



<b>Publication Year</b>	2017
<b>Acceptance in OA</b>	2020-11-13T12:27:48Z
<b>Title</b>	From molecules to young stellar clusters: the star formation cycle across the disk of M 33
<b>Authors</b>	CORBELLI, Edvige, Braine, Jonathan, BANDIERA, Rino, Brouillet, Nathalie, Combes, Françoise, Druard, Clément, Gratier, Pierre, Mata, Jimmy, Schuster, Karl, Xilouris, Manolis, Palla, Francesco
<b>Publisher's version (DOI)</b>	10.1051/0004-6361/201630034
<b>Handle</b>	<a href="http://hdl.handle.net/20.500.12386/28318">http://hdl.handle.net/20.500.12386/28318</a>
<b>Journal</b>	ASTRONOMY & ASTROPHYSICS
<b>Volume</b>	601

# From molecules to young stellar clusters: the star formation cycle across the disk of M 33<sup>★</sup>

Edvige Corbelli<sup>1</sup>, Jonathan Braine<sup>2</sup>, Rino Bandiera<sup>1</sup>, Nathalie Brouillet<sup>2</sup>, Françoise Combes<sup>3</sup>, Clément Druard<sup>2</sup>, Pierre Gratier<sup>2</sup>, Jimmy Mata<sup>2</sup>, Karl Schuster<sup>4</sup>, Manolis Xilouris<sup>5</sup>, and Francesco Palla<sup>1,★★</sup>

<sup>1</sup> INAF–Osservatorio Astrofisico di Arcetri, Largo E. Fermi, 5, 50125 Firenze, Italy  
e-mail: [edvige;bandiera]@arcetri.astro.it

<sup>2</sup> Laboratoire d’Astrophysique de Bordeaux, Univ. Bordeaux, CNRS, B18N, allée Geoffroy Saint-Hilaire, 33615 Pessac, France

<sup>3</sup> Observatoire de Paris, LERMA (CNRS: UMR 8112), 61 avenue de l’Observatoire, 75014 Paris, France

<sup>4</sup> Institut de Radioastronomie Millimétrique 300 rue de la Piscine, Domaine Universitaire 38406 Saint-Martin-d’Hères, France

<sup>5</sup> Institute for Astronomy, Astrophysics, Space Applications & Remote Sensing, National Observatory of Athens, P. Penteli, 15236 Athens, Greece

Received 8 November 2016 / Accepted 24 March 2017

## ABSTRACT

**Aims.** We study the association between giant molecular clouds (GMCs) and young stellar cluster candidates (YSCCs) to shed light on the time evolution of local star formation episodes in the nearby galaxy M 33.

**Methods.** The CO ( $J = 2-1$ ) IRAM all-disk survey was used to identify and classify 566 GMCs with masses between  $2 \times 10^4$  and  $2 \times 10^6 M_{\odot}$  across the whole star-forming disk of M 33. In the same area, there are 630 YSCCs that we identified using *Spitzer*-24  $\mu\text{m}$  data. Some YSCCs are embedded star-forming sites, while the majority have GALEX-UV and  $H\alpha$  counterparts with estimated cluster masses and ages.

**Results.** The GMC classes correspond to different cloud evolutionary stages: inactive clouds are 32% of the total and classified clouds with embedded and exposed star formation are 16% and 52% of the total, respectively. Across the regular southern spiral arm, inactive clouds are preferentially located in the inner part of the arm, possibly suggesting a triggering of star formation as the cloud crosses the arm. The spatial correlation between YSCCs and GMCs is extremely strong, with a typical separation of 17 pc. This is less than half the CO ( $2-1$ ) beam size and illustrates the remarkable physical link between the two populations. GMCs and YSCCs follow the HI filaments, except in the outermost regions, where the survey finds fewer GMCs than YSCCs, which is most likely due to undetected clouds with low CO luminosity. The distribution of the non-embedded YSCC ages peaks around 5 Myr, with only a few being as old as 8–10 Myr. These age estimates together with the number of GMCs in the various evolutionary stages lead us to conclude that 14 Myr is the typical lifetime of a GMC in M 33 prior to cloud dispersal. The inactive and embedded phases are short, lasting about 4 and 2 Myr, respectively. This underlines that embedded YSCCs rapidly break out from the clouds and become partially visible in  $H\alpha$  or UV long before cloud dispersal.

**Key words.** galaxies: individual: M 33 – galaxies: star formation – galaxies: ISM – infrared: ISM – ISM: molecules

## 1. Introduction

The formation of giant molecular clouds (hereafter GMCs) in the bright disks of spiral galaxies requires the onset of instabilities and the ability of the gas to cool and fragment. Within gravitationally unstable clouds, the process of fragmentation continues to smaller scales, yielding a distribution of clumps and prestellar cores that later collapse to form stars. The evolution of these clouds is then driven by the young stellar clusters (hereafter YSC) that have formed. As the YSC evolves, it could disrupt the cloud or trigger new episodes of star formation. Our Galaxy is the natural laboratory where these processes have been studied in detail because observations can be carried out with an unbeatable spatial resolution. Galactic observations suffer from limitations because we reside within the star-forming disk. Moreover, galaxies with different masses, morphologies, metal contents, and in

different environment or cosmic time might transform the gas into stars over different timescales and with different efficiencies. The molecular gas fragments, from GMCs to protostellar clumps, do not necessarily follow the same mass spectrum as Galactic clouds, nor do they share their characteristics. These considerations have triggered great interest in studying molecular clouds in nearby galaxies, with the support of millimeter telescopes, which are steadily improving in resolution and sensitivity (e.g., ALMA and NOEMA). Validating a cloud formation and evolution model requires an unbiased survey of molecular clouds in a galaxy and of star-forming sites.

All-disk surveys of the  $^{12}\text{CO } J = 1-0$  or  $J = 2-1$  line emission in nearby galaxies are most commonly used to provide a census of molecular complexes down to a certain sensitivity limit. The Large Magellanic Cloud in the southern hemisphere and M 33 in the northern hemisphere have been targets of several observing campaigns of molecular gas emission because they are nearby, gas rich, and with active star formation (Fukui et al. 1999, 2008; Engargiola et al. 2003; Heyer et al. 2004; Gratier et al. 2010; Druard et al. 2014). The NANTEN

\* Full Tables 5 and 6 are only available at the CDS via anonymous ftp to [cdsarc.u-strasbg.fr](ftp://cdsarc.u-strasbg.fr) (130.79.128.5) or via <http://cdsarc.u-strasbg.fr/viz-bin/qcat?J/A+A/601/A146>

\*\* Deceased

group studied the LMC and identified 168 clouds (Mizuno et al. 2001; Fukui et al. 2001, 2008). Wilson & Scoville (1990) surveyed the inner 2 kpc of M33 and detected 38 GMC with a 7'' synthesized beam. The BIMA  $^{12}\text{CO } J = 1-0$  survey of the M33 disk with a 13'' synthesized beam has provided a catalogue of 148 clouds out to  $R = 6$  kpc, complete down to  $1.5 \times 10^5 M_{\odot}$  (Engargiola et al. 2003; Rosolowsky et al. 2003, 2007). Gratier et al. (2012) analyzed the IRAM CO  $J = 2-1$  survey of a large area of the disk of M33 and catalogued 337 GMCs. The star-forming disk of M33 was observed in the CO  $J = 1-0$  and  $J = 3-2$  line with an angular resolution of 25'' by Miura et al. (2012), who identified 71 GMCs, 70 of which were already listed in the Rosolowsky et al. (2007) or Gratier et al. (2012) catalogues. The data used here are from the deep CO ( $J = 2-1$ ) whole-disk survey carried out with the IRAM 30 m telescope (Druard et al. 2014) at 12'' resolution (49 pc), from which 566 clouds were identified.

Molecular clouds are not located around or very close to optically visible stellar clusters, as we show in this paper. In the early phases of star formation (hereafter SF), however, protostars and stars are embedded in the gas and can be detected via imaging with infrared telescopes because of the high extinction provided by the molecular material. The *Spitzer* Space Telescope has surveyed the LMC and M33 in the mid-infrared (hereafter MIR) with sufficient spatial resolution to provide a detailed view of where the hot dust emission is located. Emission from hot dust, typically detected at  $24 \mu\text{m}$ , is an excellent tracer of star-forming sites where massive or intermediate-mass stars have formed. The detection of protostars and of the earliest phases of star formation (hereafter SF) at radio or far-infrared wavelengths is not yet feasible in M33 since it requires surveys with far higher resolution than what is currently available. Searches for embedded SF are limited to less compact sites, that is, to stellar clusters within clouds that are visible at  $24 \mu\text{m}$  in the *Spitzer* survey of the galaxy (Verley et al. 2007, 2009). A catalogue of MIR-emitting sites in M33 is now available (Sharma et al. 2011), and a list of candidate star-forming sites has been selected to investigate their distribution in mass and the spatial distribution in the disk of M33. After formation, stars may clear some of the gas such that, depending on viewing angle, optical counterparts to the MIR emission can be found. The YSC is still compact (i.e. individual members cannot be resolved) and may still suffer from modest extinction. In this context, we study the relationship between young stellar cluster candidates (hereafter YSCCs, used here only for sources that were selected via their MIR emission and which are listed in Table 6), and the molecular cloud population in M33. A detailed analysis of the spatial correlation between GMCs and YSCCs is carried out across the whole star-forming disk of M33. Using essentially the same classification scheme as in Gratier et al. (2012), we define classes of GMCs in terms of their star formation. Combined with the classification and age determination of YSCCs, it is possible to estimate durations for the various phases of the star formation process and the GMC lifetime in this nearby galaxy. This is important in order to link the physical conditions within GMCs to the processes that regulate star formation, its efficiency, and possible time variations (Elmegreen 2007; Murray 2011).

The paper plan is the following. In Sect. 2 we describe the new GMC catalogue and introduce the cloud classification scheme. In Sect. 3 we describe the MIR-source catalogue, introduce the YSCC classification scheme, and discuss the association between the GMCs and YSCCs. In Sect. 4 we examine the association between GMCs and optically visible YSCs and other sources related to the SF cycle. The catalogues of GMCs and

YSCCs and their classification are provided in Tables 5 and 6. The properties of molecular cloud classes and of YSCC classes are presented in Sects. 5 and 6. Molecular cloud lifetimes across the M33 disk are analyzed in Sect. 7. Section 8 summarizes the main results.

## 2. Molecular clouds: properties and classification

M33 is a low-luminosity spiral galaxy, the third most luminous member of the Local Group, with a well-determined distance  $D = 840$  kpc (Freedman et al. 1991; Gieren et al. 2013). The GMC catalogue is a product of the IRAM 30 m all-disk CO  $J = 2-1$  survey of M33 presented in Druard et al. (2014) using a modified version of the CPROPS package, originally developed by Rosolowsky & Leroy (2006), which is described in detail by Gratier et al. (2012). The CO data cube has a spatial resolution of 12 arcsec which, for the adopted distance of M33, corresponds to a physical scale of 49 pc (hence, GMCs are not well resolved). The spectral resolution of the data cube is  $2.6 \text{ km s}^{-1}$ , and the pixel size is 3 arcsec (i.e., 12 pc). CPROPS identifies continuous regions of CO emission in the data cube, and details of our use of the program can be found in Gratier et al. (2012). In Fig. 1 we show the location of the 566 identified GMCs on the CO ( $J = 2-1$ ) map.

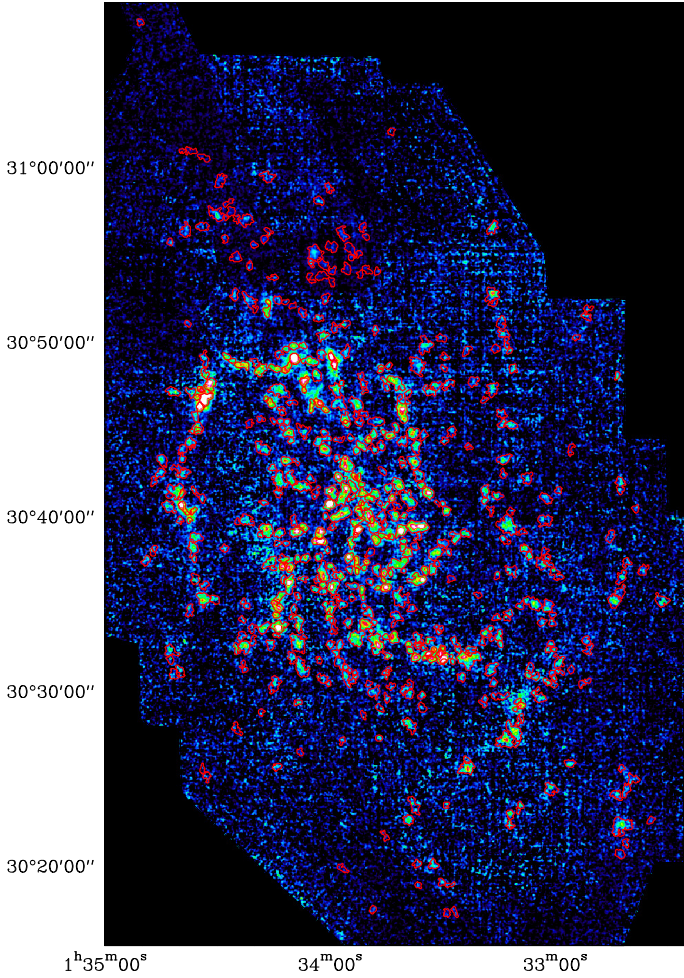
The cloud mass is computed either by converting the total CO line luminosity of the cloud into mass, referred to as luminous mass, or using the virial relation, referred to as virial mass. The luminous mass of the GMC,  $M_{\text{H}_2}$ , is computed for a hydrogen fraction  $f_{\text{H}}$  of 73% using a CO-to- $\text{H}_2$  conversion factor  $X = N(\text{H}_2)/I_{\text{CO}(1-0)} = 4 \times 10^{20} \text{ cm}^{-2}/(\text{K km s}^{-1})$  (Gratier et al. 2017), twice the standard Galactic value, and an intrinsic line ratio  $R_{1-0}^{2-1} = I_{2-1}/I_{1-0} = 0.8$  (Druard et al. 2014). The luminous mass is a function of the CO ( $J = 2-1$ ) line luminosity,  $L_{\text{CO}}$ , which is the CO  $J = 2-1$  line intensity integrated over the cloud. Another way of estimating cloud masses is to assume virial equilibrium. The virial mass,  $M_{\text{H}_2}^{\text{vir}}$ , is a function of the deconvolved effective cloud radius,  $r_{\text{e}}$ , and of the CO ( $J = 2-1$ ) line dispersion (Solomon et al. 1987; Rosolowsky & Leroy 2006). The line dispersion is measured by fitting a Gaussian function to the cloud integrated line profile. If  $\Delta V_{FWHM}^{\text{gau}}$  is the full width at half maximum of the fitted line profile, corrected for the finite channel width (by subtracting in quadrature  $2.6 \text{ km s}^{-1}$ ), then  $\sigma_v^{\text{gau}} = \Delta V_{FWHM}^{\text{gau}}/\sqrt{8 \ln 2}$ . We apply the following relations, which include chemical elements heavier than hydrogen:

$$\frac{M_{\text{H}_2}}{M_{\odot}} = \frac{19.1}{R_{1-0}^{2-1}} \frac{X}{4 \times 10^{20}} \frac{2 m_{\text{p}}}{f_{\text{H}}} \frac{L_{\text{CO}(2-1)}}{\text{K km s}^{-1} \text{ pc}^2} \quad (1)$$

$$= 10.9 \frac{L_{\text{CO}(2-1)}}{\text{K km s}^{-1} \text{ pc}^2} \quad (2)$$

$$\frac{M_{\text{H}_2}^{\text{vir}}}{M_{\odot}} = 1040 \frac{r_{\text{e}}}{\text{pc}} \left( \frac{\sigma_v^{\text{gau}}}{\text{km s}^{-1}} \right)^2. \quad (3)$$

