

Evolved Massive Stars at Low-metallicity I. A Source Catalog for the Small Magellanic Cloud

Ming Yang (杨明)¹, Alceste Z. Bonanos¹, Bi-Wei Jiang (姜碧为)², Jian Gao (高健)², Panagiotis Gavras³, Grigoris Maravelias¹, Yi Ren (任逸)², Shu Wang (王舒)⁵, Meng-Yao Xue (薛梦瑶)⁴, Frank Tramper¹, Zoi T. Spetsieri¹, and Ektoras Pouliaxis¹

¹ Institute for Astronomy, Astrophysics, Space Applications & Remote Sensing, National Observatory of Athens, Vas. Pavlou and I. Metaxa, Penteli 15236, Greece
e-mail: myang@noa.gr, bonanos@noa.gr

² Department of Astronomy, Beijing Normal University, Beijing 100875, P. R. China

³ Rhea Group for ESA/ESAC, Camino bajo del Castillo, s/n, Urbanizacion Villafranca del Castillo, Villanueva de la Canada, 28692 Madrid, Spain

⁴ International Centre for Radio Astronomy Research, Curtin University, Bentley, WA 6102, Australia

⁵ Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, P. R. China

April 6, 2019

ABSTRACT

We present a relatively clean, magnitude-limited ($IRAC1$ or $WISE1 \leq 15.0$ mag) multiwavelength source catalog for the Small Magellanic Cloud (SMC) with 45,466 targets in total, intending to build an anchor for the future studies, especially for the massive star populations at low-metallicity. It contains data in 50 different bands including 21 optical and 29 infrared bands, retrieved from SEIP, VMC, IRSF, AKARI, HERITAGE, Gaia, SkyMapper, NSC, Massey et al. (2002), and GALEX, ranging from the ultraviolet to the far-infrared. Additionally, radial velocities and spectral classifications are collected from the literature, as well as the infrared and optical variability information derived from WISE, SAGE-Var, VMC, IRSF, Gaia, NSC, and OGLE. The catalog is essentially built upon a 1'' crossmatching and a 3'' deblending between the Spitzer Enhanced Imaging Products (SEIP) source list and Gaia Data Release 2 (DR2) photometric data. Further constraints on the proper motions and parallaxes from Gaia DR2 allow us to remove the foreground contamination. We estimate that about 99.5% of the targets in our catalog are likely to be the genuine members of the SMC. By using the evolutionary tracks and synthetic photometry from MESA Isochrones & Stellar Tracks and the theoretical $J - K_s$ color cuts, we identify 1,405 red supergiant, 217 yellow supergiant and 1,369 blue supergiant candidates in the SMC in five different color-magnitude diagrams (CMDs). We rank the candidates based on the intersection of different CMDs. A comparison between the models and observational data shows that, the lower limit of RSGs population may reach to 7 or even $6M_{\odot}$, making RSGs an unique population connecting the evolved massive and intermediate stars, since stars with initial mass around 6 to $8M_{\odot}$ are thought to go through a second dredge-up to become asymptotic giant branch stars. We encourage the interested reader to further exploit the potential of our catalog.

Key words. Infrared: stars – Magellanic Clouds – Stars: late-type – Stars: massive – Stars: mass-loss – Stars: variables: general

1. Introduction

Massive stars are stars thought to be born with initial masses $\gtrsim 8 M_{\odot}$. They are relatively rare compared to the large number of low-mass stars due to the scaling of luminosity with initial mass. However, as a result of the intensive interior energy transportation and radiative output, they are responsible for some of the most extreme astrophysics in the Universe, including supernovae (SN), black holes, gravitational waves and long gamma-ray bursts, and critical for the stellar evolution, star formation and chemical evolution throughout the cosmic time (Humphreys & McElroy 1984; Woosley et al. 2002; Massey 2003; Meynet et al. 2011; Maeder & Meynet 2012; Massey 2013). The first generation of stars in the early Universe is expected to be massive and one of the main contributors of dust content in the high-redshift galaxies. Therefore, as a key role in the chemical evolution of the early metal-poor Universe, understanding of the physical properties, evolution and mass loss of massive stars may help to reveal the formation of the primi-

tive cosmic structures (Gall et al. 2011; Smith 2014; Zhang et al. 2018).

Unfortunately, such early Universe can not be directly observed up to now. The alternative way is to study the analogs of early Universe in our own cosmic backyard. A particularly growing interest of massive stars in the metal-poor environment has been risen in recent years, due to the rapid development of instrumental technology, and the needs to extrapolate the astrophysics from the local Universe to the early cosmic epochs. As a consequence, metal-poor star-forming dwarf irregular (dIrr) galaxies, or even the transition phase from dIrr to the dwarf spheroidal (dSph) galaxies (dIrr/dSph), in the local Universe may serve as an ideal laboratory for investigating the evolution and mass loss of massive stars in low-metallicity, since they mimic the behaviours of galaxies in the early Universe (Kunth & Östlin 2000; McConnachie 2012).

To understand the evolution of massive stars in the low-metallicity environment, both theoretical and observational constraints are needed. Among many physical parameters of mas-

sive stars, one deterministic parameter is the mass loss, which has profound impact on a star's lifetime, luminosity (L), effective temperature (T_{eff}), radiation field, and the end fate as SN. However, the most important modes of mass loss are the most uncertain (Smith 2014). It has been known that the mass-loss rates (MLR) adopted in modern stellar evolution codes for standard metallicity-dependent winds of hot main-sequence stars are overestimated by a factor of 2–3, due to the clumped and inhomogeneous stellar wind (Puls et al. 2008). Thus, the removal of hydrogen envelope has to rely on the post-main sequence phase, that is to say, the stellar winds, pulsations, rotation, convections and eruptions of evolved supergiants, as well as binary mass transfer. Meanwhile, it has also been recognized that the unsteady modes of mass loss, like the episodic mass-loss events, are more important than previously thought (Smith & Owocki 2006; Ofek et al. 2013). Consequently, observation in the infrared wavelengths, especially the relatively longer wavelengths which are mainly dominated by the dust emission, has great impact on the understanding of MLR of massive stars.

Among the Local Group galaxies, the Large and Small Magellanic Cloud (LMC and SMC) are particularly intriguing due to their low-metallicity environments (about half and one-fifth of the Milky Way; Russell & Dopita 1992; Rolleston et al. 2002; Keller & Wood 2006; Dobbie et al. 2014; D'Onghia & Fox 2016) and close distances, for which individual star can be resolved and allow detailed analysis of massive star populations in variety of ways (Barba et al. 1995; Massey & Olsen 2003; Evans & Howarth 2008; Bonanos et al. 2009, 2010; Neugent et al. 2010; Yang & Jiang 2011, 2012; Bouret et al. 2013; Kourniotis et al. 2014; Hainich et al. 2015; Castro et al. 2018; Yang et al. 2018). In this paper, we focus on the evolved dusty massive star populations in the SMC, based on the infrared detection related to the MLR, astrometric solution regarding to the membership of the SMC, time-series data concerns about stellar variability, and evolutionary models in relation to the evolutionary stages, aiming to build a comprehensive anchor for the future studies.

The paper is structured as follows: the multiwavelength source catalog and time-series data are presented in §2 and §3, respectively. The identification of evolved massive star candidates is described in §4. The summary is given in §5.

2. Multiwavelength Source Catalog

Since our goal is to focus on the evolved dusty massive stars in the SMC, the Spitzer Enhanced Imaging Products (SEIP) source list is a good starting point. SEIP source list contains sources detected with a high signal-to-noise ratio (S/N; 10-sigma level) in at least one channel among 12 near-infrared (NIR) to mid-infrared (MIR) bands of J (1.25 μm), H (1.65 μm), K_S (2.17 μm), IRAC1 (3.6 μm), IRAC2 (4.5 μm), IRAC3 (5.8 μm), IRAC4 (8.0 μm), MIPS24 (24 μm), WISE1 (3.4 μm), WISE2 (4.6 μm), WISE3 (12 μm) and WISE4 (22 μm), from Two Micron All Sky Survey (2MASS; Skrutskie et al. 2006), Spitzer (Werner et al. 2004) and Wide-field Infrared Survey Explorer (WISE; Wright et al. 2010)¹. We retrieve the initial infrared data from SEIP source list with $3^\circ \leq R.A. \leq 25^\circ$, $-75.5^\circ \leq Decl. \leq -70^\circ$ and $IRAC1$ or $WISE1 \leq 15.0$ mag, which covers almost the whole SMC and also a small part of the Magellanic Bridge (MB). The magnitude cut of $IRAC1$ or $WISE1 \leq 15.0$ mag is justified based on a drop-off (~ 14.85 mag) in the number

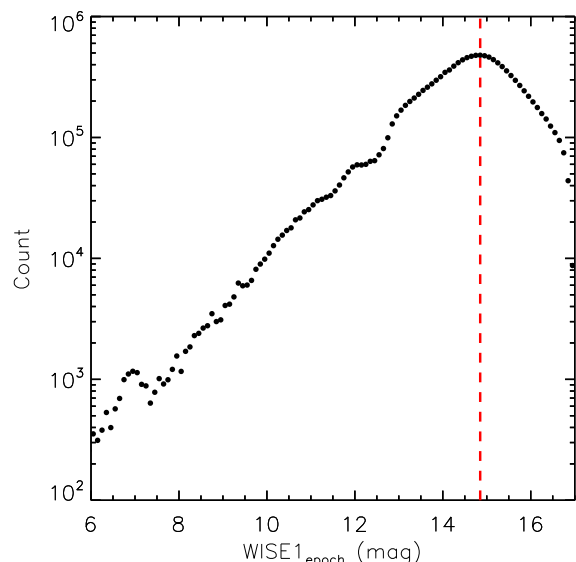


Fig. 1. Histogram of ALLWISE WISE1 single-epoch measurements in the SMC region, where a drop-off around 14.85 mag is shown by the red dashed line.

counts for 12,748,156 ALLWISE WISE1 single-epoch measurements in the same region as shown in Figure 1. Considering the lower angular resolution of WISE ($\sim 6''$) and 2MASS ($\sim 5''$) compared to Spitzer ($\sim 2''$), and WISE sources within $3''$ of a SEIP source list source are reported, a self-crossmatch deblending is performed with a search radius of $3''$, where targets having neighbors within $3''$ are excluded and results in 131,233 targets. Other than the infrared detection, the reliable membership of the SMC is also a crucial factor for our study, for which the astrometric solution from Gaia Data Release 2 (DR2) is vital (Gaia Collaboration et al. 2016, 2018). Given the very small offset between Gaia DR2 and 2MASS (median value of $\sim 0.120'' \pm 0.157''$ for targets with $G \leq 18$ mag in the same SMC region), as well as SEIP source list and 2MASS (median value of $\sim 0.089'' \pm 0.237''$), we crossmatch the self-cleaned SEIP data and Gaia DR2 with a search radius of $1''$ to fix the position, and remove any SEIP targets with multiple counterparts to eliminate the blending. Then, a $3''$ crossmatching is performed and SEIP targets with multiple counterparts are removed again, which results in 74,237 targets. Since the effective angular resolution of the Gaia DR2 source list has improved to $\sim 0.4''$, with incompleteness in close pairs of stars starting below about $2''$ (Gaia Collaboration et al. 2018; Arenou et al. 2018), we may reasonably believe that the vast majority of the blending is removed at the resolutions of both Spitzer and Gaia.

After crossmatching and deblending between SEIP and Gaia data, Gaia DR2 astrometric solution is used to eliminate the foreground contamination. The selection of SMC members is restricted to targets with errors less than 0.5 in proper motions (PMs) and parallax. The first two panels of Figure 2 show the errors versus Gaia PMs in R.A. (left) and Decl. (middle), respectively. We fit a Gaussian profile to PM at each dimension as 0.695 (peak) ± 0.240 (σ) mas/yr in R.A. and -1.206 ± 0.140 mas/yr in Decl., and calculate the limits of $\pm 5\sigma$ shown as the vertical dashed lines. The last panel (right) of Figure 2 shows the errors versus Gaia parallaxes. Similarly, a Gaussian profile fitting is adopted for the parallax as -0.009 ± 0.066 mas, while an additional elliptical constraint is also applied with the 5σ limits of $PM_{R.A.}$ and $PM_{Decl.}$ taken as the primary and secondary radii, respectively. The same criteria of $\pm 5\sigma$ is calculated for the

¹http://irsa.ipac.caltech.edu/data/SPITZER/Enhanced/SEIP/docs/seip_explanatory_supplement_v3.pdf

parallax. In brief, the membership of SMC is constrained by the Gaussian profile in parallax with additional elliptical constraint derived from PM_{RA} and $PM_{Decl.}$, which results in 45,466 targets. For targets having Gaia radial velocity (RV) measurements, we also limit the RV to be larger than ~ 90 km/s as shown below. Figure 3 shows PM_{RA} versus $PM_{Decl.}$, for which the separation of selected SMC members, NGC104 and NGC362 is clearly shown. Meanwhile, based on the median number density of this image and the association between SEIP source list and 2MASS point source catalog (SEIP sources without valid 2MASS measurements), we estimate the contamination of remaining foreground sources and the possible non-point/background sources (background point sources like quasars or blue compact dwarfs cannot be rejected at this stage, and the non-detection in the 2MASS may also due to the saturation or faintness of the targets) for the SMC, which are around 0.2% ($\sim 98/45,466$) and 0.3% ($\sim 131/45,466$), respectively, and can be ignored. This indicates that about 99.5% of the targets in our source catalog likely to be the genuine members of the SMC. We notice that the strict constrain on the astrometric solution and the previous deblending procedure may cause target loss and incompleteness in our catalog to a certain extent, but ensure that we select the true SMC targets. This also can be seen from the histogram of Gaia RV in Figure 4, where the separation of Milky Way and SMC is clear, and the vast majority of targets with Gaia RV larger than ~ 90 km/s are selected with minimal value of ~ 95 km/s (targets with both large PMs and RVs are not necessarily the members of SMC, but also could be the hypervelocity stars in the Milky Way, runaway stars from the Milky Way or the SMC, or the free-floating stars between the Milky Way and the SMC). Figure 5 illustrate the Gaia color-magnitude diagram (CMD) before (gray) and after (red) the astrometric constraining, where the large amount of foreground contamination of bright yellow and faint red stars is swept out.

