrate on the center of the cluster candidate with two masks.

Using the program *liris_mask_design*, written in IDL, it is possible to optimize the mask design. Each of the slits is required to be between 8.5 and 10 arcsec long and 0.85 arcsec wide for a resolving power R= ($\lambda/\delta\lambda \sim 2500$). Thus the maximum number of stars per mask without overlap is 25. We must consider also at least three reference stars for mask alignment and pointing. In our design we included one mask with 12 targets and one with 13. The stars in each mask have a magnitude difference (difference between the brightest and faintest stars) smaller than 2.5 mag.

An important consideration is the wavelength range covered. In the central mask position, the wavelenght coverage is 1.55 to 1.85 μm in the H-band and from 2.06 to 2.40 μm in the K-band. These ranges vary according to the position of the star along spatial axis of the detector. This is crucial when deciding which star may be included in the mask: stars closer to the edge of the mask may not include in their final spectra the lines used for OB star classification, (HeI 1.70 μm, HeI 2.11 μm, HeII 2.57 μm, HeII 1.69 μm and $Br\gamma$). The task *liris_mask_design* takes all these considerations into account. An example of a final configuration (i.e the maximum numbers of stars close to the middle of the detector) is shown in Figure 3.3.

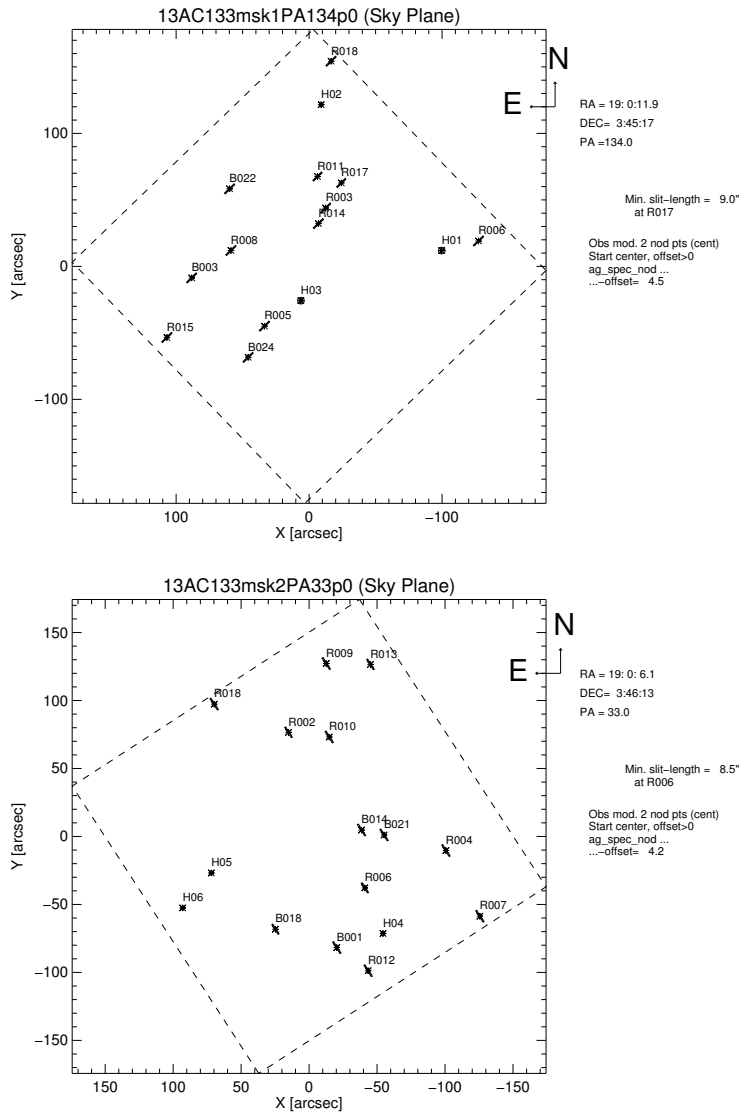

Figure 3.3: Final design of the multi-object masks. H labels correspond to the stars used for mask alignment, the *B* and *R* labels to the priority groups. On top of each mask design we included each mask internal name, assigned by the *liris_mask_design* task.

Figure 3.4 shows the distribution of a stars including in both masks, on the field. The area covered in the follow-up is larger than the candidate radius estimated by MASCA.

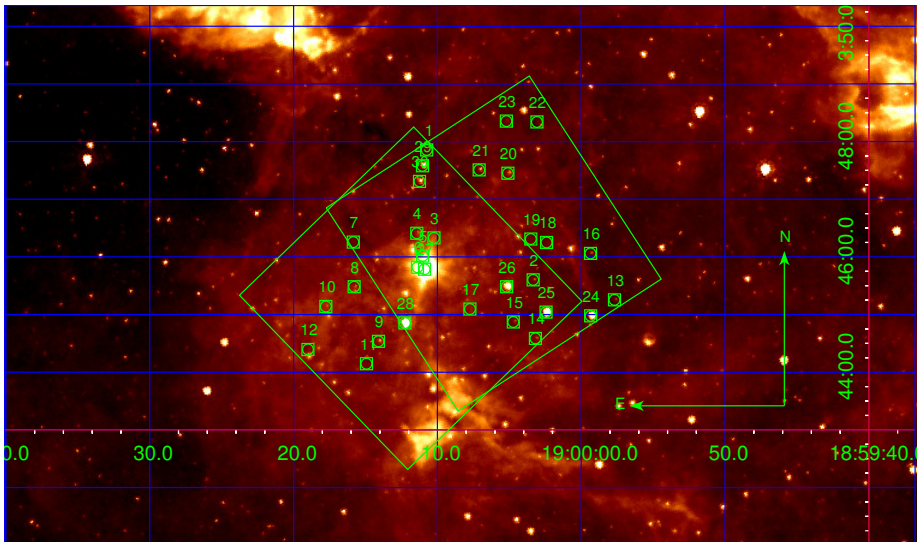

Figure 3.4: Global view of the observation using both masks over the 8μm IRAC image. The big green squares represent the field of view of LIRIS ($4.27' \times 4.27'$ arcmin^2) and the small green squares represent the stars with the identification used in Table 3.2.

liris_mask_design may simulate the position of a specific spectral line, determined by the user, to check the wavelength coverage in the final spectra. In figure 3.5 we show the simulated wavelength range for both mask, and two masks at 21,000 Å and 23,000 Å are including for guiding.

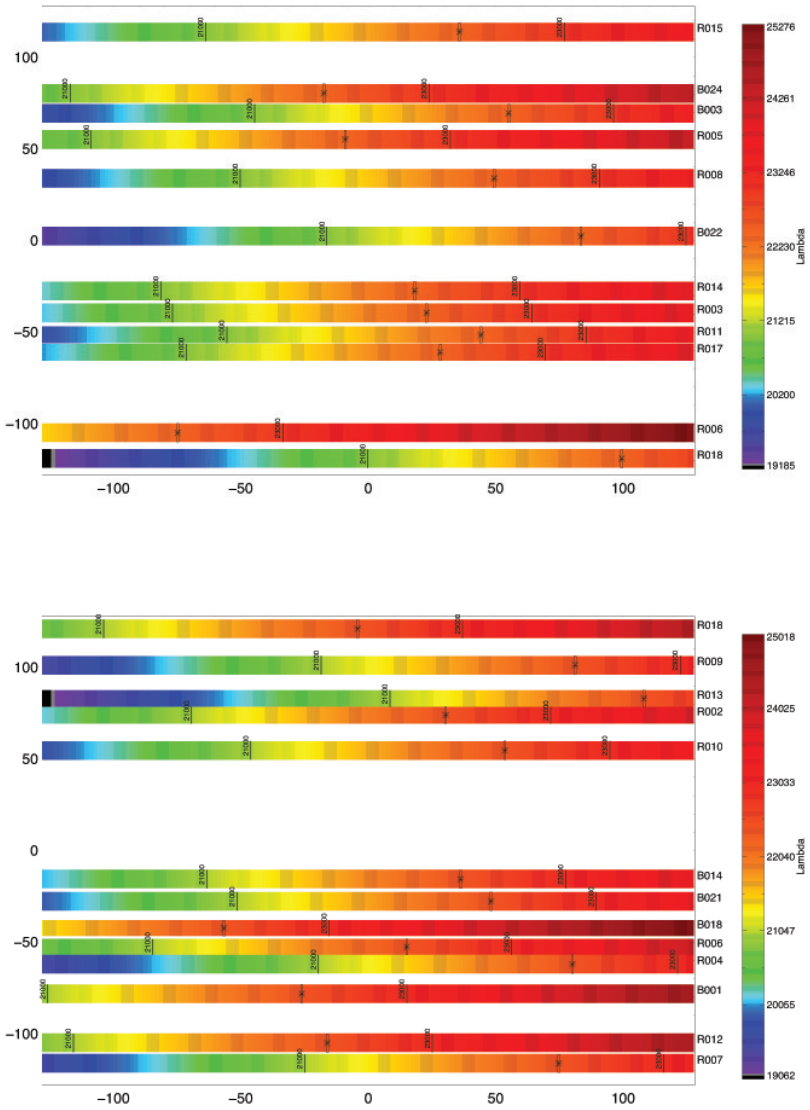

Figure 3.5: Wavelength coverage for each star in mask1 (top) and mask2 (bottom), using the K-band as reference. Each star position is displayed as an asterisk within their slit. The numbers in both axes give the position on the detector in pixels. The right color legend provide the wavelenght coverage in Å.

3.1.3 Observations and data reduction.

All data were acquired with LIRISWHT using the multi-object (MOS) and long-slit spectroscopic modes with the HK pseudogrism. For long-slit observations the candidates were observed together to obtain similar S/N using the slit of 1.00 [arcsec].

Observations with mask and long-slit were both performed following an ABBA telescope nodding pattern. This procedure places the star in two positions along the slit, separated by 3" for the mask and 12" for the long-slit. This nodding pattern is used to remove the sky emission lines from the science spectra. We observed each mask using 30 minutes of integration time, in each spectral band (H and K). The stars selected for long-slit spectroscopy were observed during 15 minutes of integration time. All data were obtained on 2013 June 29 and 30.

The near-infrared spectroscopic reduction was performed within the LIRIS IRAF package LIRISDR same as the previous chapter, plus the *ltrimspec* and *ltrimspec* tasks (the MOS mode). The were extracted using *apall*.

The telluric correction was done using MOLECFIT (Smette et al. 2015; Kausch et al. 2015), a software that corrects the atmospheric absorption lines by fitting a synthetic spectra estimated from the column density of atmospheric molecules and the physical parameters during the observation. MOLECFIT can be run in two modes: command-line or through a GUI. The advantage of using the GUI mode is to include a parameter file, which can be edited by the user using the information found in the file headers and from the observatory site.

On the parameter file we can find three types of parameters: ambient (observing date, UTC, altitude angle of telescope, relative humidity, pressure, ambient and primary mirror temperature, elevation above sea level, longitude and latitude), instrumental (slit width and pixel scale), and atmospheric parameters (reference atmospheric profiles, grid of layer heights, upper mixing height, and precipitable water vapour). Additionally, a selection of atmospheric molecules (water, carbon dioxide, methane, nitrous oxide, and ozone) can be specified in the parameter file.

MOLECFIT reads the science spectrum and the parameter files, and compiles an atmospheric profile. This profile[1], together with the target airmass, are used as input to the radiative transfer code (line-by-line radiative transfer model, LBLRTM v12.2; Clough et al. 2005). Flux scaling, wavelength grid correction, and convolution are applied to match the observad spectrum. The final atmospheric model allows telluric reduction with similar quality than XTELLCOR (Vacca et al. 2003), an IDL program traditionally used for telluric correction in our group (Ramírez Alegría et al. 2012, 2014). Because of this, and after testing both programs on a reduced sample of observed spectra, we decided to correct the atmospheric absorption telluric lines for MASGOAMS-6 using MOLECFIT.

The resulting frames were wavelength calibrated using Argon and Xenon lamps. By fitting a polynomial function we eliminate the continuum contribution in the lamp image to obtain with clarity the lamp spectrum to enhance the lamp emission lines in both H and K bands.

[1]The profile is created from standard atmospheric profiles, measures by the ESO Meteo Monitor (EMM), and data at any point on Earth from the Global Data Assimilation System (GDAS), created by the Air Resources Laboratory (ARL) of the National Oceanic and Atmospheric Administration (NOAA).

3.1.4 Spectral classification

We first divided the observed of spectra in two groups, one with stars that present the CO bands and another that do not. The absence of CO bands implies early type stars. The classification was based on the Hanson et al. (1996, 2005) and Ivanov (2004) atlases. The late type stars, which present the characteristic CO bands, were classified using the Meyer et al. (1998) atlas.

Using the hydrogen Brackett series (H), and helium (HeI and HeII) lines it is possible to distinguish the O, B and A spectral types and also assign the luminosity class for each star by comparing by eye with the atlas spectra. For the second group we used the metallic lines (i.e., CIV, FeII, AlI and MgI) plus the intensity of the CO bands to determine the luminosity class, again comparing by eye with the atlases.

For the assigned spectral types we assume an error of ±2 subtypes, similar to Hanson et al. (2010), Negueruela et al. (2010) and Ramírez Alegría et al. (2012, 2014). Table 2.5 summarizes the observed stars and their classification.

Table 3.2 lists the spectroscopically observed candidates, either in MOS or long-slit mode. They are ordered by an ID number, which corresponds to the order of the reduction from top to bottom for the mask 1 and then consecutively for mask two. After that, we show the stars observed in long-slit. Figure 3.6 and 3.7 illustrate the spectra from the masks. The position of the star in the mask determines the observed wavelength range.

Table 3.2: Parameters for spectroscopically observed candidates in MASGOMAS-6 field. The table describe the ID of the star, the method of observation (LS correspond to the long-slit), photometric data (J, H and K) from UKIDSS, distance and redderning derived from the intrinsic colors asigned by spectral classification.

resizebox15cm!

ID	Obs method	RA (J2000.0)	DEC (J2000.0)	J	H	K	Distance [kpc]	A_K	Spectral type
1	mask1	19 00 10.75	+03 47 51.2	16.909	14.174	12.639	$6.77^{+0.92}_{-0.78}$	2.96	O9III
2	mask1	19 00 03.34	+03 45 36.1	15.052	12.532	11.213	$6.66^{+1.29}_{-0.13}$	2.61	O9I
3	mask1	19 00 10.23	+03 46 19.7	16.403	13.933	12.483	$7.42^{+1.09}_{-0.60}$	2.74	O8III
4	mask1	19 00 11.43	+03 46 24.5	15.986	13.440	11.952	$8.50^{+0.22}_{-0.61}$	2.81	O6I
5	mask1	19 00 11.02	+03 46 00.9	14.941	12.262	10.698	$4.54^{+0.04}_{-1.08}$	2.93	O95I
6	mask1	19 00 11.38	+03 45 49.1	15.517	13.652	12.643	$3.07^{+0.44}_{-0.23}$	1.53	G8III
7	mask1	19 00 15.84	+03 46 15.3	14.530	13.648	12.986	$4.82^{+0.42}_{-0.35}$	0.68	G5III
8	mask1	19 00 15.78	+03 45 28.9	14.855	12.706	11.360	$2.87^{+1.10}_{-0.97}$	1.77	K3III
9	mask1	19 00 14.09	+03 44 32.1	15.568	12.653	10.996	$4.71^{+0.04}_{-1.42}$	3.15	O95I
10	mask1	19 00 17.75	+03 45 08.3	12.764	11.713	11.072	$4.17^{+1.22}_{-1.24}$	1.25	O9V
11	mask1	19 00 14.92	+03 44 08.6	14.348	13.672	13.006	$4.53^{+0.45}_{-2.08}$	0.89	B3V
12	mask1	19 00 18.99	+03 44 23.5	15.363	13.446	12.449	$4.35^{+1.14}_{-1.03}$	1.44	K2III
13	mask2	18 59 57.69	+03 45 15.3	15.394	12.754	11.184	$3.69^{+0.50}_{-0.43}$	2.88	O9III
14	mask2	19 00 03.18	+03 44 35.1	16.415	13.589	12.023	$--$	–	–
15	mask2	19 00 04.72	+03 44 52.2	12.506	11.413	10.738	$6.08^{+0.82}_{-0.70}$	1.30	O9III
16	mask2	18 59 59.36	+03 46 03.6	14.736	12.178	10.752	$8.54^{+1.54}_{-0.85}$	2.59	A0I
17	mask2	19 00 07.74	+03 45 05.7	14.733	13.372	12.569	$3.30^{+0.28}_{-0.24}$	1.09	G5III
18	mask2	19 00 02.40	+03 46 15.0	14.055	13.340	12.899	$4.21^{+3.82}_{-0.44}$	0.82	B3V
19	mask2	19 00 03.50	+03 46 18.6	13.428	12.599	12.089	$2.72^{+2.46}_{-0.29}$	0.95	B3V
20	mask2	19 00 05.09	+03 47 27.0	16.462	13.641	11.825	$4.77^{+2.72}_{-1.17}$	2.83	WN6
21	mask2	19 00 07.09	+03 47 30.5	16.343	12.628	10.601	$2.03^{+0.20}_{-0.17}$	3.13	K8III
22	mask2	19 00 03.06	+03 48 20.5	15.541	13.351	12.150	$4.10^{+1.20}_{-1.22}$	2.37	O9V
23	mask2	19 00 05.24	+03 48 21.1	16.347	13.826	12.393	$3.90^{+1.15}_{-1.16}$	2.73	O9V
24	mask2	18 59 59.35	+03 44 58.6	9.901	8.077	7.000	$2.71^{+1.20}_{-1.06}$	1.1	M5III
25	mask2	19 00 02.45	+03 45 02.5	9.600	7.482	6.447	$1.41^{+0.72}_{-0.23}$	1.30	M1III
26	mask2	19 00 05.19	+03 45 28.9	9.984	8.148	7.076	$2.8^{+1.24}_{-1.10}$	1.1	M5III
27	LS1_1	19 00 10.89	+03 45 47.1	11.953	9.1152	7.0585	10.0	–	B8I
28	LS1_2	19 00 12.26	+03 44 51.2	11.165	8.4688	6.7070	$10.10^{+0.04}_{-0.04}$	2.18	M5I
29	LS2_1	19 00 11.04	+03 47 34.7	13.791	10.806	9.1231	$4.40^{+1.30}_{-2.26}$	3.14	WN8-9h
30	LS2_2	19 00 11.25	+03 47 18.6	10.054	8.8926	8.2312	$0.88^{+0.61}_{-1.17}$	0.71	K2III

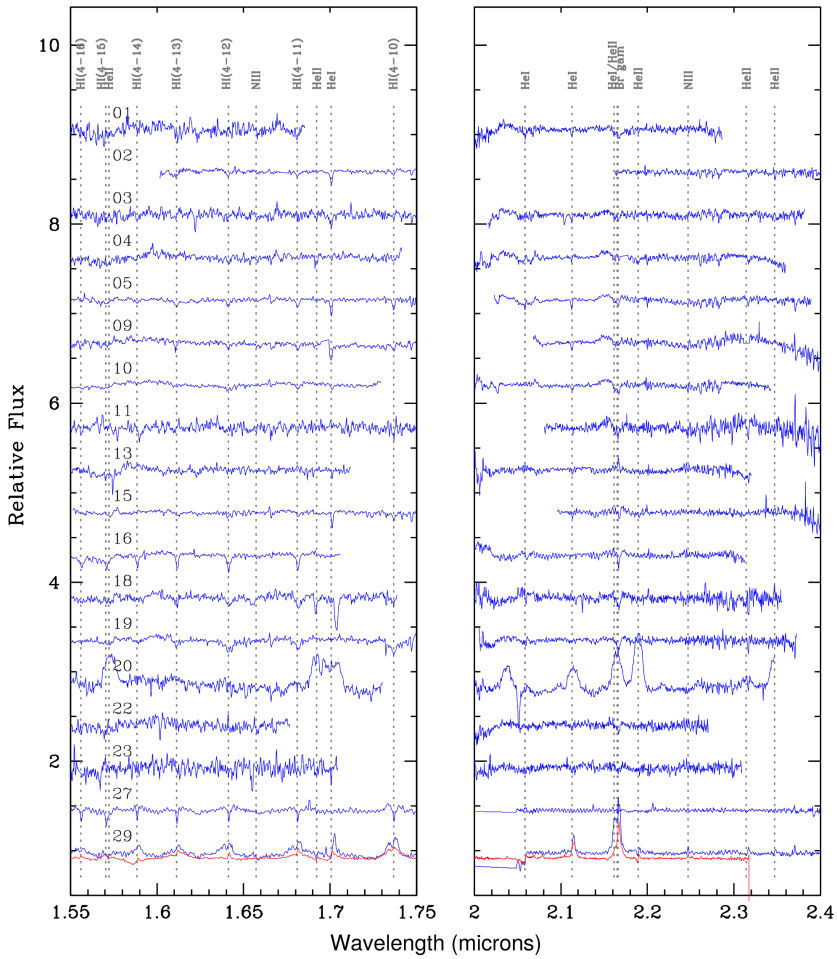

Figure 3.6: Individual H-band (left) and K-band spectra (right) for MASGOMAS-6 early-type-stars. The bottom red spectrum plotted over spectrum 29 corresponds to WR62-2, a WN8-9h object (Chené et al. 2015) shown for visual comparison.

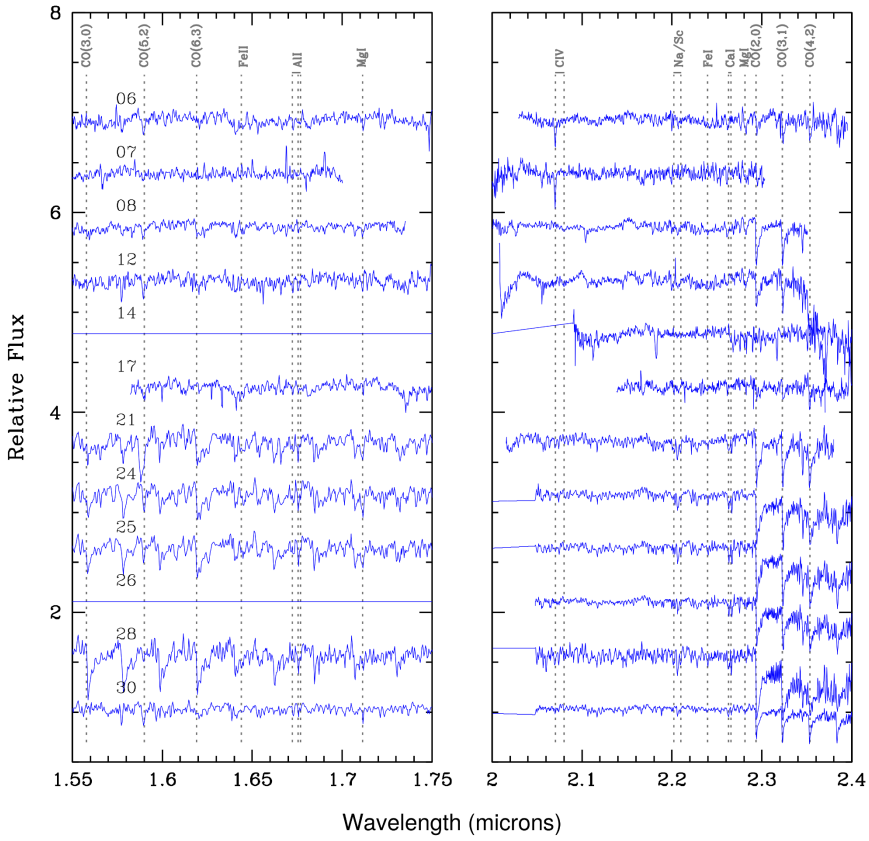

Figure 3.7: Individual H-band (left) and K-band spectra (right) for MASGOMAS-6 late-type stars.

Stars number 10, 22, and 23 are classified as O9 V, and stars number 11, 18 and 19 are classified as B3 V according to the shape of its Brackett series and the absence of helium lines.

Stars number 1, 3, 13 and 15 are classified as early-type giants. The earliest giant star in this group is number 3 (O8 III). The spectral features have similar depth and shape as those observed in stars HD 37043 (O9 III; Hanson et al. 1996), HD 36861 (O8 III; Hanson et al. 1996), and HR 1899 (O9 III; Meyer et al. 1998). For the spectra in this group the classification was based on the presence and shape of the He I line at 1.70 μm and the Brackett series. In the case of star 13, the H-band spectrum is too noisy to identify spectral features. For this star, the classification was only based on the K-band spectra.

The group of supergiants observed with the multi-object mask includes both blue (stars number 2, 4, 5, 9, 15) and yellow (star number 16) supergiants. The blue supergiants are mostly late O-type stars (O9-O9.5 I), and characterized by a narrower Brackett series and deeper He I lines at 1.70 and 2.11 μm, compared with dwarfs and giants spectra. Star number 4 has a spectral type earlier than O6 I. Its spectrum shows a HeII 1.69 μm to HeI 1.70 μm ratio similar to that observed for an O5 I (for example, Cyg OB2 8C; Hanson et al. 2005). But He I at 2.11 μm and N III at 2.115 μm do not completely fit this spectral type (He I is too strong and the N III emission, too weak). The HeI/NIII ratio fits better an O6.5 I spectral type. The central Brackett emission may be nebular. Star number 16 shows a narrow Brackett series both in H and K-band, which is characteristic of A-supergiants.

Star number 20 shows broad emission lines and corresponds to the Wolf-Rayet WR1583-B73, first reported by Faherty et al. 2014. We can clearly observe in our spectra the He I and He II emission lines in both H and K-band spectra, without a clear sign of the Brackett series in emission. All this features allow us to confirm the WN6 classification for this object, by Faherty et al.

The long-slit spectra include very bright stars, RSG candidates and bright OB-type candidates which could not be included in the mask. From this group of 7 spectra we classify 3 as massive evolved objects (stars 27, 28, and 29).

Star 25 shows a deeper CO band at 1.62 μm, compared with stars 24 and 26. The Al II, Ca I, and Mg I lines fit a supergiant (we used as example, star M1 I-HD14404 from IRTF[2] library) and a giant luminosity class (for example, star M5 III HD175865). For this star, the equivalent width of the $^{1}2$CO (2,0) band in the region between 2.294 and 2.304 μm is EW_{25} = 23.29 Å. In the relation given by Davies et al. (2007), this star could be classified as mid-M giant or early-M supergiant. Because we can not clearly estimate the luminosity class for this star, we prefered to discard it from the cluster analysis.

The spectrum of star 27 shows the Brackett series with a central narrow absorption profile mixed with a broader emission profile. The depth of the absorption line component is similar to that observed in B8 supergiants (for example, HR1713), but the emission component could affect the natural depth of the series. We understand that the emission component is linked with the bright emission observed in mid-infrared. We also observed in emission two Na I lines at 2.206 and 2.209 μm, and one Mg II at 2.137 μm.

This object, at the center of our field, is the LBV candidate IRAS 18576+0341, a.k.a. AFGL

[2]Collection of 0.8-5.0 m mostly stellar spectra observed at a resolving power of R = $\lambda/\Delta\lambda \sim 2000$ with the medium-resolution spectrograph, SpeX, at the NASA Infrared Telescope Facility (IRTF) on Mauna Kea.

2298 (Ueta et al. 2001; Pasquali & Comerón 2002; Clark et al. 2003). Based on the dust infrared emission, Ueta et al. place the star at about 10 kpc and, using this distance, Clark et al. (2003) classify it as a cool, extremely luminous blue B supergiant, which agrees with our classification. Model atmosphere analysis allowed Clark et al. (2009) to determine a change in its bolometric luminosity of a factor of 2 during the period 1998-2009. We classify this star as B8 I and, because of the emission contaminating the photometry, we prefer to discard this star from the distance and extinction estimations.

Star 28 has clear and deep CO-band in both the H and K spectra. The depth of both CO-bands implies a supergiant luminosity class. Comparison with IRTF spectra indicates that star 28 is later than M5 I. The equivalent width of the ^{12}CO (2,0) band is EW_{28} = 34.66 Å, in the region between 2.294 and 2.304 μm. The EW and the shape of the spectrum point to a spectral type of M6 I for this star.

The spectra of star number 29 spectrum has a broad Brackett series in emission. The He I (at 1.70, 2.11, and 2.16 μm) and N III (at 2.11 and 2.25 μm) lines are evident in emission. The He II can be observed in absorption clearly at 1.70 and 2.19 μm (this last one present a P-Cygni profile). The spectrum is similar to a WN object (O4- 6If+/WN9 VVVCL 73-2 by Chené et al. 2013, WN8-9h WR 62- 2 by Chené et al. 2013 or F2 and F7 by Martins et al. 2008), without a narrow C IV emission line detected in our spectrum. We classify star 29 as WN8-9h.

The contaminants in this group are characterized by CO bands in both the H and K-spectra. But the depth and shape of the CO-bands, and the observed metallic lines, discard them as supergiant objects. The shape of the CO-band at 1.619μm of star 24 resembles a M-late giant spectrum (for example, the M5 III HD175865 from IRTF library). In the case of star 26, we only have a K-band spectrum. The shape and depth of its CO-band and the Ca I lines indicate a late M-type giant (same as star 24). The spectrum of star number 30 also shows $^{12}CO\Delta v = 3$ and $^{12}CO\Delta v = 2$ bands, but with depths similar to a giant luminosity class. We assigned for this star the spectral type K2 III.