Typical cloud line widths are on the order of  $6-10 \text{ km s}^{-1}$  with  $\sigma_v \sim 3-4 \text{ km s}^{-1}$ . Cloud radii vary between 10 and 100 pc, which is typical of GMCs and complexes. However, given the spatial resolution of the survey, cloud radii have large uncertainties, and in some cases, they might be overestimated. As a consequence, virial masses often turn out to be higher than luminous masses, and scaling relations may apply only to smaller clouds within a GMC complex. We use the luminous mass definition when we refer to the GMC mass, unless stated otherwise. The algorithm finds 566 GMCs (see Fig. 1) with luminous masses

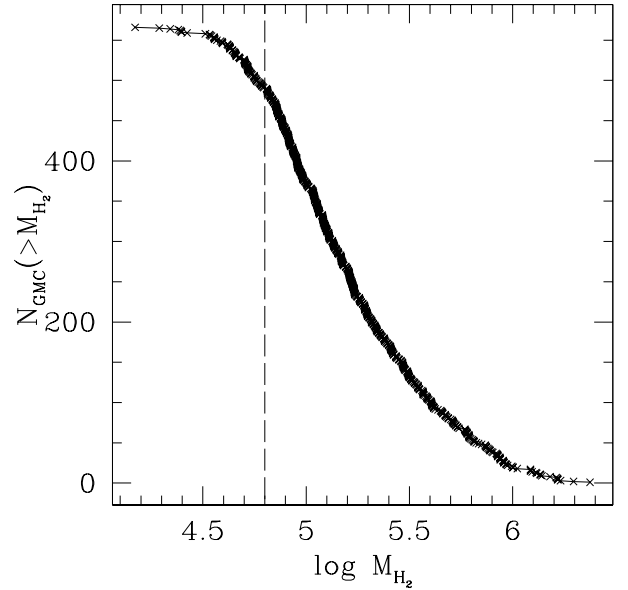


**Fig. 1.** Molecular clouds in the catalogue plotted over the CO  $J = 2-1$  integrated intensity emission map. The map rms is  $0.2 \text{ K km s}^{-1}$  and the intensity scale reaches a maximum of  $4 \text{ K km s}^{-1}$  in antenna temperature units. The thick red contour of each cloud corresponds to the deconvolved size and shape as determined by CPROPS (Rosolowsky & Leroy 2006).

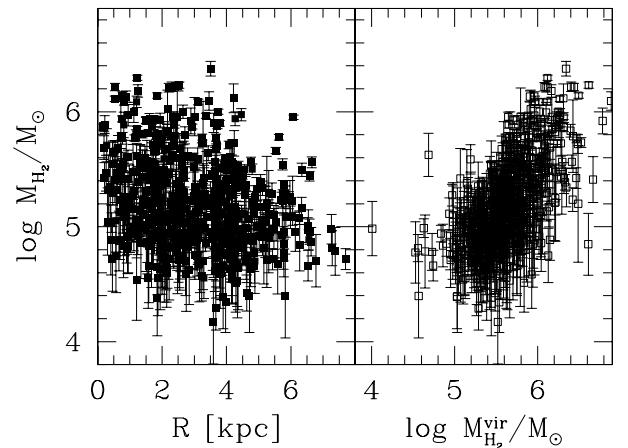
between  $2 \times 10^4$  and  $2 \times 10^6 M_{\odot}$  and virial masses between  $2 \times 10^4$  and  $6 \times 10^6 M_{\odot}$ .

The completeness limit for the luminous masses is about two-thirds of what has been estimated by Gratier et al. (2012) because of the lower rms noise of the Druard et al. (2014) full-disk survey and because the telescope efficiency has been revised. Hence we estimate a completeness limit of  $5700 \text{ K km s}^{-1} \text{ pc}^2$ . Given the assumed CO-to- $\text{H}_2$  conversion factor and  $J = 2-1/J = 1-0$  line ratio, this corresponds to a total cloud mass completeness limit of  $6.3 \times 10^4 M_{\odot}$ , including He. Figure 2 shows the cumulative distribution of the 566 GMC masses.

In the left panel of Fig. 3 we show the radial distribution of the luminous masses of the GMCs, while in the right panel we plot the luminous and virial masses of the clouds. The average luminous mass decreases with radius because the CO cloud luminosity is a decreasing function of galactocentric radius (Gratier et al. 2012). There are fewer massive cloud complexes beyond 4.5 kpc. The virial mass shows a marginal dependence on galactocentric radius because the velocity dispersion (line width) is weakly anticorrelated with the galactocentric distance of the cloud (Gratier et al. 2012; Braine et al. 2010). Clearly, the luminous and virial masses are correlated,



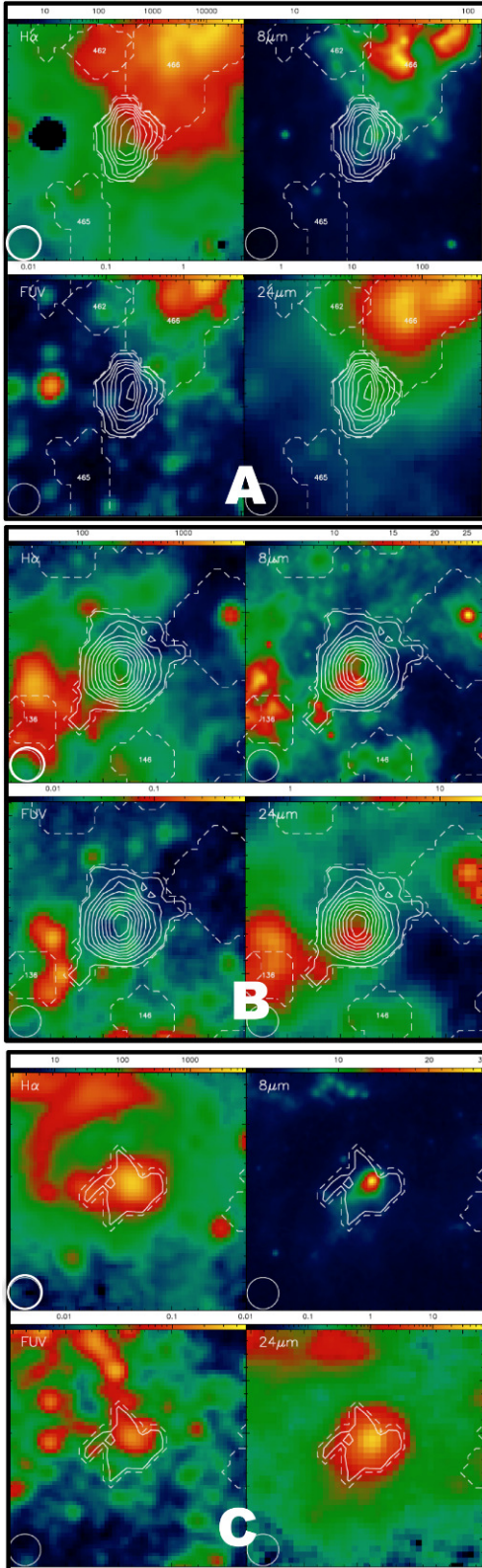
**Fig. 2.** Cumulative distribution of the GMC luminous masses.  $N_{\text{GMC}}$  is the number of GMCs with mass greater than  $M_{\text{H}_2}$ . The dashed line has been placed at the estimated survey completeness limit.



**Fig. 3.** *Left panel:* cloud luminous mass  $M_{\text{H}_2}$  as a function of the galactocentric radius in kpc. *Right panel:* we show the luminous versus the virial masses of the clouds.

even though the dispersion is non-negligible. The right panel of Fig. 3 seems to suggest that GMCs may simply be gravitationally bound entities that are not necessarily in virial equilibrium.

Molecular clouds are classified into three broad categories: clouds without obvious star formation (A), clouds with embedded star formation (B), and clouds with exposed star formation (C). Clouds with embedded or exposed SF are identified from the emission at  $8$  or  $24 \mu\text{m}$ , which in C-type clouds is associated with  $\text{H}\alpha$  and often with far-UV (FUV) emission peaks, while B-type clouds have no optical counterpart. There are a few ambiguous cases that were classified as D-type. This classification follows the spirit of the Gratier et al. (2012) procedure, in which maps at four wavelengths were made of each cloud and the region surrounding it (see the catalogue in Gratier et al. 2012), as displayed in Fig. 4. The wavebands were chosen to probe a variety of optical depths, at a higher angular resolution than the CO data in order to locate the SF region within the molecular cloud.



**Fig. 4.** Four images of three selected areas of the M 33 disk that we used to classify three GMCs: cloud 461 as inactive or A-type (*top*), cloud 147 with embedded SF or B-type (*middle*), and cloud 15 with exposed SF or C-type (*bottom*). The CO  $J = 2-1$  integrated intensity contours for each GMC (solid white line, first contour at  $80 \text{ mK km s}^{-1}$ ) are plotted on maps of H $\alpha$  (*upper left panel*), *Spitzer*  $8 \mu\text{m}$  (*upper right panel*), GALEX FUV (*lower left panel*), and *Spitzer*  $24 \mu\text{m}$  (*lower right panel*). The CO  $J = 2-1$  beam is plotted in the lower left corner of each panel and it is marked with a thicker line in the upper left panel.

The 566 clouds were visually inspected using the maps, as shown in Fig. 4. Strict automatic criteria are extremely difficult to use in a reliable way as the general flux levels decrease greatly with galactocentric radius, making a common threshold impractical because regions of star formation in the outer disk are missed. Furthermore, crowded regions (center, arms) are difficult to analyze without visual inspection. It is not possible to know whether continuum emission comes from the cloud or simply along the line of sight through the disk. When the emission comes from a region near the center of the cloud, or when H $\alpha$  emission is present at the same location of the continuum, which is within cloud contour, the association is assumed to be real. In Fig. 4 we show examples of the images used to classify each of the 566 GMCs. The CO  $J = 2-1$  integrated intensity contours for each GMC (solid white line, first contour at  $80 \text{ mK km s}^{-1}$  and following stepped by  $330 \text{ mK km s}^{-1}$ ) are plotted on maps of H $\alpha$  (*upper left panel*, units give the emission measure in  $\text{pc cm}^{-6}$ ), *Spitzer*  $8 \mu\text{m}$  (*upper right panel*, MJy/sr), GALEX FUV (*lower left panel*, counts), and *Spitzer*  $24 \mu\text{m}$  (*lower right panel*, MJy/sr). The flux levels are different for each image. The scaling used is plotted as a color bar at the top of each panel. For GMC 461 (*upper four panels*) there are no visible sources in any band, and hence it has been classified as A-type. The GMC 147 (*middle four panels*) has a weak infrared source visible at 8 and  $24 \mu\text{m}$  without an H $\alpha$  or FUV emission peak at the MIR peak location, and hence it has been classified as B-type. This cloud is located close to the M 33 center, in a crowded region, and the weak MIR source is missing in the [Sharma et al. \(2011\)](#) catalogue, as discussed in the next section. This underlines the need for visual inspection for a reliable cloud classification. At the location of GMC 15 (*bottom four panels*) a source is visible at all selected wavelengths (that corresponds to a MIR source in the catalogue of [Sharma et al. \(2011\)](#)), and hence GMC 15 has been classified as a C-type cloud.

A star-forming region on the distant side of a cloud is classified as B, although an observer, in M 31, for instance, might see the cloud as exposed (C). Thus, a cloud could be more evolved than its classification shows, either because of the geometry (the case above) or if only low-mass star formation is taking place. We estimate that the maps are sensitive enough to detect a single B0 main-sequence star. There are some differences with respect to the [Gratier et al. \(2012\)](#) classification. A slightly higher fraction of A clouds is found based on the classification of some clouds with very weak and diffuse IR emission as A rather than B or C. A slightly higher fraction of C-type clouds is also found because with the idea that an apparently embedded cloud could be exposed from a different vantage point, we pushed ambiguous B/C cases into the C class. However, the agreement with [Gratier et al. \(2012\)](#) is generally excellent.

