RESEARCH ARTICLE Journal of Multidisciplinary Academic

Special Issue in Multidisciplinary Academics related Astronomy Background
ISSN/e-ISSN: 2541 – 0369/2613 – 988X
Vol. 3, No. 2, 2019, Printed in the Indonesia

The Implementation of Markov Chain Monte Carlo on the Study of Star Cluster Membership in Open Cluster NGC 3766 and Globular Cluster 6366

R.Darma[1,*] M.I.Arifyanto[1], and Y.A.Hidayat[1]

[1]*Department of Astronomy, Faculty of Mathematics and Natural Sciences,*
Institut Teknologi Bandung, Ganeca Rd. no. 10, Bandung, 40132, Indonesia

The membership study of star cluster is a fundamental topic in order to learn more deeply about the evolution of star and dynamical evolution of star cluster itself in the galactic environment. Here, we introduce the membership study of open cluster NGC 3766 and globular cluster NGC 6366 by implementing Markov Chain Monte Carlo (MCMC) algorithm in parametric method. We implemented MCMC algorithm in the fitting processes of stellar proper motion distribution and isochrones model with HR-diagram for both star clusters. We found that MCMC algorithm has been successfully distinguished member stars with membership probability larger than 50%. The convergence processes of fitting parameters were going sufficiently fast with small uncertainties. We obtain that NGC 3766 is (25 ± 7) Myr which containing (882 ± 5) member stars with distance modulus (V-M_V)= (11,95 ± 0,32) and colour excess E(B-V)= (0,30 ± 0,01). These parameters are corresponding to the distance (1,60 ± 0,12) kpc from the Sun with extinction A_v= (0,93 ± 0,18). Besides that, the age of NGC 6366 is (11,70 ± 1,10) Gyr and containing (0,30 ± 0,01) member stars with (V-M_V)= (14,98 ± 0,09) and E(B-V)= (0,77 ± 0,09). Our computations shows that the distance of NGC 6366 is (3,30 ± 0,19) kpc from the Sun with A_v= (2,39 ± 0,28). In this work, our computations of physical parameters are in a good agreement with previous studies. Furthermore, through this work we also suggest a new computation of star cluster physical parameters with MCMC algorithm.

Keywords: Star cluster, Membership study, Physical parameters, Numerical method.

1. INTRODUCTION

Star cluster is one of the most interesting objects in the field of astrophysics. As the building blocks of the galaxy [1], star clusters have been the main source of star formation both for single and binary star [2]. This object has now been known as the best source to understand about the stellar evolution and the dynamical evolution in the widely context [3, 4]. However, the membership of most of star clusters is not well studied. Membership study is an important aspect in astrophysics to provide us large opportunity to learn more deeply about star cluster. Generally, the membership study is divided into two methods, i.e. parametric and non-parametric. Parametric is the justification through a set of fixed parameters to determine probability of the model. It is ordinary assuming normal distribution in the justification of membership. To contrast with parametric, non-parametric does not use any set of fixed parameters and there is no assumption of distribution in the justification. Many previous studies attempted to investigate the membership of star cluster using parametric method with different algorithms [5, 6, 7, 8, 9, 10, 11]. It is familiar to adopt due to its simple justification than non-parametric method. But, different algorithm in this method provides different result which depends on the assumptions used in that algorithm. In this work, we would like to introduce the implementation of Markov Chain Monte Carlo (MCMC) algorithm in order to solve parametric method in the case of star cluster membership. Strictly assumptions are not necessary in the MCMC algorithm and we only need to consider the log-likelihood for the distribution of data we have (explained

[*]*Email Address: darmarendy@gmail.com*

more in the next section). Therefore, this algorithm is sufficiently good to solve the aforementioned issue. Here, we used two star clusters in our Galaxy, i.e. open cluster NGC 3766 and globular cluster NGC 6366, which are not well studied yet. In addition, we also would like to introduce the MCMC algorithm to compute the physical parameters for both star clusters from its Hertzsprung-Russell (HR) diagram.

2. METHODOLOGY

A. Data

In this work, we adopted proper motion and photometric of stars for NGC 3766 and for NGC 6366. The radial velocity measurement for stars around star cluster is still difficult to do. Therefore, in this work we did not consider the radial velocity of stars for the first approximation. The data from were selected statistically based on its relative error both for proper motion and photometric. Therefore, we have 2468 stars of NGC 3766 and 2530 stars of NGC 6366 adopted in this work [12, 13].

B. Markov Chain Monte Carlo

In order to investigate the membership of star clusters, MCMC algorithm was implemented in the fitting process of stellar proper motion distribution. Here we use Equation (1) on the fitting process which it is a double Gaussian function that consists of member stars distribution and field stars distribution. Member stars are the stars that belong to star cluster, while field stars are not the member stars.