3.1.5 Discussion

To estimate the individual distances for the stars we compared the apparent and intrinsic magnitudes corresponding to each spectral type. We assumed the Rieke & Lebofsky (1985) extinction law, with R_V = 3.09. The intrinsic magnitudes and colors for O-type stars, for all luminosity classes, are from Martins & Plez (2006). For stars later than O9.5 V, we used the intrinsic magnitudes and colors from Cox (2000). Distance errors are dominated by the spectral type uncertainty, and we estimated them by deriving the individual distance assuming ± 2 spectral subtypes.

For the Wolf-Rayet objects, we used the absolute magnitude calibration by Rosslowe & Crowther (2015). From a sample of 126 Wolf-Rayet with known distances, the authors derive near-infrared JHK absolute magnitudes for nitrogen, carbon and oxygen-type WRs.

In Table 3.2 we present the individual extinction and distance determinations for the spectroscopically observed stars. We distinguish two groups of individual distances for the massive stars: one close group of stars (with distances between 4 and 7 kpc) and one distant group (with individual distances larger than 8 kpc). Figure 3.8 illustrates this situation.

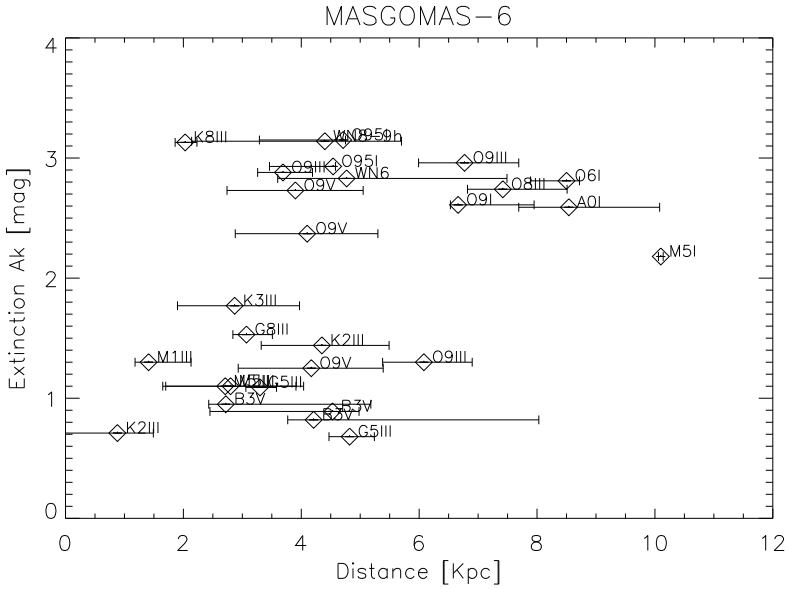

Figure 3.8: Individual distance and extinction estimates for spectroscopically observed stars.

The nearby group includes all the observed dwarf stars, three of the four observed giants, and the two Wolf-Rayet (stars number 20 and 29).

The distant group includes the supergiants (stars number 4, 16, and 28) and the LBV candidate (star 27, classified by us as B8 I). We can not estimate an individual distance for star number 27 as it is saturated in UKIDSS. According to Clark et al. (2009), this object is located at 10 kpc, in agreement with the estimated distances for the supergiants. The similar distances indicate the presence of an association of evolved massive stars at about 8 kpc. If we assume the cluster radius as the smaller circle containing all 5 stars (i.e., 1.8 arcmin), the cluster or association would have a diameter of 4.8 pc.

The individual distances trace the spiral structure of the Milky Way. In the direction of $l = 38°$, we cross the Sagittarius arm at 2 and 8 kpc, touch tangentially the base of the Scutum-Centaurus arm at 4-5 kpc and cross the Perseus arm at $\sim$ 11 kpc. Our close massive population would be part of the Scutum-Centaurus arm, and the distant population of supergiants is expected to be part of the Sagittarius arm. The analysis of MASGOMAS-6 will be included in a paper which is in preparation

3.1.6 Conclusions

The final conclusions about our work in MASGOMAS-6 may be summarized as follows:

- The preliminary version of MASCA provide us with a new massive stars cluster.
- Using broad photometric color cuts, derived from UKIDSS, we identified 30 stellar candidates for spectroscopic follow up (MOS & long-slit)
- The spectra allow the identification of massive stars in several evolutionary stages (dwarfs, giants, supergiants, and Wolf-Rayet). The number of contaminants was close to one third of the sample.
- We have identified 18 early type stars, from which 6 are supergiants. We have found a new WR (star 29) and confirm the WN6 classification of 1583-B73 (Faherty et al. 2014). The individual distance and extinction estimated for the young massive stars indicate the presence of two massive populations in the same line of sight at different distances: one at 4-4.5 kpc, in the close end of the Galactic bar and a second, more distant, in the Sagittarius arm at about 8 kpc. This second massive population appears around a distant LBV candidate (IRAS 18576+0341; (Ueta et al. 2001; Pasquali & Comerón 2002; Clark et al. 2003).
- The estimation of mass and ages for each association was not possible, due to the mixed populations.

3.2 MASGOMAS-10

3.2.1 Candidate selection

The cluster candidate MASGOMAS-10 is the second most prominent candidate from the last version of MASCA in the region of $l = 33° \pm 3°$ $b = 0° \pm 3°$ (MASGOMAS-1 was the first one candidate). For the detection of this candidate we used improved photometric cuts $K \epsilon [9, 12.5]$, $(J - K_S) \epsilon [Ec, 3.5]$ where Ec is the value given for an O9V by the adopted extinction curve (see Figure 2.12 in Chapter 2) and $Q_{IR} \epsilon [-0.2, 0.2]$.

MASGOMAS-10 was selected by the automatic program and as MASGOMAS-6, it may host Wolf-Rayets without a previous report of an associated cluster (Mauerhan et al. 2011).

With its present parameters MASCA may select an association a maximum of five times (i.e, five different optimal combinations of D_S and $Nmin$). This candidate presented four detections (although some stars are detected only once) and three main detections of the central part of the association (i.e, all the stars in the core were detected three times) as belonging to an association. Figure 3.9 shows the different detections for this candidate, figure 3.10 shows the output from the automatic method and Table 3.3 summarizes the association candidate properties.

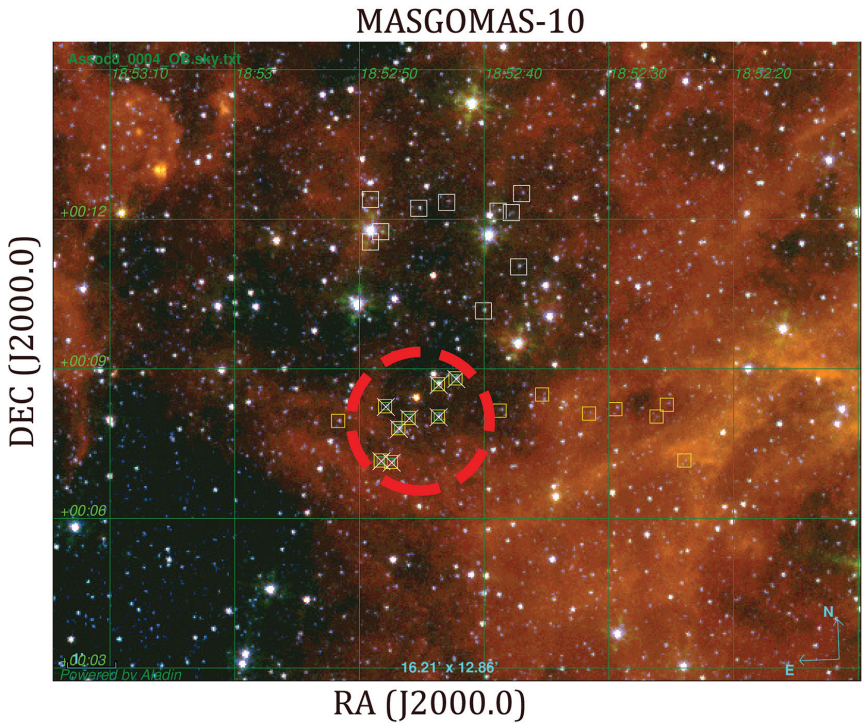

Figure 3.9: GLIMPSE false colour (blue=IRAC I1, green=IRAC I3, red=IRAC I4) image for MASGOMAS-10. The different symbols correspond to the different detections of the association candidate (i.e, orange boxes, white boxes, white crosses and green boxes). The core of the association, marked by the large red circle, was detected three times.

Table 3.3: Candidate association MASGOMAS-10 and the parameters obtained by the last version of MASCA. Coordinates (J2000.0) are given in degrees and radius and search distances in arcmin.

#	ID	RA	DEC	l	b	# stars	Radius [’]	Nmin	*Duse*
MASGOMAS–10	3329	283.190	0.132	33.201	-0.234	8	1.25	7	0.94

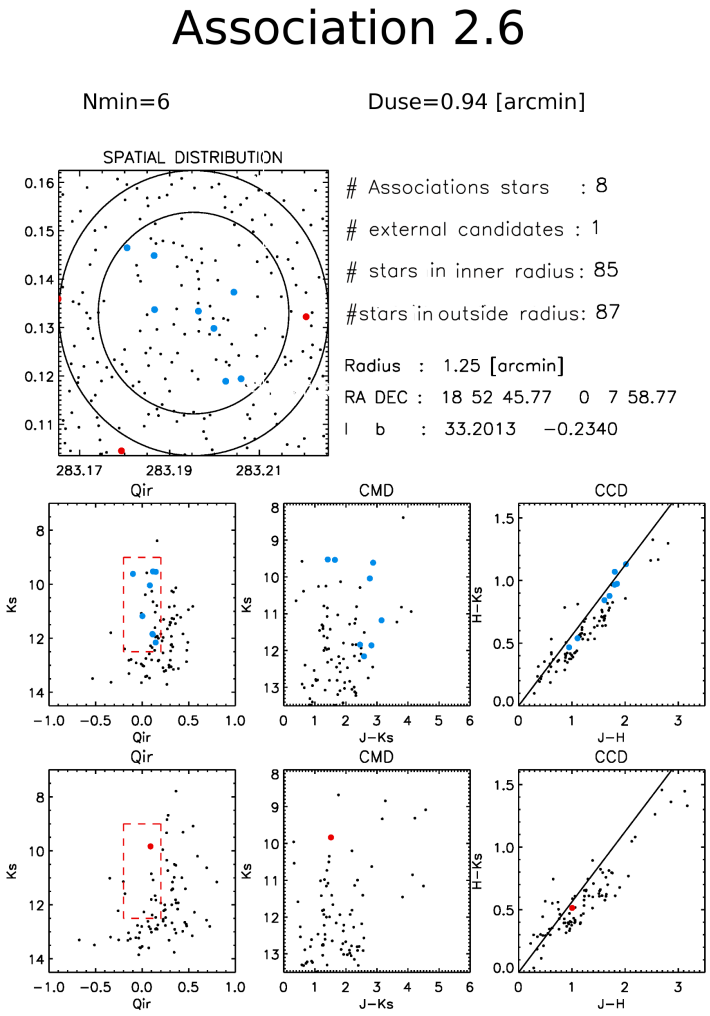

Figure 3.10: Output information for the core of MASGOMAS-10. Blue symbols are stars which fullfil the photometric cuts and are considered as an association; Red symbols are stars which fulfill the photometric cuts, but are not part of an association. The star candidates with $(J - H) \sim 1$ were classified as foreground stars, according to the observations presented in the next section.

As we mentioned before, the association candidate is not only formed by a central detection. Two additional groups, selected only once by MASCA, surround the association, and we consider all of them as a single association candidate. Figure3.9 shows the central zone of the cluster candidate detected several times, and the two extensions detected only once with the maximum *Nmin* and their corresponding *Duse*.

3.2.2 Observation and data reduction

Observations were made with LIRIS@WHT under the Director Discretionary time (DDT). For this candidate we obtained long-slit spectra in K-band. We paired stars with similar magnitudes to obtain similar S/N. The slit width was 1.00 [arcsec] and we observed with the HK pseudo-grism, to obtain a resolving power of $\lambda/\delta\lambda \sim 1700$. An A0 Vn star was used as telluric standard (HIP90967) and, as in MASGOMAS-6, we observed using an ABBA nodding mode to correct for sky emission.

We estimate that the total number of stars in the core of the association candidate is 8. Considering the assigned time, we observed 7 from the central association candidate, including the WR star, with four slits. In addition we also observed the surroundings with two slits, to include stars which fulfill the photometric cuts, but were not detected as an association by MASCA (three of them resulted as late-type stars, mentioned in table 3.5 and figure 3.13), in total we observed 12 candidates. They were observed during three nights on 2015 October 2, 6 and 7. There was thin cirrus over much of the sky during the observations. Seeing was 0.8-1.0 arcsec throughout.

In addition, we have the stars that were observed as part of the Association 118 in Chapter 2 (see Table 2.6), although most of the stars observed then were classified as late-type contaminants. We reduced and calibrated using LIRISDR in the sames procedures as MASGOMAS-6 reduction. The A0Vn star HIP90967 was used to remove telluric lines using the XTELLCOR package (Vacca et al. 2003) (due MOLECFIT was still on test over MASGOMAS-6).

3.2.3 Spectral classification

Following the spectral classification used for MASGOMAS-6, we fist separated the spectra according the presence of the CO bands. The second criteria was the detection of the HeI/II lines at 2.11 μm and 2.18 μm respectively. In Chapter 2, we could study part of this association, getting clues of the presence of a massive stellar population in the region, particularly the spectra for star 31 (labeled in this chapter as ext7).

In the core of the association (red circle in figure 3.9), we find two groups. The first is the late-type stars 1 and 2. Comparing with the template spectra from the Meyer et al. (1998) atlas we classify them as M3 III and K2 III respectively. They present the characteristic CO band, with modest equivalent with (EW) below 22 Å, confirming that they are giant stars.

The second group includes the early type stars. Star number 3 and 4 have similar spectra and we classify them as O8V based on their clear HeI 2.11 μm and relatively weak HeII 2.18μm. Star ext8 displays a weak HeII absorption line and was classified as O9V. Star 5 does not show HeII, its spectral features are similar with the B0V HR6165 lines and we assigned it this spectral type. The spectra of star 6 and ext7 show also the characteristics of an O8 star, and they are

classified as luminosity class III because of their narrower Brγ and the better agreement with the template spectrum star with this luminosity class

For the WR star, reported by Mauerhan et al. (2011), we confirmed the WN8 classification. The relation between the Brγ and HeI 2.11 lines supports the spectral classificacion when comparing with spectra from the Crowther & Smith (1996) catalog. We used the intrinsic colors from this catalog to estimate individual distances and extinctions. We adopted the absolute magnitude for this star from Nugis & Lamers (2000). Figure 3.11 shows the spectra for the core stars plus the two extra O-type stars (ext7 and ext8) founded in the surroundings of the core of MASGOMAS-10 and Table 3.4 lists the spectroscopically observed candidates.

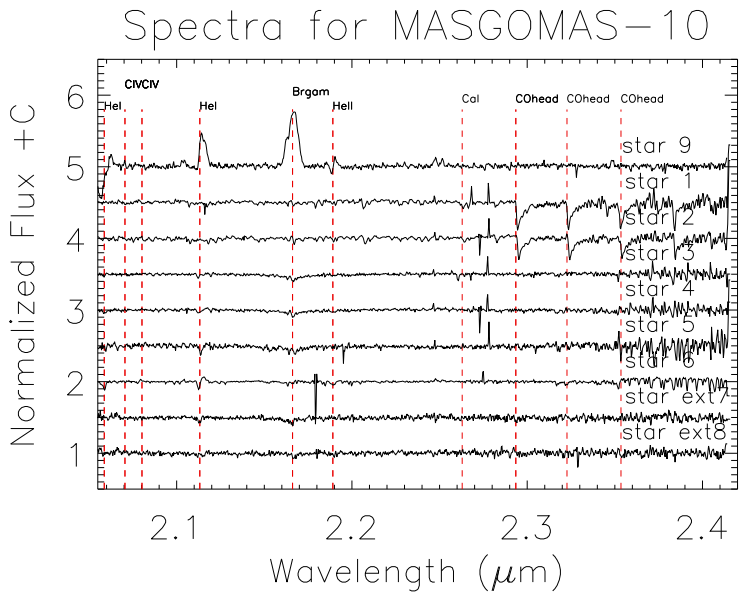

Figure 3.11: LIRIS K-band spectra for MASGOMAS-10. With the exception of ext7 and ext8, all stars are found in the core region (red circle in Figure 3.9). Stars number 3, 4, 5, 6 (from the core) ext7 and ext8 (outside the core) are OB-type stars. Together with the WR, they clearly characterize MASGOMAS-10 as a massive cluster candidate. The two late-type stars, also in the core region, were marked as low probability candidates because their ($J - Ks$) color was bluer than that of the other stars.

Table 3.4: MASGOMAS-10 classification. Extinction and distances were estimated using the intrinsic colors from Martins & Plez (2006) and Cox (2000).

ID	2MASS	RA	DEC	J	H	K	QFLAG	S-type	Ak	Distance
3	18524788+0008143	18:52:47.88	+0:08:14.3	14.681	12.833	11.859	AAA	O8 V	1.999	4.853
4	18524601+0008002	18:52:46.02	+0:08:00.2	14.294	12.683	11.840	AAA	O8 V	1.756	5.379
5	18524745+0007080	18:52:47.45	+0:07:08.0	14.325	12.309	11.178	AAA	B0 V	2.180	2.713
6	18524682+0007473	18:52:46.83	+0:07:47.4	12.812	11.008	10.041	AAA	O8 III	1.958	3.472
ext7	18525284+0010066	18:52:52.84	+0:10:06.7	12.862	11.468	10.660	AEA	O8 III	1.583	5.487
ext8	18524910+0012241	18:52:49.10	+0:12:24.2	13.022	11.660	10.886	ABA	O9 V	1.534	3.360
9	18524369+0008415	18:52:43.69	+0:08:41.5	12.490	10.684	9.614	AAA	WN8	1.652	8.364
1	18524227+0008473	18:52:42.28	+0:08:47.4	10.940	9.993	9.526	AAA	M3 III	0.181	7.889
2	18524371+0008013	18:52:43.72	+0:08:01.4	11.183	10.077	9.538	EEA	K2 III	0.590	1.696

3.2.4 Discussion

The aim of these observations was to confirm the presence of a population of young massive stars. A deep characterization of the association is challenging considering the limited spectral information of this observations. Nevertheless we provide a rough estimation of the association distance and mass.

Figure 3.12 shows the approximate values for the extinction and distance for the stars in the core which seem to concentrate around 5.5 kpc, at the distance of the Scutum-Centaurus arm. In spite of the spread distances, the relative radial velocities appear to be similar for all stars in the group, locating them in the Sectum-Centaurus arm. At this distance, the angular radius of 1.15 arcmin obtained for the association represent a linear radius of 2.0 parsecs.

Also considering the number of massive stars in this small region we expect a lower limit for the cluster/association total mass of $4 \times 10^3 M_{\odot}$.

Figure 3.13 summarizes the results of MASGOMAS-10 follow-up. Some of the stars selected by the method as part of the association turned out to be late-type giants. However, these stars were spotted only by one combination of *Nmin*, *Duse* (7 and 1.19 arcmin), which actually strengthens the reliability of the method where more than one detection is present. The Table 3.5 includes the observed parameters of observed giant- and late-type-supergiant stars.

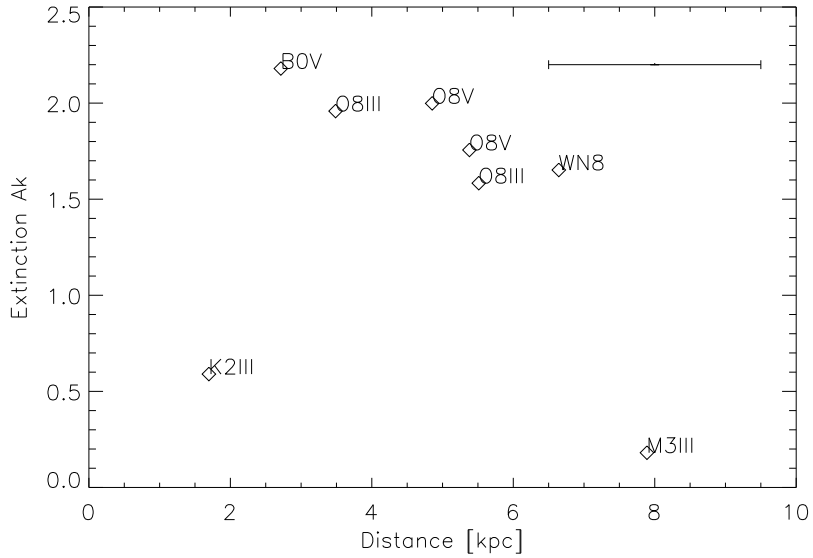

Figure 3.12: Extinction versus distance for MASGOMAS-10 stars with spectroscopic follow-up. The spread in distance is expected and can be due to the rough classification. The similar extinctions point to a single association. Top right is displayed the average error bar for the distance.

Table 3.5: Observed stars in the surroundings of the core of MASGOMAS-10, considered as contaminats plus two supergiants. The stars observed for the association 118 (described in Chapter 2), have the letter "e" plus the ID from that observation.

ID	RA	DEC	J	H	K	Spectral type
e9	18:52:37.21	+00:11:02.6	12.764	11.080	10.224	K2III
e24	18:52:31.93	+00:11:31.8	13.475	12.123	11.460	K2III
e20	18:52:36.94	+00:12:29.9	13.032	11.706	10.948	K2III
e29	18:52:49.65	+00:12:46.0	12.408	11.317	10.778	K2III
e33	18:52:48.24	+00:11:43.7	14.273	12.184	11.065	K2III
5	18:52:39.75	+00:08:05.5	13.422	12.034	11.141	M4III
8	18:52:51.67	+00:07:56.0	11.354	10.349	9.835	K2III
10	18:52:50.17	+00:10:02.3	12.612	11.536	10.821	K2III
12	18:52:35.26	+00:08:29.0	14.017	12.539	11.775	K0III
11	18:52:40.92	+00:10:13.1	14.661	12.875	11.742	G8Ib
eB5	18:52:38.85	+00:12:09.9	14.110	11.824	10.629	A0Ia

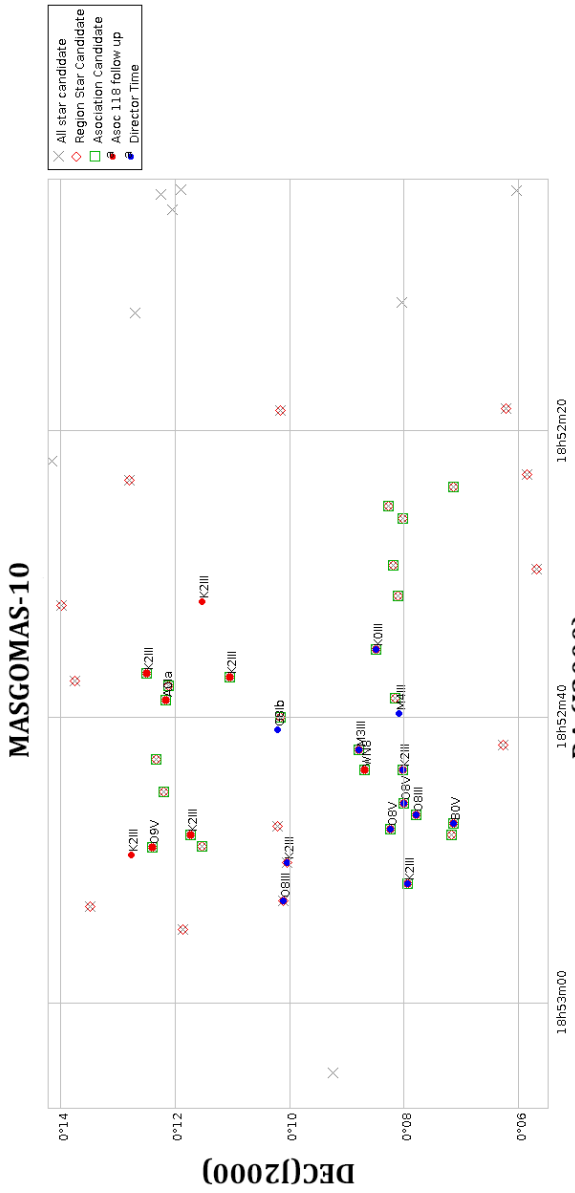

Figure 3.13: Summary of MASGOMAS-10 follow-up, including the exploratory observations from Chapter 2. Blue points mark the observation described in this chapter, which include the association center, detected three times, with the largest concentration of OB-type stars. Red points correspond to the exploratory observation. These were detected only once in the last version of MASCA, confirming that more than one detection is a positive sign for a bona-fide candidate.

3.2.5 Conclusions and future work

- By comparing the population in the individual associations we conclude that an association detected using several values of D_S and *Nmin* is most probably a real association of massive stars. The majority of the OB-type star population is the *core*, detected several times by the code, while the old-late type population are part of the peripheral associations only detected once.
- Selecting all the candidates that pass the photometric cuts in the proximity zone of a several times detected association, may increase the number of confirmed massive stars, but with the high risk of confirming contaminants.
- We detected a total of 7 massive early type stars, including a Wolf-Rayet. A preliminary estimation indicates a radius of 2.0 kpc and a mass of few thousands of solar masses for the possible association

This confirmed massive population makes MASGOMAS-10 an interesting object to study. It is necessary to review the extension of its stellar population, and improve object characterization using new spectra with higher resolution. The preliminar distance estimated in this work place MASGOMAS-10 in the base of the Scutum-Centaurus arm, in the close end of the Galactic bar, a region characterized for the presence of massive clusters with massive young and evolved population.

The plan is to complete an extended spectroscopic follow-up using EMIR at GTC. This new follow-up will including the observed stars, will increase the spectral resolution and the number of candidate stars in the area. It will allow us to accurately characterize the association, and confirm the efficiency of the MASCA code.

4

Analyzing Near-Infrared spectra

Contents

The quantitative analysis of the stars observed in our MASGOMAS clusters present an important problem: spectra have to be analyzed using only information obtained in the near-infrared. Yet, most of our expertise is based on optical studies. In this chapter we study how consistent is the information that we obtain from both wavelength ranges and address one of the physical processes that may constitute an important difference between both wavelength ranges: the presence of wind density inhomogeneities known as clumping.

4.1 Introduction

Spectra of massive stars are difficult to analyze due to the physical processes taking place in their atmosphere. The relatively high temperatures and low densities produce deviation from the equilibrium (NLTE).

Another important property is their high luminosity (above $10^4 L_\odot$), that combined with the high temperature implies an intense radiation field that ionizes the interstellar medium. This luminosity drives a strong stellar wind, extending the atmosphere that has to be treated in spherical geometry and carrying out enormous quantities of material. The proportion of expelled material is related to the metallicity of the star and the molecular cloud from which it was born (Kudritzki & Puls 2000). This metallicity has a strong impact on the opacity of the atmosphere producing an important line blanketing effect that affects the effective temperature. Thus, the atmospheres of hot massive stars have to be treated in NLTE, with spherical geometry and mass-loss and a good treatment of opacity at all wavelengths. Several codes have treated this problem with different approximations (see Sect. 1.5.1).

An additional physical process occurs in the stellar wind, in particular where the optical and near-infrared lines are formed, thus affecting its emerging profile. The wind is not always smooth and continuous, but presents density inhomogeneities. Several studies have been carried out to explore massive stars winds, focused on this particular behavior, called clumping (see e.g., Puls et al. 2008, Kudritzki & Puls 2000, Sundqvist & Owocki 2013 and references therein). Using updated model atmosphere codes it has been possible to confirm the existence of clumping by modelling their effects on the line profiles and comparing with observed spectra (Puls et al. 2008 and Najarro et al. 2011). The presence of clumping is particularly important, as it has a direct impact on the determined stellar mass-loss rate (because the usual diagnostics are proportional to the density or the square of the density, see Sect. 4.5). Thus determining the real mass-loss rates implies the use of diagnostics independent of clumping, or the knowledge of the clumping distribution in the stellar wind. The other way around, the presence of clumping modifies the stellar parameters derived from optical and near-infrared diagnostics as compared to those derived for an homogeneous wind.

Optical and near-infrared lines may be formed in different layers in the atmosphere, and thus may be affected in different forms by the clumping as it develops in the wind. This may introduce a difference in the stellar parameters derived from both atmosphere ranges. Moreover, the behavior of the infrared lines depends on upper atomic levels and has been less studied than that of the optical lines. This may again introduce differences between the derived stellar parameters.

We explore the use of the near-infrared range for the determination of stellar parameters and the effects of clumping using the H and Ks bands. These two bands show the characteristic lines of OB stars and thus they offer a good opportunity to explore the hidden massive star population. But before using the near-infrared, we need to test if we recover similar stellar paraneters as those derived from the optical spectrum.