Table 1 summarizes the classification scheme and the number of GMCs in each class. In Table 5 we list for each cloud the cloud type and the following properties: celestial coordinates, galactocentric radius  $R$ , cloud deconvolved effective radius  $r_e$ , CO (2-1) line velocity dispersion from CPROPS  $\sigma_v$  and its uncertainty, line velocity dispersion from a Gaussian fit  $\sigma_v^{\text{gau}}$  (corrected for the finite channel width), CO luminous mass  $M_{\text{H}_2}$  and its uncertainty, and virial mass from a Gaussian fit  $M_{\text{H}_2}^{\text{vir}}$ . The uncertainties on the velocity dispersion from CPROPS can be considered upper limits to the uncertainties on the velocity dispersion from a Gaussian fit. Using these and the uncertainties on the cloud radius, we can verify the large uncertainties on the virial mass estimates. There are no estimates of  $\sigma_v^{\text{gau}}$  and  $M_{\text{H}_2}^{\text{vir}}$  when the Gaussian fit to the cloud-integrated profile results in a

**Table 1.** GMC classification scheme.

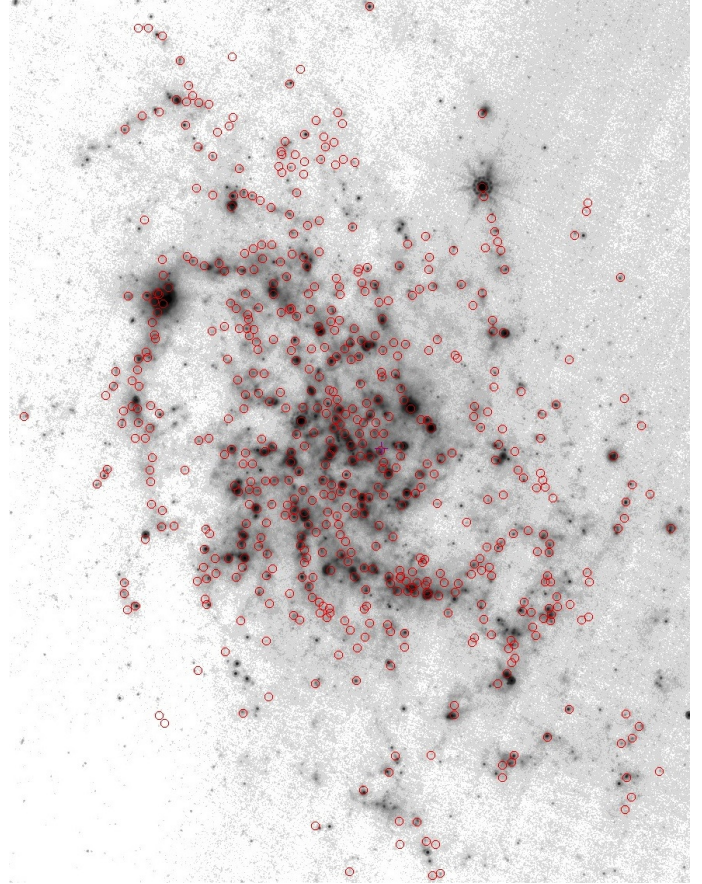
Class	Main properties	#GMCs
A	inactive	172
B	embedded SF, MIR sources only	87
C	exposed SF, MIR+FUV+H $\alpha$ sources	286
D	ambiguous	21

full width at half maximum that is comparable to spectral resolution. In addition, we give in Table 5 the identification number of the YSCCs associated with the GMC (i.e. that lie within the 80 mK km s<sup>-1</sup> GMC boundary). The YSCCs, described in the next section, are listed in Table 6. Clouds of A-type should not host any source, and only five clouds in this category have an associated YSCC that lies at the cloud boundary. The cloud classification was done by seven testers without knowledge of the position of YSCCs (selected via MIR emission peaks). At a later time, the possible presence of sources at the GMC positions has been checked by inspecting the whole M 33 maps at 8  $\mu$ m, 24  $\mu$ m, H $\alpha$ , and GALEX-FUV, each with a uniform contrast. Finally the spatial correlation of catalogued YSCCs and GMCs has been analyzed statistically (see next section).

### 3. Young stellar cluster candidates in the disk of M 33 and their association with GMCs

In the star formation cycle, YSCs form out of molecular gas; as the cluster evolves, stellar activity removes the molecular material, and light starts to escape from the clouds. Eventually, shocks due to massive stars may compress the nearby gas and trigger star formation anew. Before the cold gas is removed from the stellar birth place, newborn stars heat the dust in the surroundings with a consequent emission in the MIR. Therefore we expect the presence of YSCCs at the location of MIR sources that are associated with the M 33 disk and a spatial correlation between YSCCs and star-forming GMCs. The association between YSCCs and GMCs is established following three different methods with different levels of accuracy; each method is described in the following subsections.

Using the *Spitzer* satellite 24  $\mu$ m data, Sharma et al. (2011) have selected 915 MIR sources in the area covered by the M 33 disk. Complementing the mid- and far-infrared *Spitzer* data with UV data from the GALEX satellite and with H $\alpha$  data, it has been possible to build up the spectral energy distribution (SED) for most of the sources. The optical images cover a smaller region than the *Spitzer* images, and for about 60 MIR sources at large galactic radii, the SED could therefore not be derived. The presence in the sample of a few AGBs, the weakness of the emission in the H $\alpha$  or UV bands of some sources, and some large photometric errors have further limited the number of MIR sources with available SEDs to 648 (Sharma et al. 2011). We find that 738 out of the 915 MIR sources catalogued by Sharma et al. (2011) are within the area of the CO survey. By visually inspecting these 738 sources in several bands and with available stellar catalogues, we eliminated obvious AGB and Milky Way stars as well as background galaxies. The SDSS has been used to inspect the source optical morphology or the photometric redshift when the associated H $\alpha$  emission is weak, below  $\sim 10^{36}$  erg s<sup>-1</sup>. We have excluded sources with a reliable redshift determination (with a  $\chi^2 < 3$  as given by SDSS), and a few MIR sources with X-ray counterparts since these are probably background sources (QSOs, galaxies, etc.).

**Fig. 5.** Positions of CO clouds (red circles) on the 24  $\mu$ m *Spitzer* map.

We have a final sample of 630 MIR sources that are strong candidates for being star-forming sites in the area covered by the IRAM CO all-disk survey, and we refer to the young stellar clusters associated with these MIR sources as YSCCs. The purpose of identifying YSCCs in this paper is to associate them with GMCs to study cloud and star formation properties across M 33. The YSCCs may have an optical counterpart or may be fully embedded and detected only in the infrared, while stars are still forming. Soon after stars of moderate mass are born, the dust in the surrounding molecular material absorbs almost all the UV and optical emission of the recently born stars and re-emits the radiation in the MIR and FIR. Hence, MIR sources without optical or UV counterparts might indicate recently born stars still in their embedded phase. Furthermore, we note that small star-forming sites might be below the critical mass to fully populate the initial mass function (IMF) and only occasionally form a massive stellar outlier with H $\alpha$  or FUV luminosity above the detection threshold. This implies that for the purpose of this paper (i.e. to associate YSCCs with molecular clouds) both MIR sources with and without UV or H $\alpha$  counterparts are of interest. Ages and masses are available for 506 YSCCs with UV and H $\alpha$  emission, and have been determined by Sharma et al. (2011).

#### 3.1. Association between GMCs and YSCCs: filamentary structures across M 33

In Fig. 5 we plot the position of the 566 GMCs on the 24  $\mu$ m *Spitzer* image of M 33 to compare their large-scale distribution with the location of MIR emission peaks. Even though the correspondence between MIR peaks, which are YSCCs, and GMCs

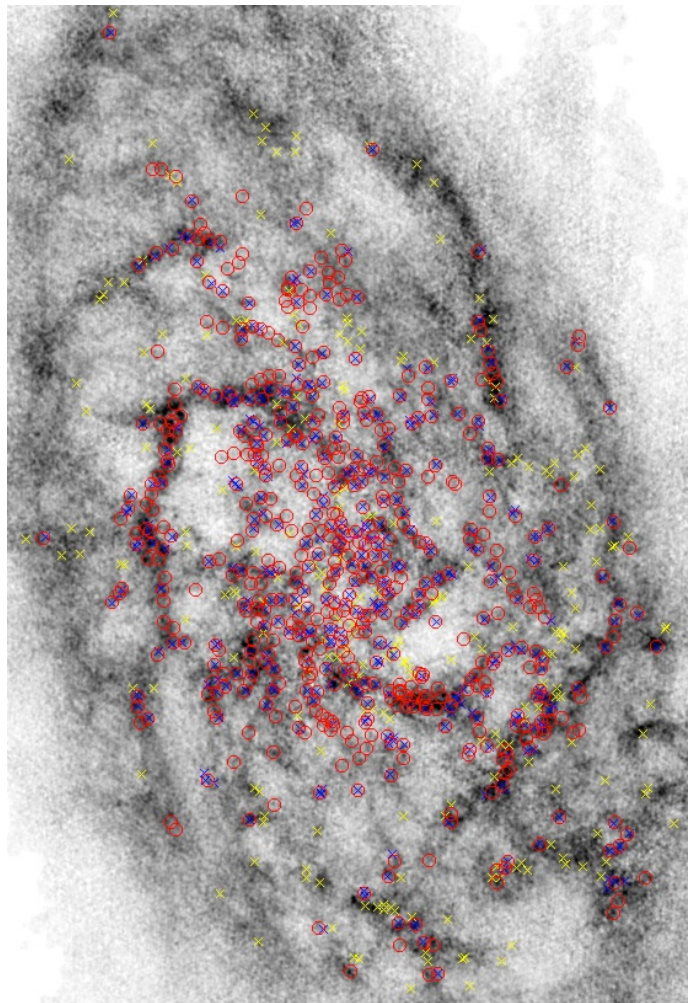
is not a one-to-one correspondence, the majority of GMCs lie along filaments traced by the MIR emission. There are no GMCs in areas devoid of MIR emission. There are some regions, however, where MIR filaments are present but no GMCs have been found. Similarly, some GMCs are present along tenuous and diffuse MIR filaments, but do not overlap with emission peaks, that is to say, with compact sources such as those detected by the Sharma et al. (2011) extraction algorithm. Even using the *Spitzer* 8  $\mu\text{m}$  map at 3 arcsec resolution (better than the 6 arcsec resolution of 24  $\mu\text{m}$  map), some of these clouds seem associated only with diffuse MIR emission.

In Fig. 6 we plot the GMC positions over the HI map at a spatial resolution of 10 arcsec, very similar to the CO map resolution. The 21 cm map presented in Corbelli et al. (2014) is obtained by combining VLA and GBT data. There is an extraordinary spatial correspondence between the GMCs and the distribution of atomic hydrogen overdensities, underlined also by Engargiola et al. (2003) and quantified by Gratier et al. (2012). This correspondence seems to weaken at large galactocentric radii. Here we note bright HI filaments in areas devoid of GMCs. This may be due to a decrease in CO  $J = 2-1$  line brightness far from the galaxy center because of CO dissociation, or due to a gradient in metallicity (Magrini et al. 2010) or in the gas density (which implies a lower CO  $J = 2-1/J = 1-0$  line ratio). Another possibility is that fewer GMCs are formed in the absence of spiral arms. Spiral arms may favor the growth of GMCs by collisional aggregation of smaller clouds. In the absence of the arms, only individual molecular clouds of lower mass than GMCs may be found, which are not detected by the survey. In the outer regions most of the CO  $J = 2-1$  emission is in fact diffuse at the 12'' resolution of our CO survey, which may be due to low-mass clouds. For  $R < 4$  kpc most of the detected CO emission comes from GMCs in the catalogue (Druard 2014). Furthermore, as has been pointed out by Engargiola et al. (2003), a high HI surface density is a necessary condition but not sufficient for the formation of molecular clouds: the atomic gas might just not be converted into molecules if hydrostatic pressure and the dust content decrease, as is the case when we move radially outward in a spiral disk. In Sect. 7 we analyze the drop in the number density of GMCs in the outer disk of M 33 in more detail by examining the association of GMCs with YSCCs and the GMC lifetime.

### 3.2. Association between GMCs and YSCCs: a close inspection of cloud boundaries and the YSCC classification

The spatial correspondence between the position of GMCs and that of YSCCs can be studied by accurately inspecting the area that is covered by each GMC. We start by searching for YSCCs located within 1.5 cloud radii of all GMCs listed in Table 5. Since GMCs are often not spherical, we used a search radius larger than the cloud radius, and we subsequently checked the association by visually inspecting the GMC contours drawn over the M 33 *Spitzer* images at 8 and 24  $\mu\text{m}$  (as for cloud classification). If a YSCC and a GMC overlap in projection, we claim that they are associated.

We searched for optical or UV counterparts to YSCCs by analyzing the  $H\alpha$  and GALEX FUV images of M 33 and by checking the SDSS image at the location of each MIR source. Taking into account whether a YSCC is associated with a GMC and whether it has an optical counterpart, we place the 630 YSCCs into four different categories. We briefly describe



**Fig. 6.** Positions of GMCs and YSCCs analyzed in this paper plotted on top of the 21 cm map at 10 arcsec resolution. Red circles indicate the positions of GMCs, blue and yellow crosses indicate the position of YSCCs that are associated with clouds and those that are not, respectively.

these categories below and discuss them in more detail in the following sections.

- Class b: YSCCs associated with GMCs without an optical counterpart (not identifiable in SDSS and without  $H\alpha$  emission).
- Class c: YSCCs with an optical (SDSS and/or  $H\alpha$ ) counterpart.
  - c1: YSCCs associated with GMCs with coincident  $H\alpha$  and MIR emission peaks, but FUV emission peaks are spatially shifted or absent;
  - c2: YSCCs associated with GMCs with coincident  $H\alpha$ , FUV, and MIR emission peaks;
  - c3: YSCCs not associated with GMCs, but star forming with optical and FUV counterparts; these often have weak  $H\alpha$  emission.
- Class d: YSCCs associated with GMCs that are ambiguous for b or c1/c2 class.
- Class e: YSCCs not associated with GMCs without  $H\alpha$  emission and no or a weak red optical counterpart in SDSS; some FUV may be present.

The lowercase letters used for YSCC categories are such that when GMCs are associated with them, they are mostly placed

**Table 2.** YSCC classification scheme.

Class	Main properties	# YSCCs
b	GMC, embedded SF	97
c	exposed SF (c1 + c2 + c3)	(410)
c1	GMC, H $\alpha$ , MIR coincident	55
c2	GMC, H $\alpha$ , MIR, FUV coincident	216
c3	no GMC, with FUV, weak H $\alpha$	139
d	GMC, ambiguous	19
e	no GMC, no optical counterpart	104

in the corresponding capital letter class. This is why we do not have a-type MIR sources: A-type GMCs are not associated with YSCCs. Similarly, we do not have E-type GMCs because the e-type YSCCs are not associated with GMCs. After an automated search we inspected the images and checked that a YSCC that is associated with a GMC lies within or at the cloud boundary and whether it has H $\alpha$ , FUV, and/or optical counterparts. We find that 243 YSCCs lie beyond any catalogued cloud borders, and 104 of these have no H $\alpha$  counterpart and weak or no emission in the UV. We place these sources in class e. Some of them might not be star-forming sites but foreground or background objects; however, in class e we may also find some small embedded star-forming region whose associated GMC brightness is below the survey detection threshold. We place the remaining 139 YSCCs that are not associated with GMCs in the c3 class. Optically, these look like SF regions. Some might be YSC associated with smaller clouds, which are not in the catalogue, and others might be associated with more evolved sources whose original cold gas reservoir has been mostly dissipated. By examining the ages and masses for c3-type YSCCs, we find ages similar to those for c1- or c2-type YSCCs, while masses are lower on average. Hence it is likely that the majority of the YSCCs of class c3 are associated with molecular clouds of lower mass that are not detected by the survey.