By using Equation (1), we can split field stars and member stars through MCMC algorithm. As the results, we obtained the best values of eleven fitting parameters, i.e. n_f, ρ_f, ρ_c, μ_{x0c}, μ_{x0f}, μ_{y0c}, μ_{y0f}, σ_{xc}, σ_{xf}, σ_{yc}, and σ_{yf}.

Here n_f, ρ_f, and ρ_c are the fraction of field stars, correlation coefficient of field stars and of member stars, respectively. The mean of stellar proper motion distribution of field stars are defined as μ_{x0f} and μ_{y0f}, as well as for the member stars, i.e. μ_{x0c} and μ_{y0c}. While the standard deviation for stellar proper motion distribution of field stars and of member stars are given as $\sigma_{xf}, \sigma_{yf}, \sigma_{xc}$, and σ_{yc}, respectively. All those parameters are presented in Equation (1) as the probability distribution function $\Phi(\mu_x, \mu_y)$.

$$\Phi(\mu_x, \mu_y) = (1 - n_f)\Phi_c(\mu_x, \mu_y) + n_f \Phi_f(\mu_x, \mu_y) \qquad (1)$$

Where:

$$\Phi_c(\mu_x, \mu_y) = \frac{1}{2\pi\sigma_{xc}\sigma_{yc}C_c} exp\left\{\frac{-1}{C_c}\left[{A_c}^2 + {B_c}^2 - 2\rho_c A_c B_c\right]\right\}$$

$$\Phi_f(\mu_x, \mu_y) = \frac{1}{2\pi\sigma_{xf}\sigma_{yf}C_f} exp\left\{\frac{-1}{C_f}\left[{A_f}^2 + {B_f}^2 - 2\rho_f A_f B_f\right]\right\}$$

$$A_c = \frac{\mu_x - \mu_{xoc}}{\sigma_{xc}}, B_c = \frac{\mu_y - \mu_{yoc}}{\sigma_{yc}}, C_c = \sqrt{1 - {\rho_c}^2}$$

$$A_f = \frac{\mu_x - \mu_{xof}}{\sigma_{xf}}, B_f = \frac{\mu_y - \mu_{yof}}{\sigma_{yf}}, C_f = \sqrt{1 - {\rho_f}^2}$$

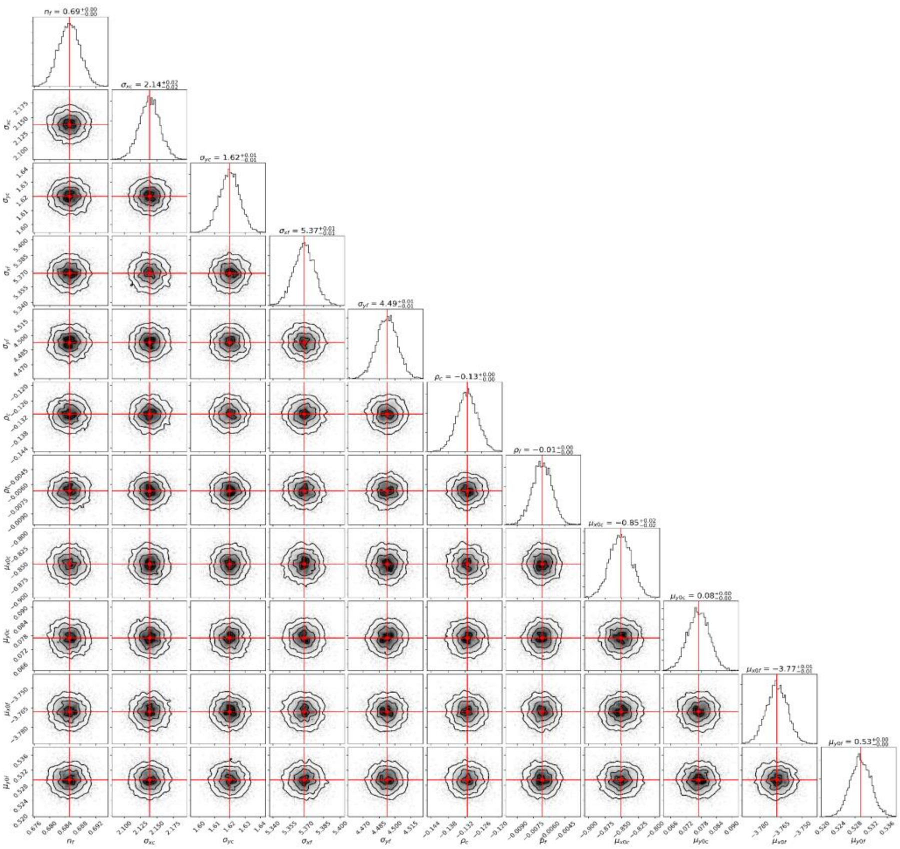

Figure 1. The triangle plot of the eleven fitting parameters after 90% burn-in process for NGC 3766. It was made with 50 bins. This plot is apparently have approximately normal distribution with mean and 1σ confidence level are corresponding to Table II. A few of 1σ confidence levels in this figure are apparently in zero value due to the rounding to 2 decimal places. The red lines and squares pointing at the mean value of every fitting parameters.