The first part of this chapter is focused in the comparison of the stellar parameters obtained from the optical and those obtained using exclusively near-infrared lines, to test the possibility of using only near-infrared spectra. We will analyze stars observed in the optical and near-infrared at high resolution and S/N and compare the results. Different to previous studies (i.e., Repolust et al. 2005 or Najarro et al. 2011), we will use an automatic method to determine the stellar parameters. To this aim, we will generate a full grid of model atmopsheres using FASTWIND (Puls et al. 2005) and then will use the *iacob_gbat* tool (Simón-Díaz et al. 2011; developed to work in the optical range), adapting it to the analysis of near-infrared lines.

The second part focuses on the particular behavior of the strong wind stars, and recreates the clumping scenario also using the FASTWIND code. The goal of this section is to asses the effects of clumping and different clumping laws on the stellar parameters and show that clumping may explain some of the differences found in the previous section between parameters derived from the optical and the near-infrared.

4.2 The Data

We selected those O stars from the near-infrared catalogue by (Hanson et al. 2005) that are also present in the IACOB Spectroscopic Database (Simón-Díaz et al. 2011; see Table 4.1). Hanson et al. spectra were obtained with the Infra-red Camera and Spectrograph (IRCS) mounted at the Cassegrain focus of the 8.2 m Subaru Telescope at Mauna Kea, Hawaii, in the H and K bands with R∼12000 and S/N∼200-300. IACOB spectra were obtained with the Fibre-fed Echelle Spectrograph (FIES) attached to the Nordic Optical Telescope (NOT) with S/N ≥ 150 and a resolution $R \sim 46000$. The IACOB-Project also provides us with better estimations of the $V \sin i$ and so-called macroturbulence velocity (Simón-Díaz & Herrero 2014). The sample covers the range of O spectral types from O4 to O9.5 and all luminosity classes. Although some combinations (like mid-type supergiants or cool late-type dwarfs) are underrepresented, the sample as a whole will provide a good test of the global behaviour of O-type stars.

4.3 Methodology

To determine optical and near-infrared parameters and their uncertainties, two main tools are required: a full grid of synthetic optical and near-infrared spectra and an automatic tool which allows us to determine the parameters for a large sample of stars. We generate the first one using the code FASTWIND (Puls et al. 2005), covering the range of massive OB star parameters, with a grid of 100000 models. To create this grid of models we used the distributed computation system HTCondor[1]. The second one is the already developed *iacob_gbat* (Simón-Díaz et al. 2011,Sabín-Sanjulián et al. 2014) an automatic tool (explained in detail in sec 4.3.2), that allows one to fit the observed H-He spectrum, returning the best model fit, i.e. the parameters of the star and their corresponding errors (based on H-He lines). This algorithm was running only for the optical range, and we expanded it to the near-infrared.

[1] http://research.cs.wisc.edu/htcondor/

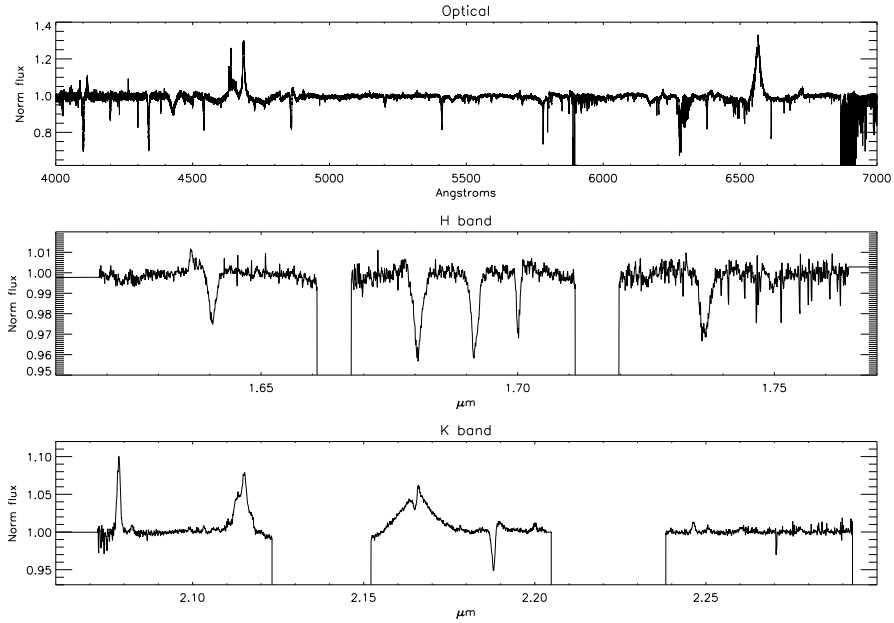

Figure 4.1: An example of the spectra available in our sample (HD 15570). Upper panel: optical spectrum from the IACOB database; lower panels: near-infrared spectra from the Hanson et al. catalog.

The supercomputer facility HTCondor@Instituto de Astrofisica de Canarias provides us with a cluster of 914 cores, each of them capable to run in parallel. We wrote a script to run FASTWIND with HTCondor. This way it was possible to reduce the total amount of time to create a full grid of model from months to only few days (typically one week). During the process it was also necessary to monitor the progress of the model grid calculation, particularly because some cores from HTCONDOR may not work properly (for example, interrupting the calculation when some priority conditions are fulfilled), which added some extra time to the grid calculations. Even so, the optimization of time using HTCondor was enormously helpful.

4.3.1 FASTWIND near-infrared grid

The FASTWIND code includes NLTE, spherical symmetry, mass loss and line blanketing. A model atmopshere is mainly defined by its effective temperature, gravity, stellar radius, microturbulent velocity, β-value (the exponent of the wind velocity law), helium abundance (by number), wind terminal velocity, mass-losse rate and clumping factors. It is clear that the combination of all these parameters would result in a huge amount of models. To reduce them, we limit the stellar radius according to the other stellar parameters and use the wind strength pa-

Table 4.1: O stars selected for the analysis. $V \sin i$ and V_{mac} are from Simón-Díaz & Herrero (2014)

#	STAR ID	Spectral Type	$V \sin i$ [Km/s]	V_{mac} [Km/s]
1	HD46223	O4 V((f))	52	97
2	HD15629	O4.5 V((fc))	70	69
3	HD46150	O5 V((f))z	69	107
4	HD217086	O7 Vnn((f))z	382	104
5	HD149757	O9.2 IVnn	290	290
6	HD190864	O6.5 III(f)	65	90
7	HD203064	O7.5 IIIn((f))	315	98
8	HD15570	O4 If	38	120
9	HD14947	O4.5 If	117	49
10	HD30614	O9 Ia	115	72
11	HD210809	O9 Iab	76	79
12	HD209975	O9.5 Ib	52	95

rameter, Q (defined as Q= $\frac{\dot{M}}{(v_\infty R_*)^{3/2}}$) that combines mass-loss, stellar radius and wind terminal velocity. The emission in H_α (and other wind diagnostics lines) is the same for the same value of Q (or very similar, see Puls et al. (1996)).

Thus for our grid of synthetic (unclumped) H-He spectra we have the following free parameters: Teff, log g, ϵ (He abundance by number, defined as $\epsilon = \frac{N(He)}{N(H)+N(He)}$[2]), wind strength (log Q), wind velocity law exponent (β) and mictro-turbulence as shown in Table 4.2. Figure 4.2 illustrates the distribution of grid models in the $\log T_{eff}$ v/s $\log g$ diagram together with the Geneva evolutionary tracks for 05, 10, 15, 20, 25, 40, 50, 60, 85, 120 and 150 $M_\odot$ (Z=0.014) by Ekström et al. (2012). The final grid contains a total of 107 547 models.

Table 4.2: Parameter ranges for the grid models

Parameter	Range of value
T_{eff}	55000 ≥ 22000 K (step 1000 K)
log(g)	[2.6-4.3] (step 0.1 dex)
Micro	5,10,15,20 km s^{-1}
ϵHe	0.06, 0.09, 0.013, 0.017, 0.020, 0.023
log(Q)	-15.0, -14.0, -13.0, -12.7, -12.5, -12.3, -12.1, -11.9, -11.7
β	0.8, 1.0, 1.2, 1.5, 1.8

[2]due to the use of different tools, we sometimes use the above definition of ϵ, and sometimes $Y(He) = \frac{N(He)}{N(H)}$. Only for large helium abundances there is an important difference, as $\epsilon = \frac{Y(He)}{1+Y(He)}$. For example, the cosmic helium abundance by number is ϵ= 0.09 or $Y(He)$= 0.10

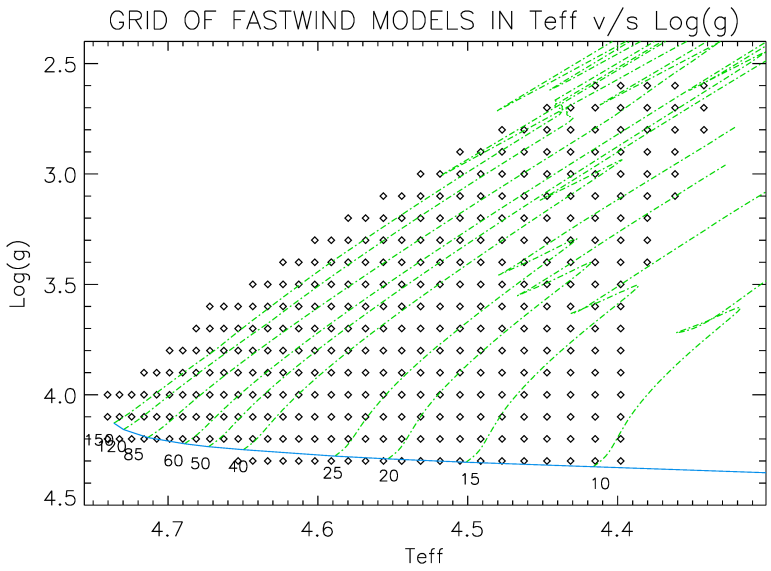

Figure 4.2: Location of the models on the FASWTIND grid in the Teff v/s log g plane. Evolutionary tracks from Geneva, without rotation, are plotted in green and the blue line defines the corresponding ZAMS. The numbers give the tracks initial mass (Ekström et al. 2012)

Moreover, in Sec. 4.5 we recalculate the same grid with the clumping laws and factors described there.

The model grids in our group are usually calculated for the optical range. For our work here we needed to extend them to the near infrared. To this aim we reviewed the atomic models, to be sure that we consider the appropriate levels when calculating the atomic populations, as the spectral lines may depend on the population of levels not directly involved in their formation. Figure 4.3 illustrates this situation, showing that HeI 4471Å in the optical and HeI 1.70 μm in the near-infrared share the upper level. Thus the population of the lower level of HeI 4471 will influence the formation of the HeI 1.70 μm line. Table 4.3 gives the list of H and He lines included in the synthetic spectra. The table also indicates the different components blending with the main spectral line considered. The Br$_\gamma$ line for example includes four different components.

4.3.2 Automatic fitting and extension to the near-infrared.

The determination of stellar parameters is carried out by comparing the observed spectrum with the synthetic ones in the model grid. The best fit model (under predefined criteria) is considered the optimal one and the parameters used to calculate it are adopted as the parameters of the observed star. Uncertainties are determined by inspecting the parameter distribution of models

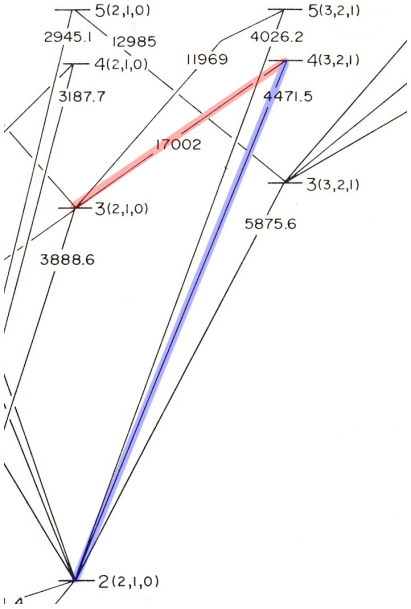

Figure 4.3: Gotrian diagram illustrating the relation between the HeI transitions at 1.70μm (red line) and 4471Å(blue line). The numbers give the main quantum number, with the different J values in parenthesis

with a fit close to the best one (again, with predefined criteria).

iacob_gbat

iacob_gbat is a grid-based automatic tool (Simón-Díaz et al. 2011) developed to fulfill the task of comparing a huge amount of synthetic spectra with the observed one, providing us with the best fit and thus the stellar parameters and errors. Before running the tool it is important to determine the rotational and macroturbulent velocities (V sini and V_{mac}) as they broaden the spectrum affecting the shape of the spectral lines. A wrong determination of these velocities can alter the value obtained for the stellar gravity (the main line broadening agent through collisional broadening processes) and through it all other stellar parameters. The values of the rotational and macroturbulent velocities are obtained with the *iacob_broad* tool, developed by our group (see Simón-Díaz & Herrero 2014).

Once we have determined these two velocities, we run *iacob_gbat* for the determination of the six free parameters defining our models. The following points give an outline of its workflow (a more detailed description may be found in Holgado et el., 2017, submitted.

1. First, some interaction is needed. We define interactively the wavelength range of the lines

Table 4.3: Optical and near-infrared lines included in the FASTWIND line formation calculations

Line	wavelength Å	number of components & transitions involved
		Optical
Hα	6562	2 - HI(2-3) & HeII(4-6)
Hβ	4861	2 - HI(2-4) & HeII(4-8)
Hγ	4340	2 - HI(2-5) & HeII(4-10)
Hδ	4101	2 - HI(2-6) & HeII(4-12)
Hϵ	3970	2 - HI(2-7) & HeII(4-14)
HeI 4387	4387	1 - HeI(2p1-5d1)
HeI 4922	4922	1 - HeI(2p1-4d1)
HeI 4026	4026	2 - HeI(2p3-5d3) & HeII(4-13)
HeI 4471	4471	1 - HeI(2p3-4d3)
HeI 6678	6678	2 - HeI(2p1-3d1) & HeII(5-13)
HeII 4200	4200	1 - HeII(4-11)
HeII 4541	4541	1 - HeII(4-9)
HeII 4686	4686	1 - HeII(3-4)
HeII 6527	6527	1 - HeII(5-14)
		H
Br_{10}	17360	1 - HI(4-10)
Br_{11}	16810	1 - HI(4-11)
Br_{12}	16410	1 - HI(4-12)
HeI 1.70	17000	1 - HeI(3p3-4d3)
HeII 1.69	16900	1 - HeII(2-12)
		Ks
Brγ	21660	4 - HI(4-7), HeI(4d1-7f1), HeI(4d3-7f3) & HeII(8-14)
HeI 2.11	21120	2 - HeI(3p3-4s3) & HeI(3p1-4s1)
HeII 2.18	21880	1 - HeII(7-10)

we want to use for the analysis and correct from radial velocity, perform a slight correction of the local continuum slope or clip nebular lines.

2. The program compares the observed and the predicted fluxes for each line in each model. We define the quantity

$$\sigma^2_{line,model} = \sum_{1}^{N_{pixels}} \frac{(F_l - F_o)^2}{N_{pixels}} \tag{4.1}$$

where F_l and F_o are the synthetic and observed line fluxes (normalized) and N_{pixels} is the number of points in the line. We have a value for each line in each model. This quantity is used to estimate the fit quality for each individual line. It is smaller when the line has a

good model fit and the observed spectrum has a good quality, as the differences between F_l and F_o may be due both to wrong stellar parameters and to observational noise. For each line, we adopt $\sigma_{line} = min(\sigma_{line,model})$. Note that different lines may be optimally fitted by different models.

3. A χ^2 is calculated for each line in each model as

$$\chi^2_{line,model} = \sum_{1}^{N_{pixels}} \frac{(F_i - F_o)^2}{\sigma^2_{line}} \tag{4.2}$$

Note that for each line, this value will be minimum for the model that best fits the corresponding line, independently of how the other lines are fitted in that model. Our grid does not allow us to fit a given line better than this.

4. For each model, a χ^2 value is constructed as

$$\chi^2_{model} = \sum_{1}^{N_{\#line}} \chi^2_{line,model} \tag{4.3}$$

This favours models having a large number of well-fitted lines, even if they have not the best possible fit in the grid. The minimum of χ^2_{model} indicates the model that deviates the less when considering all individual lines weighted by its best possible fit within the grid. This is our initial guess of the best model. But it has been constructed using the σ_{line} values, that correspond to different models.

5. We now adopt the $\sigma_{line,model}$ of the estimated best fitting model. Let's call these values $\sigma_{line,bestmod}$ and construct a new χ^2 for each model.

$$\chi^2_{model} = \sum_{1}^{N_lines} \sum_{1}^{N_{pixels}} \frac{(F_i - F_o)^2}{\sigma^2_{line,bestmod}} \tag{4.4}$$

If our initial guess is the best model, the new χ^2 will be minimum for the same model. If this is not the case, a different model will have now the minimum χ^2_{model} and we iterate the procedure until it converges or until a given number of maximum iterations is reached.

This way, a distribution of χ^2 values is obtained that can be used to estimate the uncertainties for each stellar parameter. In Figure 4.4 we plot the distribution of χ^2 versus temperature for one of the analyzed objects. The minimum χ^2 value gives us the best model and the uncertainty is estimated by adding +1 to that value. For an χ^2 distribution this would correspond to 1 σ uncertainty.

The final parameters estimation, is calculated using the average value from all the points included in the +1 distribution. However, in some cases the distributions present some difficulties. Figure 4.5 illustrates some of them: cases in which we cannot determine a given parameter with

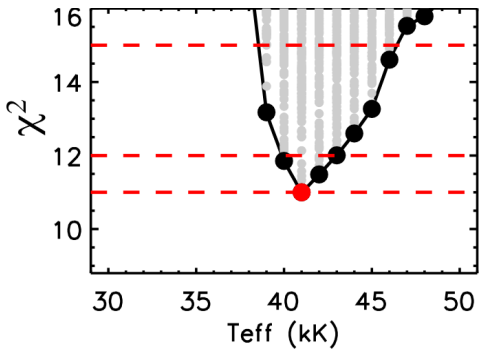

Figure 4.4: An example of the distribution of χ^2 versus effective temperature. The ordinate axis corresponds to the χ value which dependes on the numbers of lines used to determine the adjust.

accuracy or values that are at the edge of the parameter search range. Thus the final output has to be always individually examined and treated to solve (or at least minimize) these cases. A more detailed discussion of the different problems that can appear may be found in Sabín SanJulián (2014).

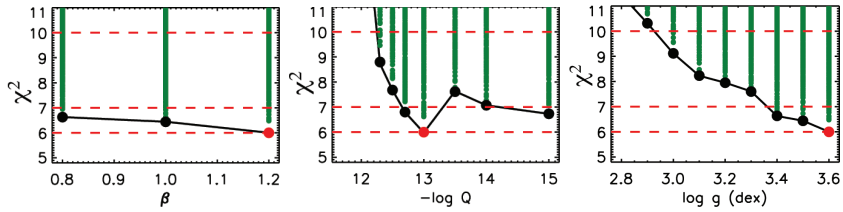

Figure 4.5: Illustration of three different problems found in the χ^2 distribution of different parameters. Left: the value of β remains unrestricted, as all models within the grid give acceptable fits. Middle: the value of log Q is badly constrained, because some models give marginally acceptable fits; often this is related to a poor sensitivity to the parameter in certain parameter regions and demands a close inspection of the model fits. Right: the explored region should be moved to higher gravity values; sometimes, this problem is found at the grid edges, implying some peculiarity (i.e., a composite spectrum).The ordinate axis in all plots, corresponds to the χ value which dependes on the numbers of lines used to determine the adjust.

Extension to the near infrared

To extend the *iacob_gbat* we wrote several modules inside the code. In addition, to include all the near-infrared lines from Table 4.3 in the determination of the best fit model, we also updated the error estimations and graphic outputs of the program.

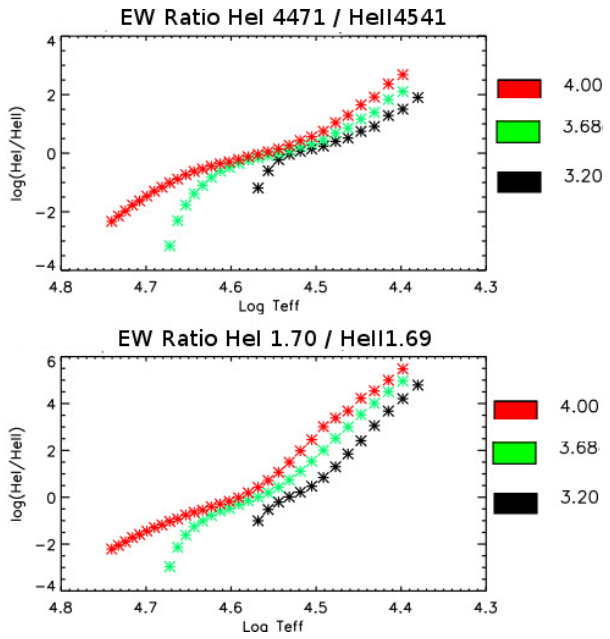

Figure 4.6: Equivalent width ratios in the model grid as function of effective temperature for optical and near-infrared HeI/HeII lines. It is easy to notice that the behavior is similar for both wavelength ranges

Having included the near-infrared lines, we carried out several test calculations to check their behavior.

It is known that the ratio HeI4471/HeII4541 is a good temperature diagnostic in the optical range. The same role can be played by the ratio HeI1.70/HeII1.69 in the near-infrared as shown in Figure 4.6. There we show the equivalent width ratio for both pairs of lines for a series of models ranging from 25 000 to 55 000 K and for three values of the gravity. We see that the near-infrared lines present in the H-band can be as sensitive to the temperature as the optical lines and with a very similar behavior.

Similarly to Balmer in the optical, the shape and wings of the Brackett series in the infrared are sensitive to gravity. However, there is an important difference. In the optical, where we use transitions with relatively small upper quantum numbers, the line cores are dominated by Doppler broadening. In the infrared, because the effect of Stark broadening increases rapidly with the upper quantum number of the transition, also the cores of the Brackett lines are very sensitive to gravity (see Section 4.1 in Repolust et al. 2005 for a more detailed discussion). We realized during the test calculations an anomalous behavior of the different Br lines, that we could not fit simultaneously to the observed spectra. We carried out then a series of tests,

grouping the lines in pairs (Br10&Br11, Br11&Br12, Br10&Br12, Br_γ being less affected by the problems described above) i.e., leaving one of the lines out of the parameter determination. This way we checked which pair was more consistent with the rest of the near-infrared lines. Our tests indicated that Br12 gave the poorest agreement (expected, as it has the highest lying upper level in the series) and thus it is not used for the rest of the work.

4.4 First Results: parameter determinations

Our stellar sample can be divided in three groups, depending on the luminosity class for the stars (i.e [I-II], [III] and [IV-V]). Each of the three groups presents a particular behavior. w.r.t. the fits obtained. Dwarf stars show the best fits to the observed spectrum, whereas the fit difficulties increase for giants and are usually larger for the luminosity class I stars, those with the strongest winds.

4.4.1 Stellar parameters from the optical spectrum

We first determine the stellar parameters using only the optical spectra secured in the IACOB database. We adopt the $V\,sini$ and V_{mac} values determined by (Simón-Díaz & Herrero 2014;see Table 4.1) and run the *iacob_gbat* tool. Table 4.4 summarizes the parameters obtained from our optical analysis. We see that β and the microturbulent velocities are the parameters presenting a larger number of upper or lower limits. This evidences that the information contained in the optical spectra is usually not sufficient to constrain these parameters accurately.Errors were given by the *iacob_gbat* as described before, but following the arguments by Sabín-Sanjulián et al. (2017) we set a lower limit of 0.1 in log Q when the formal error is lower (sometimes, the *iacob_gbat* tool may deliver unrealistically low errors, as it does not take into account uncertainties like the continumm normalization). Figure 4.7 offers the comparison between the observed profiles and the synthetic lines from the best fit model.

Table 4.4: Stellar parameters obtained from the optical analysis. Helium abundances are given by number, with Y(He)= N(He)/N(H). Note that the Y(He) values in the table are multiplied by 100, and that log Q values are actually negative.

#	Star ID	spectral type	Teff(kK)	logg(dex)	−logQ	Y(He)×100	micro (km/s)	beta
1	HD46223	O4 V((f))	43.0 ± 1.2	3.76 ± 0.07	12.8 ± 0.2	10.3 ± 1.1	> 9.1	1.0 ± 0.2
2	HD15629	O4.5 V((fc))	41.4 ± 1.4	3.71 ± 0.11	12.7 ± 0.2	12.1 ± 3.4	< 19.9	1.0 ± 0.2
3	HD46150	O5 V((f))z	41.2 ± 1.0	3.78 ± 0.07	13.0 ± 0.3	9.0 ± 1.6	> 5.0	> 0.8
4	HD217086	O7 Vnn((f))z	37.0 ± 1.0	3.60 ± 0.11	13.9 ± 1.1	11.4 ± 2.9	12.4 ± 7.4	< 1.2
5	HD149757	O9.2 IVnn	32.5 ± 0.8	3.84 ± 0.17	14.1 ± 0.9	11.0 ± 3.1	12.2 ± 7.2	< 1.2
6	HD190864	O6.5 III(f)	37.1 ± 0.7	3.58 ± 0.05	12.7 ± 0.1	12.1 ± 2.7	15.1 ± 3.4	0.9 ± 0.1
7	HD203064	O7.5 IIIn((f))	34.9 ± 0.7	3.54 ± 0.11	12.7 ± 0.1	10.4 ± 1.2	> 15.2	0.9 ± 0.1
8	HD15570	O4 If	40.1 ± 0.9	3.75 ± 0.18	12.0 ± 0.1	10.5 ± 1.8	> 5.0	> 1.0
9	HD14947	O4.5 If	38.1 ± 0.5	3.61 ± 0.05	12.0 ± 0.1	15.4 ± 2.2	> 9.5	> 1.2
10	HD30614	O9 Ia	29.4 ± 0.8	2.96 ± 0.09	12.2 ± 0.1	14.3 ± 3.2	> 15.9	> 0.8
11	HD210809	O9 Iab	31.1 ± 0.3	3.17 ± 0.07	12.4 ± 0.1	11.7 ± 2.2	> 16.2	> 1.0
12	HD209975	O9.5 Ib	31.3 ± 0.4	3.23 ± 0.05	12.7 ± 0.1	9.8 ± 1.1	> 12.2	> 1.1

From the fits shown in Figure 4.7 we obtain the following conclusions:

- Early dwarfs (HD 46 223, HD 15 629 and HD 46 150) show excellent fits. The mid-type and late type dwarfs (HD 217 086 and HD 149 757) are fast rotators and do not show any significant problem, in spite of potential effects not considered, like gravity darkening or geometrical deformation.
- The two giants we have in our sample are mid-types and one of them is again a fast rotator. HD 190 864 shows a small difference with the best model when fitting the cores of the HeI lines, with slightly shallow theoretical profiles for HeII 4200 and 4541 complemented by a slightly deep profile for HeII 4686. The fast rotator, HD 203 064, has a poor fit to Hα and to a lesser extent to HeII 4686. Hγ shows also a poor fit in the blue wing, but this may be due to metal blends (compare it to its counterpart HD 190 864)
- The supergiants display the largest fitting problems, particularly in Hα, sometimes together with problems in Hβ and HeII 4686 (much less, however), which points to some wind influence. The largest difficulties are found for the Hα P-Cygny profile of the late types, HD 30 614 (of Ia luminosity class) and the well known non-radial pulsator HD 210 809(λ Cep, Iab). In both stars the HeII 4686 core shows a shift to the red. The best fit in this group is obtained for the less luminous supergiant, HD 209 975 (Ib). Early-type supergiants have an intermediate behavior in Hα(in spite of showing emission), although they present some difficulties for the red wing of Hβ that are not seen in late-type supergiants.

Finally we compare in Figure 4.8 our parameters with those derived by Repolust et al. (2004). Since these authors published their analysis, a few updates have been introduced in FASTWIND (see Puls et al. 2005), which adds to the differences between the by-eye technique used by Repolust et al. and our automatic fitting procedure. We see that the agreement in the effective temperature and gravity values is excellent, with a small difference at the highest temperatures, which can be explained by the effect of the improved line-blanketing (which produces lower effective temperatures). The agreement in the log Q values (not shown in the Figure) is also excellent. The largest differences appear in the derived helium abundances, where we obtain He abundances that show a trend towards lower values, although they are within the errors. Comparison with other results in the literature like Martins et al. 2015 or Bouret et al. 2012 (carried out with the CMFGEN code, in the second case also taking UV data into account) indicate similar good agreement within errors in spite of some small differences (two early stars in common with Martins et al. show slightly lower values for the gravity in our case, whereas one star in common with Bouret et al. shows a slightly larger value). We conclude that our parameters are consistent with previous results in the literature.

4.4.2 Stellar parameters from near-infrared

As we explained at the beginning of this chapter, we aim at deriving stellar parameters only from the infrared spectrum, because we expect that the stars found in obscured clusters in the MASGOMAS searches will be heavily extincted, without any possibility to observe them in the

optical spectrum. While this exercise has already been done by other authors (like for example Repolust et al. (2005) or more recently authors fitting simultaneously optical and infrared spectra, like Najarro et al. (2011) or Bestenlehner et al. (2014)) we have to check that our automatic procedure extendeds to the near-infrared gives us reliable stellar parameters.

In this section we derive the stellar parameters for the same stars as in the previous one, now using the near-infrared spectra secured by Hanson et al. (2005). We adopt again the rotational and macroturbulent velocities from (Simón-Díaz & Herrero 2014 ;note that these values are derived from optical observations) and discard Br12 for the analysis, as justified earlier.