We have 387 YSCCs (61% of the total) with a high probability of being linked to catalogued GMCs since they are within the cloud boundary. A few sources are associated with more than one cloud (since they are spatially overlapping with two clouds that are at different velocities or are at the boundary of two adjacent clouds). We classified 368 of these 387 YSCCs as c1, c2, or b-type according to the presence of an optical counterpart. We place the remaining 19 YSCCs, ambiguous between b- and c-type sources, in class d. We find an optical counterpart to the majority of YSCCs associated with GMCs: 271 out of 368. Only 97 YSCCs, that is, 26% of the YSCCs associated with GMCs, do not have an optical counterpart and are candidates for being YSCC still in their fully embedded phase. The 8 and 24  $\mu\text{m}$  images of one b-type YSCC can be found in the central panels of Fig. 4. Each of the 271 YSCC associated with a GMC has been classified as c1 or c2 type according to whether the FUV emission peak is absent or shifted with respect to the PSF of the 24  $\mu\text{m}$  source, or is overlapping with it. An example of a c2-type YSCC is shown at the center of the four bottom panels of Fig. 4 at various wavelengths. The summary of the YSCC classification is given in Table 2.

The 630 YSCCs that are within the CO all-disk survey map boundary are listed in Table 6 with the type (from b to e) as described in the previous paragraph, and the identification number of the GMCs associated with them, if any. In this case, we also list the corresponding cloud class (A, B, C, or D). In addition, for each YSCC we give the celestial coordinates, the bolometric, total infrared, FUV, and H $\alpha$  luminosities, the estimated

mass and age, the visual extinction, the galactocentric radius, the source size and flux at 24  $\mu\text{m}$ . The estimates of all quantities given in Table 6 and their uncertainties are discussed in Sharma et al. (2011). Photometric errors on source luminosities are smaller than 0.1 dex. However, since for a source of a given size we perform surface photometry with a fixed aperture at all wavelengths, the errors due to nearby source contamination can increase the uncertainties in a crowded field (FUV in the center and spiral arms, for example) and are hard to quantify. The uncertainties of the 24  $\mu\text{m}$  flux, as given by SExtractor, are available in the online tables of Sharma et al. (2011). As described by Sharma et al. (2011), we apply an average correction to stellar masses of low-luminosity YSCC for the IMF incompleteness. While uncertainties on the distribution of cluster masses and on the mass of bright individual YSCCs are on the order of 0.1 dex, the IMF incompleteness implies larger uncertainties on individual YSCC masses when  $L_{\text{bol}} < 10^{40}$  erg s $^{-1}$  (Sharma et al. 2011, see their Fig. 11).

As expected, YSCCs are associated with GMCs of B, C, or D type. We have only five YSCCs that lie at the boundary of A-type GMCs, and the cloud testers considered that these peaks were probably not associated with GMCs. We expect a correspondence between b-type YSCCs and B-type GMCs, or c-type YSCCs and C-type GMCs. This is indeed what happens, even though the cloud classification was not based on the correspondence with MIR sources in the Sharma et al. (2011) catalogue. However, there are a few exceptions, and the apparent non-correspondence between the cloud and the associated YSCC classification needs some explanation. Sometimes a cloud hosts more than one source: when a b-type YSCC and a c-type YSCC are associated with one cloud, then the cloud is classified as C-type. Sometimes a YSCC is identified with a MIR source that effectively is a blend of two or more sources, and only one of them has an optical counterpart. In this case, the cloud and the YSCC might not belong to a similar class. There are also a few clouds classified as B-type or C-type without an associated YSCC. Clearly, for most of these cases, the MIR emission at 24  $\mu\text{m}$  was too weak for the emitting area to be classified as a “source”. In a few cases, blending or confusion with nearby sources might have caused the failure of the source extraction algorithm (see Sharma et al. 2011, for references and a description of the SExtractor software). Only 18 out of 87 GMCs of B-type are not associated with YSCCs, and 46 GMCs out of 286 of C-type do not host a YSCC.

We can summarize that 332 GMCs are associated with at least one YSCC. We have 58 clouds with a weak non-catalogued MIR source (with an H $\alpha$  counterpart) and 176 clouds that do not have associated emission in the UV, optical, or MIR and are considered inactive.

### 3.3. Association between GMCs and YSCCs: the spatial correlation function

In order to quantify the link between GMCs and YSCCs, we statistically analyze the association by computing the positional correlation function of the two distributions. We then compare this to what we expect for a random distribution. This approach is fully justified provided that the distribution of each class of sources, taken separately, is spatially homogeneous. This is not true for two reasons: (i) the density of objects changes with galactocentric distance; and (ii) GMCs and YSCCs are mostly located along the spiral arms, which implies that even the clouds that have not yet formed stars are expected to lie closer to a YSCC with respect to a randomly distributed population in the

disk. However, since the average distance between sources of these two classes is much smaller than these large-scale variations, the analysis that we present below is reasonably justified. In our treatment we take the density dependence on the galactocentric distance into account, but we do not apply any correction for the spiral pattern modulation, we only briefly discuss it in our analysis. In the next section we examine the positional correlation of GMCs and other populations in the disk, in order to check that the correlation with YSCCs, emitting in the MIR, is indeed the strongest.

To compute the expected positional correlation function of YSCCs and GMCs in a galaxy disk, we consider  $f_{\text{YSCC}}(R)$ , the radial density distribution of YSCCs. This is an azimuthal average of the deprojected local surface density, that is, of the surface density on the galaxy plane. We can then compute  $N(d, R)$ , the average number of YSCC within a circle of radius  $d$  centered on a randomly selected GMC, and  $P(d, R)$ , the probability of finding the closest YSCC to a GMC at a distance  $d$ . Having  $P(d, R)$ , it is then straightforward to retrieve the cumulative probability function  $C(d, R)$ , which is the probability of finding a YSCC within a distance  $d$  of a GMC, or equivalently, the fraction of GMCs that have at least one YSCC within a distance  $d$ . We can then compare  $C(d, R)$  with what is observed in a galaxy disk.

Since we have a limited number of YSCCs and GMCs in the M 33 disk, we take the density dependence on the galactocentric distance into account by dividing M 33 into three radial intervals: 1)  $R < 1.5$  kpc; 2)  $1.5 \leq R < 4$  kpc; and 3)  $R \geq 4$  kpc. There are 105, 290, and 236 YSCCs in zones 1, 2, and 3, respectively, and in each zone we compute the mean density of YSCCs as the ratio between the number of YSCCs and the disk area of the zone. We refer to the mean YSCC density using the symbol  $\langle f_{\text{YSCC}} \rangle_i$  where  $i = 1, 2,$  and  $3$  for zones 1, 2, and 3, respectively, and the angle brackets indicate that it is an average over the zone. We define  $\bar{d}$  as the radius of a circle inside which there is on average one YSCC for a randomly distributed population. The length scale  $\bar{d}$  is a typical separation length for YSCCs that are randomly distributed in a plane, and in general, it is a function of  $R$ . For each zone we determine  $\langle \bar{d} \rangle_i = 1/\sqrt{\pi \langle f_{\text{YSCC}} \rangle_i}$ , an average value of the separation length, and find the following values: 146 pc, 218 pc, and 374 pc for zones 1, 2, and 3 respectively. It is useful to define this length scale because reasoning in terms of normalized distances allows us to better compare results obtained for the different subsamples.

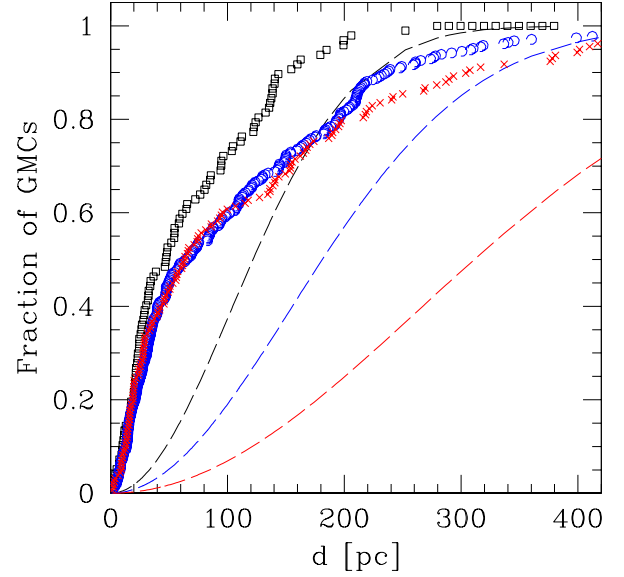
Therefore, in the absence of correlations, the average number of YSCCs within a circle of radius  $d$  centered on a randomly selected GMC,  $N(d, R)$ , can be approximated in each zone as

$$N_i(d) = \pi d^2 \langle f_{\text{YSCC}} \rangle_i = \left( \frac{d}{\langle \bar{d} \rangle_i} \right)^2. \quad (4)$$

The probability  $P(d, R)$  of finding the closest source to a GMC at a distance  $d$  and the corresponding cumulative probability  $C(d, R)$  can be retrieved in general from  $N(d, R)$  as

$$P(d, R) = e^{-N(d, R)}; \quad C(d, R) = 1 - e^{-N(d, R)}. \quad (5)$$

In the absence of correlations, we use Eqs. (4) and (5) to compute the mean expected positional correlation functions of GMCs and YSCCs for random distributions in each zone. They are shown as dashed lines in Fig. 7 as a function of the separation between a GMC and a YSCC given in parsecs. In the same figure we plot the observed fraction of GMCs with at least one YSCC at a separation  $d$  using black squares for the inner region, open blue circles for the middle region, and red crosses for the outer region. Assuming that a YSCC and a GMC are closely associated



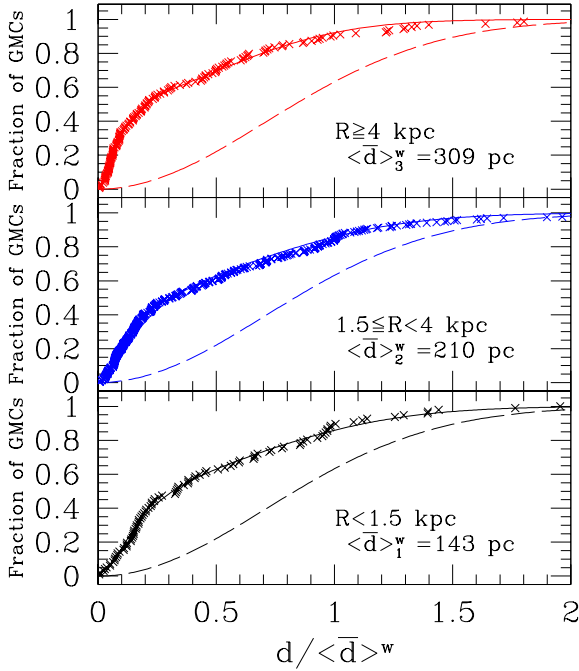
**Fig. 7.** Positional correlation of GMCs with YSCCs. We show the fraction of GMCs with at least one YSCC at a separation  $d$ , given in pc, using black squares for the inner region, open blue circles for the middle region, and red crosses for the outer region. The dashed lines are the fractions expected for a random association. A significant clustering is found.

when they are separated by no more than a cloud radius, which is typically  $r_e \simeq 50$  pc, we expect to find a fraction of GMCs on the order of  $r_e^2/\langle \bar{d} \rangle_i^2$  that have a YSCC closer than 50 pc if YSCCs are randomly distributed. On average, we should find 12%, 5%, and 2% of GMCs with a YSCC within  $r_e$  for zones 1, 2, and 3, respectively. Instead we observe GMC fractions of 53%, 45%, and 43% for zones 1, 2, and 3, respectively. This implies that we have about 4 times more YSCCs in zone 1, 9 times more YSCCs in zone 2, and 21 times more YSCCs in zone 3 that are in close association with GMCs than what might be expected if GMCs and YSCCs were randomly distributed.

We can also display the distributions of Eq. (5) as a function of the normalized distance  $d/\langle \bar{d} \rangle_i$ , and in this case, only one curve is necessary to determine the randomly distributed population, independently of the disk zone. Before doing, this we would like to introduce some weighted average quantities to better take radial variations of GMC densities into account as we move from the crowded central areas of M 33 to the disk outskirts. If YSCCs and GMCs are randomly but non-uniformly distributed in the disk, we can compute the mean separations and densities of YSCCs as seen by the GMC population. In this case, a good approximation to the density of YSCCs in each zone is a weighted mean, with the weights given by the GMC number densities. We call this weighted mean  $\langle f_{\text{YSCC}} \rangle_i^w$  (with  $i = 1, 2,$  and  $3$  for zones 1, 2, and 3, respectively). Analogous to  $f_{\text{YSCC}}(R)$ , we define  $f_{\text{GMC}}(R)$ , the radial density distribution of GMCs in the galactic plane. The weighted average density of YSCCs in the surrounding of GMCs for a random distribution can then be estimated as

$$\langle f_{\text{YSCC}} \rangle_i^w = \frac{\int_{R_{\min,i}}^{R_{\max,i}} R f_{\text{GMC}}(R) f_{\text{YSCC}}(R) dR}{\int_{R_{\min,i}}^{R_{\max,i}} R f_{\text{GMC}}(R) dR}, \quad (6)$$

where  $R_{\min,i}$  and  $R_{\max,i}$  are the radial boundaries of each disk zone. We can then estimate the mean separation between YSCCs and GMCs for a random distribution as  $\langle \bar{d} \rangle_i^w = 1/\sqrt{\pi \langle f_{\text{YSCC}} \rangle_i^w}$



**Fig. 8.** Positional correlation of GMCs with YSCCs for the three selected radial ranges in the M 33 disk as a function of  $d/\langle \bar{d} \rangle^w$ . The innermost and central regions have a quite similar positional correlation function when the distances are scaled with  $\langle \bar{d} \rangle^w$ . The continuous lines are the fits, with our model accounting for a correlation, while the dashed curves represent the expectation in the absence of a correlation.

and use this to normalize  $d$ . We have  $\langle \bar{d} \rangle_1^w = 145$  pc,  $\langle \bar{d} \rangle_2^w = 210$  pc, and  $\langle \bar{d} \rangle_3^w = 309$  pc. Thus, the difference with the earlier calculations is small and mostly relevant for zone 3. Using the weighted average YSCC density for each zone, as in Eq. (6), we can compute the quantities  $N(d)$ ,  $P(d)$  and  $C(d)$  from Eqs. (4) and (5). In Fig. 8 the random distribution is shown as a function of  $d/\langle \bar{d} \rangle^w$ . The fraction of GMCs with at least one YSCC at a distance  $\langle \bar{d} \rangle^w$  for a random distribution is 0.63, and almost all YSCCs are at a distance  $d < 2\langle \bar{d} \rangle^w$  from a GMC. The crosses in Fig. 8 show the observed cumulative functions in the three disk zones. As stated before, the true distributions are far from random since about half of the YSCCs are within  $0.25 d/\langle \bar{d} \rangle^w$  of a GMC.

