New method for estimating molecular cloud distances based on Gaia, 2MASS, and the TRILEGAL galaxy model

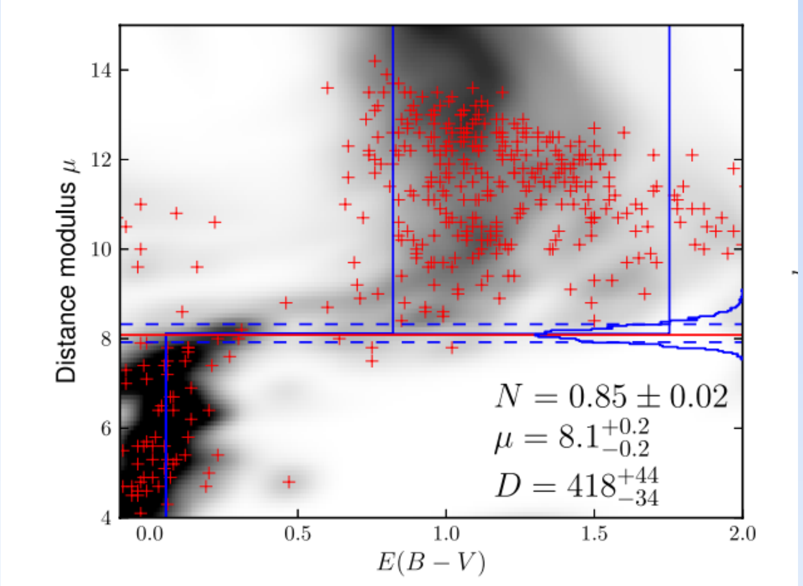
Mei et al. 2024

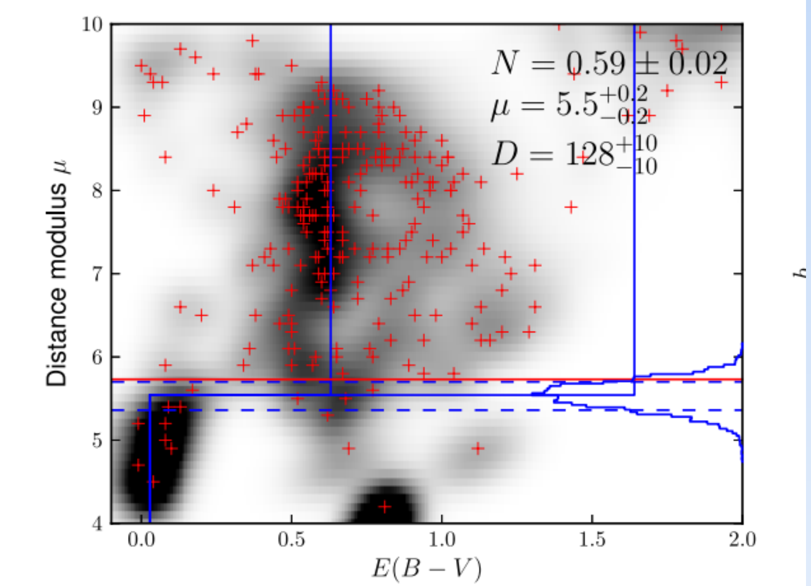
Helong Guo (郭贺龙，hlguo@ynu.edu.cn)
Yunnan University

2024-6-14

Introduction

- Molecular clouds are the birthplaces of stars. The distance is an important parameter of a molecular cloud when the intrinsic physical properties (mass and size) are to be estimated.
- It is always a challenging task to determine the distances of molecular clouds. Several methods were applied to estimate the distances of molecular clouds.
    - kinematic distances. The distances of molecular clouds can be estimated from their radial velocities (V lsr), which are assumed to be attributed to the large-scale rotation of the spiral arms of the Milky Way.
    - The cloud distances can also be estimated by identifying objects associated with or within the molecular cloud, such as OB-associations, young open clusters, H II regions, young stellar objects (YSOs), and masers, whose distances are measured from the trigonometric parallaxes or photometry of stars.
    - The extinction method relies on accurately estimating the distance and extinction of numerous stars.





Schlafly et al. 2014

Data

- Molecular clouds sample. Yan et al. (2021) applied DBSCAN to identify molecular clouds from the Milky Way Imaging Scroll Painting (MWISP) survey by the Purple Mountain Observatory (PMO) 13.7 m millimeter telescope, MWISP ^{12}CO J = 1–0 line emission between l = 25°–50° and in the velocity range –6 to 30 km s^{-1}. They extracted 359 molecular clouds and measured distances for 27 molecular clouds using the background-eliminated extinction-parallax II (BEEP-II) method.
- Astrometric and photometry catalog : Gaia EDR3 and 2MASS. They merged the Gaia and 2MASS catalogs using a crossmatch radius of 1″ and took the J–Ks color from the 2MASS catalog and the distance from Bailer-Joneset al. (2021).