Table I. The initial guessed value and given range for every prior used for NGC 3766 and NGC 6366.

No	Parameter	Initial Guessed Value		Range
		NGC 3766	NGC 6366	
1	n_f	0,8	0,05	$0 < n_f < 1$
2	ρ_f	–0,43	–0,43	$-1 < \rho_f < 1$
3	ρ_c	0,5	0,5	$-1 < \rho_c < 1$
4	μ_{x0f}	5,2	4,2	$-10 < \mu_{xof} < 10$
5	μ_{y0f}	3,4	0,4	$-10 < \mu_{yof} < 10$
6	μ_{x0c}	1,3	3,3	$-10 < \mu_{xoc} < 10$
7	μ_{y0c}	5,12	4,12	$-10 < \mu_{yoc} < 10$
8	σ_{xf}	5,43	7,44	$\sigma_{xf} > 0$
9	σ_{yf}	1,53	3,53	$\sigma_{yf} > 0$
10	σ_{xc}	5,25	3,25	$\sigma_{xc} > 0$
11	σ_{yc}	2,35	6,35	$\sigma_{yc} > 0$

These eleven fitting parameters were used as the prior in the MCMC algorithm, where we have to define any value (as initial guessed value) in any given probable range for every parameter, see Table I for both star clusters. At every step of iteration, we generate a new value for every parameter from previous one with addition a normally random number. The normally random number is set with proper mean and standard deviation. Therefore, we have a set of new value of prior for every step. We then calculate the posterior at every step as presented in Equation (2):

$$P(\theta|D,M) \propto P(D|\theta,M)P(\theta|M) \quad (2)$$

where $P(D|\theta,M) = \sum_{i=1}^{N} log\Phi(\mu_{xi},\mu_{yi})$.

Here θ is the eleven fitting parameters. While D and M are the probability distribution function in Equation (1) and the proper motion of stars, respectively. Where posterior $\left(P(\theta|D,M)\right)$ shows the probability of a set of prior $\left(P(\theta|M)\right)$ at every step of iteration. While $P(D|\theta,M)$ is the log-likelihood function which depends on the probability distribution function as mentioned before. For simplicity, in order to reach the convergence results from MCMC algorithm, we adopted the following method at the i-th step of iteration (metropolis algorithm):

- if $P(\theta|D,M)_{i+1} \geq P(\theta|D,M)_i$, the prior is accepted. It is then used to generate the next set of prior.
- if $P(\theta|D,M)_{i+1} < P(\theta|D,M)_i$, the prior is rejected. The previous set of prior is used to generate the next set, instead of using the rejected prior.

That process is repeated recursively depends on the number of iteration we set. The higher number of iteration can produce the better convergence. However, it also spends many CPU time.