In what follows we model the observed positional correlation function of Fig. 8 and determine the correlation length. For a simplified approach, we propose the following form for the observed average density of YSCCs at a distance  $d$  from a randomly chosen GMC:

$$F(d) = \frac{c_0}{\pi \bar{d}^2} + \frac{c_1}{2\pi \lambda_c^2} e^{-d/\lambda_c}, \quad (7)$$

where  $\lambda_c$  plays the role of a correlation length, and the length scale  $\bar{d}$  is, as stated before, the typical separation length for YSCCs randomly distributed in a plane. In the absence of correlation, we should expect  $c_0 = 1$  and  $c_1 = 0$ ; in the case of a positive correlation, we instead expect to have  $c_1 > 0$ , as well as  $c_0 < 1$  to balance the average density on randomly chosen positions. The quantity  $N(d)$  can be computed by integration, giving

$$N(d) = \int_0^d 2\pi d' F(d') dd' = c_0 \frac{d^2}{\bar{d}^2} + c_1 \left( 1 - \left( 1 + \frac{d}{\lambda_c} \right) e^{-(d/\lambda_c)} \right). \quad (8)$$

**Table 3.** Fits to GMCs and YSCCs positional correlation functions.

Zone	$\langle \bar{d} \rangle^w$ [pc]	$c_0$	$c_1$	$\lambda_c$ [pc]
$R < 1.5$ kpc	143	1.23	0.73	15.8
$1.5 \leq R < 4.0$ kpc	210	1.17	0.67	17.7
$R \geq 4.0$ kpc	309	1.82	0.73	17.9

Using the values of the weighted mean YSCC separation  $\langle \bar{d} \rangle_i^w$  in Eq. (8), we derive the expected  $P(d)$  and  $C(d)$  for each zone as usual. By comparing the modeled  $C(d)$  to the observed fractions of GMCs that have at least one YSCC within a given value of  $d/\langle \bar{d} \rangle^w$ , we retrieve the average values of  $c_0$ ,  $c_1$ , and  $\lambda_c$  in each zone. A least-squares fitting method of the data is used to determine  $c_0$ ,  $c_1$ , and  $\lambda_c$  in the modeled  $C(d)$ .

The values of  $\langle \bar{d} \rangle^w$  and all parameters derived from our fits for the three zones are listed in Table 3. The continuous lines in Fig. 8 shows the very good quality of the fits, and for comparison, also the expectation in the absence of correlation (dashed line). It is easy to see that by introducing a correlation, we obtain fits far better than without it. The correlation length is very similar in the three zones: 15.8, 17.7, and 17.9 pc, which seems to outline the physical relation between the GMC and YSCC within the same star-forming region. There is a highly statistically significant clustering of GMCs and YSCCs at larger distances. The relative density contrast of the correlated pairs within a circle of radius  $3\lambda_c$  (around 50 pc in all three cases) is 82.3%, 87.7%, and 91.4%, respectively. The reason why  $c_0$  is larger than unity in all cases, while it should have been lower than this, is most likely an effect of radial density variations within each zone. The mean separations vary across the disk and have been averaged by combining regions with rather different local source densities. This is especially true at large galactocentric distances justifying why the highest value of  $c_0$  is found for the outer region (here the radial density of sources drops and density variations are higher).

The steeper increase of the positional correlation function around 100–200 pc in the three zones is probably due to spiral arms or gas-rich filaments. These filaments host many CO clouds and YSCCs and have thicknesses of 100–200 pc. The positional coincidence of YSCCs and GMCs in the outer disk is extraordinary, and chance alignments are much less likely than in the inner disk. It would be interesting to analyze the effects of both GMC and YSCC crowding on the spiral arms. The present treatment is not suited to investigate such effects quantitatively, however, because the correlation is taken to be isotropic, while in the presence of a spiral pattern this assumption should be released. We plan to devise a 2D extension of the present analysis in a future work, aimed at outlining also the correlation pattern at intermediate distances (on the order of 100–200 pc). The short-range correlation presented here is not appreciably affected by spiral arm crowding, however, since it corresponds to distances that are much smaller than the typical width of the arms.

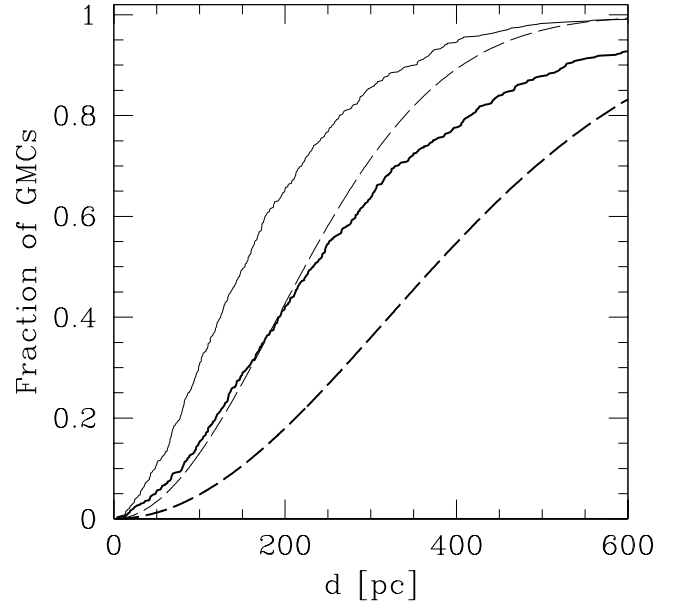
When we subdivide the clouds into a high-mass and a low-mass sample, according to whether the GMC has a luminous mass higher or lower than  $2 \times 10^5 M_\odot$ , we find that 69% of the high-mass GMCs have a YSCC within 50 pc, while this is true only for 44% of the low-mass GMCs. This implies that it is rarer to find inactive GMCs of high mass than of low mass. For a random association the percentage is much lower, on the order of 3%.

#### 4. Association between GMCs and other types of sources in the disk

In nearby galaxies the spatial resolution of current observations is high enough to identify the various products of the gas-star formation cycle, such as  $H\alpha$  regions, massive stars, and embedded or optically visible stellar clusters. Engargiola et al. (2003) used the BIMA disk survey of the inner disk of M 33 to correlate the position of 148 GMCs with HII regions identified through  $H\alpha$  emission. They found a significant clustering especially for the high-mass GMCs. Given the high number of identified HII regions (about 3000 in the innermost 4 kpc), the difference in the number of HII regions closer than 50 pc to a GMCs with respect to the random distribution is only a factor 2. Hence the clustering of our sample of YSCCs around GMCs is much stronger than that of HII regions, and this can be easily understood since the YSCCs have been identified through MIR emission, which is present in the early phases of star formation, while less compact  $H\alpha$  sources, such as shells and filaments, are formed at a later stage during the gas dispersal phase.

Miura et al. (2012) analyzed the location of 65 massive GMCs in the inner disk of M33 with respect to massive stars, which they identified through the optical surveys of Massey et al. (2006) in various optical bands. The aim was to study the evolution and lifetime of GMCs by estimating the ages of the closest bright stars using stellar evolution models. In particular, the authors confirmed a scenario of recursive star formation since dense molecular gas, the fuel for the next stellar generation, is found around previously generated massive stars. The clusters studied by Miura et al. (2012), with estimated stellar masses in the range  $10^{3.5}-10^{4.7} M_{\odot}$  and ages between 4 and 31 Myr, are on average older and more massive than YSCCs associated with MIR sources. Unfortunately, the authors did not list the positions of OB associations and clusters and did not determine the statistical significance of the clustering around GMCs.

To check whether optically visible stellar clusters lie near GMCs, we use the compilation of Fan & de Grijs (2014), who identified 707 stellar cluster candidates. They determined ages and masses of 671 of them through UBVRI photometry using archival images of the Local Group Galaxies Survey (Massey et al. 2006). Some of these clusters have ages  $\leq 100$  Myr. We find that 668 of the 707 clusters lie in the area of the CO survey, and of these only 64, about 10%, are separated by less than one cloud radius from the nearest GMC. The estimated ages of these 64 clusters vary between 5 Myr and 10 Gyr with a mean value of about 50 Myr. Since the expected fraction of GMCs within 50 pc distance from a stellar cluster is only a factor 3 smaller for a random distribution, the statistical significance of the association between GMCs and optically visible clusters is far less than with YSCCs associated with MIR sources. We also checked the association considering only optically visible stellar clusters whose ages are younger than 100 Myr (about one-third of the sample), as displayed by the thicker lines in Fig. 9. The ratio between the observed fraction of GMCs in the proximity of optically selected YSCCs and the expected fraction for a random distribution is larger than for the whole sample of optically selected clusters. Even in this case, however, optically visible clusters are less correlated with GMCs than the infrared selected YSCCs examined in this paper. Moreover, the peak of the excess of the observed fraction of GMCs with respect to a random distribution is at distances on the order of 200–300 pc, much larger than a typical cloud radius, as shown in Fig. 9. The weak correlation found is then likely driven by the location of GMCs



**Fig. 9.** Positional correlation of GMCs with optically visible stellar clusters. We show the fraction of GMCs with at least one YSC at distance  $d$  shown as a function of  $d$  in parsecs. The dashed line is the fraction expected for random association. The thicker lines refer to YSCCs with ages younger than 100 Myr.

and YSCCs along gaseous filaments or spiral arms and not by a real association.

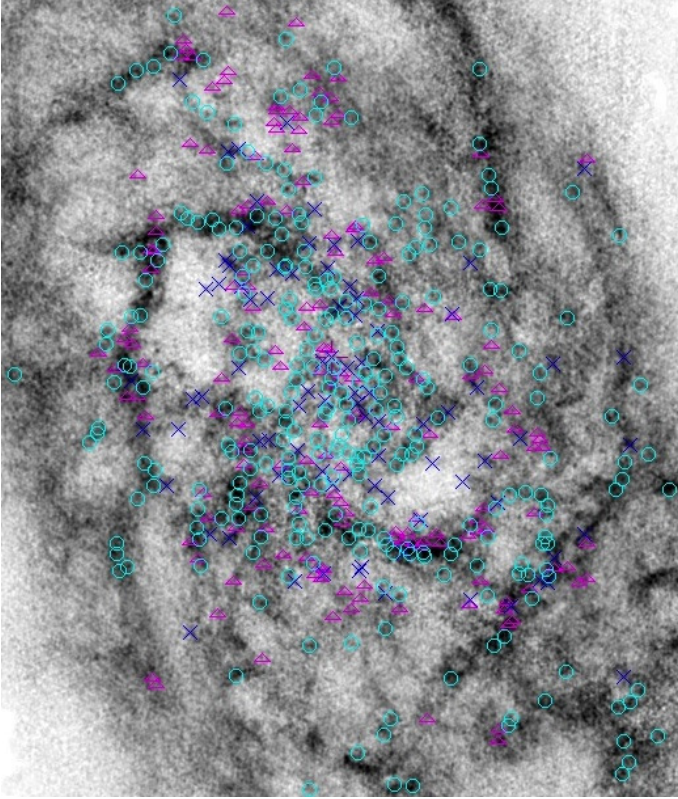
#### 5. Cloud classes and their properties

In this section we examine cloud classes as being representative of different stages of cloud evolution and their properties, such as their location across the disk and the associated CO luminosities. We have classified the population of molecular clouds in M33 into three categories according to whether they are non-active (A-type), have MIR emission without an optical counterpart (B-type), or have some MIR emission with associated FUV, optical and  $H\alpha$  emission (C-type). These three classes may correspond to three different stages of molecular cloud evolution.

The 172 clouds classified as A-type are considered inactive because they may have just formed, but have not yet fragmented to form stars. In this category we might have some clouds that are left over from a previous episode of star formation; the stars break out from the cloud, and the most massive stars could enhance the formation of molecular material close to bright  $H\alpha$  filaments by compressing the ISM through winds and expanding bubbles. Some radiation from nearby stellar associations might heat the dust and trigger some diffuse MIR emission.

We associate the 87 clouds of B-type to the early phases of star-formation, where radiation of massive stars is not yet visible in optical or UV bands because it is absorbed by the dust of the surrounding cloud material. In this case, the radiation heats the dust in the clouds that emit in the MIR in localized areas.

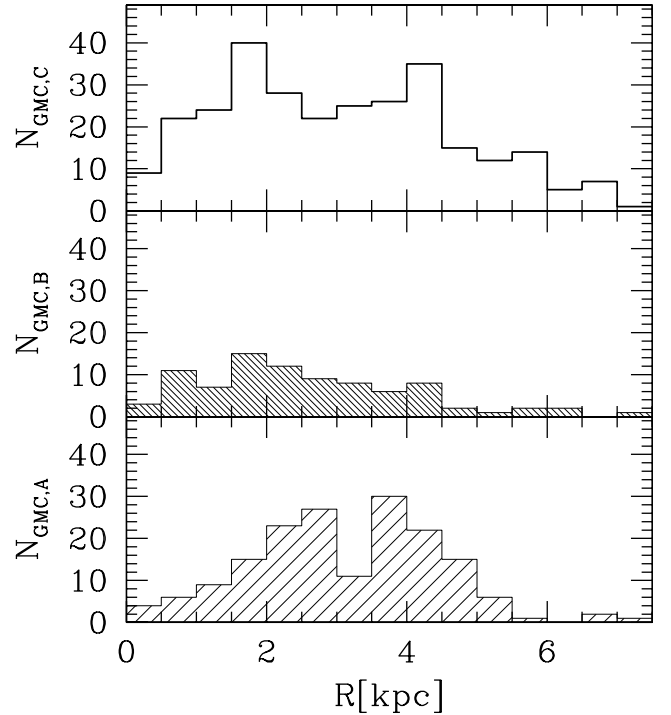
In the 286 clouds of class C, ultraviolet emission from young stars or  $H\alpha$  line emission from ionized gas is visible within the cloud contours. In this case, winds from evolved stars can sweep out gas in their vicinity and UV light escapes from the cloud. The  $H\alpha$  radiation, less absorbed by dust than UV continuum, may become detectable when the HII region is still compact. Winds from young massive stars can ultimately disrupt the parent cloud and quench star formation. If the stellar cluster is of small mass,



**Fig. 10.** Location of the three cloud classes across the M 33 disk shown over the HI image of the galaxy. Open cyan circles are C-type clouds, blue crosses are B-type clouds, and magenta triangles are A-type clouds.

massive stars are not always present (Corbelli et al. 2009). In this case, the lack of  $H\alpha$  radiation might push the classification of evolved clouds into class B and cause the MIR emission to be weak. Moreover, since clouds are not spherically symmetric and stellar clusters are not necessarily born at cloud centers, geometrical effects may play a role in slightly mixing types B and C. On the other hand, emission on the line of sight to the cloud but unrelated to the cloud (i.e., foreground or background) could result in an overestimate of the evolutionary state of the cloud. This implies that in each class there will be a few clouds whose evolutionary stage might be inappropriate.