The uncertainty of the stellar distance, $$\Delta d = \frac{1}{2}(d_{\mathrm{upper}} - d_{\mathrm{lower}}),$$

Method

- The light passing through a molecular cloud is attenuated by dust within the cloud. The average color of the background stars behind a molecular cloud is redder than that of the foregroundstars.
- The basic idea is to compare the colors of background stars to those of foreground stars, and to locate the distance at which the color difference due to cloud extinction is most prominent.
- Selection of on-cloud stars
  a) The stars lie within the irregular boundary of the molecular cloud.
  b) The relative uncertainty of the stellar distance is lower than 10%, that is, $\Delta d/d \leq 0.1$.
  c) For each cloud, They calculated a mean rms noise, σ_{int} of integrated CO emission (WCO), based on the ^{12}CO spectra of a cloud in a certain velocity range. They set a threshold COcut = $5\sigma_{int}$.

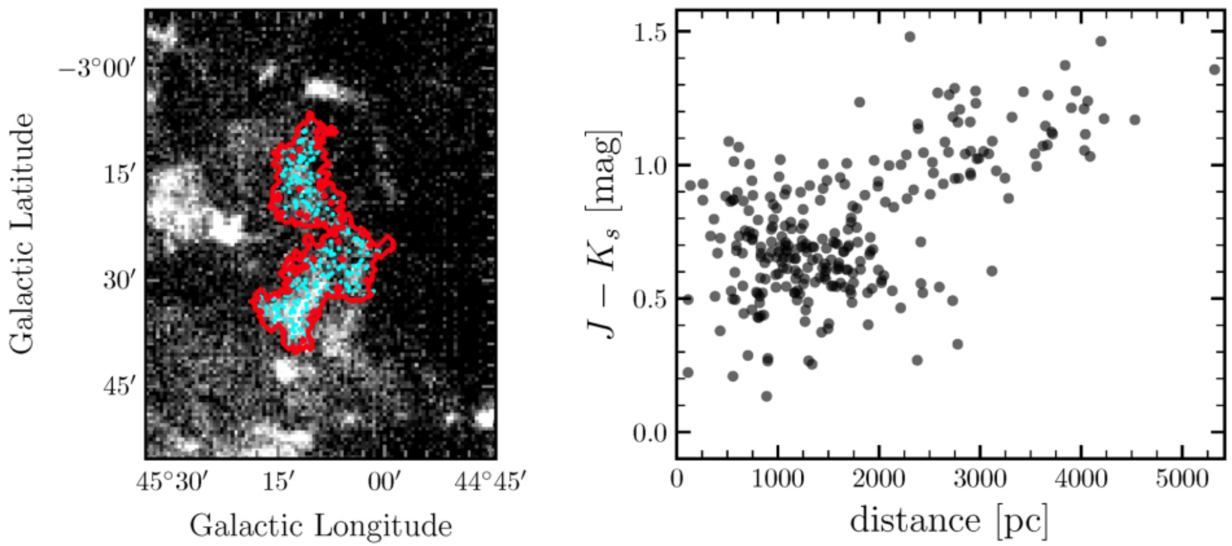


Fig. 1. Left panel: selected on-cloud stars for G045.1-03.4. The red contour shows the molecular cloud footprint, and the cyan points represent on-cloud stars. Right panel: J–K_s color vs. distance diagram. It has two branches.

Synthesized stellar population by the TRILEGAL galaxy model http://stev.oapd.inaf.it/cgi-bin/trilegal

- The distributions of on-cloud stars in the color-distance diagram were compared with the synthetic stellar population (off-cloud), which was adopted from the TRILEGAL galaxy model.
- They find a good match in the color-distance diagram between the 2MASS+Gaia combination toward these cloud-free areas and the TRILEGAL stellar population
  a. dust extinction trend of $A_V/d = 1$ mag/kpc
  b. The Sun is located at a galactocentric distance of 8700 pc and at a height of 24.2 pc
  c. thin disk, hyperbolic secant along z, $sech^2(0.5z/h_{z,d})$, where hz, d increases with age t ($h_{z,d} = z0(1 + t/t_0)^\alpha$
  d. The thick disk, hyperbolic secant along z, $sech^2(0.5z/h_{z,d})$
  e. The halo is represented by an oblate spheroid with a $r^{1/4}$ density proffle
  f. bulge is modeled as a triaxial structure