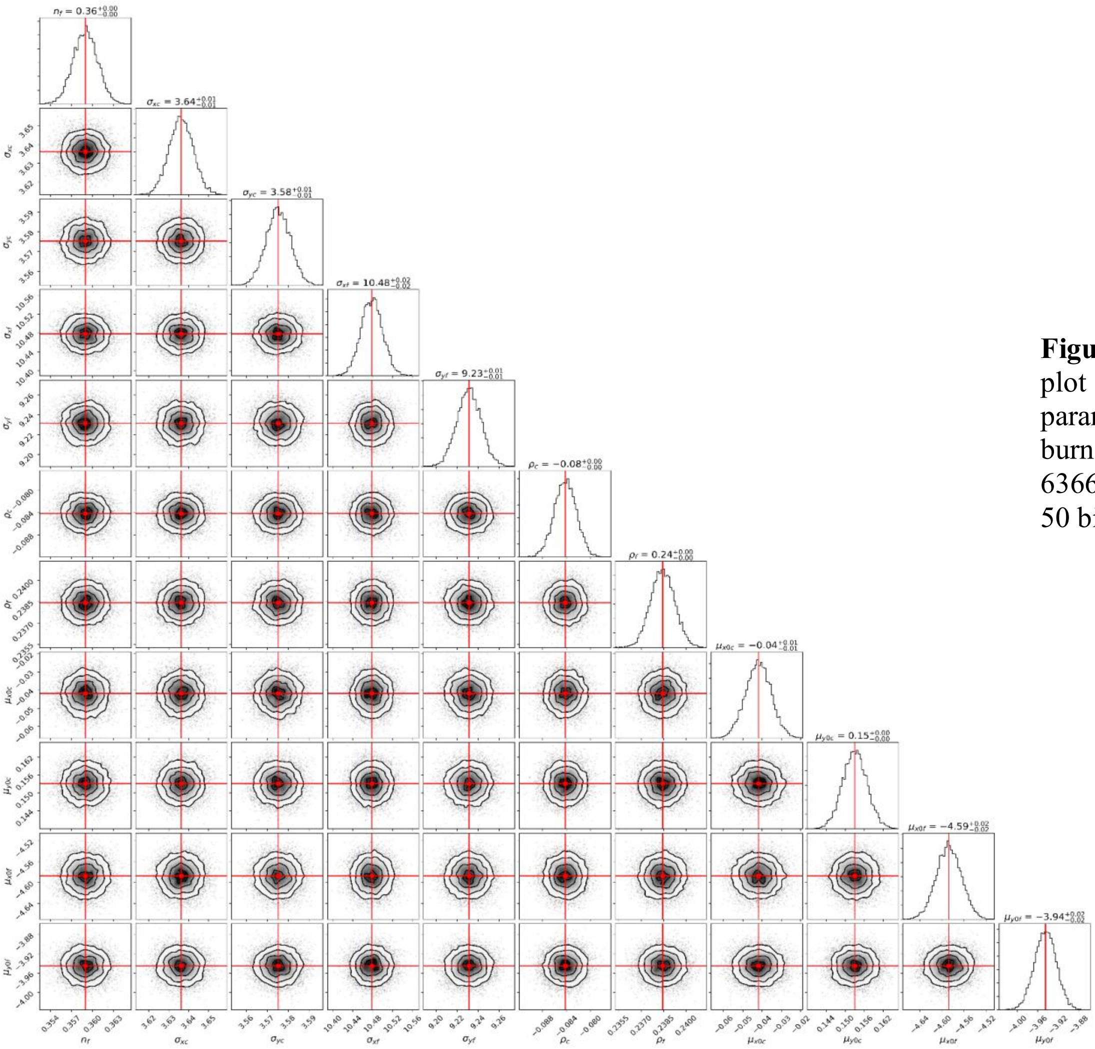

Figure 2. The triangle plot of the eleven fitting parameters after 90% burn-in process for NGC 6366. It was made with 50 bins.

Therefore, we use 1×10^5 iteration for computing process using Workstation computer and Perseus cluster computer in Astronomy Department of Institut Teknologi Bandung. It spent a week CPU time to finish the iterations. At the end of the iterations, we implemented 90% burn-in process for the prior in order to get the highly precise value of the eleven fitting parameters.

C. Membership Probability

We used the best values of the eleven fitting parameters to calculate the membership probability of every star using Equation (3) which containing parameters as explained in the previous subsection. Here $n_c = 1 - n_f$ is the fraction of member stars and $P_i(C|\mu_{xi}, \mu_{yi})$ is the membership probability of the i-th star. In this work, we categorized the stars which have membership probability greater than 50% (P_{50}), 70% (P_{70}), and 90% (P_{90}) as member stars [11, 12, 13]. While the stars with membership probability less than of 50% were categorized as field stars.

$$P_i(C|\mu_{xi}, \mu_{yi}) = \frac{n_c \Phi_c(\mu_{xi}, \mu_{yi})}{n_c \Phi_c(\mu_{xi}, \mu_{yi}) + n_f \Phi_f(\mu_{xi}, \mu_{yi})} \quad (3)$$

Table II. The best values of eleven fitting parameters over NGC 3766 and NGC 6366.

No	Parameter	Value	
		NGC 3766	NGC 6366
1	n_f	0,68515 ± 0,00276	0,35895 ± 0,00173
2	ρ_f	−0,00665 ± 0,00092	0,23844 ± 0,00080
3	ρ_c	−0,13110 ± 0,00372	−0,08418 ± 0,00184
4	μ_{x0f}	−3,76780 ± 0,00753	−4,58710 ± 0,02241
5	μ_{y0f}	0,52942 ± 0,00254	−3,94493 ± 0,02159
6	μ_{x0c}	−0,85160 ± 0,01633	−0,04184 ± 0,00654
7	μ_{y0c}	0,07685 ± 0,00434	0,15306 ± 0,00384
8	σ_{xf}	5,36783 ± 0,00942	10,47868 ± 0,02153
9	σ_{yf}	4,49252 ± 0,00984	9,23153 ± 0,01134
10	σ_{xc}	2,13898 ± 0,01614	3,63623 ± 0,00610
11	σ_{yc}	1,61996 ± 0,00664	3,57536 ± 0,00580

In order to achieve the membership results, we construct the HR-diagram of star cluster which consists of member stars and field stars. We then also implemented MCMC algorithm to extract the physical parameters from both star clusters through HR-diagram fitting. Some grids of isochrones from [14] were used on this fitting which it has been categorized based on age and metallicity of star cluster. We binned the color index $(B - V)$ and apparent magnitude (V) from HR-diagram of star cluster and isochrones model. The binned data is called as fiducially line. Note that, HR-diagram and isochrones must have the same number of bin so that the fitting can be done well. Fitting process on every isochrones model involved distance modulus $(V - M_V)$ and color excess $(E(B - V))$ as the fitting parameters, where their best values are corresponding to the smallest value of chi-squared (χ^2).