There are two evident differences in the location of the three types of clouds across the M 33 disk. The GMCs are plotted in Fig. 10 over the HI image of M 33 and in Fig. 11 as a function of galactocentric radius. The A-type clouds have a median galactocentric distance over 3 kpc with only 20% within 2 kpc of the center, where the B- and C-type clouds are more numerous (see Fig. 11). Both the A- and C-type clouds are found along HI filaments and along the northern and southern spiral arm of this flocculent galaxy. Along the southern arm, which is more regular and less disturbed than its northern counterpart, the A-type clouds are found in the inner part of the arm, while C-type clouds are more often found on the arm, where the peak of the HI and  $H\alpha$  emission is. This is in agreement with the theories of formation of molecular complexes across the spiral arms (Roberts 1969; Churchwell 2002): the atomic gas experiences a first compression in the inner part of the arm and becomes molecular, then as it enters the arm, the supersonic turbulence enhances the fragmentation and the process of star formation. This scenario is not seen for the less regular northern arm. A-type clouds are also more clustered than C-type clouds, which are more coarsely



**Fig. 11.** Number of GMCs of A-type, B-type and C-type in 0.5 kpc bins of galactocentric radius are shown in the bottom, middle, and upper panel, respectively.

placed along the gaseous filaments. Clouds of high mass are very rare at large galactocentric radii, possibly because there are no spiral arms that trigger their growth.

The B-type clouds are rarely found close to the spiral arms, on the brightest 21 cm peaks, but they are more numerous along filaments of moderate HI surface density. They are present especially at radii smaller than 4 kpc and are never a dominant population. If they represent a transition phase between inactive and active clouds, the lack of these clouds along the arm means that this transition must be rapid. Lower mass stellar clusters, which are not bright in the UV or  $H\alpha$ , might be associated with these clouds, but the star formation process may continue to increase the YSCC mass at a later stage. Although there might be some contamination among MIR sources with no optical counterpart, due to obscured distant galaxies in the background of M 33, the presence of CO emission at the same location of the MIR sources ensures that most of these are YSCCs associated with M 33. Their clustering along HI overdensities again underlines their association with the M 33 ISM.

We now examine the CO luminosities of GMCs in the three classes as well as their emission at  $24 \mu\text{m}$ . The mean GMC mass progressively increases from  $1.3 \times 10^5 M_{\odot}$  for A-type clouds to  $2.1 \times 10^5 M_{\odot}$  for B-type to  $3.6 \times 10^5 M_{\odot}$  for C-type clouds. Hence, B-type clouds are typically of intermediate mass and not necessarily of low mass. When we interpret the sequence of clouds from A-type to B-type to C-type as a time sequence, the increase in average cloud mass along the sequence implies that GMCs collect more gas as they age. The trend remains when we only consider inner disk clouds (at  $R < 3.5$  kpc) or outer disk clouds (at  $R \geq 3.5$  kpc).

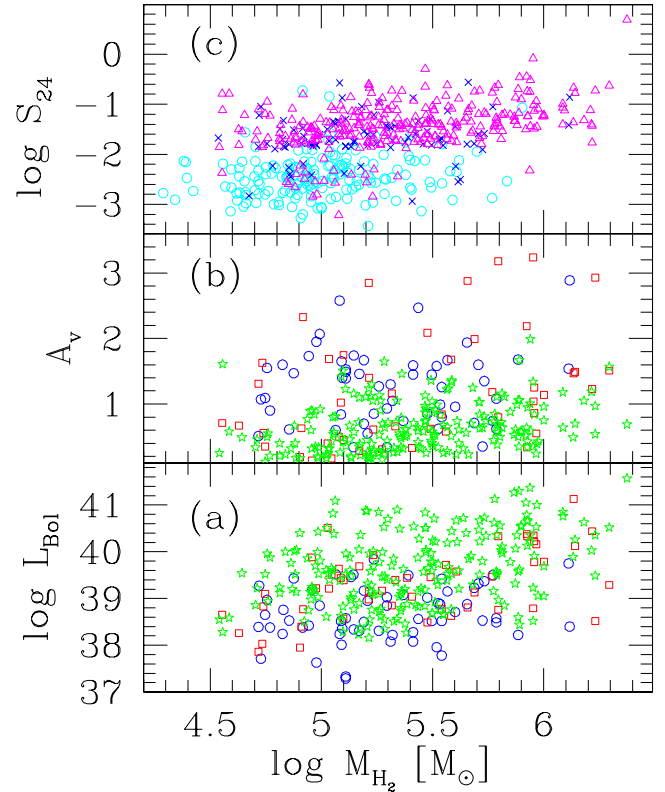
Clouds of C-type extend to higher molecular masses than B-type clouds, and Fig. 12 shows the total luminous masses of the clouds versus the YSCC bolometric luminosity (for GMCs that hosting them), extinction, and flux density at  $24 \mu\text{m}$ . We

computed the MIR flux densities of YSCC associated with B or C-type sources by dividing the flux as measured by Sharma et al. (2011) by the source area (defined at 8-sigma isophote). The A-type clouds have very low MIR flux densities, except for a few that are contaminated by nearby sources. For these and for clouds without associated YSCC, we measured the emission at  $24 \mu\text{m}$  in apertures of 3, 5, and 8 pixels ( $1 \text{ px} = 1.6 \text{ arcsec}$ ) in radius centered on the cloud. After background subtraction, we chose the aperture that gives the highest flux density; this is derived by dividing the flux in the aperture by the aperture area. With this procedure we underestimate the MIR flux density of the emitting area because we cannot measure the effective source size (the emission is indeed so weak that a source size is hard to define, and in fact flux levels are below the fluxes of sources found by the Sharma et al. 2011, procedure). The GMCs associated with catalogued sources have flux densities higher than  $0.01 \text{ mJy/arcsec}^{-2}$  ( $425 \text{ MJy/sr}$ ). This flux limit for catalogued sources, which is evident in panel (c) of Fig. 12, is a result of the source extraction algorithm (Sharma et al. 2011). GMCs without catalogued sources are mostly of A-type, but there are also a few C-type clouds with weak MIR emission, perhaps being in the process of dissolving as a result of the evolution of the newborn cluster. A correlation between the cloud CO luminous mass and the source extinction is found for sources of c2-type, as shown in Fig. 12. The same figure also shows that there is no correlation between the GMC mass and the YSCC bolometric luminosity. Similarly, there is no correspondence between GMC masses and YSCC masses when through successful SED fits, stellar masses have been determined (Sharma et al. 2011). This implies variations in the amount of gas turned into stars within giant complexes where star formation is not uniformly spread out.

There is no difference in the CO line width for clouds of class A, B, and C ( $3.5 \pm 1.5$ ,  $3.3 \pm 1.1$ , and  $3.4 \pm 1.2 \text{ km s}^{-1}$ , respectively). The slightly higher dispersion for clouds of A-type is due to a secondary peak in the width distribution at narrow line widths ( $\sigma_v^{\text{gau}} \sim 2 \text{ km s}^{-1}$ ). Similarly, there is not much difference in size between complexes of different types, except that GMCs of large size are of C-type. Sizes are distributed between 10 and 100 pc with a peak at 30 pc and a mean value on the order of 40 pc for A- and B-type clouds, and of 50 pc for C-type clouds.

## 6. YSCC classes and their properties

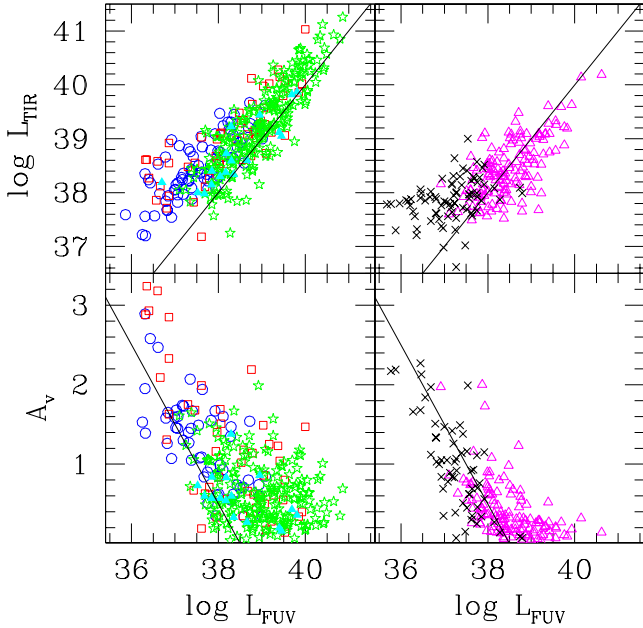
In Sect. 3.1 we defined several classes of YSCCs. YSCCs of b-, c1-, c2-, and d-type are associated with GMCs, while those of c3- or e-type are not. The YSCCs of b- and e-type do not have an optical counterpart or only very faint emission in the SDSS images. YSCCs of c-type are associated with blue light in SDSS images and have  $H\alpha$  and FUV emission. Sources of d-type are ambiguous between b- and c-type. Since sources of b-type are associated with GMCs, they presumably represent the early stages of SF, when stars are fully embedded in the cloud. e-Type YSCCs may be protoclusters in their embedded phase associated with lower mass molecular clouds that have not been detected by the IRAM-survey. Like b-type sources, the e-sources tend to lie along HI overdensities, but not along the brightest ones. The MIR sources identified as e-type YSCCs often lie near a star forming region, and it is not clear whether they host protostars or if the dust is heated by young stars in their proximity. A few are rather isolated and may be background-obscured objects. They are more numerous than b-type YSCCs in regions far from the center. The MIR emission of e-type YSCCs is on average lower than that of b-type YSCCs, and hence they might be associated



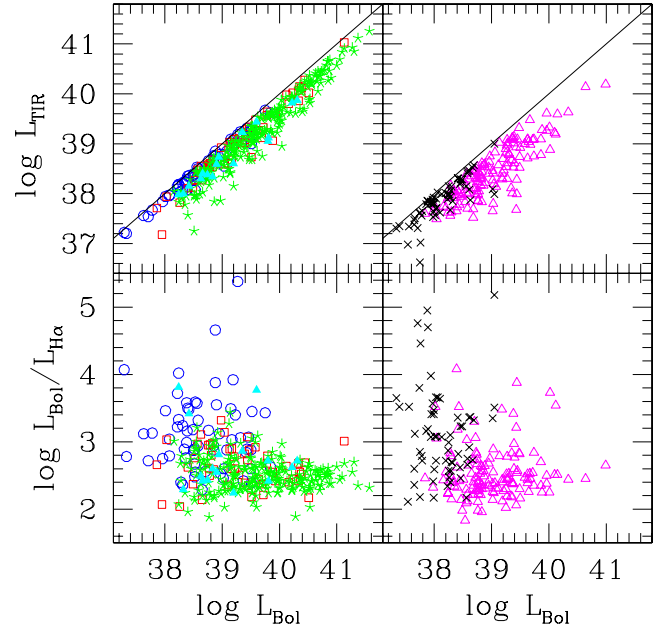
**Fig. 12.** YSCC bolometric luminosity in  $\text{erg s}^{-1}$  and visual extinction are plotted in panel a) and b), respectively, as a function of the total luminous mass of GMC hosting the YSCC. Blue circles are embedded b-type sources, red squares are YSCCs of c1-type, and green stars are YSCCs of c2-type. In panel c) the flux density at  $24 \mu\text{m}$  in units of  $\text{mJy arcsec}^{-2}$  (see text for more details) is plotted as a function of the total GMC luminous mass. Open cyan circles are C-type clouds, blue crosses are B-type clouds, and magenta triangles are A-type clouds.

with small clouds (Corbelli et al. 2011), with a lower metallicity (Magrini et al. 2010) and dust content, undetectable through the IRAM-CO survey.

If the sequence b, c1, c2, and c3, represents a time sequence in the evolution of a star-forming region, we should find progressively less extinction and older ages for the associated clusters. In Fig. 13 we show the YSCC FUV luminosity versus extinction  $A_v$ , as given by Eq. (8) of Sharma et al. (2011) and similar to the visual extinction definition used by Calzetti (2001). In Figs. 13 and 14 we can see that YSCCs of b and c1-type are the most embedded, having the highest values of  $A_v$ . There are two differences between these types of sources, however: some YSCCs of c1-type have high bolometric luminosities and hence may have already formed massive stars. On the birth-line diagram (Corbelli et al. 2009) the most luminous c1-type YSCCs occupy the same area as c2-type YSCCs, while the low-luminosity c1-type YSCCs tend to overlap with the region where b-type sources are. This implies that c1-type YSCCs are intermediate between fully embedded clusters in the process of formation (b-type) and very YSC (c-type). It is likely that c1-type YSCCs do not have FUV emission at the location of the MIR peak because of extinction. The YSCCs of c2-type have  $A_v$  values always lower than 2 and often lower than 1, with an average value of  $0.62 \pm 0.38$ . Their infrared luminosity correlates very well with the FUV luminosity. There are no YSCCs with very low total infrared (hereafter TIR) or FUV luminosity in this class (see Sharma et al. 2011, for the TIR luminosity computation).



**Fig. 13.** YSCC visual extinction and total infrared luminosity as a function of FUV luminosity. In the *lower and upper left panels* we show sources associated with GMCs, in the *lower and upper right panel* we show sources not associated with GMCs. Blue circles: embedded b-type sources; red squares: YSCCs of c1-type; green stars: YSCCs of c2-type; cyan filled triangles: unclassifiable sources with GMCs; magenta triangles: YSCCs without GMCs; and black crosses: sources of unknown nature without GMCs (e-type). Continuous lines of unity slope in all four panels are not fits to the data, but are drawn for reference.



**Fig. 14.** YSCC total infrared luminosity and bolometric-to- $H\alpha$  luminosity ratio are plotted as a function of the MIR-source bolometric luminosity. In the *lower and upper left panels* we show sources associated with GMCs, in the *lower and upper right panels* we show sources not associated with GMCs. Blue circles are embedded b-type sources; red squares are YSCCs of c1-type; green stars are YSCCs of c2-type; cyan filled triangles are unclassifiable sources with GMCs; magenta triangles are YSCCs without GMCs; black crosses are sources of unknown nature without GMCs (e-type). Continuous lines of unity slope in all four panels are not fits to the data, but are drawn for reference.

The lack or faintness of  $H\alpha$  emission for b-type sources is coherent with the picture that in these sites massive stars have not yet formed.

The YSCCs that are not associated with GMCs are of c3-type, star forming with optical counterpart, or of e-type, without optical counterpart. The YSCCs of c3-type mostly have low extinction, and this is consistent with the fact that they are not associated with GMCs. They are not of low luminosity, and hence they might be slightly more evolved than c2-type sources or associated to low-mass clouds that are highly efficient in making stars. In Fig. 15 we plot the location of all YSCCs, color-coded according to their type, superimposed on the HI map of M 33. There are some differences between the various YSCC classes: c1- or c2-type YSCCs lie over bright inner HI filaments or arms. The b-type YSCCs populate the inner disk, and there are very few beyond 4.5 kpc, but they avoid the brightest HI ridges. The c3- and e-type YSCCs are the dominant population at large galactocentric radii. It is likely that at these radii molecular clouds are of low mass to escape detection in the IRAM all-disk survey. e-Type YSCCs have more diffuse MIR emission, often mixed with faint  $H\alpha$  filaments, and hence it is unclear if at the location of these sources any star formation will ever take place.