Synthesized stellar population by the TRILEGAL galaxy model

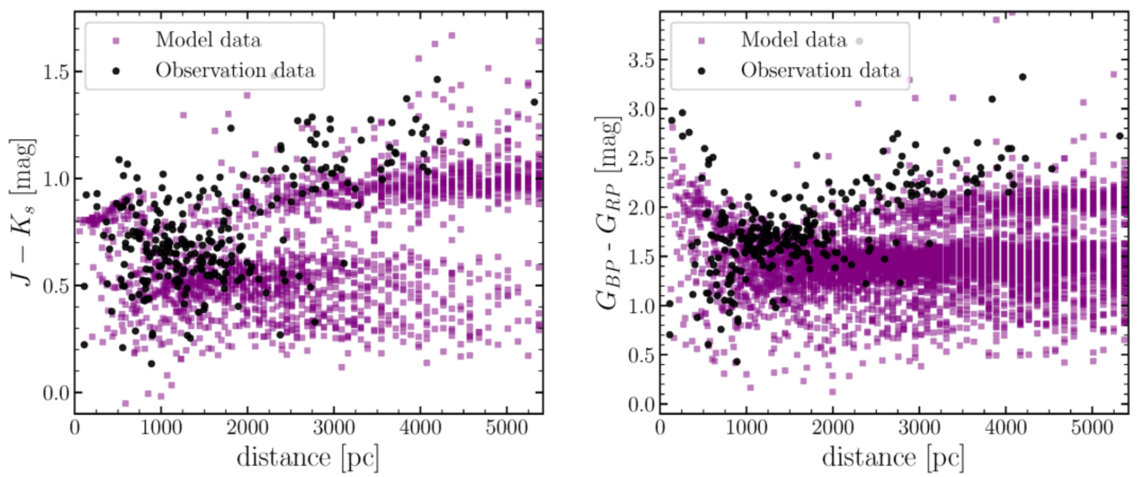
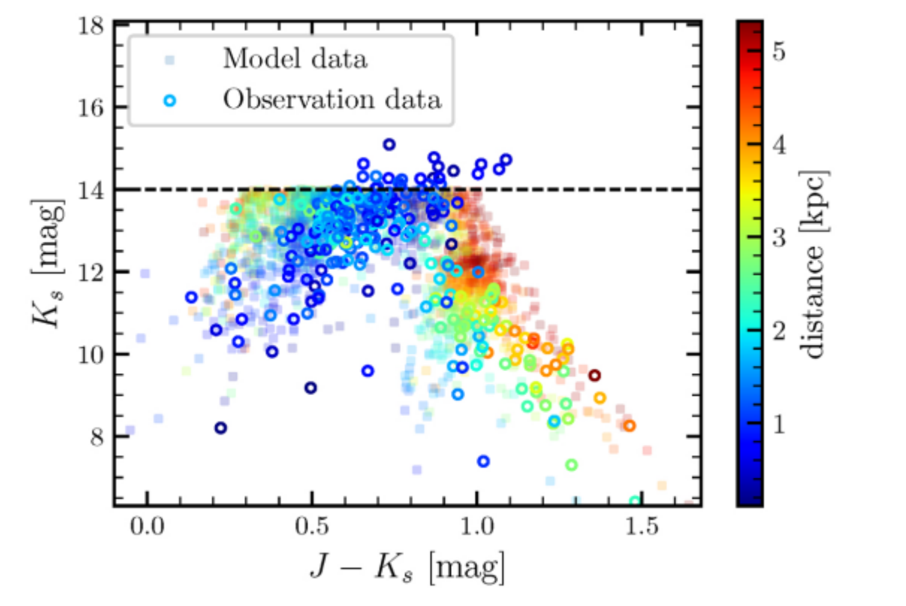
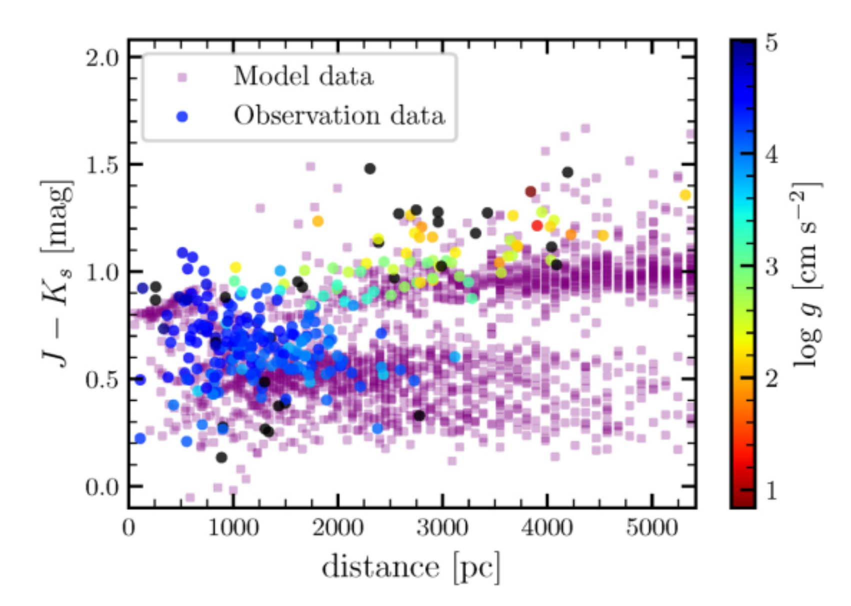


Fig. 2. Left and right panels: J–K_s color vs. distance and $G_{\rm BP}-G_{\rm RP}$ color vs. distance diagram for G045.1-03.4, respectively. The purple squares represent the model data, and the black points represent the observation data.