Here we also emphasize that instead of building a complete sample, we mainly focus on the reliable detection of infrared sources in the SMC. In addition, further constrain on the SEIP data, for example, $S/N \geq 3$ for IRAC1 ($i1_fluxtype = 1$), is investigated and only results in about 1.66% difference in the total number of sources (45,466 versus 44,712 targets) and can be ignored.

Based on this fiducial dataset of 45,466 targets, we retrieve additional optical and infrared data from following datasets with a search radius of $1''$:

- 18,641 matches (41.00%) from VISTA survey of the Magellanic Clouds system (VMC) DR4. VMC is a NIR YJK_S bands multi-epoch survey for the LMC, the SMC and the MB ($K_S \leq 20.3$ mag, $\sim 2\%$ photometric and $\sim 0.01''$ astrometric precision), by using 4-meter near-infrared optimized Visible and Infrared Survey Telescope for Astronomy (VISTA; Cioni et al. 2011).
- 28,678 matches (63.08%) from the IRSF Magellanic Clouds point source (MCPS) catalog. IRSF MCPS is a NIR JHK_S bands photometric catalog for a 40 deg² area in the LMC, an 11 deg² area in the SMC, and a 4 deg² area in the MB ($K_S \leq 16.6$ mag, photometric and astrometric accuracies for bright sources are 0.03-0.04 mag and $0.1''$, respectively), based on the data from Simultaneous three-color InfraRed Imager for Unbiased Survey (SIRIUS) camera on the InfraRed Survey Facility (IRSF) 1.4-meter telescope (Kato et al. 2007).
- 625 matches (1.37%) from AKARI SMC bright point source list. The source list represents NIR to MIR imaging and spectroscopic observations of patchy areas in the SMC ($N3 \leq$

16.5 mag, ~ 0.1 mag photometric and $\leq 0.8''$ astrometric precision), by using the Infrared Camera (IRC) aboard AKARI space telescope (Onaka et al. 2007; Murakami et al. 2007; Ita et al. 2010).

- 4 matches from HERschel Inventory of the Agents of Galaxy Evolution (HERITAGE) band-merged source catalog (units are in flux [mJy] instead of magnitude). HERITAGE catalog is derived based on the data from both Photodetector Array Camera and Spectrometer (PACS; 100 and $160\mu\text{m}$) and Spectral and Photometric Imaging Receiver (SPIRE; 250, 350, and $500\mu\text{m}$) cameras on board the Herschel Space Observatory, to identify dusty objects in the LMC and SMC (the catalog also includes the Spitzer MIPS $70\mu\text{m}$ band data, and non of our targets has been detected in the SPIRE $500\mu\text{m}$ band; Pilbratt et al. 2010; Meixner et al. 2013; Seale et al. 2014).
- 40,387 matches (88.83%) from SkyMapper DR1.1 (we apply constrain parameters of $flags < 4$, $nch_max = 1$, $nimaflags = 0$ and $class_star \geq 0.9$ to retrieve the data). SkyMapper is a six-bands southern hemisphere photometric survey (u, v, g, r, i, z ; from magnitude 8 to 18, $\sim 1\%$ photometric and $< 0.2''$ astrometric precision), by using the dedicated 1.3-meter SkyMapper telescope (Keller et al. 2007; Bessell et al. 2011; Wolf et al. 2018).
- 38,759 matches (85.25%) from NOAO source catalog (NSC) DR1 (we apply constrain parameters of $flags < 4$ and $class_star \geq 0.9$ to retrieve the data). NSC is a catalog of sources from most of the public data taken on NOAO's CTIO-4m+DECam as well as KPNO-4m+Mosaic3 (in u, g, r, i, z, Y bands; reach to ~ 23 mag in most broadband filters with ~ 1 -2% photometric precision, and astrometric accuracy of ~ 7 mas; Nidever et al. 2018).
- 11,630 matches (25.58%) from a UBV r CCD survey of the Magellanic Clouds (M2002; note that due to the existence of some extreme large errors in the catalog, we replace them, e.g., $error > 1.0$ mag, with NULL values). M2002 is a survey of 14.5 deg² region in the LMC and 7.2 deg² region in the SMC ($V \leq 18.0$ mag, $< 3\%$ photometric and $\sim 0.3''$ astrometric precision), by using 0.61-meter Curtis Schmidt telescope at CTIO (Massey 2002).
- 164 matches (0.36%) from revised GALEX source catalog for the All-Sky Imaging Survey (GUVcat_AIS). GUVcat_AIS is a science-enhanced, "clean" catalogs of GALEX ultraviolet (UV) sources with typical depth of 20.8 and 19.9 mag, position accuracy of $0.32''$ and $0.34''$ in NUV and FUV bands, respectively (Morrissey et al. 2007; Bianchi et al. 2017).

All the datasets are also self-cleaned with a search radius of $3''$, prior to crossmatching with the SEIP-Gaia dataset. Figure 6 shows the normalized transmission curves of all filters used in our study. In total, we have 50 filters including 21 optical (includes the two UV filters) and 29 infrared filters. The spatial distributions of the additional optical (left) and infrared (right) datasets are shown in Figure 7. For convenience, the GALEX and HERITAGE data are not shown in the diagram due to a paucity of matches.

Except the photometric data, additional classifications are also retrieved from literature with a search radius of $1''$, including:

- 37,375 matches (IR color classifications) from Boyer et al. (2011). It is an investigation of the IR properties of cool, evolved stars in the SMC, including the red giant branch (RGB) stars and the dust-producing red supergiant (RSG) and asymptotic giant branch (AGB) stars using observations

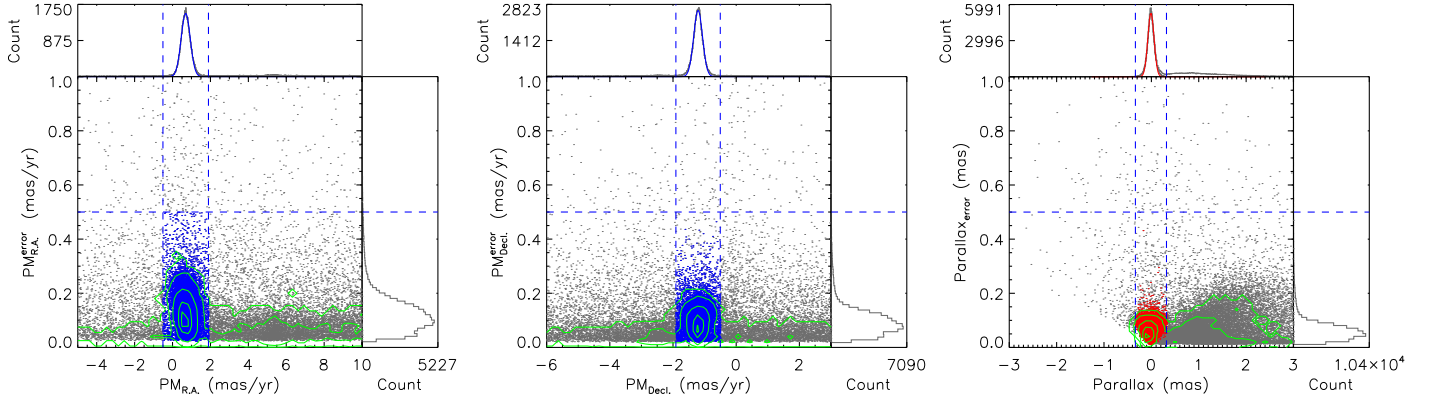


Fig. 2. Evaluation of Gaia astrometric solution. The first two panels show errors versus Gaia PMs in R.A. (left) and Decl. (middle), respectively. A Gaussian profile is fitted to PM in each dimension and the limits of $\pm 5\sigma$ is calculated (vertical dashed lines). The last panel (right) shows the errors versus Gaia parallaxes. A Gaussian fitting is adopted again for the parallax, while an additional elliptical constraint is also applied with the 5σ limits of PM_{RA} and PM_{Decl} . taken as the primary and secondary radii, respectively. The same criteria of $\pm 5\sigma$ is calculated for the parallax (vertical dashed lines).

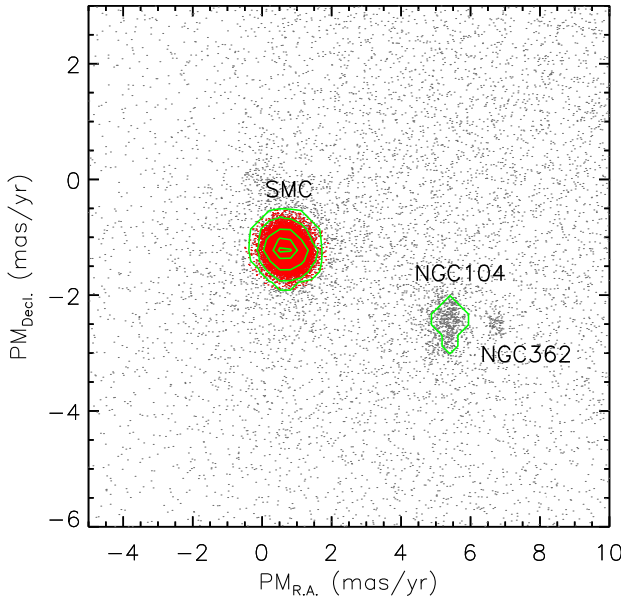


Fig. 3. PM_{RA} versus PM_{Decl} diagram, in which the separation of selected SMC members (red), NGC104 and NGC362 is clearly shown. Based on this diagram, we estimate the contamination of remaining foreground sources for the SMC is around 0.2% ($\sim 98/45,466$) and can be ignored.

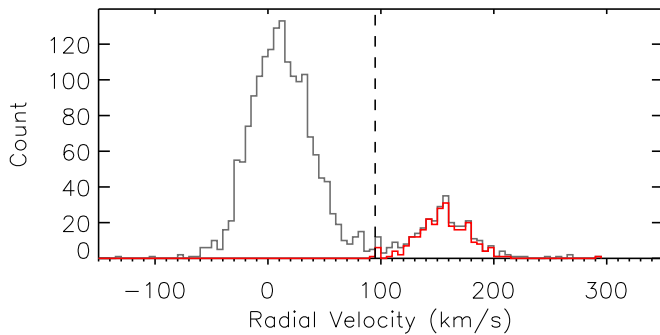


Fig. 4. Histogram of RVs from Gaia. The separation of Milky Way and SMC is clear, and the vast majority of targets with RV larger than ~ 90 km/s are selected (red) with minimal value of ~ 95 km/s (dashed line).

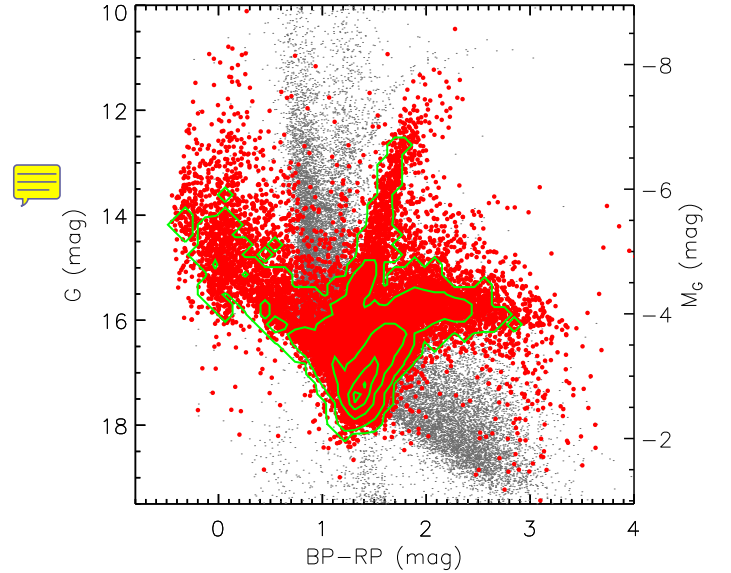


Fig. 5. G versus $BP-RP$ diagram for the Gaia data before (gray) and after (red) the astrometric constraint, where the large number of foreground contamination is swept out.

from the Spitzer Space Telescope Legacy Program entitled “Surveying the Agents of Galaxy Evolution in the Tidally Stripped, Low Metallicity SMC (SAGE-SMC)”.