3. RESULTS AND DISCUSSION

We presented our best values of the eleven fitting parameters in Table II. The best values were obtained by calculating the mean value of fitting parameters which have been filtered through 90% burn-in process while the uncertainties with 1σ confidence level were from normal distribution approximation. Such small values of uncertainties have carried us to the conclusion that our fitting parameters obtained through MCMC algorithm are sufficiently good. Besides that, we obtained that the fluctuations of fitting parameters during the fitting process are relatively not wide and the convergence processes at every fitting parameter were going sufficiently fast, i.e. $\leq 5 \times 10^4$ iteration for NGC 3766 and $\leq 2 \times 10^4$ iteration for NGC 6366. To make it more clearly, we presented the distribution of the eleven fitting parameters, after implemented 90% burn-in process, in the form of triangle plots with 50 binned data (see Figure 1 and 2). Every contour shows the distribution combined from two fitting parameters. From those figures, we can see that every fitting parameters reached a convergence area at the center of the distribution with mean and standard deviation as mentioned in Table II.

We already have divided the member stars and field stars of NGC 3766 and NGC 6366 in the Vector Point Diagram (VPD) and HR-diagram with their own distribution of membership probability (see Figure 3). We obtained (882 ± 5) and (1805 ± 3) member stars around NGC 3766 and NGC 6366, respectively. The member stars are always apparently surrounding the most probable area of where star cluster located. These member stars of every star cluster were then divided into their own membership probability categories and we found that there are no member stars with P_{90} for NGC 3766 (see Table III). It could be explained from the difference of stellar density in both star clusters, where NGC 3766 as open cluster has lower stellar density than globular cluster NGC 6366 (see the value of n_f in Table II and also n_c). Therefore, the higher stellar density affected on the gradually membership probability in member stars.

Table III**.** The number of member stars around NGC 3766 and NGC 6366 with P_{50}, P_{70}, and P_{90}.

No	Probability	Number of member stars	
		NGC 3766	NGC 6366
1	P_{50}	415 ± 5	233 ± 1
2	P_{70}	466 ± 8	712 ± 6
3	P_{90}	0	859 ± 7
4	Total	882 ± 5	1805 ± 3

We distinguished the photometric of member stars and of field stars in HR-diagram (see Figure 3b and 3e) to confirm our membership results from stellar proper motion distribution. As the consequence of our chosen membership method which used only stellar proper motion distribution without considering the radial velocity or distance from the Sun.

We found there are field stars which overlapped with member stars area, particularly in the main sequence area. It means there is contamination from background and foreground stars in this membership method. However, this method still can identify the existence of member stars with P_{50} and we think it is still sufficiently good for the first approximation.

Table IV. The physical parameters of NGC 3766 and NGC 6366 obtained from the fitting on HR-diagram of member stars with P_{50}.

No	Physical Parameter	Value	
		NGC 3766	NGC 6366
1	Age	(25 ± 7) Myr	(11,70 ± 1,10) Gyr
2	Z	0,0104 ± 0,0067	0,0030 ± 0,0017
3	$[Fe/H]$	−0,40 ± 0,40	−0,88 ± 0,30
4	$E(B-V)$	0,30 ± 0,01	0,77 ± 0,09
5	$(V-M_V)$	11,95 ± 0,32	14,98 ± 0,09
6	A_V	0,93 ± 0,18	2,39 ± 0,28
7	d (kpc)	1,60 ± 0,12	3,30 ± 0,19

We fitted the HR-diagram of all member stars with P_{50} from NGC 3766 and NGC 6366 with some grids of isochrones model, i.e. 25 isochrones models for NGC 3766 and 20 isochrones models for NGC 6366 (see Figure 4).

These isochrones were chosen according to the age and metallicity of both star clusters obtained from previous studies [12, 15, 16, 17]). The fitting on HR-diagram informed us that NGC 3766 has $(V - M_V) =$ (11,95 ± 0,32) and $E(B-V) =$ (0,30 ± 0,01) with $\chi^2 \leq$ 0,20. While NGC 6366 has $(V - M_V) =$ (14,98 ± 0,09) and $E(B-V) =$(0,77 ± 0,09) with $\chi^2 \leq$ 0,02. Another physical parameters of both star clusters obtained from the fitting can be seen in Table IV. The age of NGC 3766 obtained in this work is in a good agreement with [16], where they obtained the age of NGC 3766 was 24 Myr. In addition, $[Fe/H]$ value of NGC 3766 is also relatively in a good agreement. But it is still bias due to the large uncertainty obtained in this work. Besides that, [12] computed the distance of NGC 3766 was 2,5 kpc and this value is completely different with our result. It is mostly likely due to the fitted isochrones models containing wide range of metallicity and it sensitively affected the fitting process. To contrast with NGC 3766, our results for NGC 6366 are in a good agreement with previous studies, where [15] computed the age of NGC 6366 was 11 Gyr with $[Fe/H]$ = –0,67 and [17] found that the distance of NGC 6366 from the Sun was 3,5 kpc.