The results of the near-infrared analysis are given in Table 4.5. We see that again β and microturbulent velocities could not be well constrained, indicating that for these stars the infrared is not better suited than the optical for this task. This indicates that the difference in line formation depth between the optical and the H- and K-band spectra is not sufficient to provide new information, at least at the resolution and S/N ratio of the spectra analyzed here. The comparison of the observed profiles with those from the best fit model is presented in Figure 4.10. Inspection of these profile fits suggests us the following comments:

- The best fit quality is again obtained for the dwarfs, but now not without significant problems. The best fitted lines are HeII λ1.69 and HeI λ1.70μm. HeII λ2.18 on the other hand shows acceptable profile fits. Br10 and Br11 also fit well, with small discrepancies in the wings, although they show a bad fit to the HD15629 observations (particularly Br11). This does not seem to be a global problem, as the rest of the dwarfs do not show it. However, Brγ is not well fitted in any of the cases. The fast rotators, particularly HD 217 086, show a profile for this line different from the other dwarfs, with a strong and relatively narrow absorption in the blue half-line and a broad absorption in the red wing.

- The giant fast rotators show similar Brγ profiles as the dwarf fast rotators, and a similar poor fit. Figure 4.9 offers a zoom-in into this problem that points to some problem with the line formation not considered in our models. The fit to Br10 and Br11, however, is much better. The slower rotating giant, HD 190 864, shows also a good fit for Br10 and Br11 and a poor fit for Brγ, although without the characteristic shape of the fast rotators. The fit to the He lines is of varying quality. We find a perfect fit for the HeI λ1.70μm line of the slow rotator, which is poorer for the fast one (that also displays a small central reversal). The HeII lines, on the other hand, fit better for the fast rotator, being of modest quality for the slower one. Globaly the fits are again acceptable, except for the Br$_\gamma$ line.

- The Br lines in supergiants show a poor fit quality, particularly Brγ, except, surprisingly, for HD 30 614 (that had the largest problems in the optical) and to a lesser extent, for the low luminosity one, HD 209 975. The third late type supergiant, the non-radial pulsator HD 210 809, shows a poor fit. The early-type supergiants show the poorest fit to the Br spectrum, with the models predicting an absorption profile for Brγ while we observed an emission one. The only exception to the poor fit is the Br11 line of HD 15 570. The He lines are also poorly fitted in the late-type supergiants. The HeI λ1.70μm line departs strongly from a good fit when the luminosity increases (in the low luminosity supergiant, the fit is good) whereas HeII λ2.18μm is always stronger than predicted and HeII λ1.69μm

is very faint but only fairly reproduced. The situation is different in the early-type supergiants, where the fits to the HeII lines are far from perfect, but acceptable.

Table 4.5: Stellar parameters obtained from the near-infrared analysis. Helium abundances are given by number, with Y(He)= N(He)/N(H). Note that the Y(He) values in the table are multiplied by 100, and that log Q values are actually negative

#	Star ID	Teff(kK)	logg(dex)	−logQ	Y(He)	micro (km/s)	beta
1	HD46223	41.2 ± 1.5	3.78 ± 0.10	12.7 ± 0.2	< 9.5	> 5.0	> 0.8
2	HD15629	39.9 ± 2.3	3.73 ± 0.26	13.0 ± 1.3	< 10.5	12.4 ± 7.4	< 1.2
3	HD46150	39.7 ± 1.1	3.87 ± 0.12	12.9 ± 0.3	< 7.8	< 18.8	> 0.8
4	HD217086	36.9 ± 1.2	3.88 ± 0.12	13.0 ± 1.2	13.9 ± 6.5	> 5.0	> 0.8
5	HD149757	32.8 ± 0.9	3.69 ± 0.25	12.6 ± 0.2	14.8 ± 5.6	13.6 ± 6.2	> 0.8
6	HD190864	36.8 ± 1.0	3.66 ± 0.13	12.7 ± 0.2	> 12.6	13.6 ± 6.3	> 0.9
7	HD203064	34.5 ± 1.5	3.69 ± 0.30	12.5 ± 0.2	15.0 ± 5.4	> 5.0	> 0.9
8	HD15570	38.8 ± 1.9	3.55 ± 0.15	11.9 ± 0.1	9.7 ± 3.3	< 19.9	< 1.0
9	HD14947	41.5 ± 1.3	3.84 ± 0.15	12.2 ± 0.2	> 14.4	> 5.0	< 1.1
10	HD30614	27.6 ± 0.9	2.77 ± 0.09	12.1 ± 0.1	< 15.8	> 5.0	1.0 ± 0.2
11	HD210809	34.1 ± 1.3	> 3.37	12.5 ± 1.4	> 12.9	> 10.1	> 0.8
12	HD209975	32.7 ± 1.0	3.38 ± 0.12	> 12.6	> 11.9	> 15.8	> 0.8

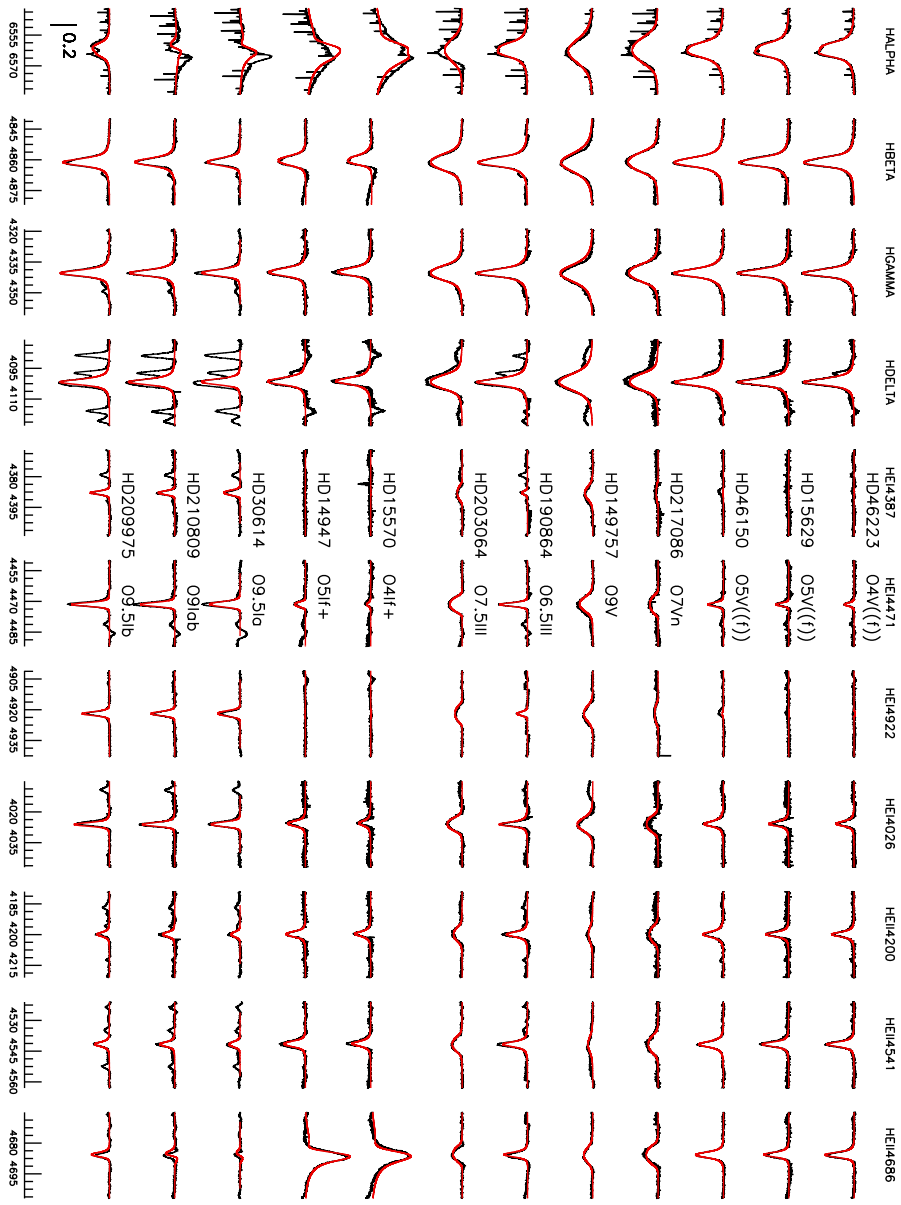

Figure 4.7: Spectral fits for the optical lines. Observations are shown in black and best fit model profiles in red. The abscissa gives the wavelength for each line, and the scale of the ordinate axis is given by the vertical bar at the lower left corner (at the bottom of the Hα column).

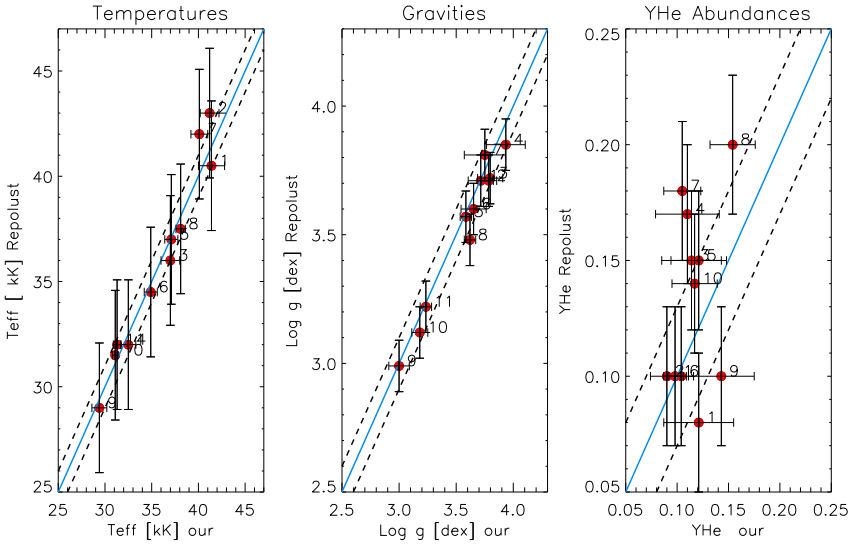

Figure 4.8: Comparison between the stellar parameters obtained in the optical by Repolust et al. (2004) and our work. Left panel: effective temperature; the dashed lines represent ± 1000K. Middle panel: logaritmic gravity; the dashed lines represent ±0.1[*dex*]. Right panel: helium abundance by number; the dashed lines represent ±0.03. Numbers indicate the stars as listed in Tab. 4.1.

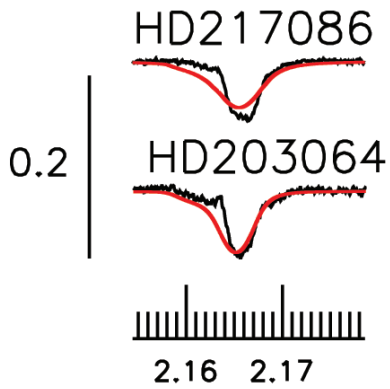

Figure 4.9: behavior. of Brγ for the dwarf and giant fast rotators HD 217086 and HD 203 064. Both show a small blue emission peak close to the core of the line, resulting in a distorted blue wing.

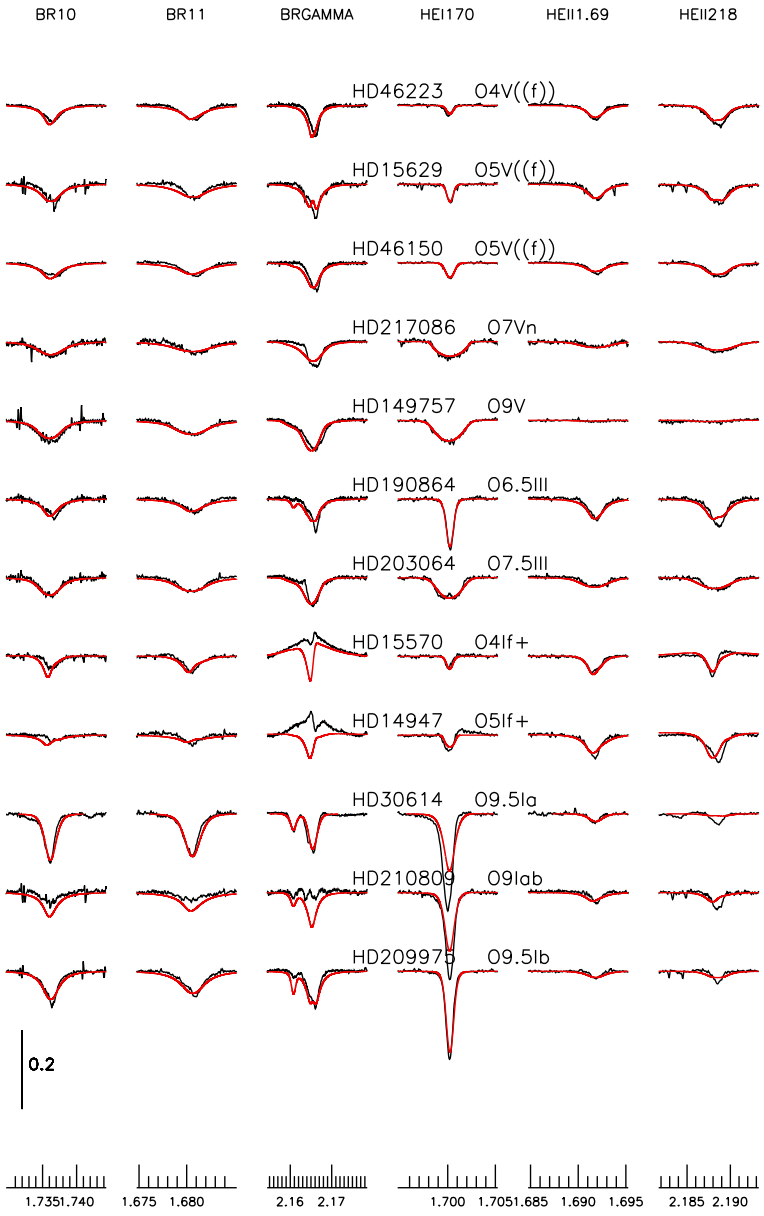

Figure 4.10: Spectral fits for the near-infrared lines. Observations are shown in black and best fit model profiles in red. The abscissa gives the wavelength for each line, and the scale of the ordinate axis is given by the vertical bar at the lower left corner (at the bottom of the HeI λ1.70μm column). Unlike the optical case, the infrared best fit models present several lines that are not well fitted, specially for the giant and supergiant stars.

We compare again with previous results in the literature, namely those from Repolust et al. (2005). These authors analyzed the infrared spectra of the same stars that Repolust et al. (2004) analyzed in the optical. Figure 4.11 illustrates the differences with our results. The similarity with the comparison in the optical presented in Figure 4.8 is clear, with a good agreement for effective temperatures and gravities and some trend for our He abundances towards lower values (also the log Q values, not shown in the figure, present a good agreement). However, we see now a larger dispersion in the comparison. We attribute this not to the effect of the improvements in FASTWIND, but to the differences in the by-eye and automatic techniques. Because the profile fits are better in the optical, the difference between the two methods is small. When the line fits are poorer, the subjective weigth given to a particular fit increases, pushing the result in a given direction, whereas the automatic procedure still forces a compromise for all considered profiles.

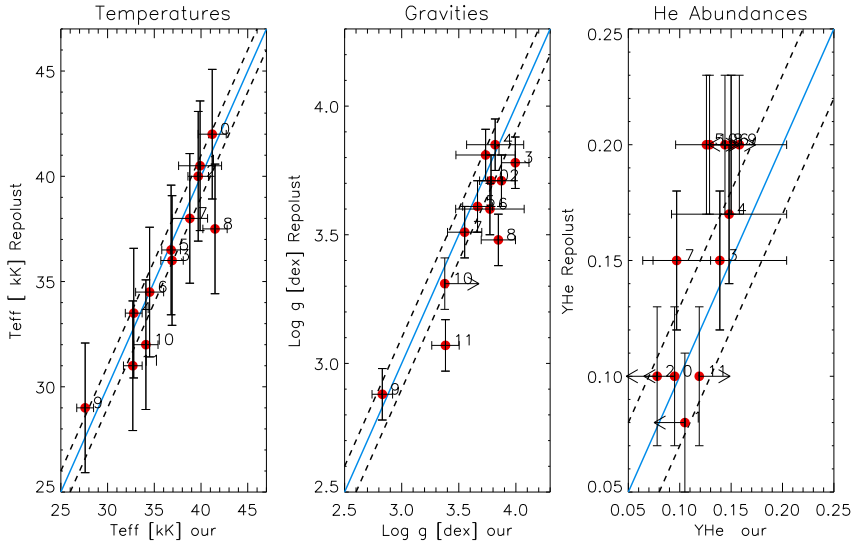

Figure 4.11: Comparison between the stellar parameters obtained in the infrared by Repolust et al. (2005) and our work. Left panel: effective temperature; the dashed lines represent ± 1000K. Middle panel: logaritmic gravity; the dashed lines represent ±0.1[*dex*]. Right panel: helium abundance by number; the dashed lines represent ±0.03.

The most relevant comparison is that of the optical versus infrared derived parameters, as this will indicate us how reliable can be our parameters when derived only from the infrared (except for the rotational and macroturbulent velocities, that we adopted from the optical spectrum). This comparison can be seen in Fig 4.12. The global agreement is fair within the errors, as shown by the mean values of the differences: $< \Delta T_{\rm eff} >= -180 \pm 1950$ K, $< \Delta log\, g >= -0.06 \pm 0.16$

dex and $< \Delta log\, Q >= -0.22 \pm 0.53$ dex. Moreover, the largest differences in log Q are found for HD 217 086 and HD 149 757. Both are fast rotating dwarf stars, which may imply that the Q parameter is particularly difficult to determine for these objects that combine a weak wind effect with a large rotational broadening. For the helium abundances the agreement is poorer ($< \Delta YHe >= -0.02 \pm 0.02$) but in this case the statistics is not as good due to the large number of upper or lower limits present in the results. Although these numbers are indicate that we can use the infrared spectra to determine stellar parameters in a similar way as we do with the optical ones, we observe some trends that have to be taken into account in future analyses.

Globally, the infrared fits are worse than the optical ones, which is reflected in larger uncertainties. These increase in 500 K for the effective temperature and in 0.07 and 0.17 dex for log g and log Q, although the last values are dominated by just a couple of stars with much larger uncertainties in the infrared (HD 203 064 for log g, HD 210 809 for log Q and HD 15 629 for both of them). Without these objects, the mean errors only increase in 0.04 dex for log g and 0.05 dex for log Q. Nevertheless, the conclusion that errors from the infrared are systematically larger remains. Moreover, inspection of the χ^2 distributions from *iacob_gbat* and the fits from Figure 4.10 indicates that this is not due to the difference in the resolution and S/N ratio of the optical and infrared spectra, but is a consequence of a less accurate modelling of the infrared lines. This is different from the result quoted by Repolust et al. (2005) who found comparable errors in both wavelength ranges, reflecting the different treatment of the error determination in both cases and also the different procedure for the fitting. Finally we may cite the relatively large number of objects for which only upper or lower limits could be obtained for the helium abundance, reflecting a lack of sensitivity of the infrared spectrum to that parameter (or a degeneracy because of the larger uncertainties involved).

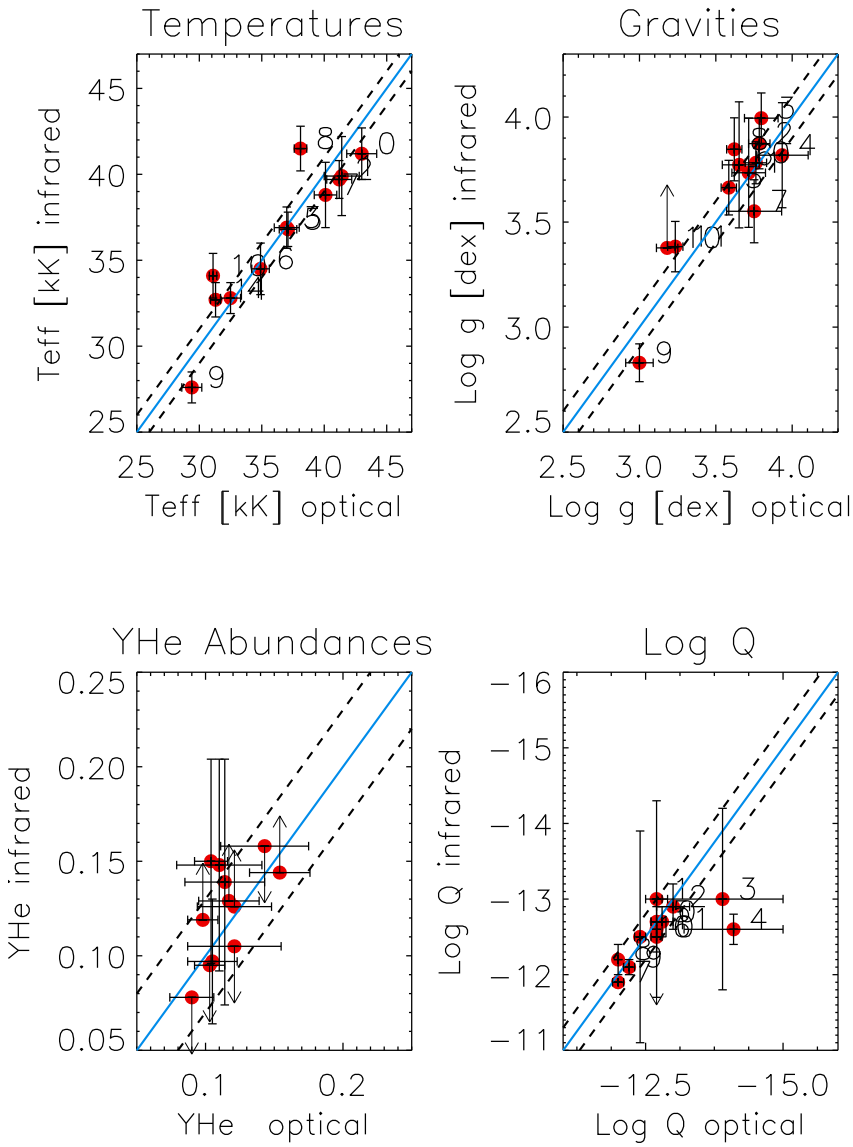

Figure 4.12: Comparison between the stellar parameters obtained in this work in the optical and in the near infrared. Upper left panel: effective temperature; the dashed lines represent ± 1000K. Upper right panel: logaritmic gravity; the dashed lines represent ±0.1[*dex*]. Lower left panel: helium abundance by number; the dashed lines represent ±0.03. Lower right panel: wind strength parameter, log Q; the dashed lines represent ±0.3[*dex*].

4.4.3 Determination of V sini and Vmac in the near-infrared

The $V \sin i$ and $Vmac$ used to broaden the profiles in the infrared analysis were taken from the optical determinations (Simón-Díaz & Herrero 2014). This may be a reason for the relatively poor fits of the near-infrared spectrum. If the actual rotational velocity is different for the near-infrared lines this may contribute to poorer fits in the near-infrared and larger differences in the parameter derived from the near-infrared compared with the optical.

Moreover, the purpose of this work is to check to what extent the quantitative analysis can be carried out using only near-infrared spectra (e.g. for the embedded cluster candidates obtained from Chapter 2, where no optical observations are possible and the only information is provided by near-infrared spectrum). To test if we obtain similar parameters as the well measured V sini and Vmac from the optical, we derived those values using *iacob_broad* (Simón-Díaz et al. 2011) on near-infrared lines.

We notice that the values of rotational and macroturbulence velocities were derived in the optical using metal lines, whose broadening is dominated by those processes. However, the metal near-infrared lines are too weak in our spectra and of relatively poor quality. For this reason we are forced to use the HeI lines, which are affected by the quadratic Stark effect, limiting our ability to measure the rotational velocity for slow rotators (or the macroturbulent velocity when this is low). H and HeII lines are even less well suited, as they are dominated by the strong linear Stark effect. Thus we decided to use the HeI $\lambda 1.70\mu$m line, which is strong enough for all the stars.

We also must take two additional points into account. First, Herrero et al. (1992) and Ramírez-Agudelo et al. (2013) established that optical HeI lines give a good estimation of the rotational velocity for V sini $\geq 100 km/s$. We adopt this value as well. The Stark effect is larger for larger wavelengths but also the rotational broadening is proportional to wavelength (as are Doppler shifts), and thus we expect the optical value to be a good approximation for the infrared. On the other hand, Simón-Díaz & Herrero (2014) showed that, for large rotational velocities the Vmac velocity is very difficult to determine, especially when the line wings are contaminated.

Figure 4.13 compares the projected rotational velocities obtained from both wavelengths ranges (filled circles), whereas Table 4.6 gives the numerical values. In the figure dashed lines indicate the region that departs ±30 km s^{-1} or ±30% (whatever is larger) from the 1:1 relationship. This band marks the region where stellar parameters are not affected beyond errors by changes in the adopted rotational velocity Sabín SanJulián (2014). We see that the differences in V sini are always within this band. In fact, most points agree reasonably well, except for some low V sini values from the optical. The only really departing point, at V sini (opt) =115 and Vsini(near-infrared)=78 km s^{-1} correspons to HD 30 614, with a really strong HeI $\lambda 1.70\mu$m line in absorption, which prevents an accurate V sini determination. This is related to the large value found V_{mac}. As anticipated, Vmac departs strongly from the 1:1 relation for some objects, especially for fast rotators. An important note is that the lowest V sini obtained from the IR is $\sim$ 80 km s^{-1}. While for the fast rotators the large V_{mac} value is a minor problem because the broadening is still dominated by rotation, for HD 30 614 (and to a less extent, for HD 201 809) the large values of V_{mac} may alter the derived rotational velocity, introducing problems during the line fitting process.

In spite of these difficulties, the stellar parameters resulting from the near-infrared fit do

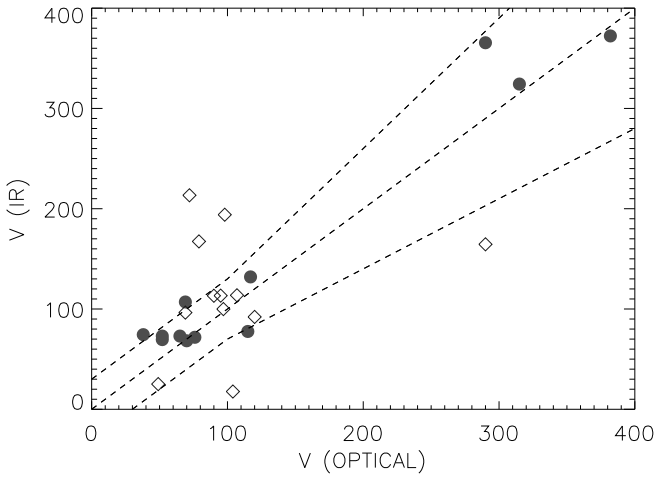

Figure 4.13: V sini (filled circles) and Vmac (open diamonds) values obtained from the optical metal lines and from the HeI λ1.70μm line in the infrared. The dashed lines give the band ± 30 km/s (for low vsini) or 30% of vsini (optical), whatever is larger.

Table 4.6: Comparison between V sini and Vmac values obtained from the optical metal lines and from the near-infrared HeI λ1.70μm line.

#	Star ID	Type	V sini OP	Vmac OP	V sini near-infrared	Vmac near-infrared
1	HD46223	O4 V((f))	52	97	69.5 ± 4.3	99.9 ± 4.3
2	HD15629	O4.5 V((fc))	70	69	68.3 ± 4.0	96.4 ± 4.0
3	HD46150	O5 V((f))z	69	107	107.0 ± 6.7	113.7 ± 6.7
4	HD217086	O7 Vnn((f))z	382	104	372.3 ± 17.7	17.7 ± 17.7
5	HD149757	O9.2 IVnn	290	290	365.5 ± 18.3	164.5 ± 18.3
6	HD190864	O6.5 III(f)	65	90	73.0 ± 3.7	113.2 ± 3.7
7	HD203064	O7.5 IIIn((f))	315	98	344.3 ± 17.2	103.3 ± 17.2
8	HD15570	O4 If	38	120	74.4 ± 3.5	92.1 ± 3.5
9	HD14947	O4.5 If	117	49	132.0 ± 6.2	25.1 ± 6.2
10	HD30614	O9 Ia	115	72	77.6 ± 6.5	213.4 ± 6.5
11	HD210809	O9 Iab	76	79	71.8 ± 4.8	167.4 ± 4.8
12	HD209975	O9.5 Ib	52	95	73.0 ± 2.5	113.3 ± 2.5

not change significantly when the rotational and macrotrubulent velocities are changed to the new values (as expected following Sabín-Sanjulián et al. (2014), because the differences in the

combined rotational and macroturbulent broadening are relatively small). Thus we conclude that it is possible to derive all stellar parameters from the near-infrared spectrum only, with just slightly larger uncertainties than from optical spectra (that can be reduced for comparable resolution and S/N spectra).