- 30 matches (IR color and SED classifications) from [Sewilo et al. \(2013\)](#). They use CMDs based on the multi-wavelength photometric data and the SED fitting to identify a population of $\sim 1,000$ intermediate- to high-mass young stellar objects (YSOs) in the SMC.
- 43 matches (MIR spectral classifications) from [Ruffle et al. \(2015\)](#). They have classified 209 point sources observed by Spitzer Infrared Spectrograph (IRS; [Houck et al. 2004](#)) using a decision tree method, based on infrared spectral features, continuum and spectral energy distribution shape, bolometric luminosity, cluster membership and variability information (all the targets from [Kraemer et al. 2017](#) are also included).
- 666 matches (optical spectral classifications) from [Bonanos et al. \(2010\)](#). It is a catalog of 3,654 massive

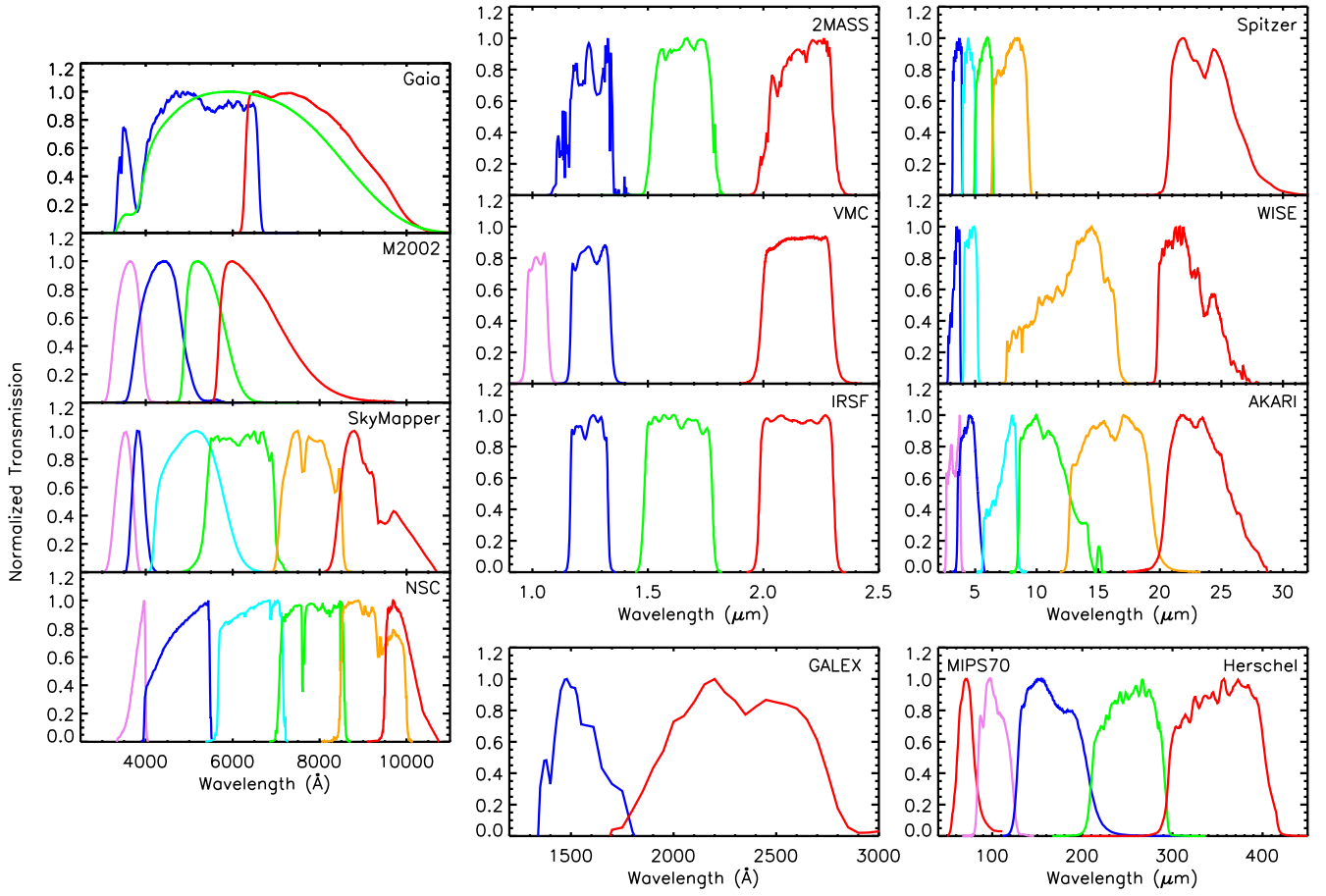


Fig. 6. Normalized transmission curves of filters used in our study. In total, we have 50 filters including 21 optical (includes the two UV filters) and 29 infrared filters. For each dataset, the filters are color coded from shorter (blueish) to longer (reddish) wavelengths. Left panel from top to bottom: BP, RP, G from Gaia, U, B, V, R from M2002, u, v, g, r, i, z from SkyMapper, and u, g, r, i, z, Y from NSC. Middle panel from top to bottom: J, H, K_S from 2MASS, Y, J, K_S from VMC, J, H, K_S from IRSF, and FUV, NUV from GALEX. Right panel from top to bottom: IRAC1, IRAC2, IRAC3, IRAC4, MIPS24 from Spitzer, WISE1, WISE2, WISE3, WISE4 from WISE, N3, N4, S7, S11, L15, L24 from AKARI, and MIPS70 (from Spitzer), PACS100, PACS160, SPIRE250, SPIRE350 from Herschel.

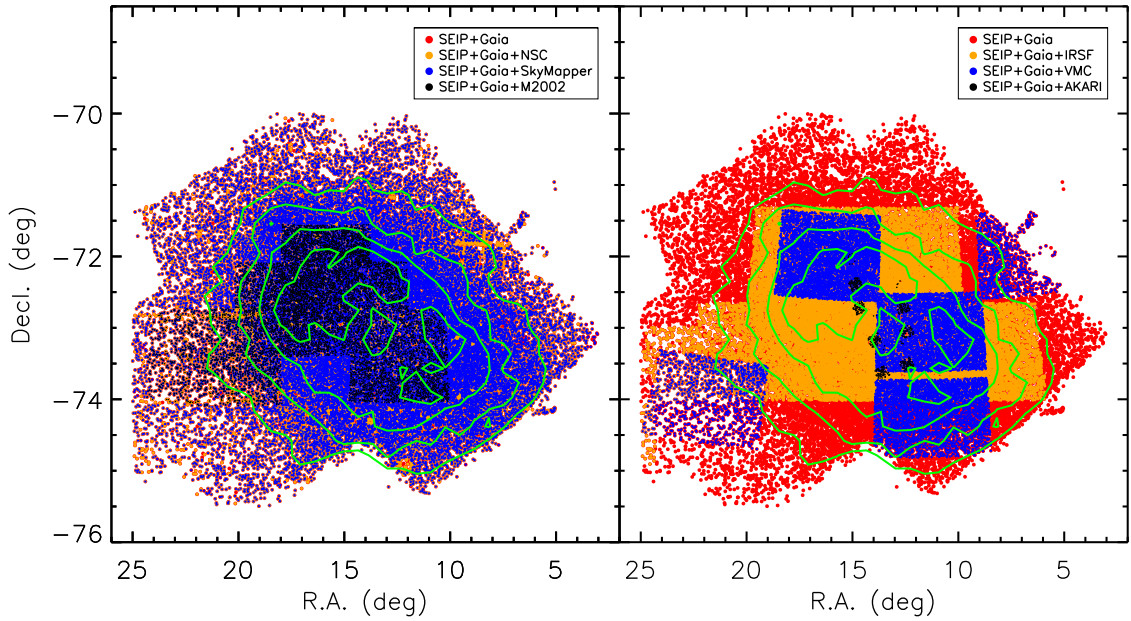


Fig. 7. Spatial distribution of the additional optical (left) and infrared (right) datasets. For convenience, the GALEX and HERITAGE data are not shown in the diagram due to a paucity of matches.

stars from the literature with accurate spectral types and multiwavelength photometry in the SMC, intending to study their infrared properties.

- 198 matches (spectral variability flag, radial velocities and optical spectral classifications; for convenience, we only keep the first spectral classification for targets with multiple measurements) from [González-Fernández et al. \(2015\)](#). They have studied physical properties of about 500 RSGs in the LMC and SMC by using NIR/MIR photometry and optical spectroscopy, aiming at exploring the fainter end of RSGs and extrapolating their behavior to other environments by building a more representative sample.
- 113 matches (radial velocities and optical spectral classifications) from [Neugent et al. \(2010\)](#). They spectroscopically observe 176 near-certain (Category 1) SMC yellow supergiant stars (YSGs) among ~500 candidates to test against the evolutionary model in the low-metallicity environment.
- 39,295 matches from Simbad ([Wenger et al. 2000](#)). The radial velocities, optical spectral classifications, main object types, and auxiliary object types are retrieved.

The unmatched targets are likely due to larger PMs, blendings, or the quality cuts in SEIP or Gaia catalog.

This multiwavelength source catalog with 45,466 targets serves as the backbone of our study. The sample is consist of “bona-fide” and “dusty” SMC targets determined by both astrometric measurement and infrared detection. Table 1 shows the absolute and relative (related to the most detected filter in each dataset) percentage of detected targets in each filter. Figure 8 shows the histograms of magnitude distribution for each dataset (for convenience, the HERITAGE data is not shown here). The bin size is 0.1 mag, except for the GALEX (0.5 mag) and AKARI (0.25 mag) data. For WISE3 and WISE4 bands, due to the fact that the majority of the targets (~75% in WISE3 and ~95% in WISE4 bands) has low S/N (< 2) and are derived with a 95% confidence brightness upper limit, the histograms only show targets with $S/N \geq 2$.

3. Multiwavelength Time-Series Data

Following [Yang et al. \(2018\)](#), the MIR time-series data of WISE1 (3.4 μm) and WISE2 (4.6 μm) bands for all the 45,466 targets are collected from both ALLWISE ([Cutri & et al. 2013](#)) and Near-Earth Object WISE Reactivation mission (NEOWISE-R; [Mainzer et al. 2014](#)), with a search radius of 1” and several parameters constrained as $qi_fact > 0$, $saa_sep > 0$, $moon_masked = 0$, $qual_frame > 0$, $det_bit = 3$, $\chi^2 \leq 10$, $S/N \geq 3$, which results in about 2.28 million measurements from ALLWISE and 6.99 million measurements from NEOWISE-R. As the NEOWISE 2018 data released, the total frame coverage is about ten major epochs spanning ~2800 days (~7.7 years) with two epochs from ALLWISE and eight epochs from NEOWISE-R separated by an approximately three-year gap. The beginning and ending of each epoch set by us are given in Table 2. We have binned the data within each epoch by using the median values of the date and magnitudes. For each epoch, we require at least five valid points to calculate the median value. Figure 9 shows examples of the original lightcurves overlapped with the binned lightcurves. The median absolute deviation (MAD) and standard deviation (SD) are used to calculate the long-term (full lightcurve) and short-term (within single epoch) variability of each target with at least five valid points, where the former is more resistant to outliers than the latter ([Rousseeuw et al. 2009](#)). Although it is possible to derive periods based on the current data, the period search may be highly contaminated by the strong

alias structures due to the very low sampling of WISE data as shown in Figure 9, that we only have ten epochs spanning ~2800 days and each epoch only covers about 5 to 10 days (we refer interested readers to [Chen et al. 2018](#) for the WISE catalog of periodic variable stars). In addition to MAD and SD, the full amplitude ($Amp = max_{mag} - min_{mag}$) is also calculated for each target. More details about the WISE time-series data reduction can be found in [Yang et al. \(2018\)](#). In total, there are 39,495 (86.87%) targets having variability information in both WISE1 and WISE2 bands and covering almost the whole area of our sample. The lack of variability information for some targets is likely due to either faint magnitudes or the quality cuts we adopted.