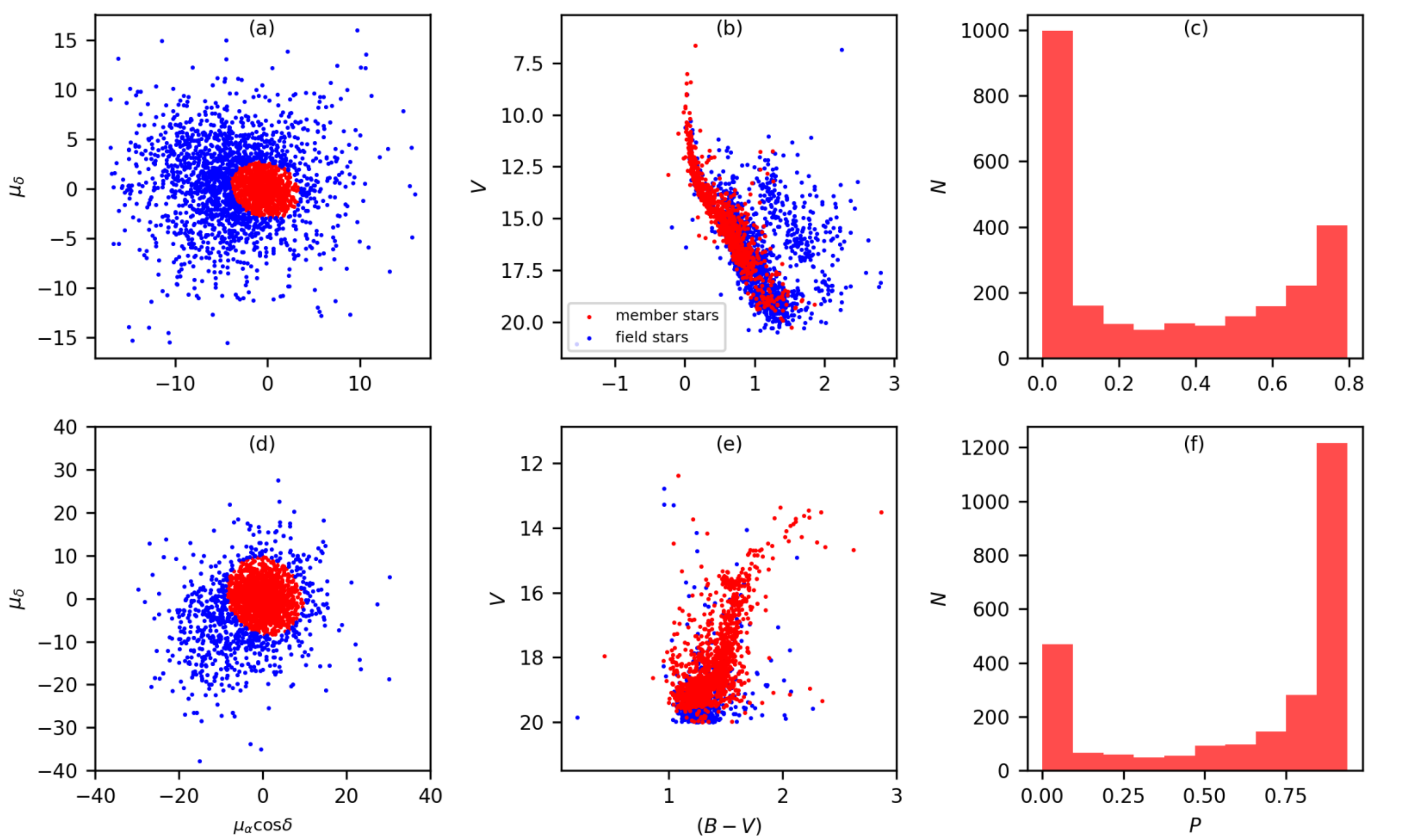

Figure 3. The VPD (a), HR-diagram (b), and distribution of membership probability (c) for NGC 3766. The VPD (d), HR-diagram (e), and distribution of membership probability (f) for NGC 6366. The red dots representing the member stars with membership probability larger than 50% around both star clusters. While the blue dots are the field stars.