4.5 Clumping

Massive stars have strong winds driven by line scattering of the intense stellar ultra-violet (UV) radiation field. This wind may present small instabilities resulting in an inhomogeneous outflow (*clumping*, e.g., Hillier & Miller (1998); Crowther et al. (2002); Hillier et al. (2003); Bouret et al. (2003); Puls et al. (2008)). These density inhomogeneities (*clumps*) modify the shape of the spectral lines formed in the wind. If the clumps formed are optically thin (which is known as $micro - clumping$) they particularly affect the emission in lines formed through recombination processes. These are the processes producing the emission in H_α or the near-infrared infrared lines used as wind diagnostics. In these processes, emission is proportional to ρ^2 (where ρ is the density of the wind). The wind mass-loss rate is obtained by deriving the wind density from the total emission in the diagnostic lines (that measures the total number of ion-electron encounters and is thus proportional to ρ^2). In homogeneous wind, the total emission gives us ρ^2 and therefore $\rho = \sqrt{\rho^2}$. In an inhomogeneous wind, the total emission will be proportional to $< \rho^2 >$. But in this inhomogeneous wind, $< \rho > \neq \sqrt{< \rho^2 >}$. In fact, $< \rho^2 >$ will be larger than $< \rho >^2$ and thus a clumped wind produces more emission for the same mean density. Alternatively, for an observed emission line, we obtain a lower mass-loss rate when adopting a clumped wind. Moreover, as clumping may be radially dependent, it may affect differently lines formed in different layers in the atmosphere, which may help to explain the differences found in the previous sections when fitting optical and near-infrared lines (partially).

We shall mention here the possibility that clumps are optically thick. This situation, known as $macro - clumping$ or *porosity*, affects processes proportional to density (like resonance UV lines) and will not be treated here, as except for extreme conditions, clumps are optically thin for optical and near-infrared line processes (e.g., Sundqvist et al. (2011)).

The wind is characterized by the clumping factor, defined as

$$F = \frac{< \rho^2 >}{< \rho >^2} \geq 1 \tag{4.5}$$

As long as F is spatially constant, the emission in lines like H_α will be equivalent for two stars having $\dot{M}_{unclumped}$ and $\dot{M}_{clumped}$ with

$$\dot{M}_{clumped} = \frac{\dot{M}_{unclumped}}{\sqrt{F}} \tag{4.6}$$

In our clumping study we have considered three different clumping laws as radial functions. Firstly, we consider a linear increase of the clumping factor from 1 to a maximum value between two points in the atmosphere. After reaching the maximum, the clumping factor remains constant. We call this the "Linear law". The second law is the one proposed by Hillier et al. (2003)

("Hillier law") in which the clumping factor follows an exponential increase until it reaches a maximum (and then keeps constant). Finally, the third law is based on Najarro et al. (2011) ("Najarro law") and is similar to Hillier law, but the clumping factor decreases exponentially after the maximum. Fig 4.14 illustrates the different laws.

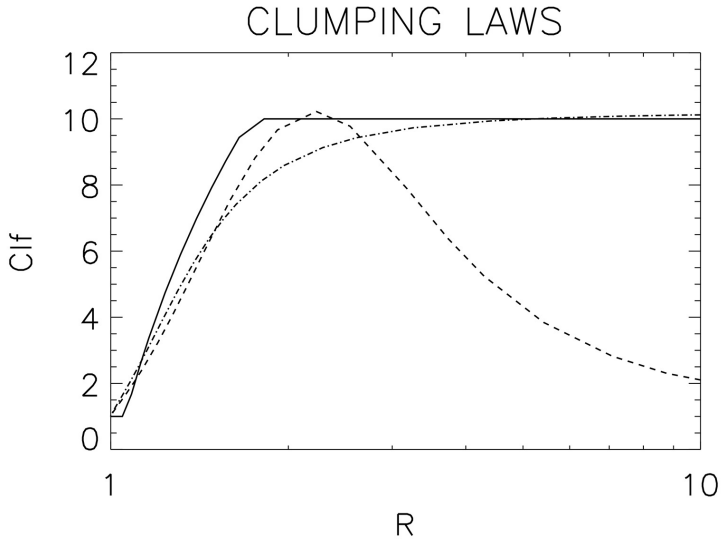

Figure 4.14: Comparison between the clumping laws: solid line is an example for the Linear law; dash-dotted line, Hillier law; dashed line, Najarro law. We plot the laws identified as Linear_05, Hillier_200 and Najarro_200 in Tab. 4.7

The laws are implemented in FASTWIND by defining some parameters. The clumping factor is labeled as F rising to the maximum level (F_{max}). The three laws use the following parameters:

- Linear law: it is characterized by three parameters, $F_{max}, v_1 \text{and} v_2$:

$$F(v, v_1, v_2) = 1 + (F_{max} - 1) \times \left(\frac{v(r) - v_1}{v_2 - v_1} \right); \text{for } v1 \leq v \leq v_2 \tag{4.7}$$

  where F_{max} is the maximum value for the clumping factor, v_1 is the wind velocity at the point where clumping begins, and v_2 that of the point where the clumping factor reaches its maximum. Before v_1 the clumping factor is set to unity, and after v_2 is set to F_{max}.

- Hillier law requires two parameters and is expressed as

$$f = f_\infty + (1 - f_\infty)e^{-\left(\frac{v(r)}{v_{cl}}\right)} \tag{4.8}$$

where f is the so-called filling factor, the inverse of the clumping factor (i.e., $F = \frac{1}{f}$, assuming the interclump medium is void). The two parameters defining the function are f_∞, the filling factor when the wind velocity reaches the terminal velocity v_∞ (which could be interpreted as $1/F_{max}$ in our tests) and v_{cl} which marks the point where clumping begins to become important and controls how fast the function reaches its maximum. In this law, clumping begins to increase directly from the bottom of the phtosphere (but will only be significant from $v = v_{cl}$ on).

- Najarro law requires

$$f(v) = f_\infty + (1 - f_\infty)e^{-\left(\frac{v(r)}{v_{cl1}}\right)} + (1 - f_\infty)e^{\frac{v_\infty - v(r)}{v_{cl2}}} \tag{4.9}$$

where f_∞ is the same as in Hillier's law and v_{cl1} has the same meaning as v_{cl}, whereas v_{cl2} defines how fast the clumping factor decreases after reaching its maximum.

Table 4.7 gives the different values adopted for our tests. We have considered five different linear laws (the reason will be explained later), two forms of the Hillier law and one of the Najarro law. The linear law is characterized by F_{max}, v_1 and v_2 defined above. For v_1 and v_2 we adopt values based on the findings by Najarro et al. (2011), who obtain different values for these parameters in different stars. We adopt a compromise and fix our values to a given fraction of the terminal velocity. This results in values (in km/s) within the ranges obtained by Najarro et al. The F_{max} parameter presents a large range in the literature, ranging from 1.0 (unclumped) to values as high as 100. We will test F_{max}= [10, 20], following Table 2 in Najarro et al. (2011). However, we warn that this approach has the drawback of considering that all stars will follow the same clumping law, not only in its radial distribution but also in the absolute values adopted, especially F_{max}. With a high probabibility, this is incorrect and thus the present work has an exploratory character. Once we understand how the profile fits and the derived stellar parameters react to clumping, we will be in a good position to consider F_{max} as a free parameter in future works. Similarly, the Hillier and Najarro laws are characterized by f_∞, v_{cl1} and v_{cl2} (in the case of Najarro law) and we adopt values resulting in similar clumping distributions to those of the linear law (in the case of the Najarro law, until the maximum clumping factor is reached). We also give the F_{max} that would correspond to each case. We point out here that the laws will result in similar but not identical runs of the clumping factor with radius (as can be seen in Figure 4.14).

Notice that for Najarro law, the combination of f_∞ v_{cl1} & v_{cl2} may provide a different clumping profile. In order to recreate a similar Linear profile (reaching the maximum clumping factor around 2 R_*, and also following the linear slope as is possible see in Figure 4.14), we addapt these parameters, especially f_∞, to represent a very likely Linear profile. Same behaviour but with less impact over f_∞ for Hillier.

Table 4.7: The clumping laws used in our analyses

Clumping Law / label	Clumping Factor	v_1/v_∞	v_2/v_∞	
$Linear_{015}$	F_{max} = 10.	0.1	0.15	
$Linear_{025}$	F_{max} = 10.	0.1	0.25	
$Linear_{050}$	F_{max} = 10.	0.1	0.50	
$Linear_{20-040}$	F_{max} = 20.	0.1	0.40	
$Linear_{20-094}$	F_{max} = 20.	0.1	0.94	
Clumping Law	Maximum Factor	Filling factor	Control Parameter 1	Control parameter 2
$Hillier_{100}$	F_{max} = 10.5	f_∞ = 0.095	v_{cl1} = 100.	v_{cl2} = 0
$Hillier_{200}$	F_{max} = 10.5	f_∞ = 0.095	v_{cl1} = 200.	v_{cl2} = 0
$Najarro_{200}$	F_{max} = 10.	f_∞ = 0.075	v_{cl1} = 200.	v_{cl2} = 100.

Table 4.8: Model notation of the (unclumped) coarse grid. Names contain the effective temperature, logarithmic gravity, helium abundance by number (using $\epsilon = \frac{N(He)}{N(H)+N(He)}$), the logarithmic wind strength parameter Q and the β exponent of the wind velocity law

	V	III	I
HOT	T420g400He09Q145b08	T420g380He09Q135b08	T420g360He09Q124b08
MID	T360g400He09Q145b08	T360g370He09Q135b08	T360g340He09Q124b08
COOL	T300g400He09Q145b08	T300g340He09Q135b08	T300g300He09Q124b08

4.5.1 FASTWIND coarse grid

Before analyzing their effects on real stars we will test how clumping behaves for different template models. We construct a coarse grid of models representing dwarfs, giants and supergiants at different temperatures (hot, mid and cool). Thus we have a small grid of 9 models covering the parameter range of O-stars, where we will explore the effects of clumping. In Figure 4.15 we have plotted the models on the $T_{\rm eff} - \log g$ diagram to illustrate the evolutionary stages represented. Table 4.8 lists the coarse grid generated with their identification labels, which contain the parameters used. All models have the same helium abundance and β value. We will calculate the models in the coarse grid with the adopted clumping law in each exercise. Note that, as explained below, when using clumped models the $\log Q$ identification will change.

4.5.2 Clumping versus no clumping

First we explore the effects of clumping on our coarse grid. Clumping will modify the transfer of radiation and the occupation numbers of the atomic levels, thus altering the ionization equilibrium of all elements and parameters derived from model fits. Its most conspicous effect will be to increase the emission in lines like H_α. To obtain a similar emission we must divide the unclumped mass-loss rate by $\sqrt{F}$ (in our case this is an approximation, due to the radial

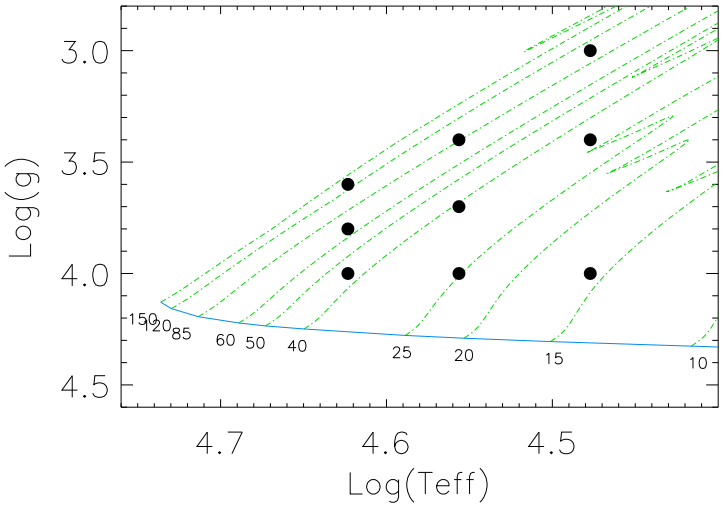

Figure 4.15: Models present in the FASTWIND coarse on the $T_{\rm eff} - \log g$ diagram. Models are chosen to represent from hot dwarfs to cool supergiants. In green similar tracks than Figure 4.2

dependence of F). This means that the wind strength unclumped Q-value ($Q = \frac{\dot{M}}{(R_* v_\infty)^{1.5}}$) will be equivalent to a clumped value $Q' = \frac{\dot{M}'}{(R_* v_\infty)^{1.5}} = \frac{\dot{M}}{\sqrt{F}(R_* v_\infty)^{1.5}}$. Figure 4.16 illustrates the strong effects that clumping may have on the H_α line emission. The figure shows the H_α profile obtained for the grid model with $T_{\rm eff}$= 36 000 K, $\log g$ = 3.40 and (unclumped) $\log Q$= -11.90 (which corresponds to the "mid-temperature supergiant" model). For the clumped model we have used the $Linear_{050}$ law (see Tab. 4.7 with $\log Q'$ = -12.4, i.e., $\log Q' = \log Q + \log \sqrt{F}$) to have an equivalent emission. Thus in our tests unclumped models will be compared with clumped counterparts having the same stellar parameters except for a mass-loss rate and $\log Q$ values reduced by a factor of $\sqrt{F}$ and $\log \sqrt{F}$ respectively (which will be reflected in the model name identification).

Figure 4.17 shows the differences between the unclumped and clumped (scaled one) in other spectral lines in the optical and near-infrared wavelength ranges. We see that our choice of parameters results in a very similar clumped/unclumped H_α emission, but other lines react differently. We notice a clear increment in the near-infrared emission for the H lines, even Brγ is affected with a significant emission enhancement in the center of the line core. Unlike near-infrared H lines, Hβ presents more absorption in the core (Hα remains mostly unchanged as selected). HeI lines are strongly affected in both wavelength ranges, with a stronger absorption

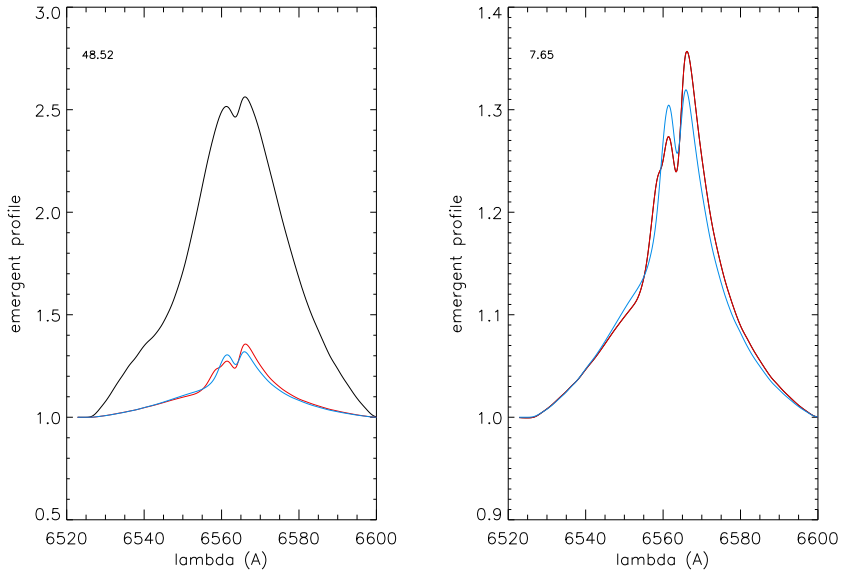

Figure 4.16: Effect of introducing the clumping in a model with $T_{\rm eff}$= 36 000 K, $\log g$ = 3.40 and (unclumped) log Q= -11.90 (mid-temperature O-supergiant). In red, the H_α profile for the model without clumping, in black the one with a clumping factor F_{max}=10 and $[v_1, v_2]$ = [0.1, 0.5] and the same mass-loss rate (i.e., the same Q-value) and in blue the clumped model with a mass-loss rate decreased by $\sqrt{F_{max}}$ (i.e., with a log Q value smaller by 0.5). Profiles have been broadened by V sini= Vmac= 50 km s^{-1} and R= 12 000. Left panel: all three models; right panel: zoom-in into the lower emission profiles. The number at the left top give the equivalent width of the black and red models

when clumping is introduced. HeII wind lines behave similarly to H wind lines, in the near-infrared HeII λ1.69 and λ2.18μm show an increased emission in the center of the line and in the optical HeII λ4686 remains mostly unchanged, as does H_α. This behaviour can be understood realizing that the dominant helium ion is HeIII, and thus HeII lines behave much as H lines when they are dominated by recombination processes. Note that in cooler winds, when HeIII is not the dominant helium ionization stage, HeII λ4686 will behave differently to H_α (see Kudritzki et al. (2006)). For other lines (or for the H and HeII lines under different conditions), the line formation regions will be modified as consequence of the different density structure in the clumped models. Thus, the increased absorption of certain lines can be explained by their formation in inner layers, before clumping plays a significant role. In those cases, the dominant effect will be the decreased mass-loss rate in the clumped model, resulting in a deeper absorption. The

stronger emission at the core of Br_γ and HeII $\lambda 2.18\mu$m in the clumped model as compared to the unclumped one may be then understood as clumping playing a more significant role (i.e., more emission) due to the difference in the formation layers at those wavelengths, without an exact compensation through the F_{max} parameter.

The models presented in Figure 4.16 and Figure 4.17 show the strongest effects in our coarse model grid. In general, supergiants (hot, mid, and cool) display the strongest effects. For giants we observe smaller changes, and they become negligible for dwarfs.

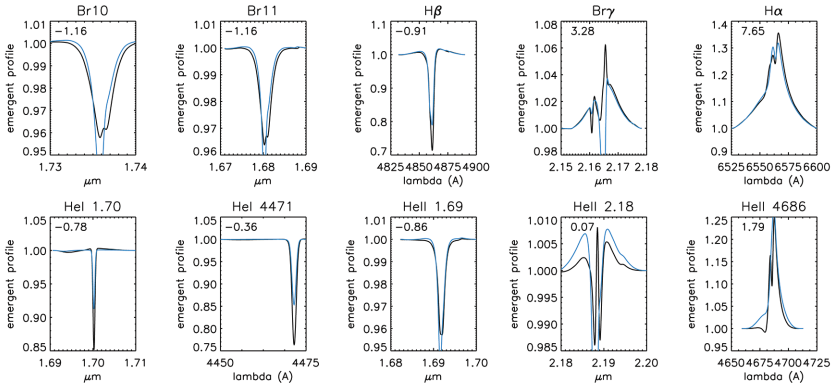

Figure 4.17: Example of the effects introduced in the spectral lines by clumping for the same model as in Figure 4.16, and with the same line broadening. Black is here clumped model with scaled mass-loss and blue the unclumped model (in Figure 4.16 they are, respectively, the blue and red profiles).

4.5.3 Which clumping law to use

Calculation of a full model grid is an expensive task. Thus, before analyzing the real spectra, we performed a series of tests using the coarse grid to see the differences among the clumping laws considered in Table 4.7. Our aim is to minimize the required calculations with the full grid.

Hillier vs Najarro

We first compare the $Hillier_{200}$ with $Najarro_{200}$ clumping laws (see Table 4.7, Figure 4.14 and the formal expressions). We see that the main difference between the laws is for the outer wind layers, after we have reached the maximum clumping factor in $Najarro_{200}$ law. After this, the clumping factor decreases to reach 1.0 (no clumping) for $Najarro_{200}$, whereas it keeps increasing asymptotically for Hillier's law to reach the maximum at the outer wind boundary.

The comparison of the line profiles obtained using these two laws shows no significant effects, in spite of the (small) differences between the laws in the inner wind (see Figure 4.14).

Figure4.18 shows the differences for H_α and Br_γ for the supergiant models. These are the models showing the largest effects, and we see that they are very similar. Giants and dwarfs show no significant differences. This can be clearly seen in Figs. 4.19 and 4.20 that show selected H-He lines for the whole grid. Inspection of these figures confirms the conclusion that there are no significant differences when using Hillier or Najarro law for the analysis of optical and near-infrared H-He spectra. The reason is simply that the lines have already been formed when the differences between both laws begin to be important. Near-infrared and optical lines are formed below $2R_*$ and the influence of the clumping law far from this point is not relevant. Nevertheless, using far-IR or radio wavelengths the lines could be affected by the outer clumping conditions and our conclusions may change. But for our purposes in the present work we can restrict ourselves to one of these clumping laws.

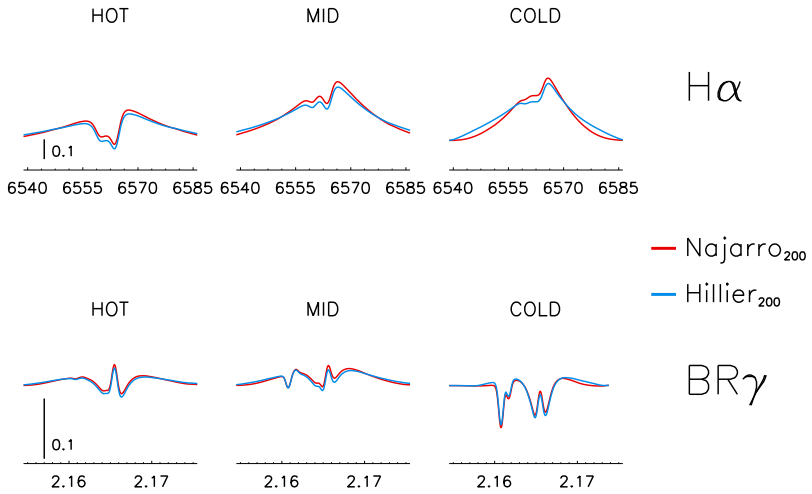

Figure 4.18: H_α and Br_γ profiles for supergiant models using $Najarro_{200}$ (blue) and $Hillier_{200}$ (red) laws. Profiles have been broadened as in Figure 4.16

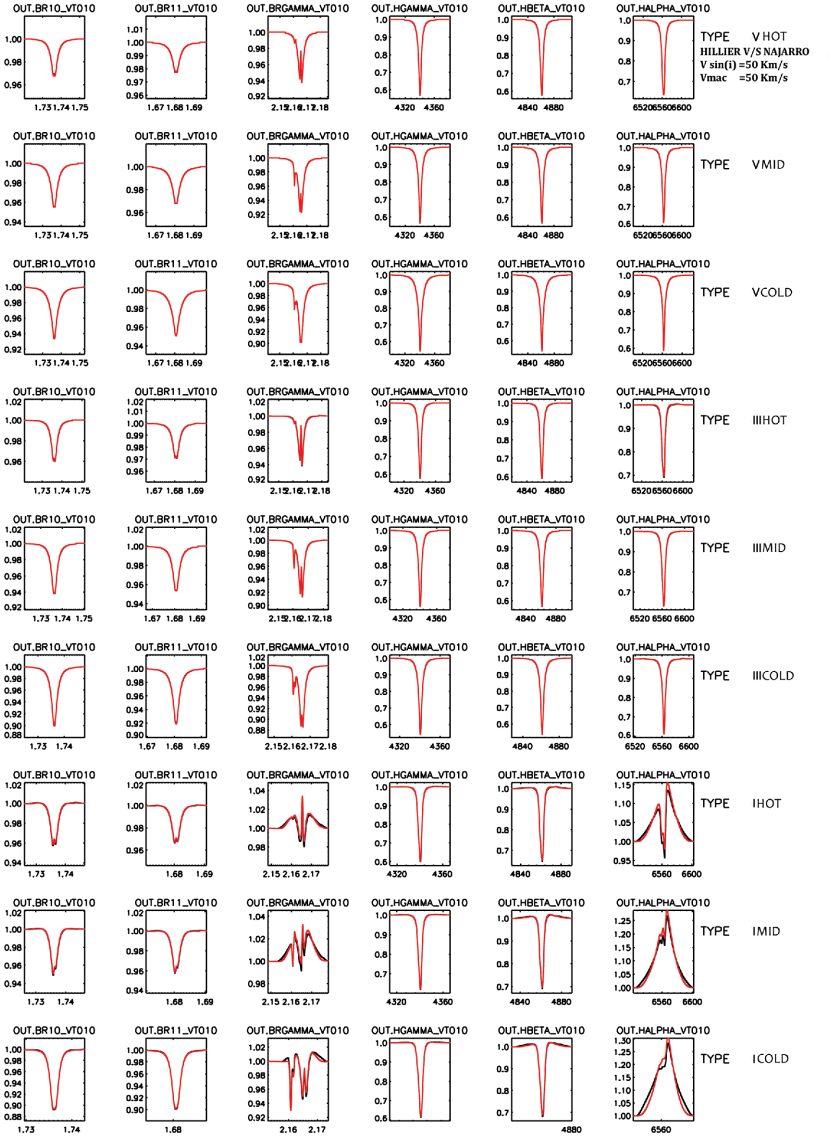

Figure 4.19: Comparison of selected H profiles obtained with the $Najarro_{200}$ (red) and $Hillier_{200}$ (black) laws in the corase model grid. Profiles have been broadened as in Figure 4.16

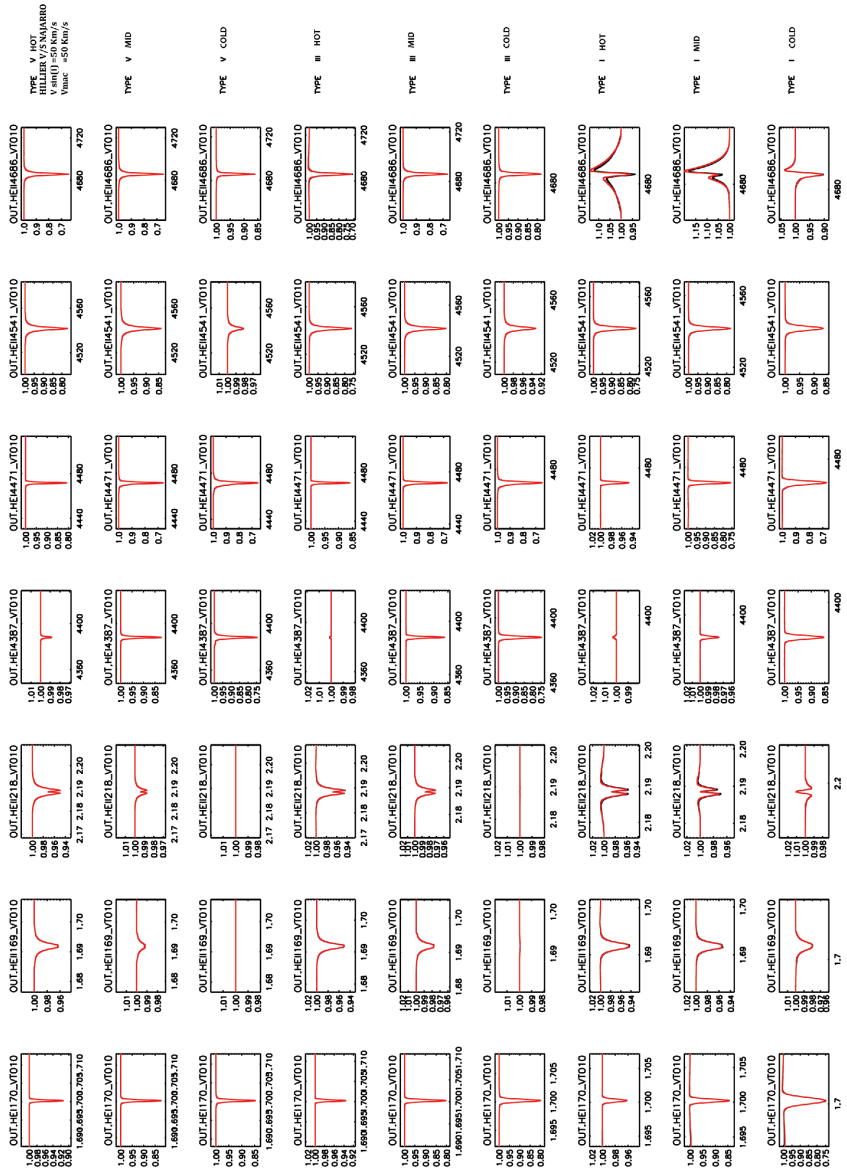

Figure 4.20: Comparison of selected H profiles obtained with the $Najarro_{200}$ (red) and $Hillier_{200}$ (black) laws in the corase model grid. Profiles have been broadened as in Figure 4.16

Hillier vs Linear

After concluding that the use of Hillier or Najarro laws does not result in any significant difference, we compare now the profiles obtained from the $Hillier_{200}$ and $Linear_{050}$ laws (see Table 4.7). Observing again Figure 4.14, we see that the differences between these two laws are now a bit larger than those between Hillier and Najarro ones, although in inner layers, where the optical and near-infrared lines are formed, they are very similar. It is thus not surprising that the largest differences between line profiles, shown in Figure 4.21, are again modest. These are found once more for the supergiant models, particularly the cool supergiant one. Again, no significant differences are found for models representing giants and dwarfs.

The small differences found can actually be compensated with small changes in the law parameters. This is shown in Figure 4.22 where we represent the same lines, now for the $Hillier_{100}$ and $Linear_{025}$ laws. We see that the profile differences are of minor importance. Thus we conclude that to study the clumping effects it is sufficient to consider one family of clumping laws, as the differences are mainly due to the relation between the line formation and the clumping distribution. Therefore we decided to use the simple linear law.