Except the WISE time-series data, we also collect other sets of time-series data from different projects, including IR data from Spitzer Surveying the Agents of Galaxy Evolution (SAGE)-Var program ([Riebel et al. 2015](#)), VMC, IRSF ([Ita et al. 2018](#)) and optical time-series data from Gaia, NSC and OGLE. SAGE-Var is a follow-up to the Spitzer legacy program SAGE ([Meixner et al. 2006](#); [Gordon et al. 2011](#)), for which six total epochs of photometric observations at IRAC1 (3.6 μm) and IRAC2 (4.5 μm) bands are obtained covering the bar of the LMC and the central region of the SMC with 15 different timescales ranging from ~20 days to ~5 yr. We have collected the SMC data and calculated the median magnitudes, MADs, SDs and Amps for targets with all six epochs available. In total, there are 7,160 (15.75%) targets in IRAC1 band and 5,894 (12.96%) targets in IRAC2 band matched with our source catalog within 1”. The left panel of Figure 10 shows the spatial distribution of those targets overlapped on our source catalog.

Since the final release (may also include the time-series data) with global photometric and astrometric calibration will be made upon completion of the survey, for the VMC data, we now only rely on the variability statistics provided by the VMC DR4². As the observational cadences for different targets and filters are irregular with the median values of cadences vary from few hours to hundreds of days, for each filter, we constrain all the values of median magnitudes, MADs, SDs, and Amps to be within the range of 0 to 99 with at least five good measurements ($*nGoodObs \geq 5$) to avoid any unphysical values. In total, there are 11,197 (24.63%) targets matched with our source catalog within 1” as shown in the middle panel of Figure 10. There are 11,197 (24.63%) targets in Y band, 11,125 (24.47%) targets in J band, and 7,574 (16.66%) targets in K_S band.

For IRSF data, we retrieve the time-series data of ~1,000 targets for each filter from [Ita et al. \(2018\)](#), where a very long-term (15 years up to now) near-infrared variable star survey towards an area of 3 deg^2 along the bar in the LMC and an area of 1 deg^2 in the central part of the SMC have been carried out with more than one hundred times repeated observations for each area. The median magnitudes, MADs, SDs, Amps are calculated for each targets. There are 160 targets in J band, 161 targets in H band, and 154 targets in K_S band matched with our source catalog within 1” as shown in the right panel of Figure 10.

Gaia DR2 provides classifications for more than 550,000 variable sources consisting of different types of variables. However, only a subset of the variable stars classified as a certain type are characterized in detail and a fraction of the classifications may well be wrong ([Gaia Collaboration et al. 2019](#); [Mowlavi et al. 2018](#)). Since the time-series data will be released in Gaia DR3, we retrieve the variability statistics (including

²http://horus.roe.ac.uk/vsa/www/VMCDR4/VMCDR4_TABLE_vmcVariabilitySchema.html#vmcVariability

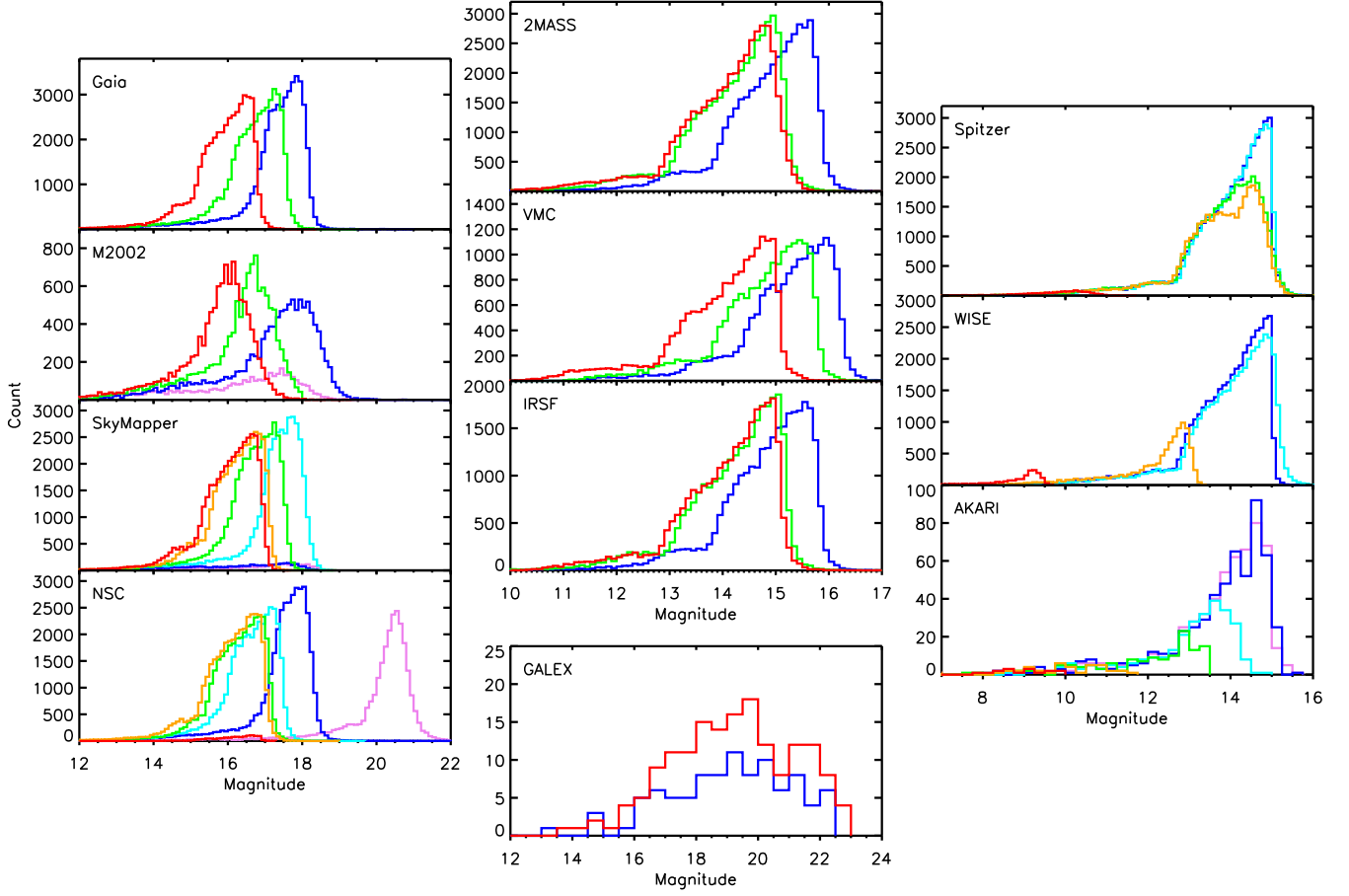


Fig. 8. Histograms of magnitude distribution in each dataset. The color convention of filters is the same as Figure 6 (same below). The bin size of magnitude is 0.1 mag, except for the GALEX (0.5 mag) and AKARI (0.25 mag) data. For WISE3 and WISE4 bands, the histograms only show targets with $S/N \geq 2$. For convenience, the HERITAGE data are not shown in the diagram.

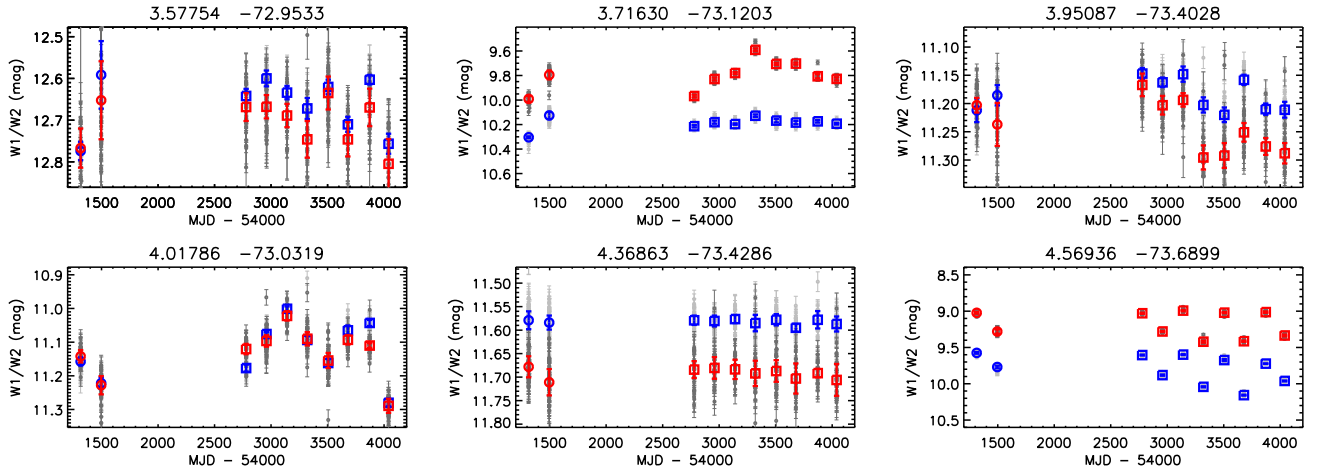


Fig. 9. Examples of original (light gray for WISE1 and dark gray for WISE2) and binned (blue for WISE1 and red for WISE2) lightcurves. There are about ten major epochs with two epochs from ALLWISE (open circles) and eight epochs from NEOWISE-R (open squares) separated by an approximately three-year gap. Coordinates of the targets are indicated on top of each panel.

MADs, SDs and Amps) and classifications from Gaia Archive³ with typically ~ 30 measurements spanning ~ 620 days. However, we notice that the MAD is scaled by 1.4826, so that the expectation of the scaled MAD at large number of measurements equals

the standard deviation of a normal distribution⁴. For consistency of the dataset, we have reversed the scaled MAD by dividing it with 1.4826. There are 1,379 (3.03%; blue) targets with variability information⁴ are matched with our source catalog within

³<https://gea.esac.esa.int/archive/>

⁴https://gea.esac.esa.int/archive/documentation/GDR1/datamodel/Ch2/phot_variable_time_series_gfov_statistical_parameters.html

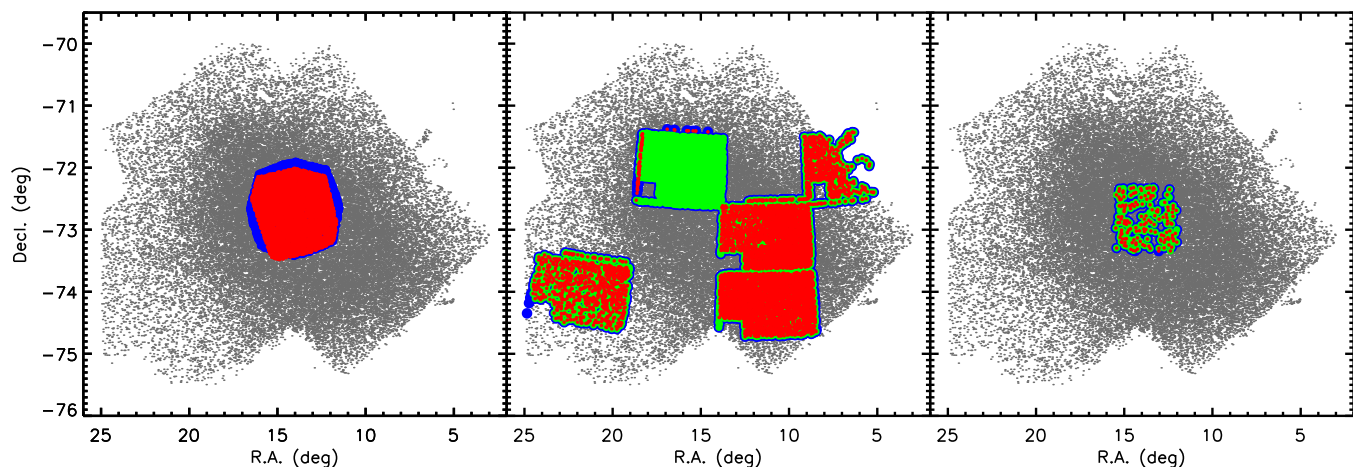


Fig. 10. Spatial distribution of targets with IR variability information matched with our source catalog within 1''. Left: targets from SAGE-Var project, where blue and red colors indicate the targets in IRAC1 and IRAC2 bands, respectively. Middle: targets from VMC DR4, where blue, green and red colors indicate the targets in Y, J, and K_s bands, respectively. Right: targets from IRSF survey, where blue, green and red colors indicate the targets in J, H, and K_s bands, respectively.

1'' as shown in the left panel of Figure 11, while 1,277 (2.81%; orange) of them are classified, and 868 (1.91%; red) of them having $best_class_score \geq 0.5$, for which the classification may be acceptable.