The observed and synthesized stellar populations both start to split into two branches at a distance of ≳1 kpc in the color–distance diagram.
The synthesized stellar population at a distance of ≳1 kpc with J–Ks > 0.8 is dominated by RGs with log g < 3.45, while MSs are generally of bluer colors J–Ks < 0.8 at the same distances.

The log g are from Gaia DR3 (Gaia Collaboration2023; Andrae et al. 2023),





Color baselines of the off-cloud stars

- They classiffed the TRILEGAL stellar population into MSs (log g > 3.45) and RGs (log g < 3.45) based on the modeled log g of each simulated star.
- Cloud extinction leads to higher J–Ks colors for MSs, resulting in the degeneracy between reddened MSs and RGs on colors.

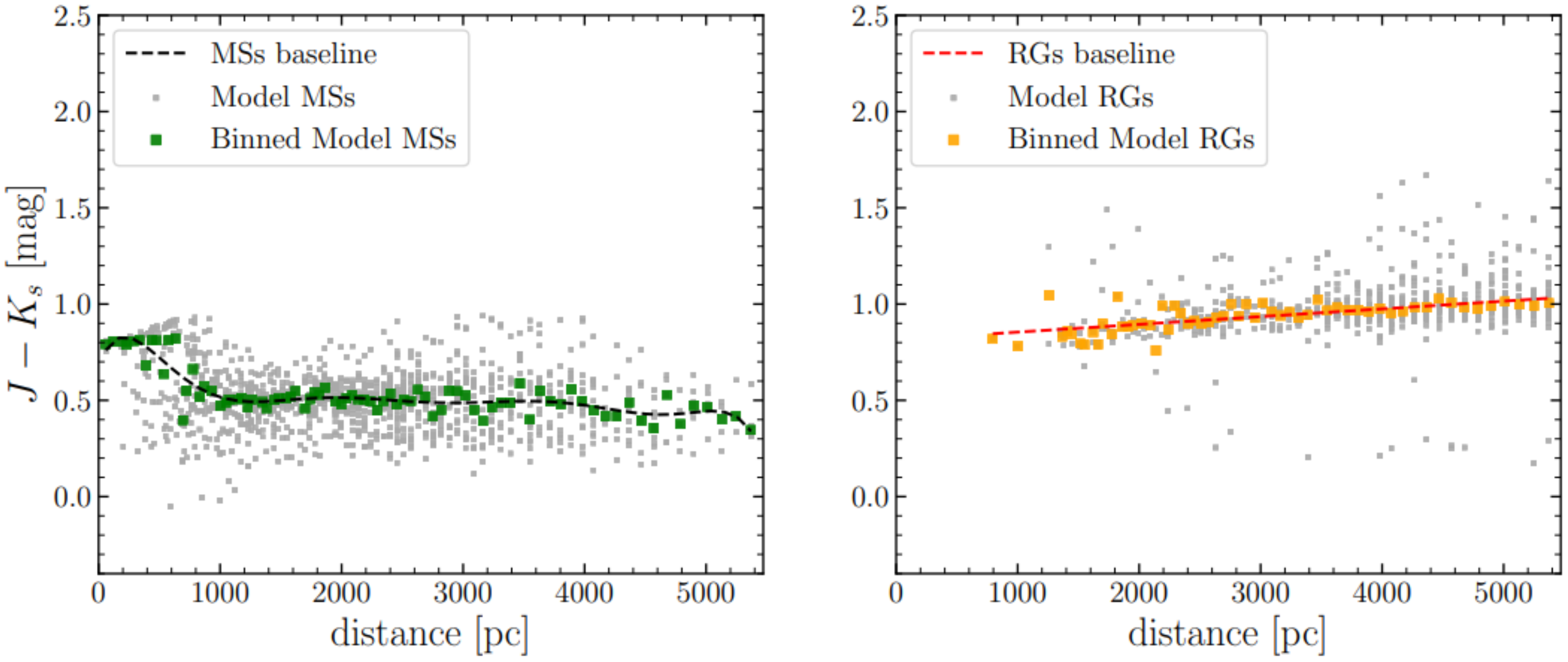


Fig. 5. Left and right panels: fit baselines for MSs (dashed black line) and for RGs (dashed red line) in the model data, respectively. The gray squares present the model data. The green and orange squares present the binned MSs and RGs stars (every 50 pc), respectively.

Color baselines of the off-cloud stars

- The stars with distances greater than 1 kpc and colors exceeding the 3σ lower bound line of RGs baseline are more likely to be RGs,
- Considering the degeneracy between reddened MSs and RGs on colors, they further distinguished reddened MSs and RGs according to the fractions of MSs and RGs in TRILEGAL.