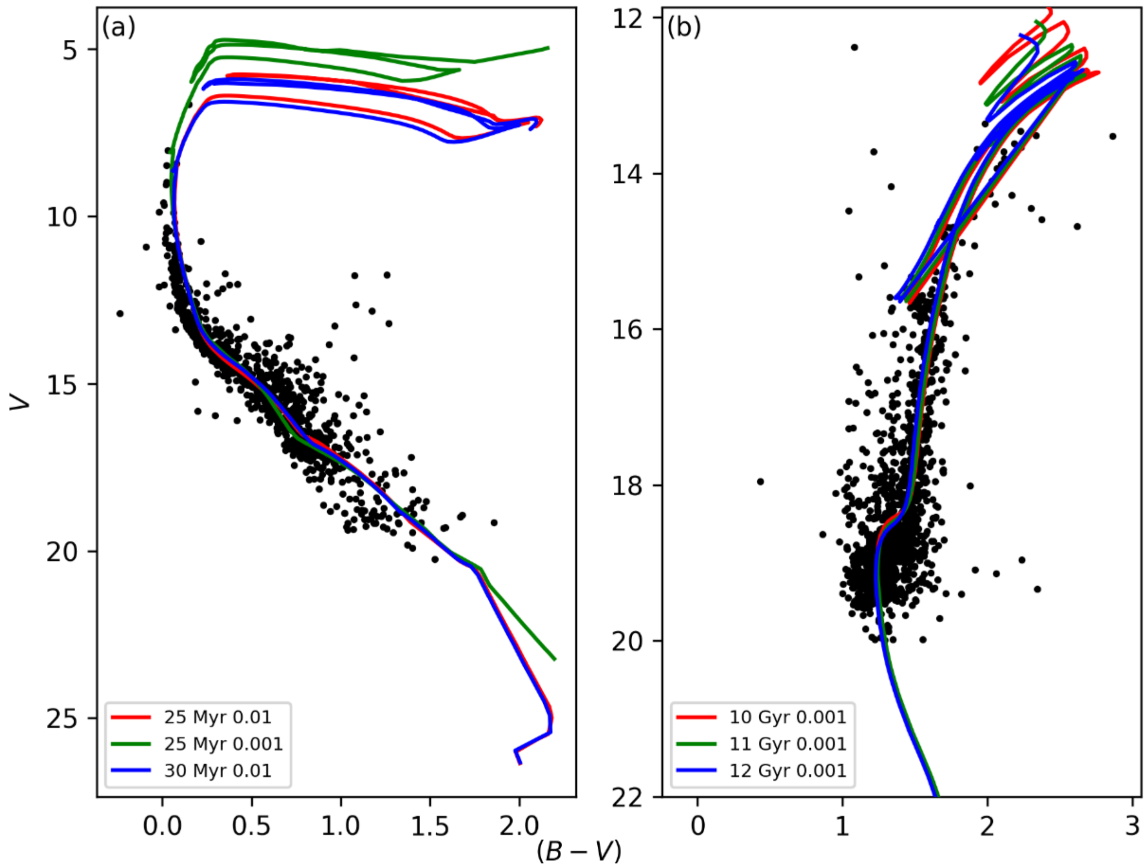

Figure 4. The HR-diagram of NGC 3766 (a) and NGC 6366 (b) which consist of member stars with P_{50}. Different colors of solid lines represent three samples of isochrones models with different age and metallicity which have $\chi^2 \leq$ 0,20 for NGC 3766 and $\chi^2 \leq$ 0,02 for NGC 6366.

It is well accepted that star cluster with highly stellar density is more easily to investigate its membership and physical parameters. However, the uncertainty of physical parameters for NGC 3766 are sufficiently small, $[Fe/H]$ as the exception. Therefore, our results are still statistically good. Some aspects are necessary to consider in the future work. The distance or radial velocity of stars should be involved in order to obtain 3D spatial membership. Besides that, we suggest to use Affine Invariance algorithm to reach better convergence in the MCMC algorithm [18]. This algorithm also can reduce the CPU time even for highly number of data used in our work. Another way is the emcee module [19] in python programming language can be used to solve CPU time issue. Affine Invariance algorithm is implemented within it.

4. CONCLUSIONS

We have implemented our MCMC algorithm on the membership study of NGC 3766 and NGC 6366. The algorithm has provided best values of the eleven fitting parameters with small uncertainties. We found that 31,49% of stars around NGC 3766 are member stars and the rests were supposed to be field stars. While NGC 6366 contains 64,11% of member stars surrounding the cluster. We clearly concluded that the MCMC algorithm has been successfully investigated the membership of NGC 3766 and NGC 6366 with proper range of given prior and sufficiently fast convergence. We noted that, the selected range value of prior depends on the rate of convergence process. We have also implemented the MCMC algorithm on the fitting of HR-diagram in order to obtain the physical parameters of both stars clusters, as can be seen in Table IV. Our results are in a good agreement with previous studies and it carried us to the conclusion that MCMC algorithm is as powerful as another method in the membership study of star clusters.