The NSC time-series data of the SMC are mainly from the Survey of the Magellanic Stellar History (SMASH; Nidever et al. 2017). We retrieve the data from NOAO Data Lab⁵ with the same sky coverage and constrain parameters ($flags < 4$ and $class_star \geq 0.9$) and crossmatch with our source catalog by using a search radius of 1'', which results in 132,342 measurements in u-band, 355,823 measurements in g-band, 497,443 measurements in r-band, 210,496 measurements in i-band, 404,087 measurements in z-band, and 16,916 measurements in Y-band. However, due to the irregular sampling in different filters, the observational timescales vary from tenths of a day to more than one thousand days with several to hundreds measurements. Thus, we require at least five valid measurements for each target in individual filter in order to calculate the MAD, SD and Amp. Here we need to emphasize that as a result of the irregular sampling, and particularly, a strong systematic effect we discovered during the data processing (instead of relatively uniform distribution, different SMC regions show variety levels of average variabilities), the calculated variability may not fully represent the true variability of the target and user should be cautious when using these values. Middle panel of Figure 11 shows the spatial distribution of targets with variability information in different filters, including 17,548 (38.60%) targets in u-band, 25,897 (57.00%) targets in g-band, 29,572 (65.04%) targets in r-band, 19,632 (43.18%) targets in i-band, 35,850 (78.85%) targets in z-band, and 1,212 (2.67%) targets in Y-band.

Finally, we obtain additional data from the Optical Gravitational Lensing Experiment (OGLE; Udalski et al. 1992; Szymanski 2005; Udalski et al. 2008, 2015) by using a search radius of 1'' from both OGLE-III Catalog of Variable Stars (O³CVS) and OGLE-IV Collection of Variable Stars (O⁴CVS), which results in 8,956 (19.70%) Long Period Variables (LPVs) from O³CVS, and 482 (1.06%) Classical Cepheids (CCeps), 11 (0.02%) Type II Cepheids (T2Ceps), and 87 (0.19%) Eclipsing Binaries (EBs) from O⁴CVS (Soszyński et al. 2011, 2015;

Pawlak et al. 2016; Soszyński et al. 2018) as shown in the right panel of Figure 11. However, since the OGLE data are calculated by using Fourier analysis, which is different from our dataset, only the classifications are used in the further analysis. Figure 12 shows Gaia color-magnitude diagrams with variable classifications from Gaia (left; only shows targets with $best_class_score \geq 0.5$) and OGLE (right).

We have also checked the time-series data from SkyMapper DR1.1. However, as the first data release, there are only few measurements (~1-2) covering ~300 days, which is not suitable for robust variability calculation.

All the information about the 45,466 targets are listed in Table 3. Targets without errors indicate either a 95% confidence upper limit, or the errors are simply too large to be reliable (e.g., $> 1.0\ mag$).

4. Identify Evolved Massive Star Candidates on the Color-Magnitude Diagrams

As we focus on the evolved dusty massive stars, the primary task is to identify them. Here we utilize the evolutionary tracks and synthetic photometry from Modules for Experiments in Stellar Astrophysics (MESA; Paxton et al. 2011, 2013, 2015, 2018) Isochrones & Stellar Tracks (MIST⁶; Choi et al. 2016; Dotter 2016), which covers a wide range of ages, masses and metallicities by using solar-scaled abundance under a single computational framework, to identify evolved massive star candidates on the CMDs of our multiwavelength source catalog.

We use a canonical value of 18.95 as the distance modulus of the SMC (Graczyk et al. 2014; Scowcroft et al. 2016). Since the metallicity of SMC is about 10% to 20% solar (Russell & Dopita 1992; Dobbie et al. 2014; D'Onghia & Fox 2016), we adopt the chemical composition of -1.0 to -0.7 dex for [Fe/H]. The nonrotation and rotation ($V/V_{crit} = 0.40$) models of 7 to 40 M_{\odot} are computed with no extinction and extinction of $A_V = 1.0\ mag$, respectively (Cardelli et al. 1989; Zaritsky et al. 2002; Haschke et al. 2011; Gao et al. 2013). We choose the color-magnitude combinations based on the available synthetic photometry in MIST with relative percentage $> 90\%$

⁵<https://datalab.noao.edu/>

⁶<http://waps.cfa.harvard.edu/MIST/>

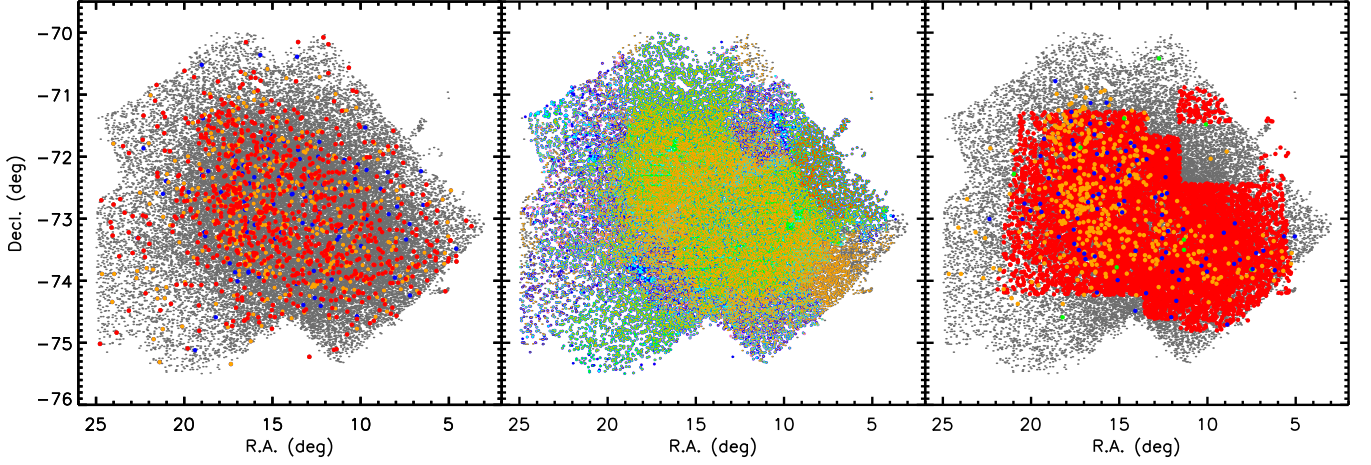


Fig. 11. Spatial distribution of targets with optical variability information matched with our source catalog within $1''$. Left: targets from Gaia, where blue, orange and red colors indicate targets without classification, with *best_class_score* < 0.5, and *best_class_score* \geq 0.5, respectively. Middle: targets from NSC, where violet, blue, cyan, green, orange and red colors indicate the targets in u, g, r, i, z, and Y bands, respectively. Right: targets from OGLE, where blue, green, orange, and red colors indicate the targets classified as EBs, T2Ceps, CCeps, and LPVs, respectively.

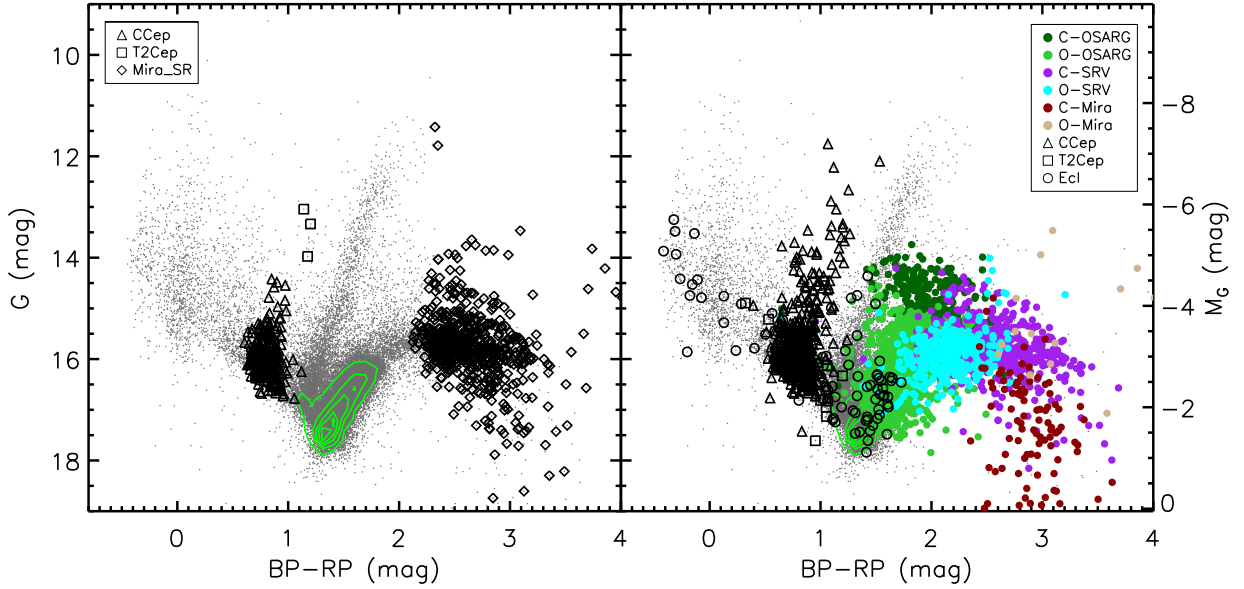


Fig. 12. Gaia color-magnitude diagrams with variable classifications from Gaia (left) and OGLE (right). Note that the left panel only shows the Gaia classifications for targets with *best_class_score* \geq 0.5.

(see Table 1), at longer wavelengths (some dusty massive stars may not be identified in the shorter wavelengths due to higher extinction and reddening), and in which models are also clearly shown the Yellow Void between blue supergiant stars (BSGs) and RSGs in order to identify yellow supergiant stars (YSGs).

Figure 13 shows multiple optical CMDs of Gaia, SkyMapper, NSC and M2002 datasets, overlapped with the evolutionary tracks of 7, 9, 12, 15, 20, 25, 32, and $40M_{\odot}$ generated from MIST synthetic photometry. The tracks are color coded based on the equivalent evolutionary phases (EEPs) from core helium burning to carbon burning, and T_{eff} of $7500\text{ K} < T_{\text{eff}}$ (blue; BSGs), $5000 < T_{\text{eff}} \leq 7000\text{ K}$ (yellow; YSGs) and $T_{\text{eff}} \leq 5000\text{ K}$ (red; RSGs) (Neugent et al. 2010). The regions of each type of evolved massive stars are outlined by the dashed lines with color and magnitude criteria listed in Table 4. The average photometric uncertainties are indicated when available.

Due to the relatively large MLR during the RSGs phase, the star could be heavily obscured by the surrounding dust envelope (Smith et al. 2001; Massey et al. 2005; Levesque et al. 2006; Yang et al. 2018). Thus, we also empirically extend the RSGs region from the reddest and faintest points of the models to the even redder but not fainter area in order to avoid the contamination from extreme AGBs (x-AGBs; Boyer et al. 2011) as shown in the diagrams. It may be that some super-AGBs (Herwig 2005; Siess 2006, 2007, 2010; Groenewegen et al. 2009; Doherty et al. 2017) are also selected by the extension. However, we expect that they can be rejected by using several methods as shown in Yang et al. (2018). Moreover, inevitably, there will be contaminations of the main sequence massive stars at the blue end, which cannot be easily disentangled. The diagrams show a clear bimodal distribution of the BSGs and RSGs candidates with few YSGs candidates in between them. We notice that the conven-

tional lower limit of the stellar mass for “massive star” is usually defined as eight solar mass. However, our initial tests show that the observational data are fitted better with the evolutionary track of seven solar mass (or even lower) as shown in the diagrams. This may be due to the treatments of parameters in the model such as convective overshooting, rotation, mixing, metallicity, and MLR, or the uncertainties of extinction and/or bolometric correction, or it is true that “massive star” does have lower limit of stellar mass, which are complicated and need further investigation.

We also notice one important observational evidence that, for the RSGs population, there is a distinct branch stretching continuously from the top of luminous cool region towards the relatively faint warm area, reaching approximately to the tip of the red giant branch (TRGB) without blending into the AGBs population. It is obviously beyond the limit of $7M_{\odot}$ track and most probably down to $6M_{\odot}$. False detection is ruled out since all the Gaia, SkyMapper and NSC data show the same tendency except the M2002 data, which is likely due to the photometric sensitivity. This may indicate the uniqueness of RSGs population, which connects the evolved massive and intermediate stars, since stars with initial mass around 6 to $8M_{\odot}$ are thought to go through a second dredge-up to become AGBs (Eldridge et al. 2007). These low mass RSGs may be also related to the intermediate luminosity optical transients (ILOTs; Prieto et al. 2008; Bond et al. 2009; Berger et al. 2009). Still, more investigations are needed to confirm the true nature of RSGs.