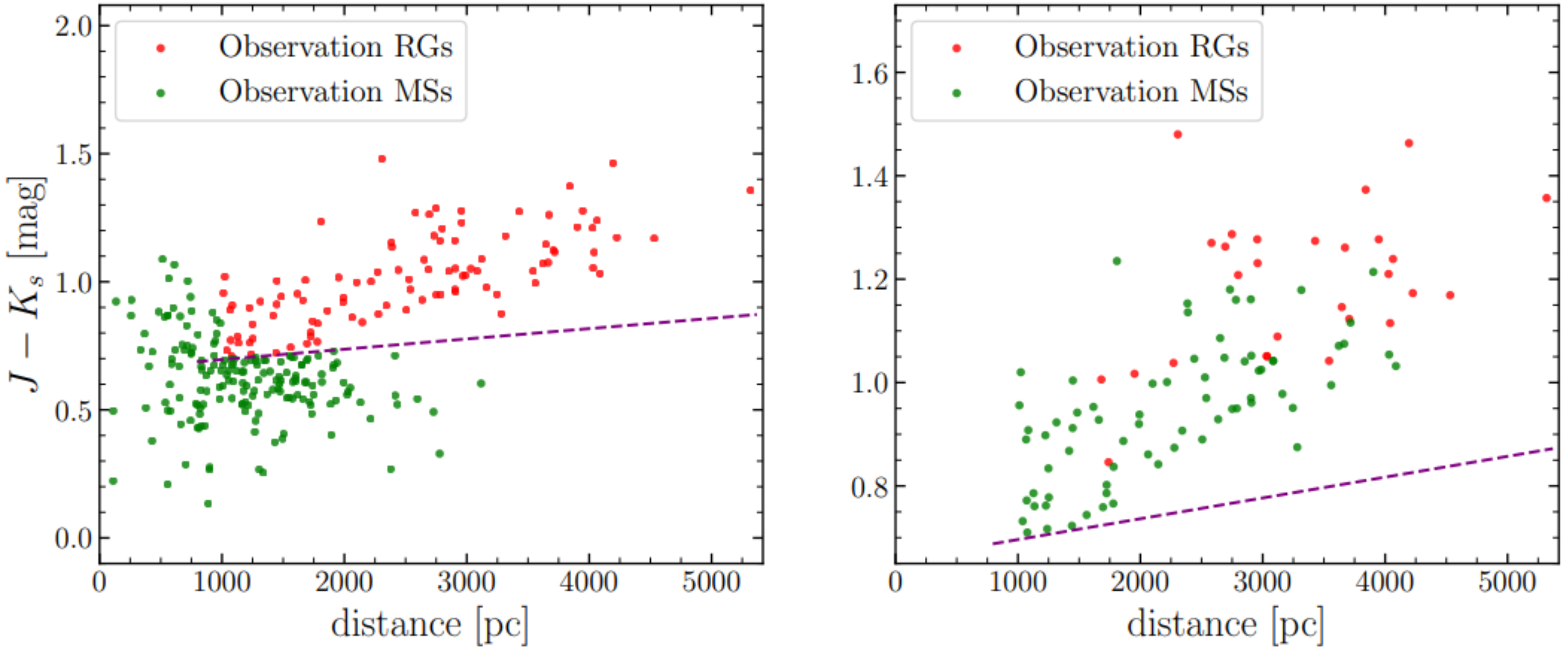


Fig. 6. Left panel: distinguished MSs and RGs in the observation data using the 3σ lower bound line of the model-fit RGs baseline (dashed purple line) and a distance of 1 kpc. Right panel: further distinction between the RGs and reddened stars for the differentiated RGs, and reddened stars are classified as MSs. The green and red points show the MSs and RGs of the observation data, respectively.

Bayesian analysis and Markov chain Monte Carlo sampling

- After subtracting the baseline, the color distribution of on-cloud stars became regular and approximately Gaussian.
- They assumed that the distribution of diffuse dust in the on-cloud region is approximately uniform and that a star is either in front of or behind a molecular cloud.
- The probability that the star is a foreground star is

$$f_i = \phi\left(\frac{D - d_i}{\Delta d_i}\right),$$

$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{\infty}^{x} e^{-t^2/2}$ is a Gaussian function

- The likelihoods of the foreground and background stars are denoted as

$$PF_i = p\left((J - K_s)_i | \mu_1, \sqrt{\sigma_1^2 + \Delta(J - K_s)_i^2}\right),$$

$$PB_i = p\left((J - K_s)_i | \mu_2, \sqrt{\sigma_2^2 + \Delta(J - K_s)_i^2}\right).$$

Therefore, the likelihood of a star is

$$p((J - K_s)_i | D, \mu_1, \sigma_1, \mu_2, \sigma_2) = f_i PF_i + (1 - f_i) PB_i.$$