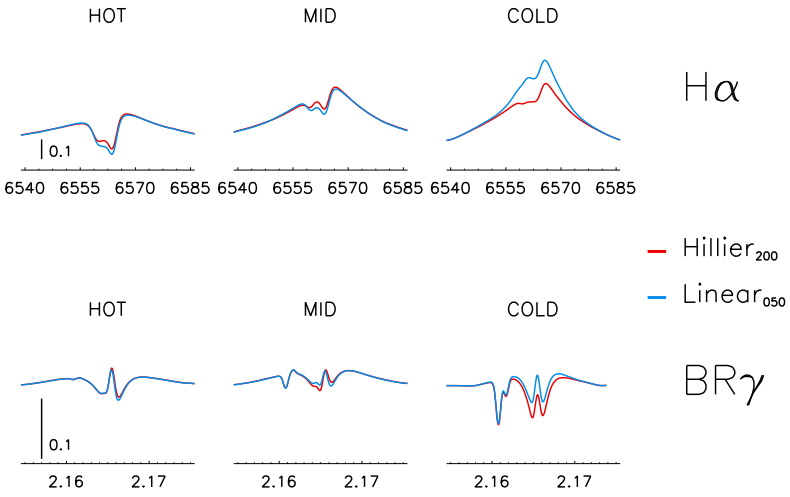

Figure 4.21: H_α and Br_γ profiles for supergiant models using $Hillier_{200}$ (red) $Linear_{050}$ (blue) laws. Profiles have been broadened as in Figure 4.16

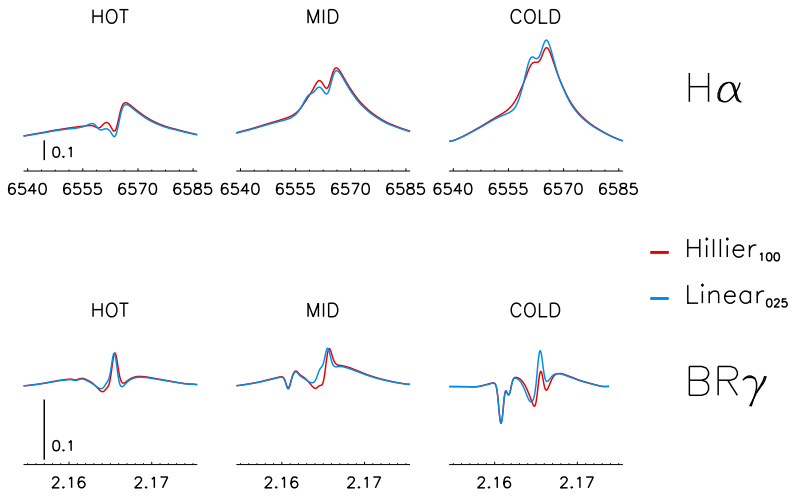

Figure 4.22: H_α and Br_γ profiles for supergiant models using $Hillier_{100}$ (red) $Linear_{025}$ (blue) laws. Profiles have been broadened as in Figure 4.16. Different as figure 4.21, the clumping reach the maximum closest to the photosphere

4.5.4 Which clumping profile to use

Once we have decided to use the linear law we explore the changes introduced by the different parameters, in particular the maximum clumping factor, F_{max}, and the point where this maximum is reached, v_2. We keep fixed the point where clumping starts, given by v_1, as it has no real impact on the results as far as it is sufficiently small (i.e., a factor 0.1-0.2 of the terminal velocity).

We consider two values for F_{max}, 10 and 20. For $F_{max} = 10$ we choose two values for the point where this maximum clumping factor is reached, v_2: $0.5\times v_\infty$ and $0.25\times v_\infty$. For $F_{max} = 20$ we choose the v_2 values in such a way that the inner runs of the clumping laws overlap until $F_{max} = 10$ is reached. This results in $v_2 = 0.40 \times v_\infty$ and $v_2 = 0.94 \times v_\infty$. Figure 4.23 shows the four linear laws, that we name $Linear_{050}$, $Linear_{025}$, $Linear_{20-040}$ and $Linear_{20-094}$. Note that the law is linear in velocity, whereas the figure plots the clumping factor versus the radius, which curves the distribution.

We will compare the line profiles obtained for the supergiants, as they show the largest effects and we have seen that they are very small for giants and dwarfs. Figure 4.24 shows the H_α and Br_γ profiles of the $Linear_{025}$ and $Linear_{20-040}$ laws, i.e., those in which the maximum clumping factor is reached in the inner wind layers. The figure also shows the profiles obtained without clumping (note that the mass-loss rates of the clumped models have been reduced by the corresponding factor $\sqrt{F_{max}}$). We see that the H_α profiles are similar for the unclumped and $Linear_{025}$ models, whereas the profile of the $Linear_{20-040}$ is a bit different for the hot and mid-temperature supergiants, with less emission in the later case. This indicates that H_α is formed in a region where the clumping is fully compensating for the lower adopted mass-loss rate in $Linear_{025}$ (i.e., beyond $v(r) = 0.25 \times v_\infty$), but where this is not yet the case for the $Linear_{20-040}$ law. Thus we conclude that the differences between both clumped models are due to the formation of H_α between $v(r) = 0.25 \times v_\infty$ and $v(r) = 0.40 \times v_\infty$.

Opposite to this behavior, the Br_γ profiles of the clumped models show strong emission, very similar to each other, presenting an increased emission with respect to the unclumped model. We conclude from here that Br_γ is formed in layers where clumping has the same influence for both laws (integrated through the line formation region in the wind). The higher density makes Br_γ appear in absorption in the unclumped model. Note the strong HeI component appearing in the blue wing of the Br_γ profiles in cooler models.

Figure4.25 presents the same comparison as 4.24, now for the clumping laws $Linear_{050}$ and $Linear_{20-094}$. Here clumping increases more slowly with radius, and thus H_α is formed before the maximum clumping factor is reached. As a consequence, the decrease in mass-loss rate and wind density produces a lower emission when considering the clumped models. The effect is stronger for the $Linear_{20-094}$ because of the larger decrease in mass-loss. Br_γ, being formed in outer layers, presents the same effect but to a lower extent. As it is closer to reach the region where the clumping fully compensates for the decreased mass-loss rate of the model, their profiles are more similar to the unclumped models than those of H_α.

A global comparison is shown in the Fig 4.26, with H_α and Br_γ shown for all models at once. We see that ;

1. we obtain more emission when the maximum clumping factor is reached in inner layers.

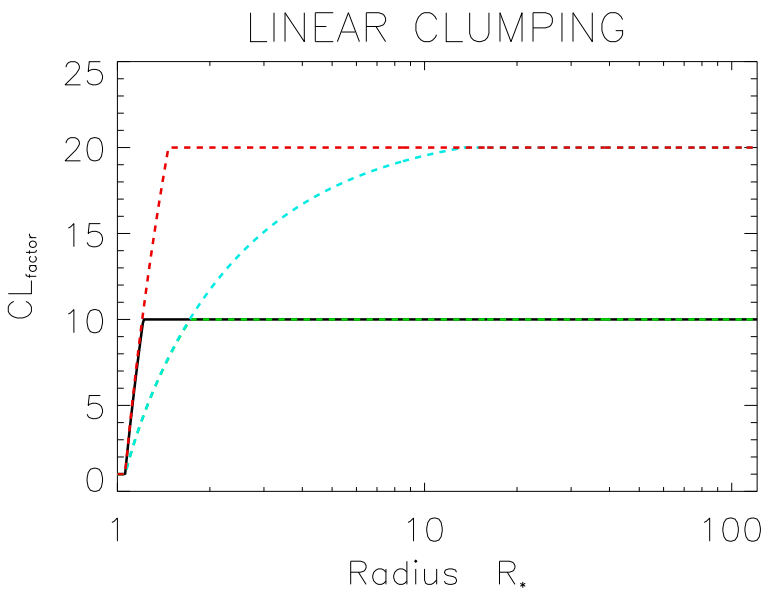

Figure 4.23: The different linear clumping laws considered for comparison. Shown is the clumping factor as function of the stellar radius (in units of the photospheric radius, R_*). Black solid line: $Linear_{025}$; dashed green line: $Linear_{050}$; dashed red line: $Linear_{20-040}$; dashed cyan line: $Linear_{20-094}$. Clumping starts at $v_1 = 0.1 \times v_\infty$ in all cases.

This way lines are formed when clumping already fully compensates for the decrease in mass-loss rate in the models.

2. when the clumping factor increases along a very extended region, its maximum value F_{max} is of less relevance. What matters is the value of the clumping factor in the line forming region, together with the global mass-loss rate.

For these reasons, and given the exploratory character of the present work, we will set $F_{max} = 10$ for the rest of our work. Again, we remember that the choice of a unique F_{max} value is a limitation of the present work, as this is most likely not realistic (see values in Table 2 of Najarro et al. (2011) and the differences in the tests above). In future work a more extended parameter range will be examined. However, we will consider the two laws $Linear_{025}$ and $Linear_{050}$ as we see that the point where F_{max} is reached may have a strong influence.

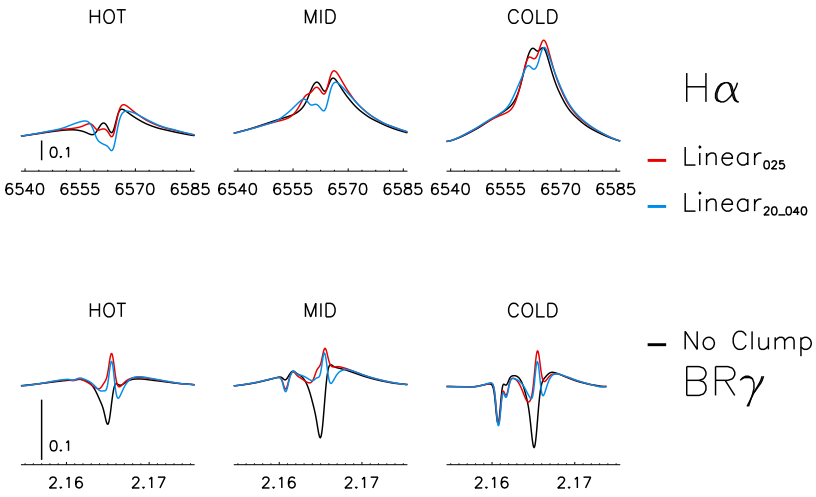

Figure 4.24: H_α and Br_γ profiles for $Linear_{025}$ (red) and $Linear_{20-040}$ (blue). In black, the profiles for the model without clumping. Profiles have been broadened as in Figure 4.16

4.6 FASTWIND Clumping Grid

Now we can calculate a full grid of models to compare with observations. For the purposes of the present work, with an exploratory character, we choose to set $F_{max} = 10$. Again, we remember that the choice of a unique F_{max} value is a limitation of the present work, as this is most likely not realistic (see values in Table 2 of Najarro et al. 2011 and the differences in our previous tests). We will consider two linear clumping laws, reaching F_{max} at 0.25 and 0.50 v_∞. These two grids were created in the same way as our previous optical-near-infrared grid. Table 4.9 shows the parameter ranges for these two grids.

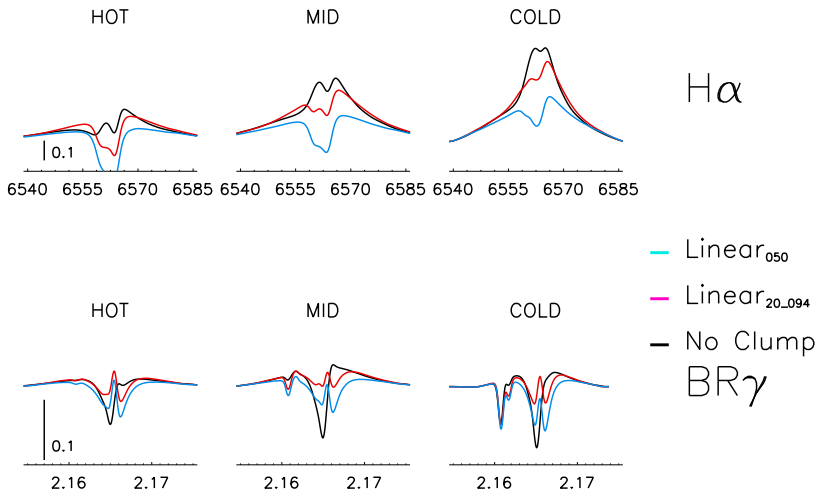

Figure 4.25: H_α and Br_γ profiles for $Linear_{050}$ (blue) and $Linear_{20-094}$ (red). In black, the profiles for the model without clumping. Profiles have been broadened as in Figure 4.16

Table 4.9: Range of the parameters used to produce the grid of syntetics models.

Parameter	Range of value
T_{eff}	55000 ≥ 22000 K (step 1000 K)
log(g)	[2.6-4.3] (step 0.1 dex)
Micro	5,10,15,20 km s^{-1}
ϵHe	0.06, 0.09, 0.13, 0.17, 0.20, 0.23
log(Q)	-15.5, -14.5, -13.5, -13.2, -13.0, -12.8, -12.6, -12.4, -12.2
β	0.8, 1.0, 1.3

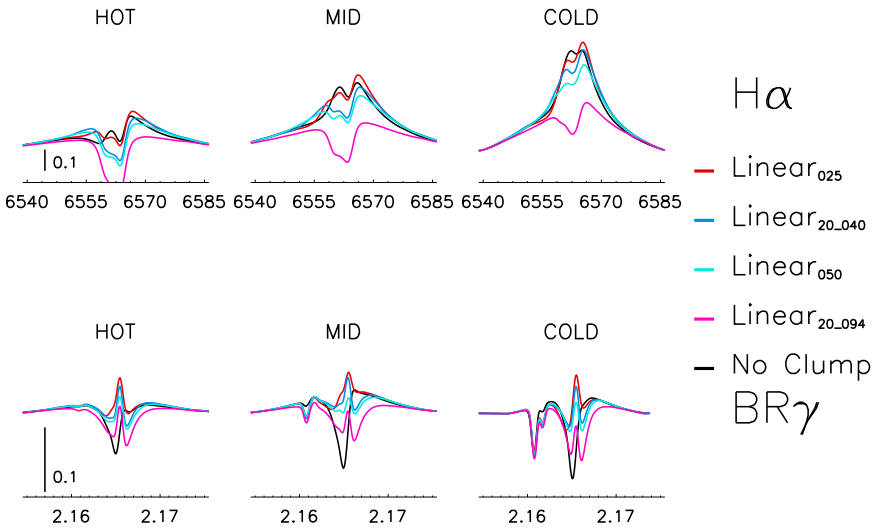

Figure 4.26: H_α and Br_γ profiles for $Linear_{025}$ (red), $Linear_{050}$ (blue), $Linear_{20-040}$ (cyan), $Linear_{20-094}$ (pink) and the model without clumping (in black). Profiles have been broadened as in Figure 4.16

4.6.1 Analysis with the $Linear_{025}$ clumping law

We have analyzed again our stellar sample, now including clumping. We find the results of the analyses with the *iacob_gbat* tool for the $Linear_{025}$ law in Tables 4.10 (for the optical spectrum) and 4.11 (for the near infrared). The corresponding fits can be seen in Figs. 4.28 and 4.29. In the figures we can see that the global fits have a similar quality to those of unclumped models (especially for giants and dwarfs, see Figs. 4.7 and 4.10), but there are some changes that are worth mentioning (we point out here that the parameters of the clumped and unclumped models are different; thus, the changes will not only be due to clumping, but also to the parameter change produced by it).

For the hot supergiant stars we observe two main changes in the optical. The first one is a distinct improvement in the fit of Hα, particularly in the case of HD15570 and HD14947, with a broader emission line (see Figure 4.27). A similar improvement is not seen for the H_β line. The second one is a deeper absorption in HeII λ4541 that improves the fit for the hot supergiants. However the good fit for HeII λ4686 without clumping is now a bit worse. The cool supergiants do not present the same global improvement in H_α, but there is a partial improvement. Moreover, the He lines, particularly HeII λ4686 also improve slightly, including a correction in the apparent shift in the line core between the observations and the unclumped profile. This differential behaviour in HeII λ4686 in hot and cool winds is expected because of the change in the dominant ionization stage of helium, as we explained earlier (see Kudritzki et al. (2006) for more details) and strengthens our warning about the use of a single clumping law for all stars. We conclude that the $Linear_{025}$ improves H_α for the hot supergiants and improves the agreement between H_α and HeII λ4686 for the cool supergiants with P-Cygni profiles (but without reaching a good fit).

In the near-infrared, the fits to the spectra of the hot supergiant HD 15 570 and the cool one HD 210 809 improve greatly for Br_γ. The rest of the line fits also improve slightly in these stars, except HeII λ2.18μm that is now significantly worse. Unlike for the optical spectra now we have changes in the line fits to giants and dwarfs, that may slightly improve (e.g., Br_{10} in HD 46 223 and HD 149 757 or HeII λ2.18μm in HD 190 864), slightly worsen (e.g., Br_γ in the same HD 46 223 or HeII λ2.18μm in HD 15 629) or not change significantly. Particularly interesting is the improvement in the Br_γ fit of HD 149 757, for which we found a large discrepancy between log Q in the optical and the infrared (see Figure 4.12; the same, but to a much less extent can be seen in the other fast rotator dwarf, HD 217 086). There is a remarkable bad fit to the HeI λ1.70μm line in the cool supergiants HD 30 614 and HD 210 809, both in models with and without clumping. Thus in the near-infrared the major improvement of using the $Linear_{025}$ law is for the Br_γ line of some supergiants, but with a significant quality loss in the fit of HeII λ2.18μm.

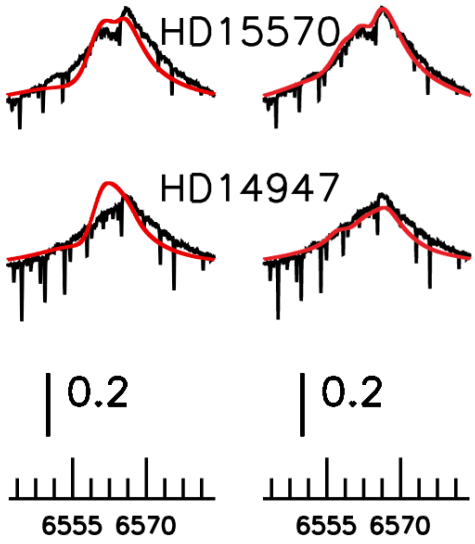

Figure 4.27: Fits to the Hα line in the hot supergiants when using the unclumped model grid (left) and the *Linear*$_{025}$ clumping law (right).

Table 4.10: Stellar parameters obtained from the optical analysis using the *Linear*$_{025}$ clumping law. Helium abundances are given by number, with Y(He)= N(He)/N(H). Note that the Y(He) values in the table are multiplied by 100, and that log Q values are actually negative

Star	Teff(kK)	log g(dex)	$-\log Q$ (cgs)	Y(He)	micro (km/s)	beta
HD46223	43.4 ±0.9	3.83 ± 0.07	13.1 ± 0.1	10.3 ± 1.2	10.2 ± 5.2	> 1.0
HD15629	42.3 ±1.8	3.78 ± 0.10	13.1 ± 0.1	12.0 ± 3.4	12.4 ± 7.4	>1.0
HD46150	40.0 ±0.8	3.80 ± 0.08	13.4 ± 0.2	9.6 ± 1.2	<11.8	>0.8
HD217086	37.0 ±1.0	3.60 ± 0.10	13.5 ± 1.1	11.2 ± 2.7	12.4 ± 7.4	1.0 ±0.2
HD149757	32.5 ±0.9	3.82 ± 0.17	14.0 ± 1.0	11.0 ± 3.2	12.0 ± 7.0	>0.8
HD190864	37.2 ±0.8	3.60 ± 0.10	13.1 ± 0.1	12.2 ± 3.1	10.4 ± 5.4	>0.8
HD203064	35.0 ±0.5	3.50 ± 0.06	13.1 ± 0.1	10.2 ± 1.1	>13.7	1.0 ±0.2
HD15570	39.8 ±0.6	3.48 ± 0.07	12.4 ± 0.1	10.3 ± 1.3	<19.9	1.1 ±0.1
HD14947	38.0 ±0.2	3.50 ± 0.03	12.5 ± 0.1	14.4 ± 2.4	<11.3	>1.2
HD30614	29.1 ±0.2	<2.83	12.6 ± 0.1	> 19.5	>18.4	1.1 ±0.1
HD210809	31.0 ±0.8	3.05 ± 0.12	12.7 ± 0.1	> 13.2	>14.9	1.1 ±0.2
HD209975	31.5 ±0.6	3.26 ± 0.09	13.1 ± 0.2	10.2 ± 1.3	<12.1	1.0 ±0.2

Table 4.11: Stellar parameters obtained from the near-infrared analysis using the $Linear_{025}$ clumping law. Helium abundances are given by number, with Y(He)= N(He)/N(H). Note that the Y(He) values in the table are multiplied by 100, and that log Q values are actually negative

Star	Teff(kK)	log g(cgs)	$-\log Q$ (cgs)	Y(He)	micro (km/s)	beta
HD46223	42.7 ±1.7	3.83 ± 0.10	14.1 ± 1.4	< 10.2	>5.0	< 1.3
HD15629	40.8 ±1.2	3.85 ± 0.10	13.0 ± 1.3	9.5 ± 3.1	<19.9	> 0.8
HD46150	39.5 ±0.8	3.85 ± 0.11	13.1 ± 0.2	<8.1	12.1 ± 7.1	>0.9
HD217086	36.8 ±1.1	3.88 ± 0.11	14.2 ± 1.3	13.4 ± 6.9	>5.0	> 0.8
HD149757	32.5 ±1.6	3.52 ± 0.22	>13.3	13.4 ± 7.0	>9.4	< 1.3
HD190864	37.5 ±1.0	3.85 ± 0.10	12.9 ± 0.1	>16.1	>11.8	> 1.1
HD203064	35.6 ±0.9	3.87 ± 0.06	12.8 ± 0.1	15.3 ± 5.1	>8.4	> 1.2
HD15570	39.1 ±0.3	3.52 ± 0.03	12.4 ± 0.1	< 8.1	<19.9	1.1 ±0.1
HD14947	41.9 ±1.0	>3.94	12.6 ± 0.1	>14.7	>5.0	> 1.1
HD30614	28.5 ±0.6	<2.90	12.7 ± 0.1	< 9.1	>10.1	1.0 ± 0.2
HD210809	34.6 ±1.3	>3.43	12.6 ± 0.1	> 7.5	>5.0	> 0.8
HD209975	32.5 ±1.0	3.39 ± 0.13	> 13.6	15.1 ± 5.3	>14.8	> 0.8

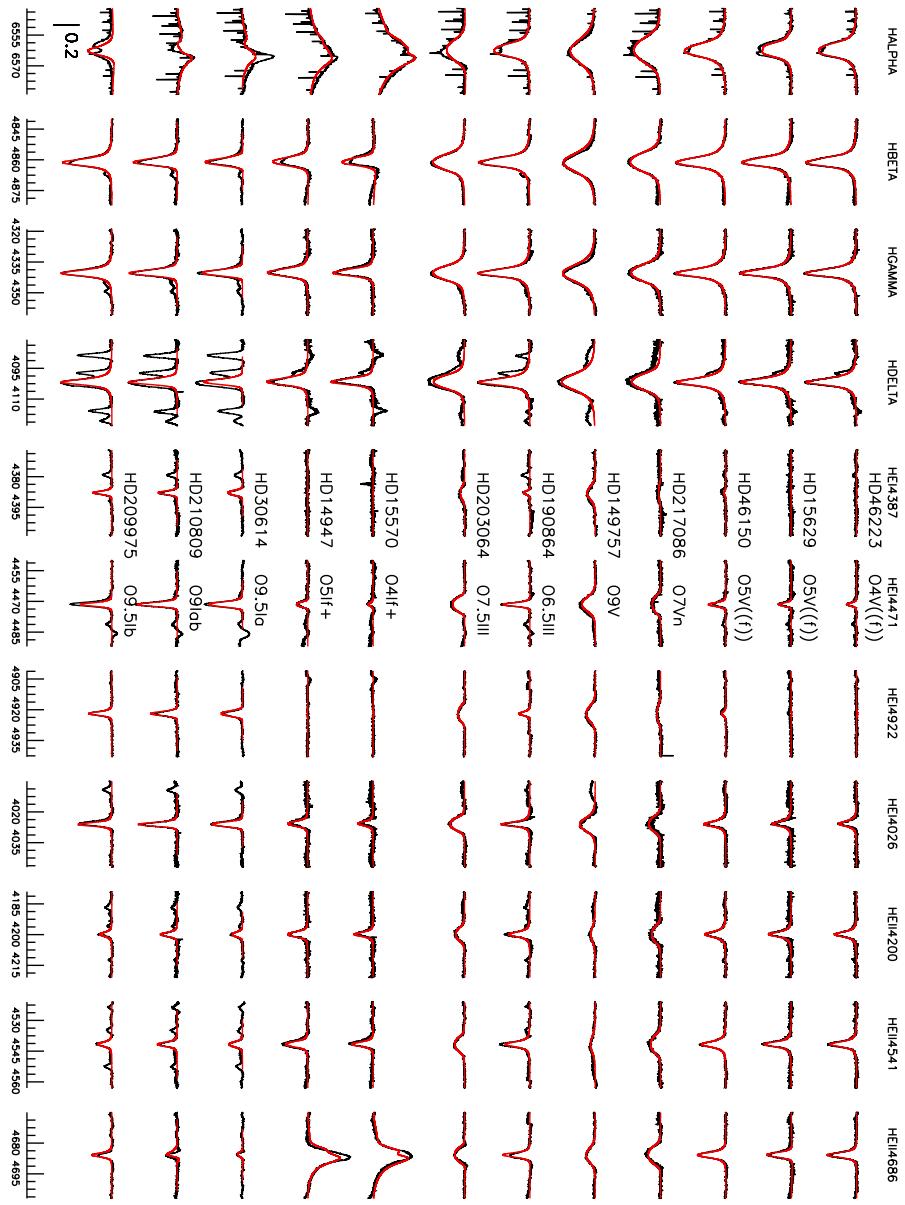

Figure 4.28: Spectral fits for the optical lines using the clumping law $Linear_{025}$. Observations are shown in black and best fit model profiles in red. The abscissa gives the wavelength for each line, and the scale of the ordinate axis is given by the vertical bar at the lower left corner (at the bottom of the Hα column).

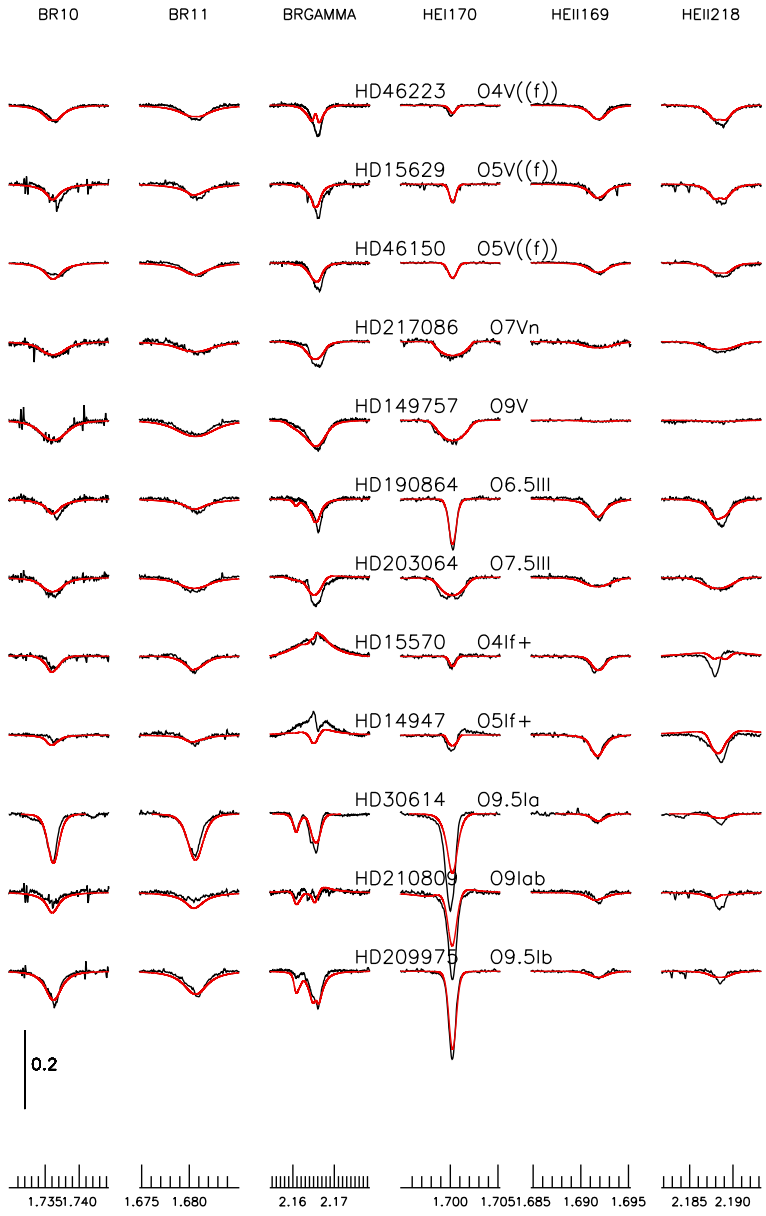

Figure 4.29: Spectral fits for the near-infrared lines using the clumping law $Linear_{025}$. Observations are shown in black and best fit model profiles in red. The abscissa gives the wavelength for each line, and the scale of the ordinate axis is given by the vertical bar at the lower left corner (at the bottom of the HeI $\lambda 1.70 \mu$m column).