It can be seen from the Figure 13 that in the optical bands, generally, massive stars evolve horizontally across the upper part of CMD, while it is slightly different in the NIR bands. The left panel of Figure 14 shows the 2MASS CMD overlapped with MIST tracks, for which the tracks extend from the fainter and bluer region to the brighter and redder area due to the displacement of T_{eff} and intrinsic emission peaks. The right panel of Figure 14 shows a different way to classify RSGs originated from Cioni et al. (2006a) and Boyer et al. (2011) (hereafter CB method), where Carbon-rich AGBs (C-AGBs) defined by $K_2 < K_S < K_0$, Oxygen-rich AGBs (O-AGBs) defined by $K_S < K_0$ and $K_1 < K_S < K_2$, x-AGBs defined by $K_S < K_S$ -band TRGB ($K_S - TRGB \approx 12.7 \text{ mag}$) and $J - K_S > 2.1 \text{ mag}$, and RSGs defined by $\Delta(J - K_S) = 0.25 \text{ mag}$ from the O-AGBs shown as the dashed line in the diagram (the distance and 0.05 mag for the metallicity between LMC and SMC are corrected; Cioni et al. 2006b). It occurs to us that for the CB method, it covers almost the whole magnitude range of RSGs population down to the K_S -TRGB, where the MIST tracks are unable to cover. However, MIST tracks are more broadened with a part of the tracks also extending to the O-AGBs region. This discrepancy between MIST tracks and CB method will be addressed in our next paper. A simple calculation by using a constant bolometric correction (BC) of $BC_{K_S} = 2.69$ (Davies et al. 2013) and $A_{K_S} = 0.1 \text{ mag}$ shows that RSGs candidates close to K_S -TRGB (e.g., 12.6 mag) only correspond to $\sim 10^{3.4}$ ($\sim 2,500$) solar luminosity (L_{\odot}).

We combine the candidates of each type of evolved massive stars from different datasets and remove the duplications, which results in 1,405 RSG, 217 YSG and 1,369 BSG candidates listed in Table 5. Since the candidates are mostly selected based on the MIST model prediction in different datasets with variety of filters, sky coverages, photometric sensitivities and qualities, it is difficult to judge how well the candidates. Therefore, we rank Rank 0, 1, 2, 3, 4, and -1) the candidates based on the intersection between different CMDs, where Rank 0 indicates that a target has been identified as the same type of evolved massive star

in all five datasets (Gaia, SkyMapper, NSC, M2002 and 2MASS) by the MIST models and so on, and Rank -1 indicates the additional RSG candidates identified by the CB method but not covered by the MIST models. Figure 15 illustrates two CMDs (Gaia and 2MASS) for all the candidates with ranks, where RSG, YSG, and BSG candidates are color coded in red, yellow, and blue colors ranging from dark (Rank 0) to light (Rank -1). Detailed information about each type of evolved massive star candidates are presented in Table 6, 7 and 8. It has to be emphasized that some candidates may have different classifications at the same time, which is likely due to either the inevitable slightly overlapping of adjacent types of massive stars, or the larger photometric errors in the fainter magnitudes. We notice that there are few RSG candidates scattered in the much fainter and redder regions in the Gaia CMD. However, simultaneous inspection of the optical and MIR CMDs shows that almost all the scattered candidates in the bottom right region show infrared excess and/or high MIR luminosity, which may indicate that the dimming in the optical band is caused by the circumstellar dust envelop. Finally, Figure 16 shows the spatial distribution of evolved massive star candidates. It is clearly shown that due to the interaction between LMC and SMC, a bunch of candidates are stretched towards the MB. Further detailed analysis of identified massive star populations will be presented in our following papers.

5. Summary

We present a relatively clean, magnitude-limited (reach to $IRAC1$ or $WISE1 \leq 15.0 \text{ mag}$) multiwavelength source catalog for the SMC with 45,466 targets in total. We intend to build our catalog as a comprehensive dataset serving as an anchor for the future studies, especially for the massive star populations at low-metallicity. It contains data in 50 different bands including 21 optical and 29 infrared bands, retrieved from SEIP, VMC, IRSF, AKARI, HERITAGE, Gaia, SkyMapper, NSC, M2002, and GALEX datasets, ranging from ultraviolet to far-infrared. Additionally, radial velocities and spectral classifications are collected from the literature, as well as the infrared variability information, including MAD, SD, and Amp, derived from WISE, SAGE-Var, VMC and IRSF, and the optical variability information derived from Gaia, NSC and OGLE.

The catalog is essentially built upon a 1” crossmatching and a 3” deblending between SEIP source list and Gaia photometric data. We further constrain the PMs and parallaxes from Gaia DR2 to remove the foreground contamination, by applying a Gaussian profile in parallax with additional elliptical constraint derived from PM_{RA} and PM_{Decl} . We estimate that about 99.5% of the targets in our catalog are likely to be the genuine members of the SMC.

By using the evolutionary tracks and synthetic photometry from MIST and also the theoretical $J - K_S$ color cuts from CB method, we identify three evolved massive star populations in the SMC, namely the BSGs, YSGs and RSGs, in five different CMDs. There are 1,405 RSG, 217 YSG and 1,369 BSG candidates, respectively. We rank the candidates based on the intersection of different CMDs, where the source with the most intersections are given the highest rank. A comparison between the models and observational data shows that, the lower limit of RSGs population may reach to 7 or even $6M_{\odot}$, making RSGs an unique population connecting the evolved massive and intermediate stars, since stars with initial mass around 6 to $8M_{\odot}$ are thought to go through a second dredge-up to become AGBs.

We encourage interested reader to further exploit the potential of our catalog, including, but not limited to, mas-

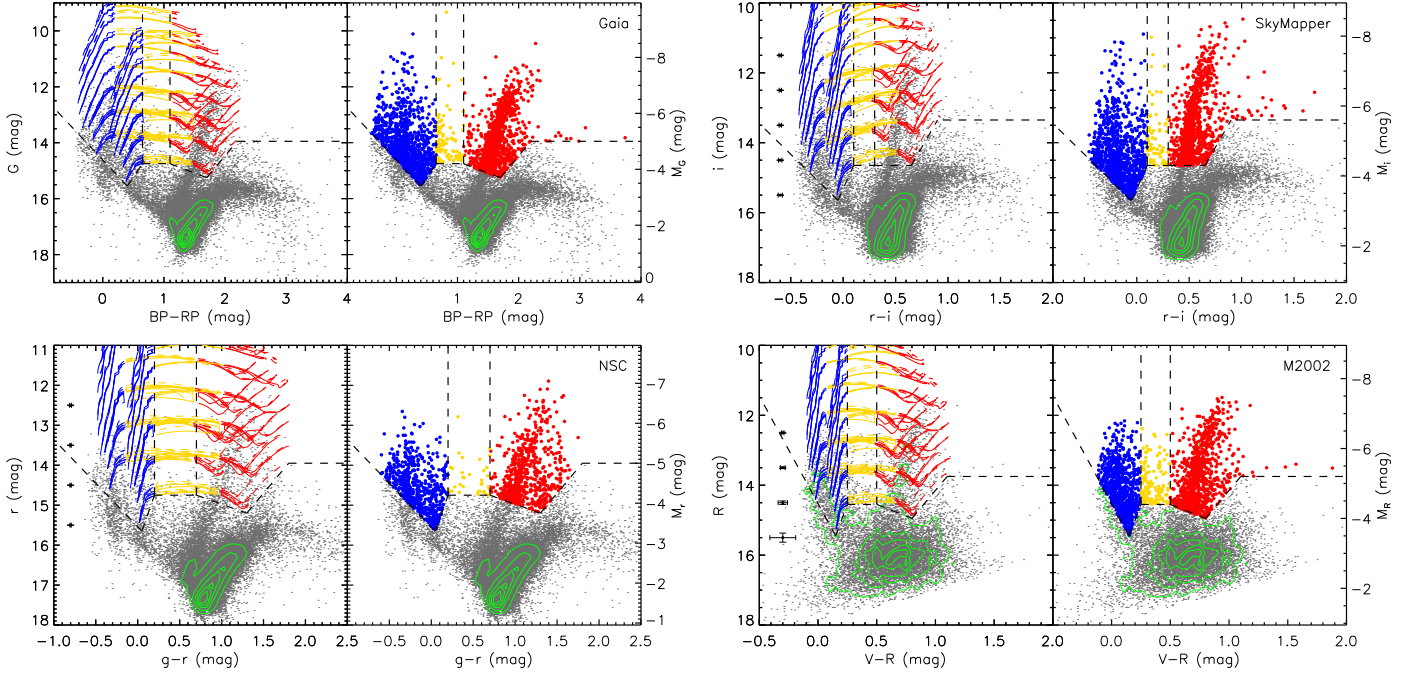


Fig. 13. Color-magnitude diagrams of Gaia (upper left), SkyMapper (upper right), NSC (bottom left) and M2002 (bottom right) datasets. In each diagram, the left panel shows the CMD overlapped with MIST evolutionary tracks of 7, 9, 12, 15, 20, 25, 32 and $40M_{\odot}$ and color coded as BSGs (blue), YSGs (yellow) and RSGs (red). The regions of each type of evolved massive stars are outlined by the dashed lines with color and magnitude criteria listed in Table 4. The average photometric uncertainties are indicated when available. The right panel shows the selected targets for each type of massive stars with the same color convention. The RSGs region is empirically extended from the reddest and faintest points of the models to even redder but not fainter area in order to avoid the contamination from x-AGBs. The diagrams show a clear bimodal distribution of the BSGs and RSGs candidates with few YSGs candidates in between. For the RSGs population, it reaches to approximately the TRGB without blending into AGBs, which is beyond the boundary of $7M_{\odot}$ shown by the model. See text for details.

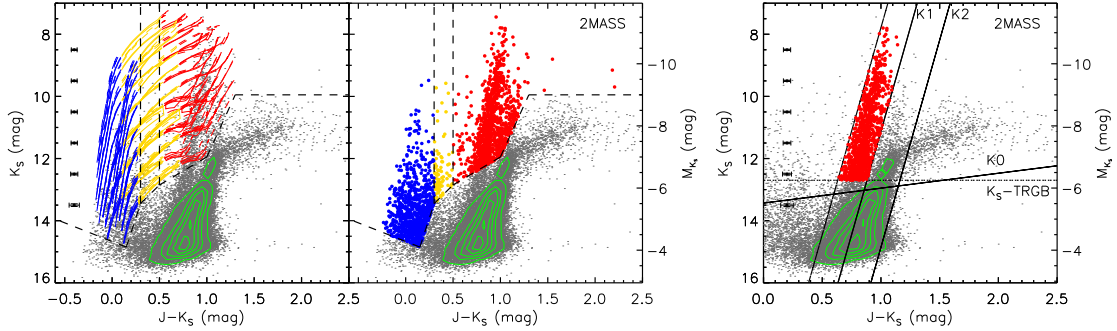


Fig. 14. K_S versus $J - K_S$ diagram for the 2MASS dataset. The left panel is similar to Figure 13 as massive star candidates selected by MIST tracks. The right panel shows the definitions of C-AGBs, O-AGBs, x-AGBs and RSGs regions by using the theoretical $J - K_S$ color cuts from Cioni et al. (2006a) and Boyer et al. (2011). Targets selected as RSGs candidates are shown as red color.

sive stars, supernova progenitors, star formation history, stellar population, stellar kinematics, chemical evolution, individual/integrated spectral energy distribution, time-domain astronomy, and so on. Further detailed analysis of identified massive star populations will be presented in our following papers.

6. Acknowledgments

This study has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement number 772086), and also from Hubble Catalog of Variables project funded by the European Space Agency (ESA) under contract No.4000112940. B.W.J and J.G. gratefully acknowledges sup-

port from the National Natural Science Foundation of China (Grant No.11533002 and U1631104).