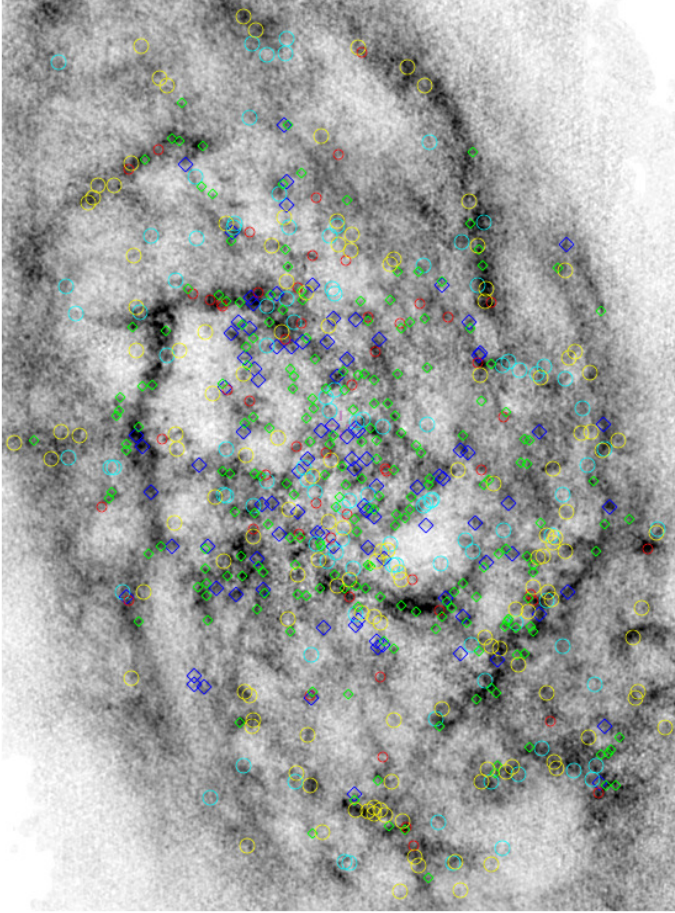
We now examine the luminosities of the YSCCs for the various classes. In Fig. 16 we plot the bolometric luminosity (Sharma et al. 2011) as a function of galactocentric radius for the various classes and the linear fit to the log distributions. We have included all YSCCs with an estimated bolometric luminosity that lie at galactocentric distances  $R < 7$  kpc. We did not consider the uncertainties on the bolometric luminosities because they are hard to quantify as they are not dominated by photometric errors, but by uncertainties related to the modeling. On

average and at all galactocentric radii, the class with the highest luminosities is the c2 class, followed by c1, c3, b, and e classes. Since YSCCs of c-type are likely to be associated with clusters that have completed the formation process, it is conceivable that they are more luminous than b- and e-type sources that are linked to the embedded phase of star formation, when not all cluster members have been formed. In particular, this finding is in agreement with non-instantaneous cluster formation theories, which predict that massive stars are fully assembled at a later stage and their quick evolution switches off the cluster formation process (Zinnecker & Yorke 2007). A similar radial trend to that shown in Fig. 16 is found for the TIR luminosity.

Through SED fits, Sharma et al. (2011) have estimated the ages of YSCCs of c-type. Their distribution is shown in Fig. 17. Given the uncertainties in age determination (on the order of 0.1 dex for bright rich YSCCs and larger for dim YSCCs, because of the IMF incompleteness) the mean ages of c1, c2, or c3-type YSCCs are consistent. It is important to underline that the vast majority of the YSCCs have ages between 3.5 and 8 Myr, and only 15 of them, less than 4%, are older than 8 Myr.

## 7. Molecular cloud lifetime and the star formation cycle

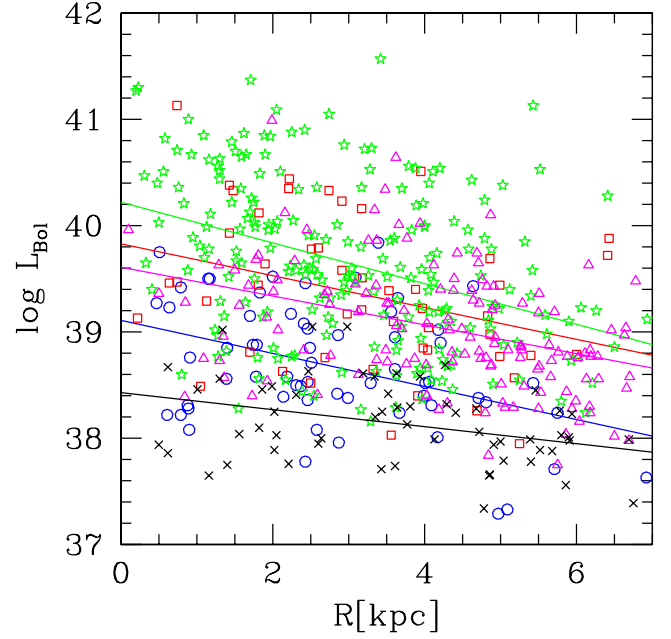
Stars form from gas and leave their imprint on the gas through cloud dispersal or gas compression for the next generation of stars. During the formation phase and for the following few Myr, the dust associated with the molecular material is heated by the energetic photons emitted by the YSC. The CO survey of the LMC at a resolution of 40 pc demonstrated the good correlation of molecular clouds with young stellar clusters of age  $\leq 10$  Myr



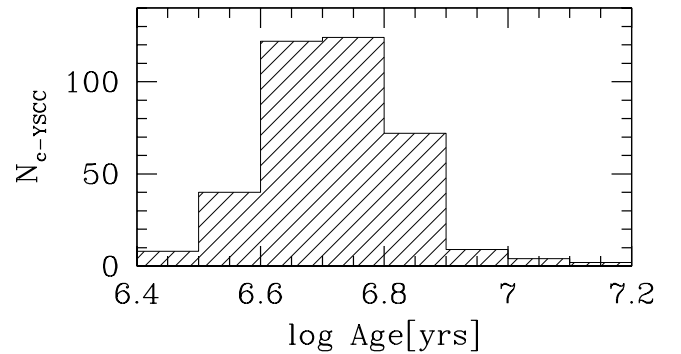
**Fig. 15.** YSCCs in the area covered by the IRAM all-disk survey color-coded and with symbols according to their class type are plotted over the HI map of M33. Rhombus symbols are for YSCCs associated with GMCs (blue for b-type, red for c1-type, and green for c2-type). Circles are for YSCCs that are not associated with GMCs (yellow for c3-type, and cyan for e-type).

(Fukui et al. 1999). The timescale of each evolutionary stage has then been estimated by subdividing GMCs into different classes corresponding to different ages of the associated YSC (Yamaguchi et al. 2001; Kawamura et al. 2009). The LMC is an irregular metal-poor galaxy, interacting with the Milky Way, and hence it is not clear that the same timescales for GMC evolution apply as for M33 or the Milky Way.

One difficulty we have to face for M33 is that despite the excellent spatial resolution of the CO survey, which gives a comparable physical scale to the LMC CO survey, there is no high-resolution radio continuum database that can be used to locate thermal and nonthermal sources to obtain information on the very early phases of stellar cluster formation. Dedicated optical surveys of the LMC have produced a catalogue of clusters and OB associations, with 137 of them being of age  $<10$  Myr. In M33 there are only 16 optically identified clusters with age  $<10$  Myr, and none of these are within the cloud contour of identified GMCs. The MIR-source catalogue of Sharma et al. (2011) with numerous YSCCs is a substitute for the lack of optically identified YSCs in M33. However, the use of these data implies that we have more information on the intermediate stages of YSC formation but less on the dissipation phase, when the YSC emerges from the GMC and is optically identified. The



**Fig. 16.** YSCC bolometric luminosities in  $\text{erg s}^{-1}$  as a function of galactocentric radius color-coded according to the various classes as in Fig. 14. Straight lines show the linear fits to the displayed distribution for each class.



**Fig. 17.** Age distribution of YSCCs of c-type (c1, c2, and c3). Only 15 YSCCs, less than 4%, are older than 7 Myr.

YSC candidates in the Sharma et al. (2011) catalogue have reliable age estimates, especially if they are bright and with coincident peaks in the various bands, such as c2-type YSCCs. All 216 c2-type YSCCs have ages  $\leq 10$  Myr, and 90% between 3.5 and 8 Myr and a marked peak around 5 Myr. Similar ages are found for c3-type YSCCs, which are not associated with GMCs, but with an optical counterpart as well. Figure 17 shows a histogram of the age distribution of c-type YSCCs. In the previous section we showed that the number of YSCC associated with GMCs drops when their age is  $>8$  Myr. If 8 Myr is the typical age of the cluster when it breaks through the cloud, we can say that phase C last 8 Myr. We hence define 8 Myr as the timescale for a GMC of C-type. During this stage, YSC are fully assembled, including massive star formation. Shortly after this stage, the YSC dissipates the associated GMCs.

We can estimate how long the inactive A-type and the embedded B-type phases last based on the number of clouds in the catalog. Considering only GMCs above our survey completeness limit, the total number of classified clouds is 474. Of these

**Table 4.** GMCs and YSCCs of various classes in different radial ranges.

$R$ [kpc]	$N_{\text{GMC}}$	A-type	B-type	C-type	$N_{\text{YSCC}}$	b-type	c1-type	c2-type	c3-type	e-type
<1.5	95	19(20%, 18)	21(22%, 19)	55(58%, 55)	98	21(21%)	10(10%)	39(40%)	9(9%)	19(19%)
1.5–4	296	106(36%, 79)	50(17%, 45)	140(47%, 133)	279	53(19%)	30(11%)	115(41%)	44(16%)	37(13%)
$\geq 4$	150	47(31%, 30)	16(11%, 15)	87(58%, 80)	231	23(10%)	15(6%)	60(26%)	85(37%)	48(21%)

127 are of A-type, 79 are of B-type, and 268 are of C-type. Assuming a continuous rate of star formation in M 33 and that the C-type phase lasts 8 Myr, we estimate that the B-type phase lasts 2.4 Myr and the A-type phase about 3.8 Myr. The shortest phase for GMCs in M 33 is the totally embedded phase, when the newborn cluster has no H $\alpha$  or optical counterpart: this phase was not considered explicitly by Yamaguchi et al. (2001) and Kawamura et al. (2009), but it is included in their inactive type-I phase. We define the GMC lifetime as the time interval between the stage when most of the GMC mass has been assembled but the cloud is still inactive, and the starting of the gas dissipation phase, when most of the molecular gas is dispersed in the interstellar medium. This lifetime includes the cloud inactive phase, the embedded star-forming phase, and the time when the YSC breaks through the cloud and has an optical counterpart. The cloud growth and dissipation times, which involve a substantial change in GMC mass, have not been considered because of the survey completeness limits. We therefore have a lifetime of about 14.2 Myr for GMCs in M 33 before they are dissipated. Considering only GMCs above a limiting mass of  $10^5 M_{\odot}$ , we have 115 clouds of A-type, 58 clouds of B-type, and 221 of C-type. Since the cluster age is not a function of the associated cloud mass, assuming that C-type clouds last 8 Myr before gas dispersal, we estimate that GMCs spend 4.2 Myr in class A and 2.1 Myr in class-B. This gives a total lifetime of 14.3 Myr, which is very close to our previous estimate for the whole sample of GMCs above the completeness limit. This is somewhat shorter than the lifetime derived by Kawamura et al. (2009), who used a sample of GMCs of masses and effective spatial resolution similar to that of our survey. The type I (similar to our A- and B-type clouds) and type II (C-type) evolutionary sequence of Kawamura et al. (2009) lasts 19 Myr in total. In the Milky Way the GMC lifetime has been estimated to be in the range 10–20 Myr (Blitz & Shu 1980; Larson 1981; Murray 2011), which is shorter than earlier estimates of 30–40 Myr (Bash et al. 1977). In M 33, the GMC lifetime, prior to gas dispersal, is comparable with estimates of GMC lifetimes in the Milky Way.

We now subdivide the galaxy in three zones, as in Sect. 3: 1)  $R < 1.5$  kpc; 2)  $1.5 \leq R < 4$  kpc; and 3)  $R \geq 4$  kpc. The number of classified clouds and sources in the three zones is given in Table 4 (unclassified objects such as D-type clouds or d-type sources are not considered here). We list in parenthesis the percentage of the various GMC types for each radial zone, and in addition, the number of GMCs in that zone whose mass is above the completeness limit. We can estimate the cloud lifetime for the three zones considering only GMCs above the completeness limit and find that in zone 2 GMCs have the longest lifetime, on the order of 15.4 Myr, because of the longer time the clouds spend in the inactive phase. In zones 1 and 3 the GMC lifetime is on the order of 13.4 Myr and 12.5 Myr, respectively. In the intermediate radial range, where spiral arms are found, molecular clouds have a longer quiescent time as more A-type clouds are found. The quiescent time is short for GMCs in the inner

regions, about half of that inferred for zone 2. In the outer regions, the embedded phase instead is shorter than elsewhere, about 1.5 Myr. However, as we discuss in the next paragraph, the large drop in GMC number density in zone 3 and the lower average stellar mass of the associated YSCCs increase the uncertainties in the estimated cloud-type fractions and YSC ages, and therefore in the GMC lifetime in the outer zone.

When we move radially outward from zone 2 to zone 3, the number density of GMCs decreases more rapidly than the number density of YSCCs. This can be explained either with an unseen population of clouds of lower mass that is undetected by the IRAM all-disk survey, or by a quicker evolution of GMCs (once stars are formed, GMCs dissolve in a shorter time as the stellar cluster evolves). The number of YSCC without any associated GMCs (c3-type sources) is in fact the dominant population in zone 3, while in the other zones the dominant YSCC population is of c2-type. We now suppose that the larger drop in the density of GMCs at large galactocentric radii with respect to YSCC density is due to an unseen population of molecular clouds, either GMCs with weaker CO emission or molecular clouds of lower mass (below the IRAM survey completeness limit). Assuming that the percentage of missing clouds is the same for each cloud class, that the scale length of the molecular gas surface density is the same at all radii (i.e., 2 kpc) and that the cloud mass spectrum is also radially constant, in zone 3 we should find a number of clouds equal to 75% the number of clouds of zone 2 (i.e. 222 clouds). Hence there are 72 clouds that escaped detection (or even more if their mass decreases with galactocentric radius). Of these, about 42 clouds will be of C-type, 8 clouds of B-type, and 22 of A-type. Consequently, the number of b-type sources would be 31 (13%), the number of c1+c2 sources would be 117 (51%), the number of c3-type would be 41 (18%), and there would be 40 e-type sources. In this case, the percentage of optically visible YSCCs associated with clouds would be the same as in zone 2, and the percentage of sources with or without clouds will also be in closer agreement. This means that if the hypothesis of the unseen population is correct, there is not much difference in the lifetime of molecular clouds across the M 33 disk.