4.6.2 Analysis with the $Linear_{050}$ clumping law

The results from the analysis of our stellar sample with the $Linear_{050}$ law (that reaches the maximum clumping factor F_{max} further out than the $Linear_{025}$ from the previous section) can be found in Tables 4.12 (for the optical spectrum) and 4.13 (for the infrared). The corresponding best fits for the optical and infrared lines can be seen in Figures 4.30 and 4.31. As in the case of the $Linear_{025}$ law, we emphasize here the main changes found when using the $Linear_{050}$ one.

As in the case of the $Linear_{025}$ law, there is an improvement in the fit of the H_α profiles of hot supergiants, that in fact show excellent agreement in all optical lines. The cool supergiants, HD 209 975 (that had an acceptable fit without clumping) and HD 30 614 (that had a poor fit to H_α and HeII λ4686) do not change significantly. But now we see an much better fit for H_α and HeII λ4686 for the cool supergiant HD 210 809 (and other He lines also show improved fits), indicating that the clumping distribution is more extended in this star than represented by the $Linear_{025}$ law. For the rest of the sample (giants and dwarfs), we get similar good fits with and without clumping.

In the near-infrared the fits to Br_γ of HD 15 570 and HD 210 809, as well as HD 149 757 improve again significantly compared to the unclumped models (again a bit less for HD 217 086), but again with poorer fits to HeII λ2.18μm for the former two. In general, the fits to Br_γ and other H and He near-infrared lines but HeII λ2.18μm improve slightly. The exception is the HeI λ1.70μm line in the hot dwarf HD 46 223. This line remains very poorly fit in the cool supergiants HD 30 614 and HD 210 809. The global line profile fits in the near-infrared are of similar quality and do not seem to be affected by the different distributions in the clumping factors.

Our conclusion is that globally, clumping sometimes has a positive impact on the fits of H_α and HeII λ4686 of supergiants. The impact may be dependent on the particular clumping law chosen, which indicates that more work is needed to determine the actual clumping distribution in these stars. Unlike in the optical, the near-infrared does not show any differential improvement when using one or the other clumping law. There is an overall improvement for near-infrared lines when using clumped models (particularly in Br_γ) except for HeII λ2.18μm.

Table 4.12: Stellar parameters obtained from the optical analysis using the $Linear_{050}$ clumping law. Helium abundances are given by number, with Y(He)= N(He)/N(H). Note that the Y(He) values in the table are multiplied by 100, and that log Q values are actually negative

Star	Teff(kK)	log g (cgs)	− log Q (cgs)	Y(He)	micro (km/s)	beta
HD46223	43.2± 0.8	3.80 ±0.06	13.1 ± 0.1	10.3±1.2	>5.0	>1.0
HD15629	42.1± 1.4	3.77 ±0.09	13.0 ± 0.1	12.0± 2.9	12.3 ± 7.3	>1.2
HD46150	40.0± 0.9	3.79 ±0.08	13.4 ± 0.3	9.6± 1.3	<19.9	1.0 ± 0.2
HD217086	37.0± 1.0	3.60 ±0.11	13.5 ± 1.2	11.3± 2.8	12.4 ± 7.4	1.0 ± 0.2
HD149757	32.5± 0.9	3.83 ±0.17	13.5 ± 1.0	11.0± 3.1	12.1 ± 7.1	<1.3
HD190864	37.0± 0.5	3.59 ±0.06	13.0 ± 0.1	11.4± 2.0	14.7 ± 3.8	>1.0
HD203064	35.0± 0.9	3.53 ±0.12	13.0 ± 0.1	10.3± 1.3	> 13.9	>0.8
HD15570	38.5± 0.9	3.45 ±0.09	12.4 ± 0.1	11.4± 1.9	12.4 ± 7.4	>1.3
HD14947	37.3± 0.9	3.35 ±0.08	12.4 ± 0.1	>14.0	12.4 ± 7.4	>1.0
HD30614	29.2± 0.3	<2.85	12.5 ± 0.1	>18.9	>17.9	>1.1
HD210809	31.0± 0.9	3.08 ±0.18	12.7 ± 0.1	>9.8	14.1 ± 5.8	1.1 ± 0.2
HD209975	31.4± 0.5	3.23 ±0.05	13.0 ± 0.1	10.0± 1.2	8.5 ± 3.5	>1.0

Table 4.13: Stellar parameters obtained from the near-infrared analysis using the $Linear_{050}$ clumping law. Helium abundances are given by number, with Y(He)= N(He)/N(H). Note that the Y(He) values in the table are multiplied by 100, and that log Q values are actually negative

Star	Teff(kK)	log g (cgs)	− log Q (cgs)	Y(He)	micro (km/s)	beta
HD46223	45.7 ± 1.2	3.82± 0.06	12.8± 0.1	<8.4	<19.9	>1.1
HD15629	40.2 ± 1.1	3.87± 0.11	13.0± 0.2	9.2 ± 2.5	<19.9	>1.1
HD46150	39.2 ± 0.7	3.83± 0.09	13.0± 0.2	<7.8	9.9 ± 4.9	> 0.9
HD217086	36.7 ± 1.0	3.87± 0.12	13.0± 0.2	14.1± 6.2	> 5.0	>1.1
HD149757	32.5 ± 1.6	3.52± 0.22	>13.2	13.4± 7.0	> 9.4	<1.3
HD190864	37.0 ± 1.0	3.86± 0.12	12.9± 0.1	16.7± 3.6	> 9.4	>1.2
HD203064	35.5 ± 1.0	3.87± 0.05	12.8± 0.1	14.5± 5.9	>5.0	>1.2
HD15570	38.3 ± 0.8	3.59± 0.05	12.4± 0.0	<9.9	<19.9	>1.2
HD14947	40.8 ± 2.1	3.86± 0.14	12.4± 0.0	>15.7	> 5.0	1.0 ±0.1
HD30614	28.5 ± 0.6	<2.90	12.6± 0.1	<9.2	> 6.9	>0.8
HD210809	34.4 ± 1.5	>3.34	12.6± 0.2	> 7.8	> 7.7	>0.8
HD209975	32.5 ± 1.0	3.39± 0.13	>13.6	15.1± 5.3	>14.8	>0.8

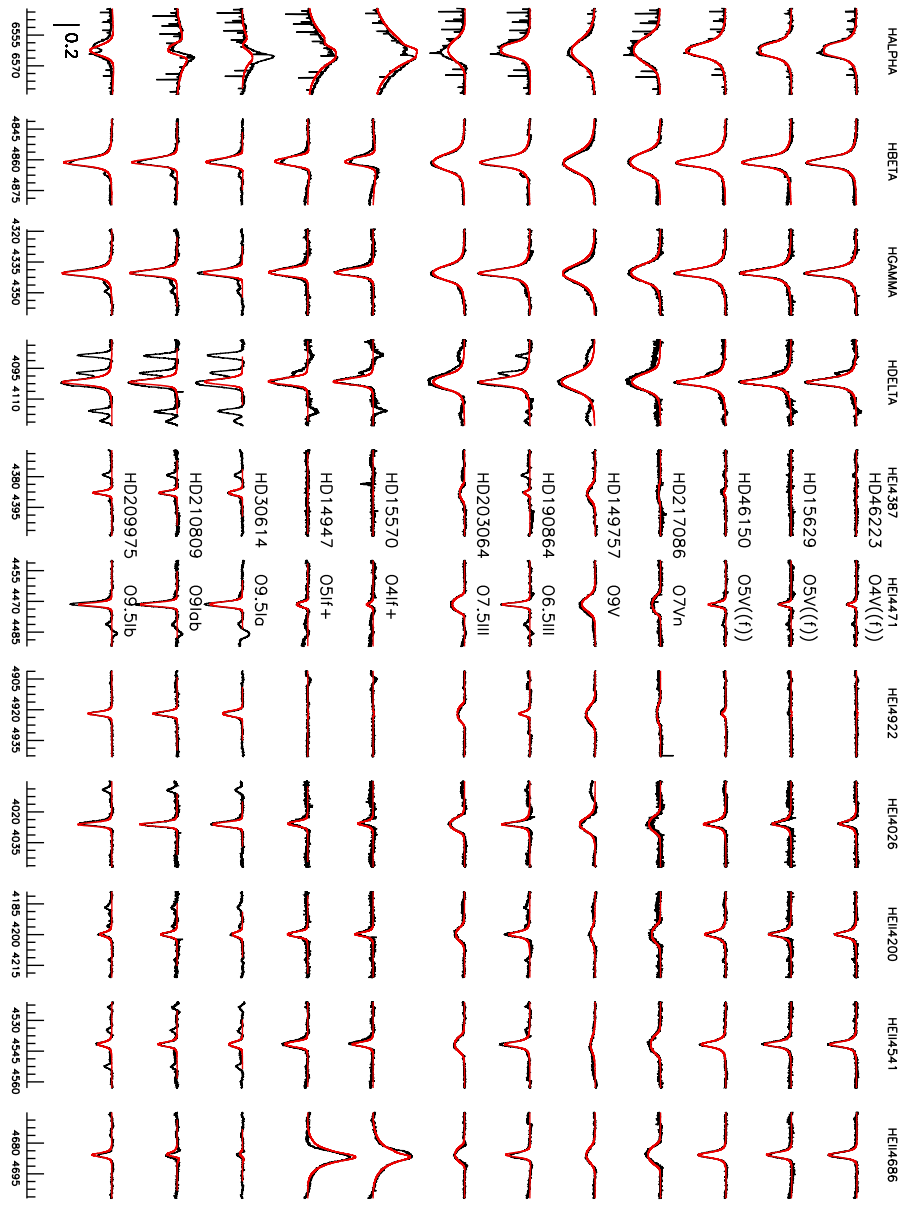

Figure 4.30: Spectral fits for the optical lines using the clumping law $Linear_{050}$. Observations are shown in black and best fit model profiles in red. The abscissa gives the wavelength for each line, and the scale of the ordinate axis is given by the vertical bar at the lower left corner (at the bottom of the Hα column).

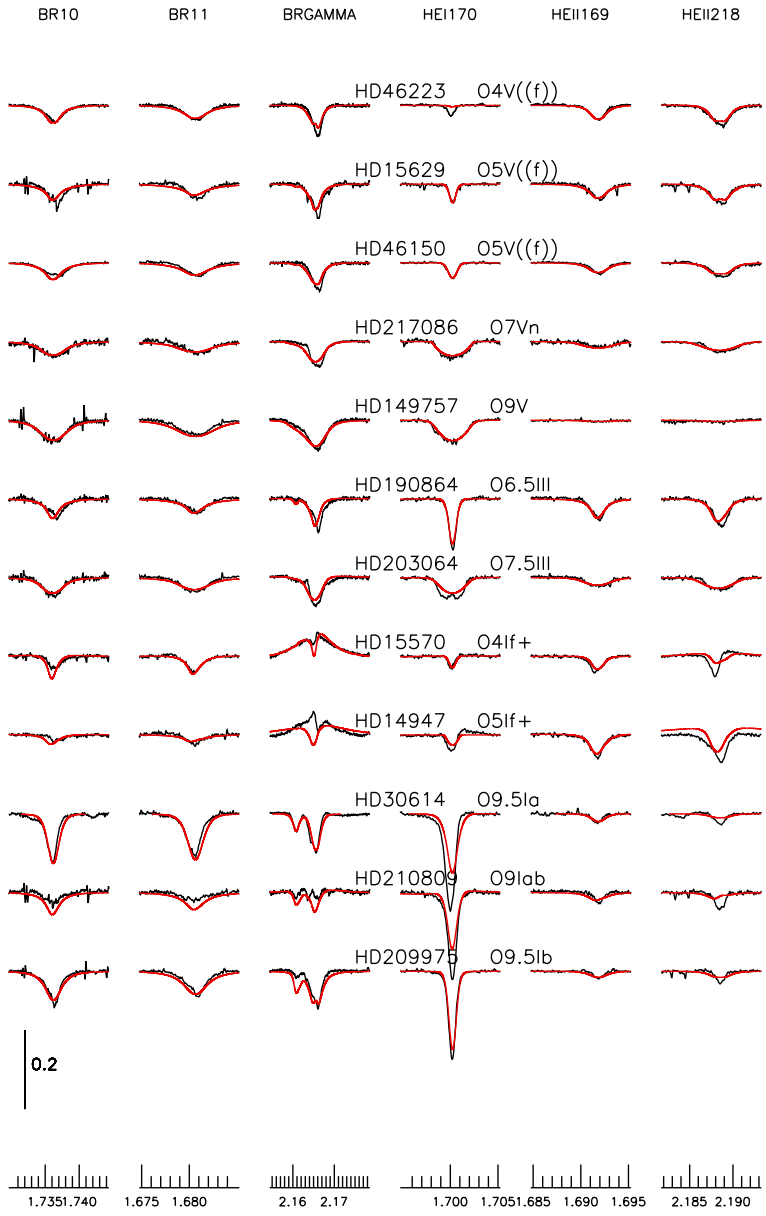

Figure 4.31: Spectral fits for the near-infrared lines using the clumping law $Linear_{050}$. Observations are shown in black and best fit model profiles in red. The abscissa gives the wavelength for each line, and the scale of the ordinate axis is given by the vertical bar at the lower left corner (at the bottom of the HeI $\lambda 1.70 \mu$m column).

4.7 Analysis and discussion

We can now compare the derived parameters to see whether the introduction of clumping modifies our determinations or improves the agreement between optical and infrared stellar parameters. We will not consider the microturbulence and β exponent, as they remain basically unrestricted in our analyses. We notice that this will introduce some scatter in our results, as they may slightly influence the preferred values for the other parameters. Nevertheless, their impact should be small and remain within adopted uncertainties. In Table 4.14 we give again the list of our stars with their ID number that will be used in the plots.

We begin by comparing the results obtained from the optical analysis without clumping and with $Linear_{025}$ and $Linear_{050}$. Figure 4.32 shows the comparison for $T_{\rm eff}$. Comparing the unclumped and $Linear_{025}$ results we see that we obtain the same results for all stars within typical uncertainties. In fact, the average difference is 0.02±0.5 kK. When we compare the results both clumping laws (right panel in Figure 4.32) we see that they are also very similar. Thus the temperature determination in the optical is not affected by the presence of clumping or by differences in the clumping distribution (as far as it concerns the laws used in this work).

The comparison for the gravities obtained from the optical analyses are presented in Figure 4.33. We see that now HD 15 570 (#8) departs clearly from the 1:1 relation, with the rest of the supergiants (except the less luminous one, HD 209 975) showing a clear trend to lower gravities when clumping is included. This is a consequence of the lower mass-loss rate implied by clumping, that affects mainly the fit to the Balmer lines (particularly H_α, as we discussed previously). This is more clear when we compare with the $Linear_{050}$ results. There both hot supergiants depart from the 1:1 relation, in agreement with the discussion that we presented for the H_α fit. Finally, except for the second hot supergiant (HD 14 947, #9) all stars give similar results with both clumping laws.

We compare the results for the wind strength $\log Q$ in Figure 4.34. As we used the same wind terminal velocity and stellar radius, this is equivalent to the mass-loss rate. We see that the unclumped models give higher mass-loss rates for most stars (except #4 and #5), as expected. The average difference in $\log Q$ for these stars is 0.4, very close to the nominal value of $F_{max} = 10$ that would be 0.5. Both clumping laws give again the same results. The stars that, surprisingly, depart from this behavior., #4 and #5, are the fast rotator late-type dwarfs HD 217 086 and HD 149 757. The error bars for $\log Q$ are very large, at least one magnitude sigma, a consequence of the low luminosity, weak wind and the very large rotational velocity. This explains also the difference between both clumping laws for one of the stars, HD 149 757.

The helium abundances are compared in Figure 4.35. All determinations (unclumped and clumping laws models) are equal within typical uncertainties, except for the cool bright supergiant HD 30 614, that gives a higher He abundance when clumping is included (with the same value for both clumping laws). This higher He abundance is the result of a better fit to the He lines after the changes produced in H_α. Thus we consider it more reliable. It is interesting to see

that the He abundance of star #1, the early dwarf HD 46 223, presents a small difference between both clumping laws, but within the typical uncertainties.

Thus we conclude that the stellar parameters derived from the optical are only affected in the case of the gravities and wind strengths determined for supergiants (except in particular cases like the helium abundance of HD 30 614 or the wind strength of the fast rotator dwarfs, that should be treated individually).

Table 4.14: Sample of stars with their identification numbers and spectral types.

ID	NAME	Spectral Type
1	HD46223	O4 V((f))
2	HD15629	O4.5 V((fc))
3	HD46150	O5 V((f))z
4	HD217086	O7 Vnn((f))z
5	HD149757	O9.2 IVnn
6	HD190864	O6.5 III(f)
7	HD203064	O7.5 IIIn((f))
8	HD15570	O4 If
9	HD14947	O4.5 If
10	HD30614	O9 Ia
11	HD210809	O9 Iab
12	HD209975	O9.5 Ib

The situation is quite similar for the temperature in the near-infrared, without significant changes when introducing clumping (Figure 4.36). For the gravity, however, the situation is different to the optical, as we derive the same values for clumped and unclumped models within the typical adopted uncertainties (see Figure 4.37). This is an important difference for the case of hot supergiants. For the wind strength we have a similar global behavior. again as in the optical: as expected, clumped models result in lower mass-loss rates (or wind strengths). In fact, the mean difference, $\Delta \log Q = 0.55$, is about the expected factor (for the $Linear_{025}$ law; it is a bit smaller, 0.30, for the $Linear_{050}$ one). Both linear laws are quite consistent in the wind strengths values, except for stars #1 and #4. The case of star #4 (HD 217 086) has already been discussed, whereas the differences between the fits obtained with both clumping laws in the case of star #1 have been already discussed in the previous section. Lastly, the helium abundances also result in similar values within uncertainties, except again HD 46 223 and HD 30 614, resulting from slightly different fits.

Our conclusion is that the impact of clumped models as compared to unclumped in the near-infrared is similar but smaller than in the optical, with a larger scatter in the global trends, although this may be a logical consequence of the lower quality (resolution and S/N ratio) of the near-infrared sample compared to the optical one. A second conclusion is that, compared to the

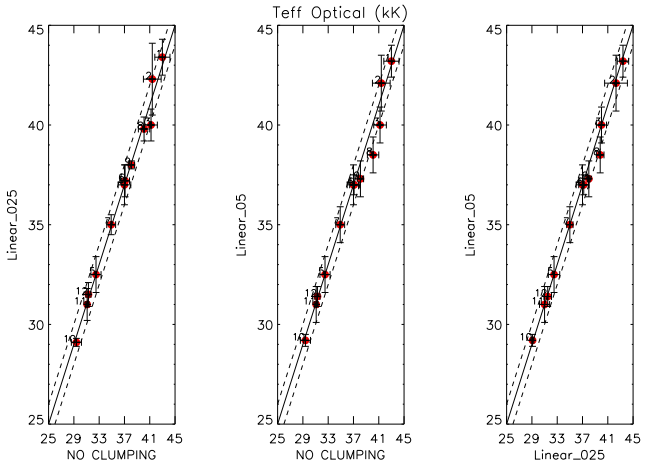

Figure 4.32: Comparison of effective temperatures (in kK) obtained with *iacob_gbat* from the optical spectra. Left: $Linear_{025}$ versus unclumped results; middle: $Linear_{050}$ versus unclumped; right: $Linear_{050}$ versus $Linear_{025}$. Numbers close to the symbols give the ID from Tab. 4.14. The solid line marks the 1:1 relation and the dashed line the ±1.0 kK zone.

optical, the near-infrared in the H and K bands do not offer us a clear advantage to characterize clumping. However, this conclusion is dependent on our adoption of clumping parameters, particularly F_{max}. A higher value for the maximum clumping would affect the near-infrared lines more than optical ones, because of the different line formation region.

We can now address the question whether the introduction of clumping improves the agreement between the optical and infrared parameter determination. Figure 4.40 shows the comparison of $T_{\rm eff}$ determinations in the optical and the infrared for the different clumping conditions explored in our experiment. We see that there is not a clear improvement in the correlation (the average value of the absolute differences for $T_{\rm eff}$, for example, decreases only from 1.4±1.1 kK for the unclumped values to 1.1±1.3 and 1.3±1.1 kK for the two clumping laws). In particular the two stars that depart more, HD 14 947 and HD 210 809 do not show a significant change in the temperature determination. However, if we exclude these two stars, the absolute differences agree better in the case of the clumping models, with average values of 1.0±1.0 kK for the unclumped models and 0.6±0.7 and 0.9±1.3 kK for the clumping ones. Thus we conclude that for the temperature the $Linear_{025}$ offers a slight improvement in the agreement between the optical and the infrared.

In Figure 4.41 we see the comparison for $\log g$ values. The agreement is poor, particularly for the supergiants and it seems to worsen when we consider clumping. However, this is strongly

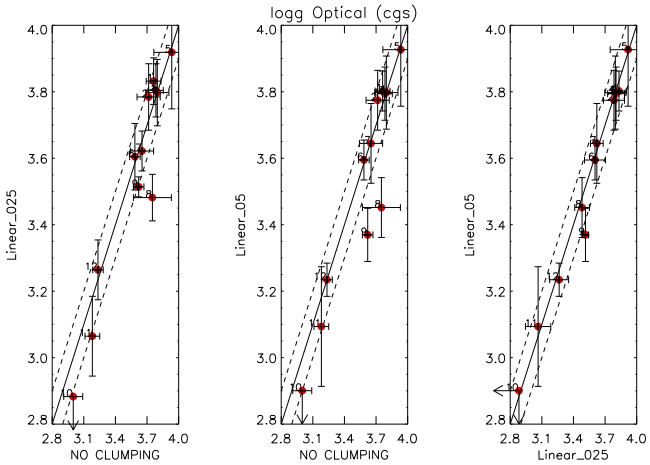

Figure 4.33: Same as Figure 4.32 for the logarithmic gravity. The solid line marks the 1:1 relation and the dashed line the ±0.10 dex zone.

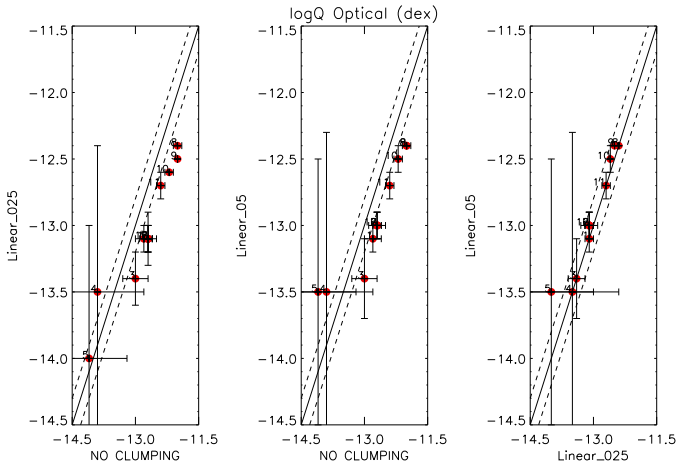

Figure 4.34: Same as Figure 4.32 for the wind strength log Q. The solid line marks the 1:1 relation and the dashed line the ±0.20 dex zone.

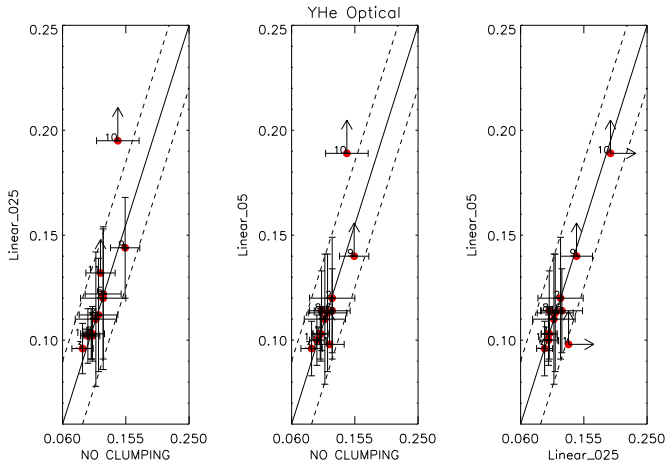

Figure 4.35: Same as Figure 4.32 for the helium abundance by number, N(He)/N(H). The solid line marks the 1:1 relation and the dashed line the ±0.03 zone.

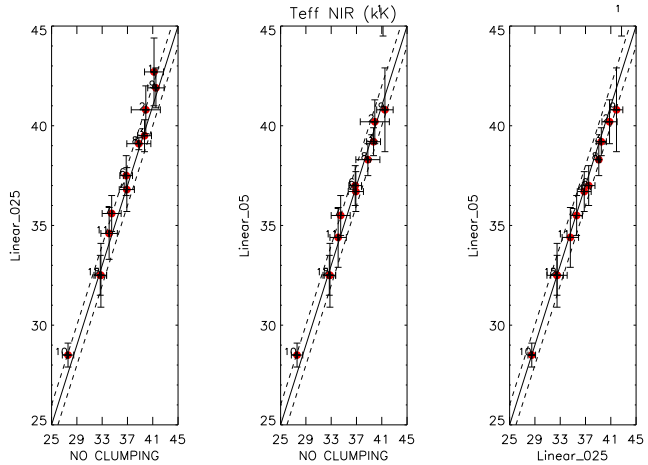

Figure 4.36: Same as Figure 4.32 for the effective temperature in the near-infrared. The solid line marks the 1:1 relation and the dashed line the ±1.0 kK zone.

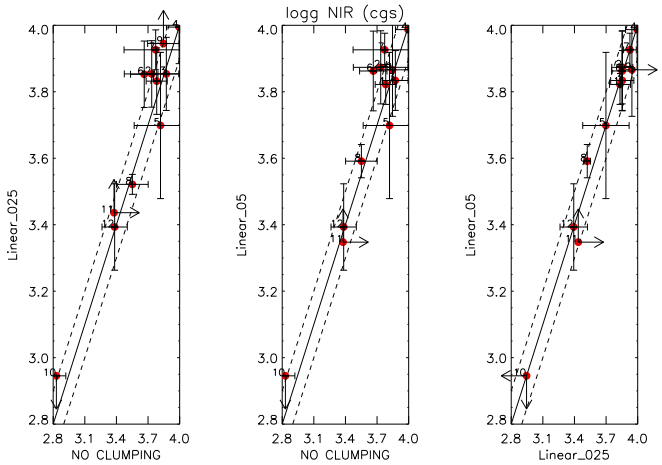

Figure 4.37: Same as Figure 4.32 for the logarithmic gravity in the near-infrared. The solid line marks the 1:1 relation and the dashed line the ±0.10 dex zone.

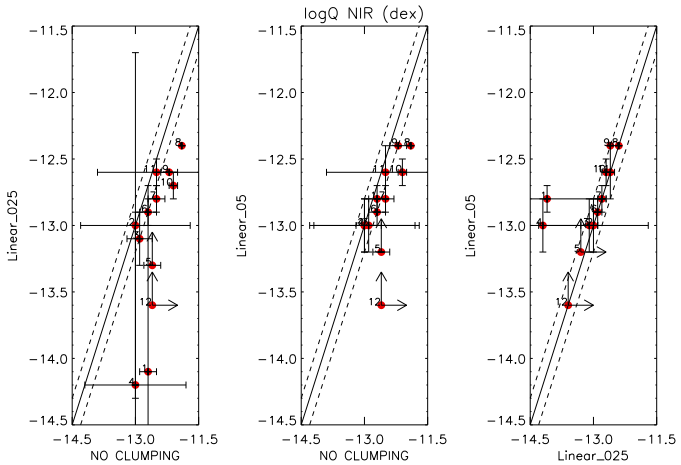

Figure 4.38: Same as Figure 4.32 for the wind strength log Q in the near-infrared. The solid line marks the 1:1 relation and the dashed line the ±0.20 dex zone.

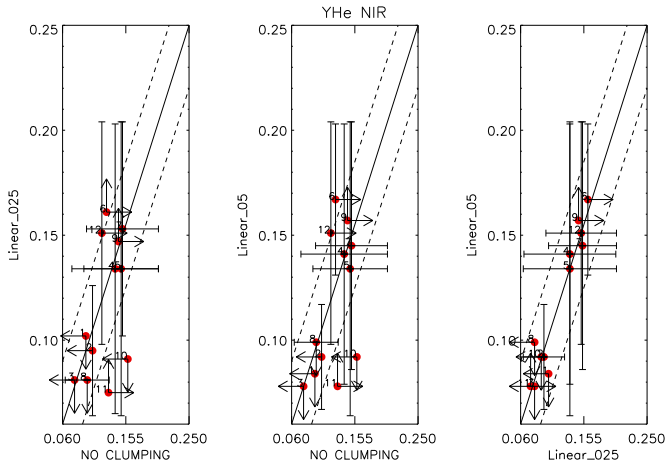

Figure 4.39: Same as Figure 4.32 for the He abundance by number, N(He)/N(H) in the near-infrared. The solid line marks the 1:1 relation and the dashed line the ±0.03 zone.

related to the presence of many upper and lower limits in the near-infrared determinations. This is a consequence of the fact that the main reaction to gravity in the Bracket lines is produced in the core of the lines (the wings being not so sensitive to gravity as are in the optical Balmer lines). Thus, very accurate observations are needed for the gravity determinations in the near-infrared. This of course adds scatter to the determination of other stellar parameters.