This publication makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation. This work is based in part on observations made with the Spitzer Space Telescope, which is operated by the Jet Propulsion Laboratory, California Institute of Technology under a contract with NASA. This publication makes use of data products from the Wide-field Infrared Survey Explorer, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology. It is funded by the National Aeronautics and Space Administra-

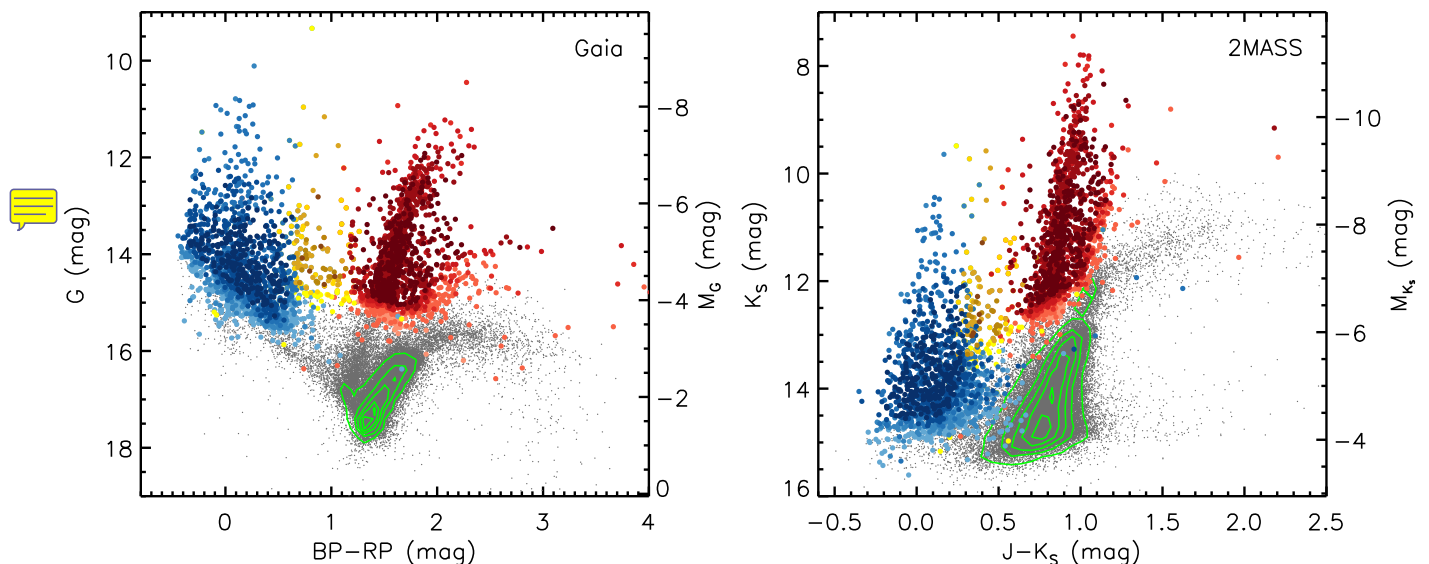


Fig. 15. Color-magnitude diagrams of Gaia (left) and 2MASS (right) with RSG (red), YSG (yellow), and BSG (blue) candidates overlapped, where the colors are coded from dark (Rank 0) to light (Rank -1) based on the ranks (Rank 0, 1, 2, 3, 4, and -1). The RSGs branch extends towards fainter magnitude with few candidates scattered in the much more fainter and redder region in the optical band, which is likely caused by the circumstellar dust envelop.

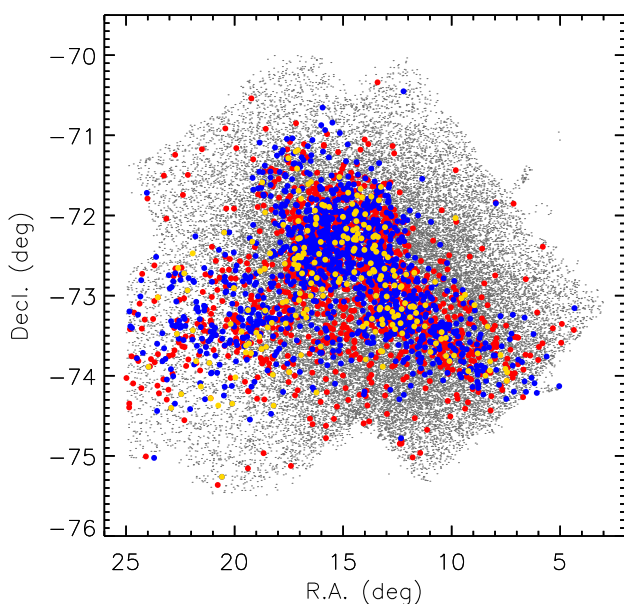


Fig. 16. Spatial distribution of evolved massive star candidates. Due to the interaction between LMC and SMC, a bunch of candidates are stretched towards the MB.

tion. This publication makes use of data products from the Near-Earth Object Wide-field Infrared Survey Explorer (NEOWISE), which is a project of the Jet Propulsion Laboratory/California Institute of Technology. NEOWISE is funded by the National Aeronautics and Space Administration. This research has made use of the NASA/IPAC Infrared Science Archive, which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

This work has made use of data from the European Space Agency (ESA) mission *Gaia* (<https://www.cosmos.esa.int/gaia>), processed by

the *Gaia* Data Processing and Analysis Consortium (DPAC, <https://www.cosmos.esa.int/web/gaia/dpac/consortium>). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.

This research uses services or data provided by the NOAO Data Lab. NOAO is operated by the Association of Universities for Research in Astronomy (AURA), Inc. under a cooperative agreement with the National Science Foundation.

The national facility capability for SkyMapper has been funded through ARC LIEF grant LE130100104 from the Australian Research Council, awarded to the University of Sydney, the Australian National University, Swinburne University of Technology, the University of Queensland, the University of Western Australia, the University of Melbourne, Curtin University of Technology, Monash University and the Australian Astronomical Observatory. SkyMapper is owned and operated by The Australian National University's Research School of Astronomy and Astrophysics. The survey data were processed and provided by the SkyMapper Team at ANU. The SkyMapper node of the All-Sky Virtual Observatory (ASVO) is hosted at the National Computational Infrastructure (NCI). Development and support the SkyMapper node of the ASVO has been funded in part by Astronomy Australia Limited (AAL) and the Australian Government through the Commonwealth's Education Investment Fund (EIF) and National Collaborative Research Infrastructure Strategy (NCRIS), particularly the National eResearch Collaboration Tools and Resources (NeCTAR) and the Australian National Data Service Projects (ANDS).

This research has made use of the SIMBAD database and VizieR catalog access tool, operated at CDS, Strasbourg, France, and the Tool for Operations on Catalogues And Tables (TOPCAT; Taylor 2005).

Based on data products from observations made with ESO Telescopes at the La Silla or Paranal Observatories under ESO programme ID 179.B-2003.

References

- Arenou, F., Luri, X., Babusiaux, C., et al. 2018, *A&A*, 616, A17
- Barba, R. H., Niemela, V. S., Baume, G., & Vazquez, R. A. 1995, *ApJ*, 446, L23
- Berger, E., Soderberg, A. M., Chevalier, R. A., et al. 2009, *ApJ*, 699, 1850
- Bessell, M., Bloxham, G., Schmidt, B., et al. 2011, *PASP*, 123, 789
- Bianchi, L., Shiao, B., & Thilker, D. 2017, *ApJS*, 230, 24
- Bonanos, A. Z., Massa, D. L., Sewilo, M., et al. 2009, *AJ*, 138, 1003
- Bonanos, A. Z., Lennon, D. J., Köhlinger, F., et al. 2010, *AJ*, 140, 416
- Bond, H. E., Bedin, L. R., Bonanos, A. Z., et al. 2009, *ApJ*, 695, L154
- Bouret, J.-C., Lanz, T., Martins, F., et al. 2013, *A&A*, 555, A1
- Boyer, M. L., Srinivasan, S., van Loon, J. T., et al. 2011, *AJ*, 142, 103
- Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *ApJ*, 345, 245
- Castro, N., Oey, M. S., Fossati, L., & Langer, N. 2018, *ApJ*, 868, 57
- Chen, X., Wang, S., Deng, L., de Grijs, R., & Yang, M. 2018, *ApJS*, 237, 28
- Cioni, M.-R. L., Girardi, L., Marigo, P., & Habing, H. J. 2006, *A&A*, 448, 77
- Cioni, M.-R. L., Girardi, L., Marigo, P., & Habing, H. J. 2006, *A&A*, 452, 195
- Cioni, M.-R. L., Clementini, G., Girardi, L., et al. 2011, *A&A*, 527, A116.
- Choi, J., Dotter, A., Conroy, C., et al. 2016, *ApJ*, 823, 102
- Cutri, R. M., & et al. 2013, *VizieR Online Data Catalog*, 2328
- Davies, B., Kudritzki, R.-P., Plez, B., et al. 2013, *ApJ*, 767, 3
- Dobbie, P. D., Cole, A. A., Subramaniam, A., & Keller, S. 2014, *MNRAS*, 442, 1680
- Doherty, C. L., Gil-Pons, P., Siess, L., et al. 2017, *Publications of the Astronomical Society of Australia*, 34, e056.
- D’Onghia, E., & Fox, A. J. 2016, *ARA&A*, 54, 363
- Dotter, A. 2016, *ApJS*, 222, 8
- Eldridge, J. J., Mattila, S., & Smartt, S. J. 2007, *MNRAS*, 376, L52
- Evans, C. J., & Howarth, I. D. 2008, *MNRAS*, 386, 826
- Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, *A&A*, 595, A1
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, *A&A*, 616, A1
- Gaia Collaboration, Eyer, L., Rimoldini, L., et al. 2019, *A&A*, 623, A110.
- Gall, C., Hjorth, J., & Andersen, A. C. 2011, *A&A Rev.*, 19, 43
- Gao, J., Jiang, B. W., Li, A., & Xue, M. Y. 2013, *ApJ*, 776, 7
- González-Fernández, C., Dorda, R., Negueruela, I., & Marco, A. 2015, *A&A*, 578, A3
- Gordon, K. D., Meixner, M., Meade, M. R., et al. 2011, *AJ*, 142, 102
- Graczyk, D., Pietrzyński, G., Thompson, I. B., et al. 2014, *ApJ*, 780, 59
- Groenewegen, M. A. T., Sloan, G. C., Soszyński, I., & Petersen, E. A. 2009, *A&A*, 506, 1277
- Hainich, R., Pasemann, D., Todt, H., et al. 2015, *A&A*, 581, A21
- Haschke, R., Grebel, E. K., & Duffau, S. 2011, *AJ*, 141, 158
- Herwig, F. 2005, *ARA&A*, 43, 435
- Houck, J. R., Roellig, T. L., van Cleve, J., et al. 2004, *ApJS*, 154, 18
- Humphreys, R. M., & McElroy, D. B. 1984, *ApJ*, 284, 565
- Ita, Y., Onaka, T., Tanabé, T., et al. 2010, *PASJ*, 62, 273
- Ita, Y., Matsunaga, N., Tanabé, T., et al. 2018, *MNRAS*, 481, 4206
- Kato, D., Nagashima, C., Nagayama, T., et al. 2007, *PASJ*, 59, 615
- Keller, S. C., & Wood, P. R. 2006, *ApJ*, 642, 834
- Keller, S. C., Schmidt, B. P., Bessell, M. S., et al. 2007, *PASA*, 24, 1
- Kourmionis, M., Bonanos, A. Z., Soszyński, I., et al. 2014, *A&A*, 562, A125
- Kraemer, K. E., Sloan, G. C., Wood, P. R., Jones, O. C., & Egan, M. P. 2017, *ApJ*, 834, 185
- Kunth, D., & Östlin, G. 2000, *A&A Rev.*, 10, 1
- Levesque, E. M., Massey, P., Olsen, K. A. G., et al. 2006, *ApJ*, 645, 1102
- Maeder, A., & Meynet, G. 2012, *Reviews of Modern Physics*, 84, 25
- Mainzer, A., Bauer, J., Cutri, R. M., et al. 2014, *ApJ*, 792, 30
- Massey, P. 2002, *ApJS*, 141, 81
- Massey, P. 2003, *ARA&A*, 41, 15
- Massey, P., & Olsen, K. A. G. 2003, *AJ*, 126, 2867
- Massey, P., Plez, B., Levesque, E. M., et al. 2005, *ApJ*, 634, 1286
- Massey, P. 2013, *New A Rev.*, 57, 14
- McConnachie, A. W. 2012, *AJ*, 144, 4
- Meixner, M., Gordon, K. D., Indebetouw, R., et al. 2006, *AJ*, 132, 2268
- Meixner, M., Panuzzo, P., Roman-Duval, J., et al. 2013, *AJ*, 146, 62
- Meynet, G., Georgy, C., Hirschi, R., et al. 2011, *Bulletin de la Societe Royale des Sciences de Liege*, 80, 266
- Morrissey, P., Conrow, T., Barlow, T. A., et al. 2007, *ApJS*, 173, 682
- Mowlavi, N., Lecoœur-Taïbi, I., Lebzelter, T., et al. 2018, *A&A*, 618, A58
- Murakami, H., Baba, H., Barthel, P., et al. 2007, *Publications of the Astronomical Society of Japan*, 59, S369.
- Neugent, K. F., Massey, P., Skiff, B., et al. 2010, *ApJ*, 719, 1784
- Nidever, D. L., Olsen, K., Walker, A. R., et al. 2017, *AJ*, 154, 199
- Nidever, D. L., Dey, A., Olsen, K., et al. 2018, *AJ*, 156, 131
- Ofek, E. O., Sullivan, M., Cenko, S. B., et al. 2013, *Nature*, 494, 65
- Onaka, T., Matsuhara, H., Wada, T., et al. 2007, *Publications of the Astronomical Society of Japan*, 59, S401.
- Pawlak, M., Soszyński, I., Udalski, A., et al. 2016, *Acta Astron.*, 66, 421
- Paxton, B., Bildsten, L., Dotter, A., et al. 2011, *ApJS*, 192, 3
- Paxton, B., Cantiello, M., Arras, P., et al. 2013, *ApJS*, 208, 4
- Paxton, B., Marchant, P., Schwab, J., et al. 2015, *ApJS*, 220, 15
- Paxton, B., Schwab, J., Bauer, E. B., et al. 2018, *ApJS*, 234, 34
- Pilbratt, G. L., Riedinger, J. R., Passvogel, T., et al. 2010, *A&A*, 518, L1
- Prieto, J. L., Kistler, M. D., Thompson, T. A., et al. 2008, *ApJ*, 681, L9
- Puls, J., Vink, J. S., & Najarro, F. 2008, *A&A Rev.*, 16, 209
- Riebel, D., Boyer, M. L., Srinivasan, S., et al. 2015, *ApJ*, 807, 1
- Rolleston, W. R. J., Trundle, C., & Dufton, P. L. 2002, *A&A*, 396, 53
- Rousseeuw P. J., Croux C., 1993, *Journal of the American Statistical Association*, 88, 424, 1273
- Ruffle, P. M. E., Kemper, F., Jones, O. C., et al. 2015, *MNRAS*, 451, 3504
- Russell, S. C., & Dopita, M. A. 1992, *ApJ*, 384, 508
- Scowcroft, V., Freedman, W. L., Madore, B. F., et al. 2016, *ApJ*, 816, 49
- Seale, J. P., Meixner, M., Sewilo, M., et al. 2014, *AJ*, 148, 124
- Sewilo, M., Carlson, L. R., Seale, J. P., et al. 2013, *ApJ*, 778, 15
- Siess, L. 2006, *A&A*, 448, 717
- Siess, L. 2007, *A&A*, 476, 893
- Siess, L. 2010, *A&A*, 512, A10
- Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, *AJ*, 131, 1163
- Smith, N., Humphreys, R. M., Davidson, K., et al. 2001, *AJ*, 121, 1111
- Smith, N., & Owoccki, S. P. 2006, *ApJ*, 645, L45
- Smith, N. 2014, *ARA&A*, 52, 487
- Soszyński, I., Udalski, A., Szymański, M. K., et al. 2011, *Acta Astron.*, 61, 217
- Soszyński, I., Udalski, A., Szymański, M. K., et al. 2015, *Acta Astron.*, 65, 297
- Soszyński, I., Udalski, A., Szymański, M. K., et al. 2018, *Acta Astron.*, 68, 89
- Szymanski, M. K. 2005, *Acta Astron.*, 55, 43
- Taylor, M. B. 2005, *Astronomical Data Analysis Software and Systems XIV*, 347, 29
- Udalski, A., Szymanski, M., Kaluzny, J., Kubiak, M., & Mateo, M. 1992, *Acta Astron.*, 42, 253
- Udalski, A., Szymanski, M. K., Soszynski, I., & Poleski, R. 2008, *Acta Astron.*, 58, 69
- Udalski, A., Szymański, M. K., & Szymański, G. 2015, *Acta Astron.*, 65, 1
- Wenger, M., Ochsenbein, F., Egret, D., et al. 2000, *A&AS*, 143, 9
- Werner, M. W., Roellig, T. L., Low, F. J., et al. 2004, *ApJS*, 154, 1
- Wolf, C., Onken, C. A., Luvaul, L. C., et al. 2018, *PASA*, 35, e010
- Woosley, S. E., Heger, A., & Weaver, T. A. 2002, *Reviews of Modern Physics*, 74, 1015
- Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, *AJ*, 140, 1868-1881
- Yang, M., & Jiang, B. W. 2011, *ApJ*, 727, 53
- Yang, M., & Jiang, B. W. 2012, *ApJ*, 754, 35
- Yang, M., Bonanos, A. Z., Jiang, B.-W., et al. 2018, *A&A*, 616, A175
- Zaritsky, D., Harris, J., Thompson, I. B., Grebel, E. K., & Massey, P. 2002, *AJ*, 123, 855
- Zhang, Z.-Y., Romano, D., Ivison, R. J., Papadopoulos, P. P., & Matteucci, F. 2018, *Nature*, 558, 260