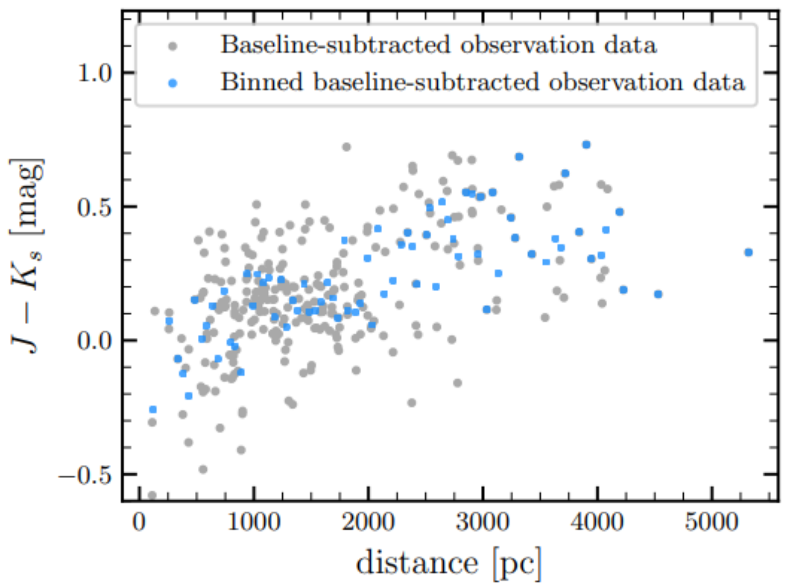
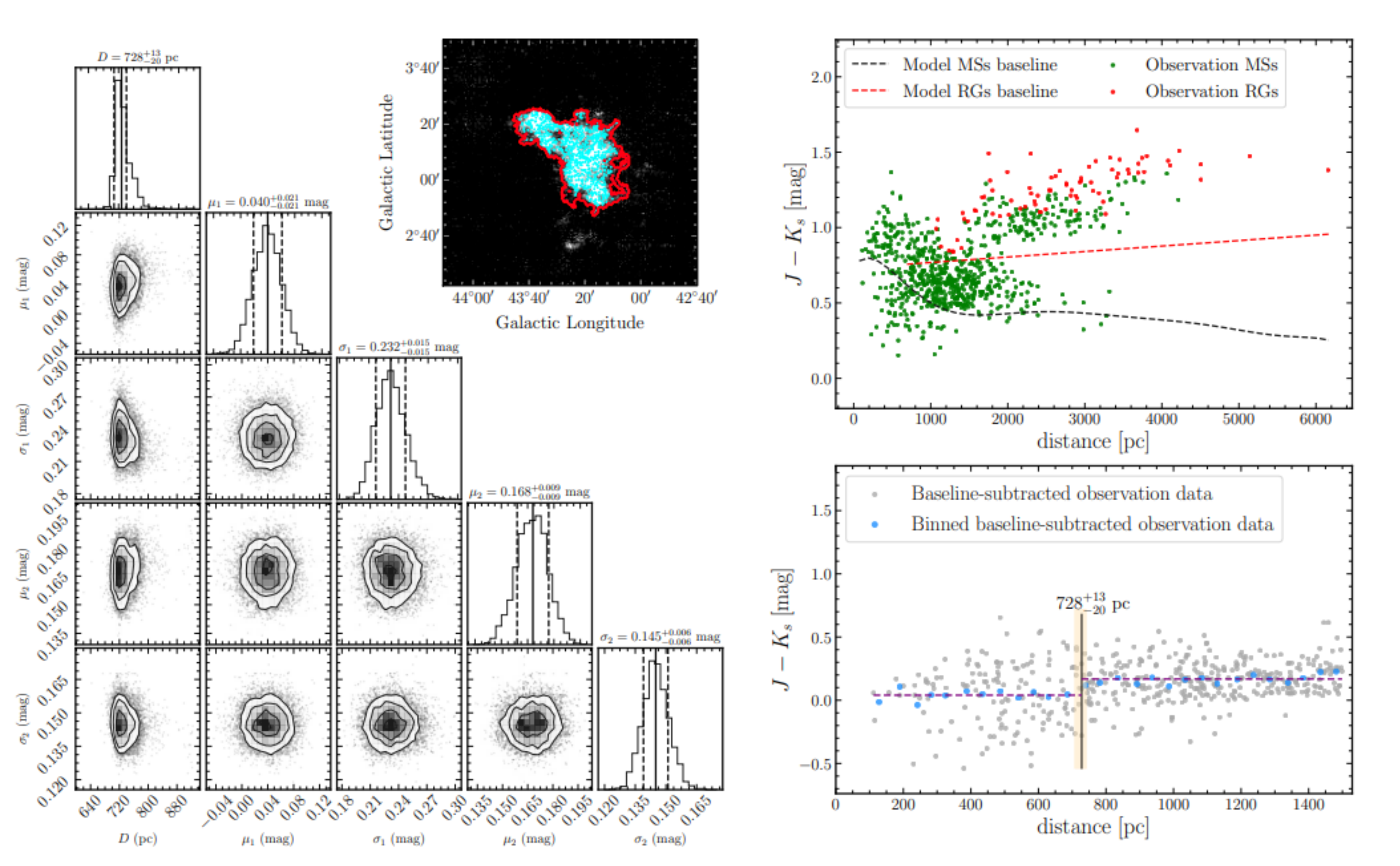


Fig. 7. Distribution of the observation data after the color is calibrated. The gray points present the baseline-subtracted observation data. The blue points plot the mean color of the binned stars (every 50 pc for the distance), which are only used to confirm the results for inspection by eye and not to calculate the distance).

$$P(\mu_1, \sigma_1, \mu_2, \sigma_2) = \begin{cases} 1 & \text{if } \begin{cases} \mu_1 < \mu_2, \\ \sigma_1 > 0, \\ \sigma_2 > 0, \end{cases} \\ 0 & \text{else.} \end{cases}$$

Result and discussion



Effect of the parameter choice

- In the selection of the on-cloud stars, They involved the COcut parameter above which on-cloud stars are selected. The choice of CO_{cut} may affect the determination of the distance.