The comparison for the log Q values is shown in Figure 4.42. In spite of the apparent scatter, all stars that depart from the 1:1 relation have either upper/lower limits or large uncertainties (beyond one order of magnitude). This indicates that, when it is possible to determine the wind strength accurately, the infrared and the optical values agree. However, we see that introducing clumping does not necessarily improves the apparent agreement between both wavelength ranges. Once more, this seems to be related to the limited accuracy with which we can determine the stellar parameters. A similar behavior. is seen for the helium abundances (Figure 4.43).

Thus our global conclusion is that we can derive stellar parameters from the infrared that agree with those in the optical within the limitations given in Sect. 4.4.2 (including rotational velocities, see Sect. 4.4.3). The inclusion of clumping in the form done here does not offer us a better agreement between the stellar parameters determined in the optical and the near infrared. However, this conclusion is mostly driven by two factors: (a) our limited ability to determine accurate parameters in the near-infrared and (b) the assumption that all stars follow the same clumping laws (both distribution and, probably more importantly, parameter values) in the inner wind, an assumption that appears unlikely (see Puls et al. (2006), Najarro et al. (2011)). Thus

further work is required to explore the possibilities of the atomic models in the infrared (for example, we didn't exclude the HeII $\lambda 2.18\mu$m line for the determination of the stellar parameters, in spite of its poor fit when clumping is included), to obtain higher quality infrared spectra (we notice here that our optical spectra are of much better quality than the infrared ones) and to extend the range of clumping parameters, particularly F_{max}.

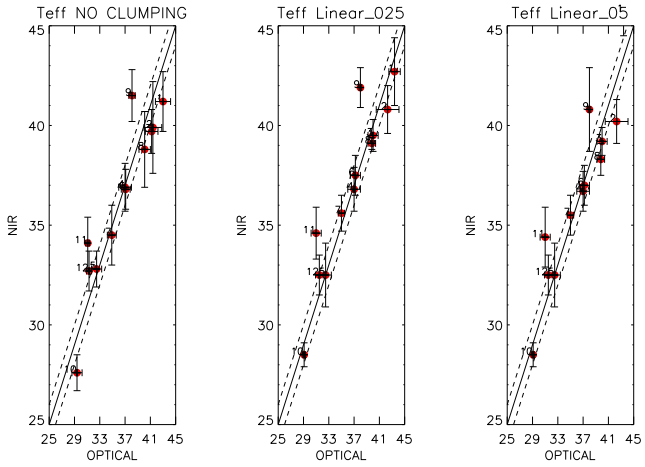

Figure 4.40: Comparison of effective temperatures (in kK) obtained with *iacob_gbat* from the optical and near-infrared spectra using different clumping laws. Left: no clumping; middle: $Linear_{025}$ law; right: $Linear_{050}$ law. Numbers close to the symbols give the ID from Tab. 4.14. The solid line marks the 1:1 relation and the dashed line the ±1.0 kK zone.

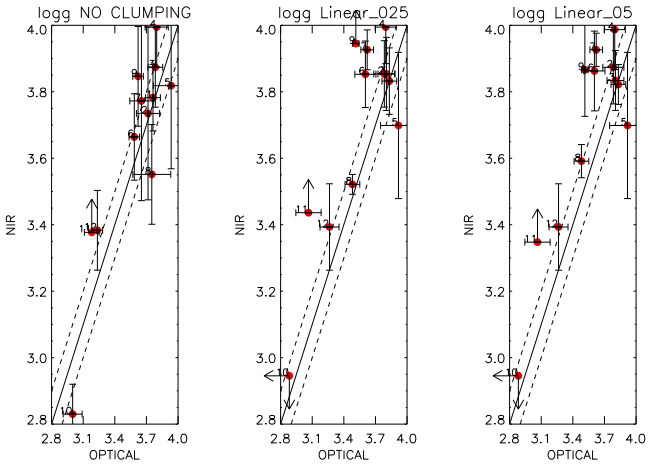

Figure 4.41: Same as Figure 4.40 for the logarithmic gravity. The solid line marks the 1:1 relation and the dashed line the ±0.1 dex zone.

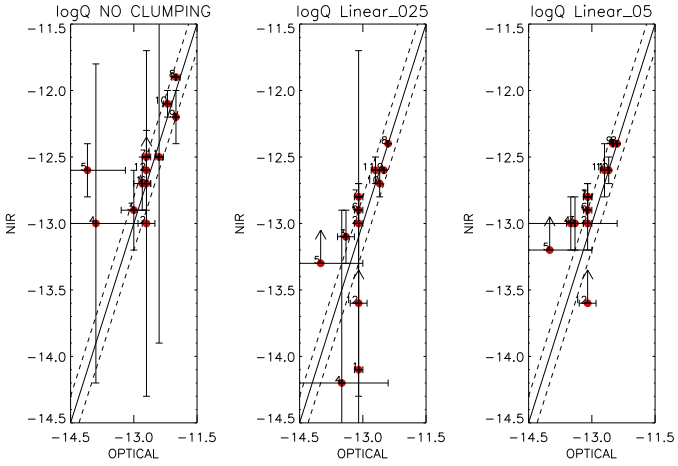

Figure 4.42: Same as Figure 4.40 for the wind strength parameter $\log Q$. The solid line marks the 1:1 relation and the dashed line the ±0.2 dex zone.

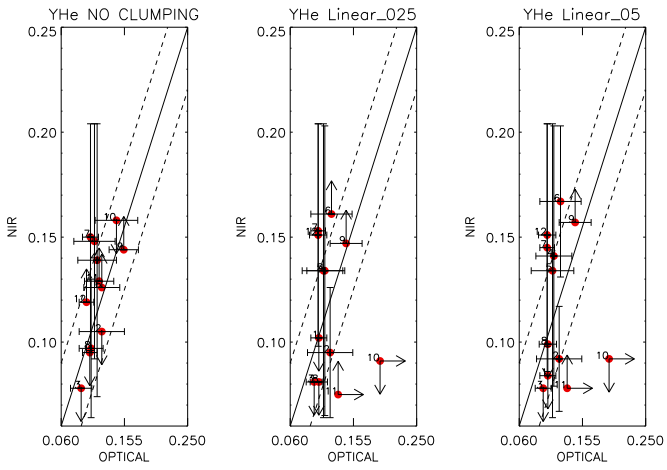

Figure 4.43: Same as Figure 4.40 for the helium abundance by number YHe= N(He)/N(H). The solid line marks the 1:1 relation and the dashed line the ±0.03 zone.

4.8 Conclusions

We have carried out a determination of stellar parameters and a study of the clumping effects in the optical and the near-infrared using the automatic methods developed in our group. Our aims were (a) to check whether we can obtain stellar parameters from the infrared, with the same or comparable accuracy to those in the optical; (b) to check that the parameters obtained were the same; (c) to study the effects of clumping in the determination of stellar parameters; and (d) check whether clumping improves the agreement between the infrared and the optical parameters. In addition, this kind of work that inlcudes many lines behaving in different ways may in the future allow the determination of clumping distributions in the wind (in contrast to studies that concentrate on pure ρ^2 processes, as done by Puls et al. (2006)). To these aims we have extended the automatic tools to include the near-infrared spectra.

The main conclusions of our study can be divided in two groups. The firs one can be considered more general while the second one is more dependent on the exploratory character of our study, i.e., on the assumption of a single clumping law and particularly on the adopted F_{max} value. As a consequence, our future work should concentrate on the extension of the F_{max} range.

Our first group of conclusions includes:

- The parameters derived from the optical and the infrared with our automatic tools are very similar.
- We can derive stellar parameters from the infrared with comparable accuracy to the optical, although with slightly larger uncertainties.
- The rotational velocities derived from the near-infrared HeI $\lambda 1.70\mu$m line agree well with those derived with higher accuracy from the metal optical lines (with the known problems due to the larger HeI broadening).
- The quality to the fits of the optical lines in absence of clumping are excellent, except for the H_α line of some supergiants.
- The effective temperature is well determined independently of the wavelength region used and the clumping conditions in the models. Gravity and wind strength (mass-loss rate) are the most affected parameters. Near-infrared helium abundances are also well correlated with optical counterparts.

And the second group of conclusions, more dependent on the adopted assumptions, are:

- We have used a coarse grid to show that clumping only has significant effects for supergiants once we take the difference in mass-loss rate as a consequence of clumping.
- We find only small differences in the synthetic line profiles among the clumping laws used, which indicate that the lines used as diagnostics are formed in layers where the differences between these laws are not critical.

- Together with the maximum clumping factor, the relevant parameter is the extension of the region of increasing clumping factor until that maximum is reached. Both together define the clumping factors in the line formation region.
- Clumping improves the fit to H_α lines of some supergiants (specifically for the hot supergiants), but makes the fit to HeII $\lambda 2.18\mu$m poorer.
- In the way we have introduced it, models with clumping do not offer a better agreement between optical and near-infrared parameters than unclumped models do. However, this conclusion may also be biased by the quality differences between the optical and the infrared spectra analyzed.

5

Conclusions and Future Work

In this chapter we present the conclusions from the creation of an automatic search of cluster algorithm MASCA and how we improve the criteria of selecting overdensities, then the new associations founded MASGOMAS- 6 and MASGOMAS-10 and finally the analysis for the O stars which lead us to study the clumping condition.

MASCA

- When we search for massive star cluster candidates we find a large number of field population contaminants. These objects were efficiently discriminated by improvements in the color cuts.
- We have investigated the detectability of overdensities in a random distribution of stars, which results in an optimization criteria for our choice of D_S and $Nmin$
- We have developed the bona-fide algorithm MASCA which allows us to automatically locate hidden clusters/associations candidates in the Milky Way.
- This algorithm was tested using 2MASS sky survey catalog, but can be used with any other near-infrared catalog.
- We have created the first bona-fide catalog of massive cluster candidates in the region of $12.5° < l < 93°$.

MASGOMAS-6

- The preliminary version of MASCA provide us with a new massive stars cluster.
- Using broad photometric color cuts, derived from UKIDSS, we identified 30 stellar candidates for spectroscopic follow up (MOS & LS)

- The spectra allow the identification of massive stars in several evolutionary stages (dwarfs, giants, supergiants, and Wolf-Rayet). The number of contaminants was close to one third of the spectral sample.
- We have identified 18 early type stars, from which 6 are supergiant. We have found a new WR (star 29) and confirm the WN6 classification of 1583-B73 (Faherty et al., 2014). The individual distance and extinction estunates fir the young massive stars indicate the presence of two massive populations in the same line of sight at different distances: one at 4-4.5 kpc, in the close end of the Galactic bar and a second, more distant, in the Sagittarius arm at about 8 kpc. This second massive population appears around a distant LBV candidate (IRAS 18576+0341; Ueta et al., 2001; Pasquali & Comerón, 2002; Clark et al.,2003).
- The estimation of mass and ages for each association was not possible, due to the mixed populations.

MASGOMAS-10

- By comparing the population in the individual associations we conclude that an association detected using several values of D_S and *Nmin* is mos probably a real association of massive stars. The majority of the OB population is the *core*, detected several times by the code, while the old-late type population are part of the peripheral associations only detected once.
- Selecting all the candidates which pass the photometric cuts in the proximity zone of a several times detected association, may increase the number of confirmed massive stars, but with the high risk of confirming contaminants.
- We detected a total of 7 massive early type stars, including a Wolf-Rayet. A preliminary estimation indicates a radius of 2.0 kpc and a mass of few thousands of solar masses for the possible association

Spectroscopic analysis and clumping

- The parameters derived from the optical and the infrared with our automatic tools are very similar.
- We can derive stellar parameters from the infrared with comparable accuracy to the optical, although with slightly larger uncertainties.
- The rotational velocities derived from the near infrared HeI $\lambda 1.70\mu$m line agree well with those derived with higher accuracy from the metal optical lines (with the known problems due to the larger HeI broadening).
- The quality to the fits of the optical lines in absence of clumping are excellent, except for the H_α line of some supergiants.

- We have used a coarse grid to show that clumping only has significant effects for supergiants once we take the difference in mass-loss rate as a consequence of clumping.
- We have no found no significant differences in the synthetic line profiles among the clumping laws used, which indicates that the lines used as diagnostics are formed in layers where the differences between these laws are not important
- Together with the maximum clumping factor, the relevant parameter is the extension of the region of increasing clumping factor until that maximum is reached. Both together define the clumping factors in the line formation region.
- Clumping improves the fit to some supergiants H_α lines (specially for the hot supergiants), but makes the fit to HeII $\lambda 2.18\mu$m poorer.
- The effective temperature is well determined independently of the wavelength region used and the clumping conditions in the models, gravity and wind strength (mass-loss rate) are the most affected. Near infrared helium abundances are also well correlated with optical counterparts.
- In the way we have introduced it models with clumping do not offer a better agreement between optical and near infrared parameters than unclumped models do. However, this conclusion may also be biased by the quality differences between the optical and the infrared spectra analyzed.

FUTURE WORK

The future work is mainly centered on the new arrival of EMIR at GTC (10.4-m Telescope) in La Palma,

- MASCA1.2 (VVV and UKIDSS version)

  In similar way that 2MASS catalog survey was used, MASCA program provide us with the open door to adapt it for new and deeper photometry surveys, such as VVV and UKIDSS. Moreover with EMIRGTC will expand the limit magnitude of $K = 12.5$ several magnitudes. We have already begin test the use of MASCA using VVV photometry .

- Full characterization of MASGOAMS-10

  The follow-up of MASGOMAS-10 is still on the way. Due to the only K-band spectra is difficult to determine with accuracy al the properties of the association. For that using EMIRGTC we expect to fully characterize the association. Moreover a secondary WR stars is found in the proximity of the association.

- Follow-up of the new catalog of massive cluster candidates

  We have already submitted an exploratory observation for 16 massive cluster candidates, from $l > 35°$ from where new MASGOMAS may be founded.

- Clumping analysis

  From the result of the clumping study, we open the possibility of realize a full quantitative analysis of stars with the two full grids created, from where the stellar parameters from the stars could be derived without having the problems of non fitting wind lines. This implies into the re-analysis of all the OB stars which present problems with fitting.

Bibliography

Abbott, D. C. 1982, ApJ, 263, 723

Acosta Pulido, J. A., Ballesteros, E., Barreto, M., et al. 2003, The Newsletter of the Isaac Newton Group of Telescopes, 7, 15

Andersen, M., Zinnecker, H., Moneti, A., et al. 2009, ApJ, 707, 1347

Arnould, M., Paulus, G., & Meynet, G. 1997, A&A, 321, 452

Barbá, R. H., Roman-Lopes, A., Nilo Castellón, J. L., et al. 2015, A&A, 581, A120

Bastian, N., Covey, K. R., & Meyer, M. R. 2010, ARA&A, 48, 339

Benjamin, R. A., Churchwell, E., Babler, B. L., et al. 2003, PASP, 115, 953

Bestenlehner, J. M., Gräfener, G., Vink, J. S., et al. 2014, A&A, 570, A38

Bica, E., Dutra, C. M., Soares, J., & Barbuy, B. 2003, A&A, 404, 223

Binney J., M. M. 1998

Borissova, J., Bonatto, C., Kurtev, R., et al. 2011, A&A, 532, A131

Borissova, J., Chené, A.-N., Ramírez Alegría, S., et al. 2014, A&A, 569, A24

Bouret, J.-C., Hillier, D. J., Lanz, T., & Fullerton, A. W. 2012, A&A, 544, A67

Bouret, J.-C., Lanz, T., & Hillier, D. J. 2005, A&A, 438, 301

Bouret, J.-C., Lanz, T., Hillier, D. J., et al. 2003, ApJ, 595, 1182

Cabrera-Lavers, A., Hammersley, P. L., González-Fernández, C., et al. 2007, A&A, 465, 825

Chené, A.-N., Borissova, J., Bonatto, C., et al. 2013, A&A, 549, A98

Chené, A.-N., Ramírez Alegría, S., Borissova, J., et al. 2015, A&A, 584, A31

Chiosi, C., Bertelli, G., & Bressan, A. 1992, ARA&A, 30, 235

Churchwell, E., Babler, B. L., Meade, M. R., et al. 2009, PASP, 121, 213

Clark, J. S., Crowther, P. A., Larionov, V. M., et al. 2009, A&A, 507, 1555

Clark, J. S., Egan, M. P., Crowther, P. A., et al. 2003, A&A, 412, 185

Clough, S. A., Shephard, M. W., Mlawer, E. J., et al. 2005, J. Quant. Spec. Radiat. Transf., 91, 233

Comerón, F., Pasquali, A., & Torra, J. 2005, A&A, 440, 163

Conti, P. S. 1973, ApJ, 179, 181

Cox, A. N. 2000, Allen's astrophysical quantities

Crowther, P. A., Hillier, D. J., Evans, C. J., et al. 2002, ApJ, 579, 774

Crowther, P. A., Schnurr, O., Hirschi, R., et al. 2010, MNRAS, 408, 731

Crowther, P. A. & Smith, L. J. 1996, A&A, 305, 541

Dame, T. M., Hartmann, D., & Thaddeus, P. 2001, ApJ, 547, 792

Davies, B., Figer, D. F., Kudritzki, R.-P., et al. 2007, ApJ, 671, 781

de la Fuente, D., Najarro, F., Borissova, J., et al. 2016, A&A, 589, A69

Dutra, C. M. & Bica, E. 2000, A&A, 359, L9

Dutra, C. M. & Bica, E. 2001, A&A, 376, 434

Dutra, C. M., Bica, E., Soares, J., & Barbuy, B. 2003, A&A, 400, 533

Ekström, S., Georgy, C., Eggenberger, P., et al. 2012, A&A, 537, A146

Epchtein, N., de Batz, B., Capoani, L., et al. 1997, The Messenger, 87, 27

Faherty, J. K., Shara, M. M., Zurek, D., Kanarek, G., & Moffat, A. F. J. 2014, AJ, 147, 115

Figer, D. F., MacKenty, J. W., Robberto, M., et al. 2006, ApJ, 643, 1166

Figer, D. F., McLean, I. S., & Morris, M. 1999, ApJ, 514, 202

Figer, D. F., Najarro, F., Gilmore, D., et al. 2002, ApJ, 581, 258

Fitzpatrick, E. L. 1999, PASP, 111, 63

Fragoso-López, A. B., Acosta-Pulido, J. A., Hernández, E., et al. 2008, in Proc. SPIE, Vol. 7014, Ground-based and Airborne Instrumentation for Astronomy II, 701468

Froebrich, D., Scholz, A., & Raftery, C. L. 2007, MNRAS, 374, 399

Garcia, M., Herrero, A., Castro, N., & José Corral, L. 2011, Bulletin de la Societe Royale des Sciences de Liege, 80, 381

Garcia, M., Herrero, A., Vicente, B., et al. 2009, A&A, 502, 1015

Glushkova, E. V., Koposov, S. E., Zolotukhin, I. Y., et al. 2010, Astronomy Letters, 36, 75

Gräfener, G., Koesterke, L., & Hamann, W.-R. 2002, A&A, 387, 244

Hamann, W.-R. & Gräfener, G. 2003, A&A, 410, 993

Hanson, M. M., Conti, P. S., & Rieke, M. J. 1996, ApJS, 107, 281

Hanson, M. M., Kudritzki, R.-P., Kenworthy, M. A., Puls, J., & Tokunaga, A. T. 2005, ApJS, 161, 154

Hanson, M. M., Kurtev, R., Borissova, J., et al. 2010, A&A, 516, A35

Hanson, M. M. & Popescu, B. 2008, in IAU Symposium, Vol. 250, Massive Stars as Cosmic Engines, ed. F. Bresolin, P. A. Crowther, & J. Puls, 307–312

Hauschildt, P. H. & Baron, E. 1999, Journal of Computational and Applied Mathematics, 109, 41

Herrero, A., Kudritzki, R. P., Vilchez, J. M., et al. 1992, A&A, 261, 209

Hillenbrand, L. A. 2009, in IAU Symposium, Vol. 258, The Ages of Stars, ed. E. E. Mamajek, D. R. Soderblom, & R. F. G. Wyse, 81–94

Hillier, D. J., Lanz, T., Heap, S. R., et al. 2003, ApJ, 588, 1039

Hillier, D. J. & Miller, D. L. 1998, ApJ, 496, 407

Hoffmann, G. 2007

Hubeny, I. & Hubeny, V. 1998, ApJ, 505, 558

Indebetouw, R., Mathis, J. S., Babler, B. L., et al. 2005, ApJ, 619, 931

Ivanov, G. A. 2004, VizieR Online Data Catalog, 4028

Ivanov, V. D., Borissova, J., Pessev, P., Ivanov, G. R., & Kurtev, R. 2002, A&A, 394, L1

Ivanov, V. D., Messineo, M., Zhu, Q., et al. 2010, in IAU Symposium, Vol. 266, Star Clusters: Basic Galactic Building Blocks Throughout Time and Space, ed. R. de Grijs & J. R. D. Lépine, 203–210

Ivanov, V. D., Piatti, A. E., Beamín, J.-C., et al. 2017, A&A, 600, A112

Izumiura, H., Deguchi, S., Hashimoto, O., et al. 1995, ApJ, 453, 837

Johnson, H. L. & Morgan, W. W. 1953a, ApJ, 117, 313

Johnson, H. L. & Morgan, W. W. 1953b, ApJ, 117, 313

Kausch, W., Noll, S., Smette, A., et al. 2015, A&A, 576, A78

Kronberger, M., Teutsch, P., Alessi, B., et al. 2006, A&A, 447, 921

Kudritzki, R.-P. & Puls, J. 2000, ARA&A, 38, 613

Kudritzki, R. P., Urbaneja, M. A., & Puls, J. 2006, in IAU Symposium, Vol. 234, Planetary Nebulae in our Galaxy and Beyond, ed. M. J. Barlow & R. H. Méndez, 119–126

Kumar, M. S. N., Keto, E., & Clerkin, E. 2006, A&A, 449, 1033

Lada, C. J. & Lada, E. A. 2003, ARA&A, 41, 57

Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, MNRAS, 379, 1599

Lee, H.-T. & Chen, W. P. 2007, ApJ, 657, 884

Lucas, P. W. & Samuel, D. 2010, Highlights of Astronomy, 15, 779

Maeder, A. 1981, A&A, 101, 385

Manchado, A., Barreto, M., Acosta-Pulido, J., et al. 2004, in Proc. SPIE, Vol. 5492, Ground-based Instrumentation for Astronomy, ed. A. F. M. Moorwood & M. Iye, 1094–1104

Marín-Franch, A., Herrero, A., Lenorzer, A., et al. 2009, A&A, 502, 559

Martins, F. 2011, Bulletin de la Societe Royale des Sciences de Liege, 80, 29

Martins, F., Hervé, A., Bouret, J.-C., et al. 2015, A&A, 575, A34

Martins, F., Hillier, D. J., Paumard, T., et al. 2008, A&A, 478, 219

Martins, F. & Plez, B. 2006, A&A, 457, 637

Martins, F., Schaerer, D., Hillier, D. J., et al. 2005, A&A, 441, 735

Massey, P., Neugent, K. F., Hillier, D. J., & Puls, J. 2013, ApJ, 768, 6

Mauerhan, J. C., Van Dyk, S. D., & Morris, P. W. 2011, AJ, 142, 40

Mercer, E. P., Clemens, D. P., Meade, M. R., et al. 2005, ApJ, 635, 560

Meyer, M. R., Edwards, S., Hinkle, K. H., & Strom, S. E. 1998, ApJ, 508, 397

Meynet, G., Maeder, A., Schaller, G., Schaerer, D., & Charbonnel, C. 1994, A&AS, 103

Minniti, D., Lucas, P. W., Emerson, J. P., et al. 2010

Morales, E. F. E., Wyrowski, F., Schuller, F., & Menten, K. M. 2013, VizieR Online Data Catalog, 356

Najarro, F., Hanson, M. M., & Puls, J. 2011, A&A, 535, A32

Najarro, F., Hillier, D. J., Puls, J., Lanz, T., & Martins, F. 2006, A&A, 456, 659

Negueruela, I., Clark, J. S., & Ritchie, B. W. 2010, A&A, 516, A78

Negueruela, I. & Schurch, M. P. E. 2007, A&A, 461, 631

Nugis, T. & Lamers, H. J. G. L. M. 2000, A&A, 360, 227

Pasquali, A. & Comerón, F. 2002, A&A, 382, 1005

Pauldrach, A. W. A., Hoffmann, T. L., & Lennon, M. 2001, A&A, 375, 161

Perryman, M. A. C., Lindegren, L., Kovalevsky, J., et al. 1995, A&A, 304, 69

Puga, E., Marín-Franch, A., Najarro, F., et al. 2010, A&A, 517, A2

Puls, J., Kudritzki, R.-P., Herrero, A., et al. 1996, A&A, 305, 171

Puls, J., Markova, N., Scuderi, S., et al. 2006, A&A, 454, 625

Puls, J., Urbaneja, M. A., Venero, R., et al. 2005, A&A, 435, 669

Puls, J., Vink, J. S., & Najarro, F. 2008, A&A Rev., 16, 209

Ramírez-Agudelo, O. H., Simón-Díaz, S., Sana, H., et al. 2013, A&A, 560, A29

Ramírez Alegría, S., Herrero, A., Marín-Franch, A., Puga, E., & Najarro, F. 2011, Bulletin de la Societe Royale des Sciences de Liege, 80, 415

Ramírez Alegría, S., Marín-Franch, A., & Herrero, A. 2012, A&A, 541, A75

Ramírez Alegría, S., Marín-Franch, A., & Herrero, A. 2014, A&A, 567, A66

Repolust, T., Puls, J., Hanson, M. M., Kudritzki, R.-P., & Mokiem, M. R. 2005, A&A, 440, 261

Repolust, T., Puls, J., & Herrero, A. 2004, A&A, 415, 349

Rieke, G. H. & Lebofsky, M. J. 1985, ApJ, 288, 618

Rosslowe, C. K. & Crowther, P. A. 2015, in IAU Symposium, Vol. 307, New Windows on Massive Stars, ed. G. Meynet, C. Georgy, J. Groh, & P. Stee, 135–136

Sabín SanJulián, C. 2014, PhD thesis, Instituto de Astrofisica de Canarias.

Sabín-Sanjulián, C., Simón-Díaz, S., Herrero, A., et al. 2017, A&A, 601, A79

Sabín-Sanjulián, C., Simón-Díaz, S., Herrero, A., et al. 2014, A&A, 564, A39

Sana, H., de Mink, S. E., de Koter, A., et al. 2012, Science, 337, 444

Santolaya-Rey, A. E., Puls, J., & Herrero, A. 1997, A&A, 323, 488

Simón-Díaz, S., Castro, N., Herrero, A., et al. 2011, Journal of Physics Conference Series, 328, 012021

Simón-Díaz, S. & Herrero, A. 2014, A&A, 562, A135

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163

Smette, A., Sana, H., Noll, S., et al. 2015, A&A, 576, A77

Smith, N. 2014, ARA&A, 52, 487

Solin, O., Haikala, L., & Ukkonen, E. 2014, VizieR Online Data Catalog, 356

Solin, O., Ukkonen, E., & Haikala, L. 2012, VizieR Online Data Catalog, 354

Sundqvist, J. O. & Owocki, S. P. 2013, MNRAS, 428, 1837

Sundqvist, J. O., Puls, J., Feldmeier, A., & Owocki, S. P. 2011, A&A, 528, A64

Tokunaga, A. T. 2007

Ueta, T., Meixner, M., Hinz, P. M., et al. 2001, ApJ, 557, 831

Vacca, W. D., Cushing, M. C., & Rayner, J. T. 2003, PASP, 115, 389

van Marle, A. J., Meliani, Z., & Marcowith, A. 2015, A&A, 584, A49

Vink, J. S., de Koter, A., & Lamers, H. J. G. L. M. 2001, A&A, 369, 574

Wainscoat, R. J., Cohen, M., Volk, K., Walker, H. J., & Schwartz, D. E. 1992, ApJS, 83, 111

Wang, S. & Jiang, B. W. 2014, ApJ, 788, L12

Zinnecker, H. & Yorke, H. W. 2007, ARA&A, 45, 481

Agradecimientos

Esta tesis no podría haber sido sin el apoyo de muchísimas personas. Cada una de ellas aporto de una manera distinta y sin esas ayuda por muy pequeas que fueran, me ayudaron en el proceso.

En primer lugar quiero agradecer a mi jefe Artemio Herrero, apoyándome, guiándome y animándome en cada paso de este camino, en especial por ese ultimo empujón que nos trajo a este momento.

Quiero agradecer a mi grupo de Masivas Sergio, Sara, Inés y Gonzalo, y tambien a los que estuvieron durante los aos de doctorado, Miriam, Sebastián y Carolina. Siempre estuvieron para ayudarme por mas pequeas que fieran mis dudas.

Gracias, por supuesto, a mis amigos y compañeros aquí, en el IAC, en especial a Jana, Felipe y Marcos. Gracias por la ayuda y esos gratos momentos.

Y los ultimos, pero no menos importantes, a mis padres, Fernando y Silvia y a mi esposa Edyta. Sin el apoyo incondicional de ustedes, independiente de si tenia que trabajar de noche, del cambio de horario, las horas de diferencia, y los trasnoches, siempre estuvieron para darme apoyo y lo que necesitara para que llegara este momento.