## 8. Summary

In this paper we presented the largest database of GMCs and candidate YSCs across a galactic disk. Using the IRAM 30 m CO  $J = 2-1$  data cube of M 33, we identified 566 GMCs; we selected 630 MIR sources from the list of Sharma et al. (2011) that are YSCCs in the early formation and evolutionary phases. We classified the GMCs as non-starforming (class A), with embedded SF (B), or with exposed SF (C). The YSCCs were sorted into classes based on their emission in the MIR, FUV, and H $\alpha$  bands and according to the association with GMCs. Most of the YSCCs with optical and UV counterparts have estimated ages and masses. The classification helps in drawing a possible evolutionary sequence and the relative timescales. The results of the classification, together with the most relevant parameters of

**Table 5.** Classification and properties of the M33 molecular cloud population.

$N_{\text{Co}}$	Type	RA	Dec	$R$ [kpc]	$r_c$ [pc]	$\delta_{\text{re}}$ [pc]	$\sigma_v$ [km s <sup>-1</sup> ]	$\delta_{\text{rv}}$ [km s <sup>-1</sup> ]	$\sigma_v^{\text{gau}}$ [km s <sup>-1</sup> ]	$M_{\text{H}_2}$ [ $M_{\odot}$ ]	$\delta_{M_{\text{H}_2}}$ [ $M_{\odot}$ ]	$M_{\text{H}_2}^{\text{vir}}$ [ $M_{\odot}$ ]	$N_{\text{YSCC}}$
1	C	23.36171	30.29141	5.7	37	20	4.1	2.5	3.3	0.84E+05	0.48E+05	0.42E+06	2
2	C	23.37007	30.28943	5.8	42	18	4.1	1.4	3.7	0.71E+05	0.29E+05	0.60E+06	...
3	A	23.30009	30.38398	4.6	36	14	6.2	2.4	2.6	0.19E+06	0.30E+05	0.26E+06	...
4	A	23.32977	30.48902	3.1	56	35	4.7	3.7	2.8	0.10E+06	0.11E+06	0.44E+06	...
5	B	23.32793	30.49364	3.1	43	35	1.8	2.0	2.3	0.11E+06	0.14E+06	0.24E+06	87
6	C	23.38505	30.33558	5.1	56	16	8.2	2.2	6.1	0.24E+06	0.41E+05	0.22E+07	13, 10
7	A	23.43679	30.49438	2.7	33	22	3.4	1.6	3.1	0.86E+05	0.38E+05	0.32E+06	...
8	C	23.31071	30.49623	3.3	55	14	4.9	1.3	7.1	0.26E+06	0.61E+05	0.29E+07	644
9	C	23.17682	30.37319	6.1	119	13	4.5	0.6	4.1	0.90E+06	0.40E+05	0.21E+07	603, 601, 602, 600
...	...	...	...	...	...	...	...	...	...	...	...	...	...

**Notes.** The full table is available at the CDS.

**Table 6.** Classification and properties of the M33 young stellar cluster candidate population.

$N_{\text{YSCC}}$	Type	$N_{\text{Co}}$	Type	RA	Dec	$\lg(L_{\text{bol}})$ [erg s <sup>-1</sup> ]	$\lg(L_{\text{IR}})$ [erg s <sup>-1</sup> ]	$\lg(L_{\text{FUV}})$ [erg s <sup>-1</sup> ]	$\lg(L_{\text{H}\alpha})$ [erg s <sup>-1</sup> ]	$\lg(M_*)$ [ $M_{\odot}$ ]	$\lg(\text{Age})$ [yr]	$A_V$	$R$ [kpc]	Size [arcsec]	$F_{24}$ [mJy]
1	c3	...	...	23.374977	30.309910	39.34	38.11	39.00	37.15	3.0	6.6	0.1	5.4	15.3	23.54
2	c2	1	C	23.361427	30.290756	38.51	37.25	38.27	35.65	2.7	6.9	0.0	5.7	3.8	0.75
3	c3	...	...	23.372110	30.301689	39.31	38.78	38.87	36.84	3.0	6.7	0.3	5.5	7.1	4.57
4	c3	...	...	23.332920	30.305302	38.65	38.35	38.02	36.12	2.7	6.7	0.6	5.5	5.1	1.75
5	e	...	...	23.335817	30.304033	38.28	37.93	37.76	35.18	0.0	0.0	0.5	5.5	3.6	0.68
6	e	...	...	23.443571	30.303825	37.75	36.62	37.27	35.57	0.0	0.0	0.1	5.8	2.4	0.29
7	e	...	...	23.448744	30.304874	37.54	36.98	36.98	35.43	0.0	0.0	0.3	5.8	2.5	0.30
8	c2	15	C	23.402447	30.336786	40.38	40.04	39.74	37.94	4.2	6.7	0.6	5.0	18.2	79.54
9	c1	14	D	23.376879	30.319651	37.95	37.18	37.61	35.88	2.2	6.6	0.2	5.2	4.2	0.79
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

**Notes.** The full table is available at the CDS.

the GMCs and YSCCs, can be found in Tables 5 and 6. Since M 33 has a non-uniform star-forming disk with structures and large variations in the population densities across the disk, we examined three distinct radial ranges:  $R < 1.5$  kpc,  $1.5 < R < 4$  kpc, and  $R \geq 4$  kpc, and we referred to these as zone 1 (inner disk), zone 2 (spiral arm dominated), and zone 3 (outer disk). Below we summarize the main results discussed in the paper.

- The GMCs catalogue comprises 566 clouds with masses between  $2 \times 10^4$  and  $2 \times 10^6 M_{\odot}$  and radii between 10 and 100 pc. The number of GMCs above the survey completeness limit ( $L_{\text{CO}} \geq 5700 \text{ K km s}^{-1} \text{ pc}^2$  or  $M_{\text{H}_2} \geq 6.3 \times 10^4 M_{\odot}$ ) is 490. By examining the 8, 24  $\mu\text{m}$ , FUV, and  $\text{H}\alpha$  emission within each cloud, 545 of them have been classified as A-, B-, or C-type clouds. The remaining 21 clouds could not be assigned unambiguous classes. More than half of the catalogued GMCs have exposed star formation (C-type) with emission peaks at several wavelengths within the cloud contours. These are the most massive clouds. About 32% of the catalogued GMCs are inactive (A-type) with no sources at any wavelengths, and only 16% have embedded or low-mass star formation (B-type) with emission peaks at MIR wavelengths alone.
- The peak of the radial distribution of A-type clouds is near 4 kpc. Beyond 2 kpc, their number density is comparable to that of C-type clouds, which is the most numerous group. Both the A- and C-type clouds are found along HI filaments or spiral arms. On the southern arm the A-type clouds are found entering the arm, while C-type clouds are found on the arm, suggesting the arm environment may play a role in triggering star formation in the absence of other perturbations (such as M31 for the northern side). The average CO luminosity increases from A- to B- to C-type clouds, and this suggests that GMC mass may continue to grow as they evolve from the inactive phase to the formation of massive stars.
- We classified 611 of the 630 YSCCs into five categories, b-, c1-, c2-, c3-, and e-type, according to the presence and location of an optical,  $\text{H}\alpha$ , or UV counterpart within a GMC boundary. The majority of these sources lie (in projection) within a GMC boundary, especially in zones 1 and 2. The largest class of YSCC has coincident peaks at in the UV and  $\text{H}\alpha$  bands, which enabled estimating the age and mass of the associated YSC. There is an extraordinary spatial correspondence between the GMCs and the distribution of atomic hydrogen overdensities in zones 1 and 2. In zone 3 there are fewer GMCs, possibly because of a steepening of the molecular cloud mass spectrum, with a larger fraction of clouds being below the survey completeness limit.
- We find that GMCs classified as B- or C-type are associated with catalogued MIR sources that are classified as YSCCs of b- or c-type with only a few exceptions. The physical association between GMCs and YSCCs is established in the three zones considering the whole sample of GMCs and YSCCs, independently of their previous classification. We analyzed the association visually and statistically by generating the positional correlation function of the two distributions, and indeed the correlation is remarkable and stronger than between GMCs and other populations in the M 33 disk. If  $\bar{d}$  is the typical separation length between YSCCs, we expect to find only 20% of the GMCs with a YSCC at a distance shorter than  $0.5\bar{d}$  if they are randomly distributed. Instead, we find fractions on the order of 60–70%. The correlation length is on the order of 17 pc, but there

is a highly statistically significant clustering out to larger distances. There is little or no correlation between the mass of the GMCs and that of the YSCCs.

- The extinction estimates are higher for b-type sources with weak or no UV/optical counterpart, which likely represent the early phases of SF. The c1-type YSCCs have visible  $\text{H}\alpha$  but not FUV emission and show on average higher extinction than the c2-type YSCCs, where FUV emission is also detected. The c1-type YSCCs may represent YSCs at an earlier stage than c2-type YSCCs, even though the YSC age determination is not precise enough to separate these two classes. The most luminous YSCCs are of c2-type and are clusters that have very likely completed the formation process.
- The estimated ages of most YSCCs are between 3.5 and 8 Myr, with a marked peak around 5 Myr, and they are associated with C-type GMCs. Using the cluster ages and the fractions of GMCs in each class (i.e. in each evolutionary phase) we estimated the GMC lifetime in M 33 to be 14.2 Myr, from the time when most of the cloud mass is assembled to the time when the YSC breaks through the cloud and becomes optically visible. Even though the lifetime may be slightly longer when the cloud dispersal time is included, our estimate for the GMC lifetime in M 33 seems comparable to GMC lifetimes in the Milky Way and somewhat shorter than the estimated GMC lifetime in the LMC. The embedded phase, where MIR emission is visible but no  $\text{H}\alpha$  or FUV emission is detected, is the shortest phase from when the cloud is assembled and inactive to the switch-off of the stellar cluster formation process.

The analysis of the largest available sample of GMCs and YSCCs and their association across the whole star-forming disk of M 33 provides reliable estimates of GMC lifetimes and evolutionary timescales necessary for understanding the gas-star formation cycle across spiral galaxy disks.

*Acknowledgements.* We would like to thank the referee, Christine Wilson, for her useful comments that improved the original version of the manuscript.

## References

- Bash, F. N., Green, E., & Peters, III, W. L. 1977, *ApJ*, 217, 464  
 Blitz, L., & Shu, F. H. 1980, *ApJ*, 238, 148  
 Braine, J., Gratier, P., Kramer, C., et al. 2010, *A&A*, 520, A107  
 Calzetti, D. 2001, *PASP*, 113, 1449  
 Churchwell, E. 2002, *ARA&A*, 40, 27  
 Corbelli, E., Verley, S., Elmegreen, B. G., & Giovanardi, C. 2009, *A&A*, 495, 479  
 Corbelli, E., Giovanardi, C., Palla, F., & Verley, S. 2011, *A&A*, 528, A116  
 Corbelli, E., Thilker, D., Zibetti, S., Giovanardi, C., & Salucci, P. 2014, *A&A*, 572, A23  
 Druard, C. 2014, Ph.D. thesis, Université Bordeaux 1, France  
 Druard, C., Braine, J., Schuster, K. F., et al. 2014, *A&A*, 567, A118  
 Elmegreen, B. G. 2007, *ApJ*, 668, 1064  
 Engargiola, G., Plambeck, R. L., Rosolowsky, E., & Blitz, L. 2003, *ApJS*, 149, 343  
 Fan, Z., & de Grijs, R. 2014, *ApJS*, 211, 22  
 Freedman, W. L., Wilson, C. D., & Madore, B. F. 1991, *ApJ*, 372, 455  
 Fukui, Y., Mizuno, N., Yamaguchi, R., et al. 1999, *PASJ*, 51, 745  
 Fukui, Y., Mizuno, N., Yamaguchi, R., Mizuno, A., & Onishi, T. 2001, *PASJ*, 53, L41  
 Fukui, Y., Kawamura, A., Minamidani, T., et al. 2008, *ApJS*, 178, 56  
 Gieren, W., Górski, M., Pietrzyński, G., et al. 2013, *ApJ*, 773, 69  
 Gratier, P., Braine, J., Rodriguez-Fernandez, N. J., et al. 2010, *A&A*, 522, A3  
 Gratier, P., Braine, J., Rodriguez-Fernandez, N. J., et al. 2012, *A&A*, 542, A108  
 Gratier, P., Braine, J., Schuster, K., et al. 2017, *A&A*, 600, A27  
 Heyer, M. H., Corbelli, E., Schneider, S. E., & Young, J. S. 2004, *ApJ*, 602, 723  
 Kawamura, A., Mizuno, Y., Minamidani, T., et al. 2009, *ApJS*, 184, 1

- Larson, R. B. 1981, [MNRAS](#), **194**, 809
- Magrini, L., Stanghellini, L., Corbelli, E., Galli, D., & Villaver, E. 2010, [A&A](#), **512**, A63
- Massey, P., Olsen, K. A. G., Hodge, P. W., et al. 2006, [AJ](#), **131**, 2478
- Miura, R. E., Kohno, K., Tosaki, T., et al. 2012, [ApJ](#), **761**, 37
- Mizuno, N., Yamaguchi, R., Mizuno, A., et al. 2001, [PASJ](#), **53**, 971
- Murray, N. 2011, [ApJ](#), **729**, 133
- Roberts, W. W. 1969, [ApJ](#), **158**, 123
- Rosolowsky, E., & Leroy, A. 2006, [PASP](#), **118**, 590
- Rosolowsky, E., Engargiola, G., Plambeck, R., & Blitz, L. 2003, [ApJ](#), **599**, 258
- Rosolowsky, E., Keto, E., Matsushita, S., & Willner, S. P. 2007, [ApJ](#), **661**, 830
- Sharma, S., Corbelli, E., Giovanardi, C., Hunt, L. K., & Palla, F. 2011, [A&A](#), **534**, A96
- Solomon, P. M., Rivolo, A. R., Barrett, J., & Yahil, A. 1987, [ApJ](#), **319**, 730
- Verley, S., Hunt, L. K., Corbelli, E., & Giovanardi, C. 2007, [A&A](#), **476**, 1161
- Verley, S., Corbelli, E., Giovanardi, C., & Hunt, L. K. 2009, [A&A](#), **493**, 453
- Wilson, C. D., & Scoville, N. 1990, [ApJ](#), **363**, 435
- Yamaguchi, R., Mizuno, N., Mizuno, A., et al. 2001, [PASJ](#), **53**, 985
- Zinnecker, H., & Yorke, H. W. 2007, [ARA&A](#), **45**, 481