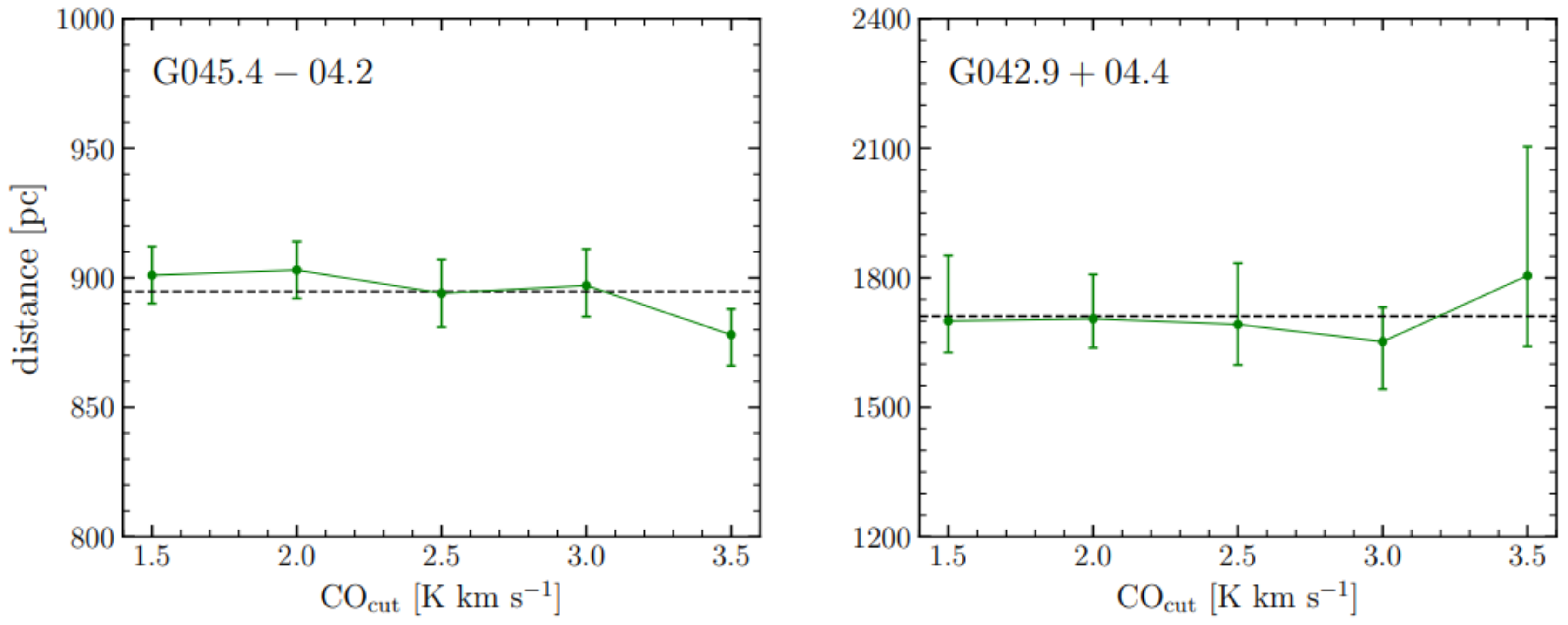


Fig. 10. Molecular cloud distances and uncertainties after changing the CO_{cut} parameter. The dashed black line is the average of the multiple distances. The left and right panels show the effect of CO_{cut} in G045.4-04.2 and G042.9+04.4, respectively. The range of CO_{cut} is $3\sigma_{int}$ to $7\sigma_{int}$, where both clouds σ_{int} are 0.5.

They display the molecular cloud distances and uncertainties after changing CO_{cut}, and the range of CO_{cut} is taken as $3\sigma_{int} \sim 7\sigma_{int}$. In general, COcut has only a weak effect on the distance estimate. Thus, the new method does not depend strongly on the boundary of the cloud.

Uncertainties in the cloud distance and comparison with previous results

- ✓ Uncertainties in the distance and color of the observation data. The systematic uncertainty of the J–Ks color is much smaller than that of the distance, and therefore, They mainly considered the distance error. The adopted stellar distances were estimated by Bailer-Jones et al. (2021), the systematic errors are smaller than 0.1 mas. The systematic uncertainty of the distance caused by the system parallax error is about 10 % at most.
- ✓ Uncertainties in the Milky Way model. As a comparison, we used another Milky Way model, the GALAXIA model. The distances estimated from the two models match each other very well, suggesting that the method proposed in this work does not depend too strongly on the model.
- ✓ Uncertainties in the Bayesian model. Ther Bayesian model is simple. A violation of these assumptions would lead to distance errors.