Acknowledgements: We acknowledged funding from Faculty of Mathematics and Natural Sciences, Institut Teknologi Bandung through *Bantuan Dana Penelitian Mahasiswa*. Thanks to Mifthanzi Ariana Sarashanti, Ni Putu Audita Placida Emas, and Tio Perdana for useful discussions during this interesting work learned in the Computational Topics course.

References

[1] Martinez-Barbosa, C. A., Brown, A. G. A., Boekholt, T., Portegies-Zwart, S. F., Antiche, E., and Antoja, T. **2016.** The Evolution of Sun's Birth Cluster and the Search for Solar Siblings in the Gaia Era. Monthly Notices of the Royal Astronomical Society, vol. 457 p.1062

[2] Lada, C. J., and Lada, E. A. **2003.** Embedded Clusters in Molecular Clouds. Annual Review of Astronomy and Astrophysics, vol.41 p.57

[3] Gieles, M., Heggie, D. C., and Zhao, H. **2011.** The Life Cycle of Star Clusters in a Tidal Field. Monthly Notices of the Royal Astronomical Society, vol.413 p.2509

[4] Baumgardt, H., and Makino, J. **2003.** Dynamical Evolution of Star Clusters in Tidal Fields. Monthly Notices of the Royal Astronomical Society, vol.340 p.227

[5] Fitriana, I. K., and Arifyanto, M. I. **2016.** Membership Determination of Open Cluster with Parametric Method: Cross Entropy. Journal of Physics: Conference Series, vol.771 p.1

[6] Priyatikanto, R., and Arifyanto, M. I. **2015a.** A New Non-Parametric Approach to Determine Proper Motions of Star Clusters. Publications of the Korean Astronomical Society, vol.30 p.271

[7] Priyatikanto, R., and Arifyanto, M. I. **2015b.** The Implementation of Binned Kernel Density Estimation to Determine Open Clusters' Proper Motions: Validation of the Method. Astrophysics and Space Science, vol.366 p.161

[8] Gao, X. H. **2014.** Membership Determination of Open Cluster NGC 188 Based on the DBSCAN Clustering Algorithm. Journal of Physics: Conference Series, vol.14 p.2

[9] Wu, Z. Y., Zhou, X., Ma, J., Jiang, Z. J., and J. S. Chen, **2006.** Membership Determination of Open Cluster M48 Based on BATC 13-Band Photometry. Publications of the Astronomical Society of the Pacific, vol.118 p.846

[10] Campana, R., Massaro, E., Gasparrini, D., Cutini, S., and Tramacere, A. **2008.** A Minimal Spanning Tree Algorithm for Source Detection in γ-ray Images. Monthly Notices of the Royal Astronomical Society, vol.383 p.1166

[11] Sabogal-Martinez, B. E., Garcia-Varela, J. A., Higuera, G., Uribe, A., and Brieva, E. **2001.** Membership in the Region of the Open Cluster NGC 2244. Revista Mexicana De Astronomia Y Astrofisica vol.37 p.105

[12] Yadav, R. K. S., Sariya, D. P., and Sagar, R. **2013.** Proper Motions and Membership Probabilities of Stars in the Region of Open Cluster NGC 3766. Monthly Notices of the Royal Astronomical Society, vol.430 p.3350

[13] Sariya, D. P., and Yadav, R. K. S. **2015.** Proper Motions and Membership Probabilities of Stars in the Region of Globular Cluster NGC 6366. Astronomy and Astrophysics, vol.584 p.A59

[14] Girardi, L., Bressan, A., Bertelli, G., and Chiosi, C. **2000.** Evolutionary Tracks and Isochrones for Low-and Intermediate-Mass Stars: From 0.15 to 7 M_{sun}, and from Z = 0.0004 to 0.03. Astronomy and Astrophysics Supplement Series, vol.141 p.371

[15] Campos, S., Kepler, S. O., and Binatto, C. **2013.** Panchromatic Fits to the Globular Cluster NGC 6366. Memorie della Societa Astronomica Italiana, vol.84, p.195

[16] Mointinho, A, Alfaro, E. J., Yun, J. L., Phelps, R. L. **1997.** CCD UBV Photometry of the Young Open Cluster NGC 3766. The Astronomical Journal, vol.113 p.1359

[17] Harris, W. E. **1996.** A Catalogue of Parameters for Globular Clusters in the Milky Way. The Astronomical Journal, vol.112 p.1487

[18] Goodman, J. and Weare, J. **2010.** Ensamble Samplers with Affine Invariance. Communications in Applied Mathematics and Computational Science, vol.5 p.1

[19] Foreman-Mackey, D., Hogg, D. W., Lang, D., and Goodman, J. **2013.** emcee: The MCMC Hammer, Publications of the Astronomical Society of the Pacific, vol.125 p.306

Received: 05 Apr 2019, Accepted: 12 May 2019