Table 1. Number of Detected Targets in Each Filter of the SMC Source Catalog

Filter	<i>Number_{detected}</i>	Absolute percentage	Relative percentage ^a
2MASS_J	45335	99.71%	100%
2MASS_H	45335	99.71%	100%
2MASS_K _S	45335	99.71%	100%
IRAC1	44712	98.34%	99.47%
IRAC2	44948	98.86%	100%
IRAC3	38272	84.18%	85.15%
IRAC4	35133	77.23%	18.16%
MIPS24	1261	2.77%	2.81%
WISE1	42460	93.39%	100%
WISE2	42460	93.39%	100%
WISE3	42460	93.39%	100%
WISE4	42460	93.39%	100%
Gaia_G	45466	100%	100%
Gaia_BP	45466	100%	100%
Gaia_RP	45466	100%	100%
VMC_Y	18637	40.99%	99.99%
VMC_J	18639	41.00%	100%
VMC_K _S	18638	40.99%	99.99%
IRSF_J	28645	63.00%	100%
IRSF_H	28592	62.89%	99.81%
IRSF_K _S	28513	62.71%	99.54%
AKARI_N3	555	1.22%	100%
AKARI_N4	551	1.21%	99.28%
AKARI_S7	274	0.60%	49.37%
AKARI_S11	134	0.29%	24.14%
AKARI_L15	40	0.09%	7.21%
AKARI_L24	18	0.04%	3.24%
HERITAGE_f70	1	—	—
HERITAGE_f100	1	—	—
HERITAGE_f160	2	—	—
HERITAGE_f250	2	—	—
HERITAGE_f350	2	—	—
SkyMapper_u	3185	7.01%	7.89%
SkyMapper_v	3711	8.16%	9.20%
SkyMapper_g	37949	83.47%	94.04%
SkyMapper_r	39930	87.82%	98.94%
SkyMapper_i	40356	88.76%	100%
SkyMapper_z	40171	88.35%	99.54%
NSC_u	26425	58.12%	68.24%
NSC_g	36700	80.72%	94.77%
NSC_r	37025	81.43%	95.61%
NSC_i	36297	79.83%	93.73%
NSC_z	38724	85.17%	100%
NSC_Y	1250	2.75%	3.22%
M2002_V	11630	25.58%	100%
M2002_B-V (B)	11630	25.58%	100%
M2002_U-B (U)	4000	8.80%	34.39%
M2002_V-R (R)	11563	25.43%	99.42%
GALEX_FUV	95	0.21%	57.93%
GALEX_NUV	164	0.36%	100%

Notes. ^(a) Related to the most detected filter in each dataset.

Table 2. Observation Epochs of ALLWISE and NEOWISE-R

Beginning (MJD-54000)	Ending (MJD-54000)
ALLWISE	
	1400
1400	1600
NEOWISE-R	
2700	2870
2870	3050
3050	3230
3230	3405
3405	3600
3600	3775
3775	3955
3955	

Table 3. SMC Source Catalog

ID	R.A.(J2000) (deg)	Decl.(J2000) (deg)	2MASS_J (mag)	e_2MASS_J (mag)	OGEL_Ecl_DP (mag)	OGEL_Ecl_DS (mag)
1	3.044541	-73.089456	15.619	0.062		
2	3.071576	-73.074336	15.096	0.047		
3	3.155707	-73.204226	15.470	0.050		
4	3.170525	-73.219597	14.814	0.034		
5	3.174998	-73.129420	14.593	0.034		
...

Notes. This table is available in its entirety in CDS. A portion is shown here for guidance regarding its form and content.

Table 4. Evolved Massive Star Candidate Selection Criteria

Group	Color Criteria	Magnitude Criteria
BSG _{Gaia}	$(BP - RP) < 0.4$	$G \leq 2.333 \times (BP - RP) + 14.617$
	$0.4 \leq (BP - RP) < 0.65$	$G \leq -3.200 \times (BP - RP) + 16.830$
YSG _{Gaia}	$0.65 \leq (BP - RP) < 1.1$	$G \leq 14.750$
RSG _{Gaia}	$1.1 \leq (BP - RP) < 1.7$	$G \leq 0.833 \times (BP - RP) + 13.833$
	$1.7 \leq (BP - RP) < 2.2$	$G \leq -2.600 \times (BP - RP) + 19.670$
	$2.2 \leq (BP - RP)$	$G \leq 13.950$
BSG _{SkyMapper}	$(r - i) < -0.05$	$i \leq 2.933 \times (r - i) + 15.797$
	$-0.05 \leq (r - i) < 0.1$	$i \leq -6.667 \times (r - i) + 15.317$
YSG _{SkyMapper}	$0.1 \leq (r - i) < 0.3$	$i \leq 14.650$
RSG _{SkyMapper}	$0.3 \leq (r - i) < 0.65$	$i \leq 14.650$
	$0.65 \leq (r - i) < 0.9$	$i \leq -5.200 \times (r - i) + 18.030$
	$0.9 \leq (r - i)$	$i \leq 13.350$
BSG _{NSC}	$(g - r) < 0.05$	$r \leq 2.182 \times (g - r) + 15.541$
	$0.05 \leq (g - r) < 0.2$	$r \leq -6.000 \times (g - r) + 15.950$
YSG _{NSC}	$0.2 \leq (g - r) < 0.7$	$r \leq 14.750$
RSG _{NSC}	$0.7 \leq (g - r) < 1.3$	$r \leq 0.750 \times (g - r) + 14.225$
	$1.3 \leq (g - r) < 1.8$	$r \leq -2.500 \times (g - r) + 18.450$
	$1.8 \leq (g - r)$	$r \leq 13.950$
BSG _{M2002}	$(V - R) < 0.15$	$R \leq 6.154 \times (V - R) + 14.527$
	$0.15 \leq (V - R) < 0.25$	$R \leq -9.000 \times (V - R) + 16.800$
YSG _{M2002}	$0.25 \leq (V - R) < 0.5$	$R \leq 14.550$
RSG _{M2002}	$0.5 \leq (V - R) < 0.8$	$R \leq 1.333 \times (V - R) + 13.883$
	$0.8 \leq (V - R) < 1.1$	$R \leq 4.000 \times (V - R) + 18.150$
	$1.1 \leq (V - R)$	$R \leq 13.750$
BSG _{2MASS}	$(J - K_S) < 0.15$	$K_S \leq 1.200 \times (J - K_S) + 14.670$
	$0.15 \leq (J - K_S) < 0.3$	$K_S \leq -9.333 \times (J - K_S) + 16.250$
YSG _{2MASS}	$0.3 \leq (J - K_S) < 0.5$	$K_S \leq -3.000 \times (J - K_S) + 14.350$
RSG _{2MASS}	$0.5 \leq (J - K_S) < 1.0$	$K_S \leq -1.800 \times (J - K_S) + 13.750$
	$1.0 \leq (J - K_S) < 1.3$	$K_S \leq -6.667 \times (J - K_S) + 18.617$
	$1.3 \leq (J - K_S)$	$K_S \leq 9.950$

Table 5. Numbers of Identified Evolved Massive Star Candidates

Dataset	Models	RSGs	YSGs	BSGs
Gaia	MIST	1029	97	952
SkyMapper	MIST	987	74	1035
NSC	MIST	627	34	606
M2002	MIST	836	133	757
2MASS	MIST	1068	93	879
2MASS	CB	1020		
Total (cleaned)		1405	217	1369

Table 6. 1,369 Blue Supergiant Star Candidates in the SMC

ID	R.A.(J2000) (deg)	Decl.(J2000) (deg)	2MASS_J (mag)	e_2MASS_J (mag)	OGLE_Ecl_DS (mag)	Rank
119	4.339399	-73.156092	14.335	0.027		1
268	5.043806	-74.129436	14.232	0.032		2
527	5.623838	-73.759717	13.849	0.029		2
817	6.076174	-74.213858	14.312	0.029		2
912	6.182415	-74.093269	14.087	0.034		2
...

Notes. This table is available in its entirety in CDS. A portion is shown here for guidance regarding its form and content.

Table 7. 217 yellow supergiant star candidates in the SMC

ID	R.A.(J2000) (deg)	Decl.(J2000) (deg)	2MASS_J (mag)	e_2MASS_J (mag)	OGLE_Ecl_DS (mag)	Rank
2249	7.441876	-73.956715	13.640	0.050		4
2278	7.458879	-74.062190	13.001	0.023		3
2361	7.521910	-73.913852	12.708	0.024		3
2726	7.746649	-73.741369	12.737	0.023		2
3584	8.228779	-73.821946	12.821	0.026		2
...

Notes. This table is available in its entirety in CDS. A portion is shown here for guidance regarding its form and content.

Table 8. 1,405 Red Supergiant Star Candidates in the SMC

ID	R.A.(J2000) (deg)	Decl.(J2000) (deg)	2MASS_J (mag)	e_2MASS_J (mag)	OGLE_Ecl_DS (mag)	Rank
124	4.368634	-73.428555	12.497	0.022		2
236	4.921237	-73.353052	11.675	0.022		4
240	4.952093	-73.588549	12.672	0.024		2
376	5.335515	-73.407538	13.358	0.026		-1
635	5.800023	-72.390737	12.920	0.029		3
...

Notes. This table is available in its entirety in CDS. A portion is shown here for guidance regarding its form and content.