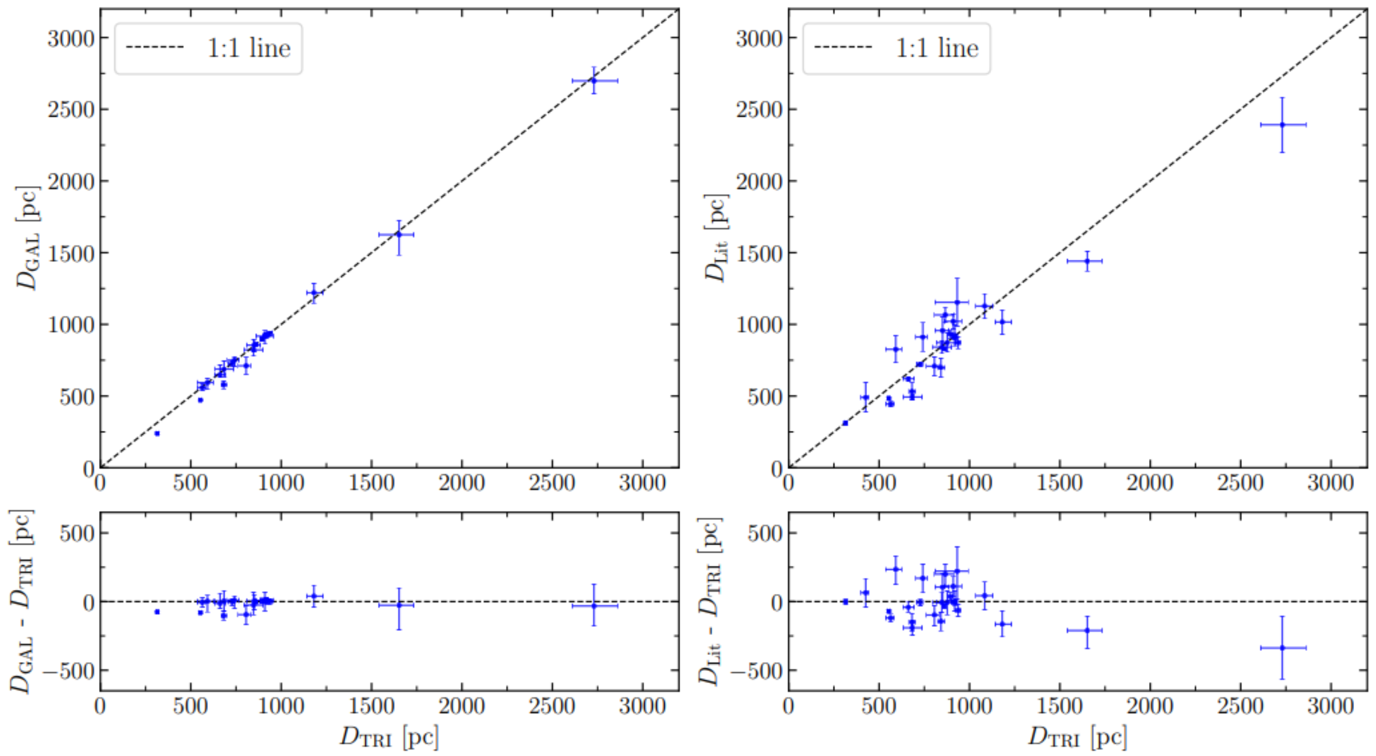


Fig. 11. Left panels: comparison between the distances of 21 clouds estimated from the TRILEGAL (D_{TRI}) and GALAXIA (D_{GAL}) models. Right panels: comparison between the distances of the 27 clouds estimated with the TRILEGAL model (D_{TRI}) and those (D_{Lit}) of Yan et al. (2021).

Limitations of our method

- ✓ It is difffcult to measure molecular clouds at too great distances (>3 kpc) because the data are limited and have large errors.
- ✓ When the number of on-cloud stars is not enough, which may mean that the distance cannot be measured.
- ✓ The reddened stars caused by molecular clouds are not sufffcient to make an obvious color jump. In this case, the distance measured may not be accurate.
- ✓ When two or more molecular clouds overlap along the LOS, multiple jumps become possible. Our method currently does not solve the multijump problem.

Conclusions

- They presented a new method to estimate the distances to 27 molecular clouds using J–Ks colors and the distances provided by 2MASS and Gaia EDR3.
- They adopted the stellar populations generated with the TRILEGAL galaxy model as the baseline instead of off-cloud stars.
- Based on the baseline-subtracted data, we successfully used Bayesian analysis and MCMC sampling to estimate the distances of the selected clouds, ranging from ~315 to ~2730 pc.
- Using the GALAXIA galaxy model, they measured the distances of 21 clouds which are consistent with the distances estimated by TRILEGAL.
- The typical statistical uncertainties of the distances are ~5%, and the systematic uncertainties from the method are ~10%.

Thanks!