Searching for new clusters in the Galactic Plane

Phil Lucas, David Bush, Basmah Riaz, David Samuel, Nithin Shajan, Mark Gallaway, Ben Burningham, Hywell Farnhill, Janet Drew, Mark Thompson
University of Hertfordshire

Dirk Froebrich	Kent
Richard de Grijs	Peking
Jura Borissova	Valparaiso
Dante Minniti	PUC
Andy Adamson	Gemini
Andy Longmore	Edinburgh
Melvin Hoare	Leeds
Simon Clark	Open University
Simon Goodwin	Sheffield

and a few others….

Optical/IR Galactic Plane meeting July 2011

Why search for new clusters?

Star formation

1) Measure variations in the IMF at low mass, e.g. η Cha cluster (Lyo et al.2006)

2) Measure the Cluster Luminosity Function

3) Search for Young Massive Clusters – rare but important

4) Cluster structure as clues to
   i) Initial conditions of cluster formation (Allison, Goodwin et al.2009, 2010)
   ii) the cluster evaporation process

5) Cluster location within Giant Molecular Clouds
   probe propagation of star formation across GMCs, in combination with tracers of imminent star formation (dense clumps, methanol masers)

Why search? (2)

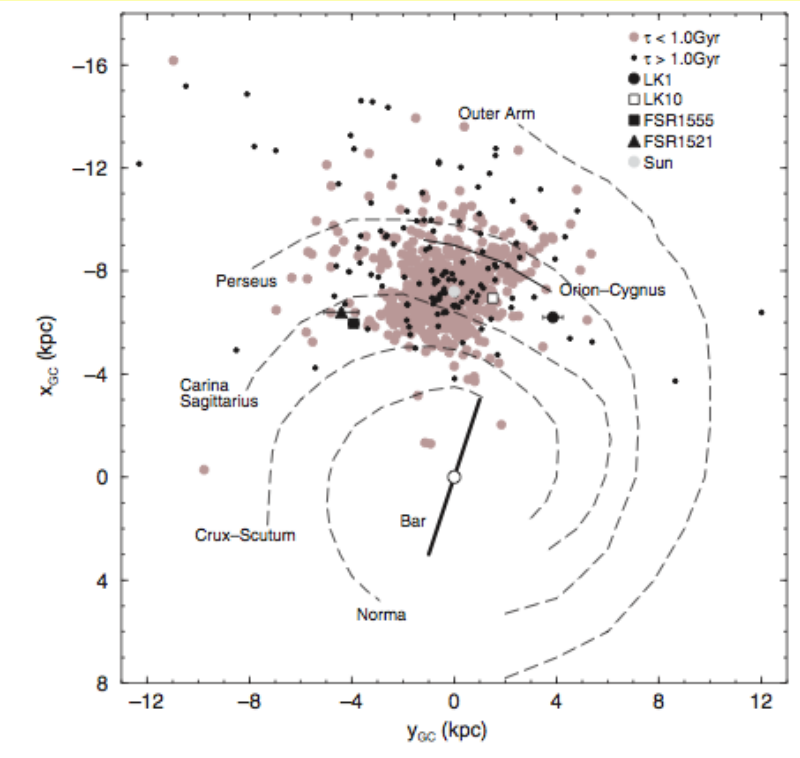
Galactic structure

1) PMS clusters – trace spiral arms and structure within arms
2) Main sequence clusters

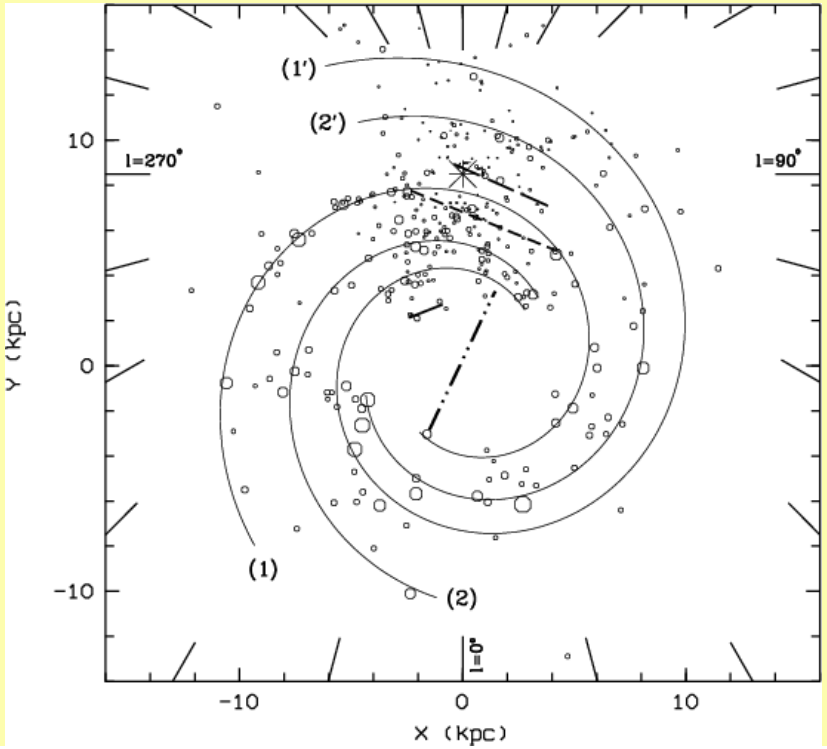
   Another probe of the edge of the Galactic disc (see Sale et al.2010, Saito et al.2011)

   Old metal poor clusters are rare. They provide clues to the origin of the Galactic disc.

Bonatto & Bica 2009, MNRAS



Russeil 2003



Cluster Detection methods

The Milky Way contains >10^4 clusters.
Large lists of clusters have been published by….

1) Bica & Dutra (2003 papers) - an eyeball search of 2MASS at targeted locations.
2) Mercer et al.(2005) using a Bayesian search of Spitzer/GLIMPSE.
3) Froebrich et al.(2007) by searching for statistical over-densities in 2MASS.

We have used several methods to search the UKIDSS GPS. Now starting on VVV.

1) A Bayesian search following Mercer et al.
2) An eyeball search of the jpeg images during quality control.
3) A search for clusters in the vicinity of methanol masers from MMB.
4) An unbiased Spitzer/GPS search for objects with the colours of YSOs,
   (see Basmah's talk tomorrow).

Bayesian search:
Expectation Maximisation (EM) Method

For each field (i.e. each 13 arcminute WFCAM array), split the array into sub-windows with 500 sources and do:

E step - determine ln(L) given cluster parameters

M step - calculate new cluster parameters by applying the current model to the data. The contribution of each star to position, cluster size etc. is weighted by the probability for each star to be a cluster member. This shifts the parameters in the direction of increased likelihood.

Iterate E and M steps to convergance.

Maximise the Bayesian Information Criterion (BIC)

$$BIC = \ln(L) - m \ln(N)$$

Tricks

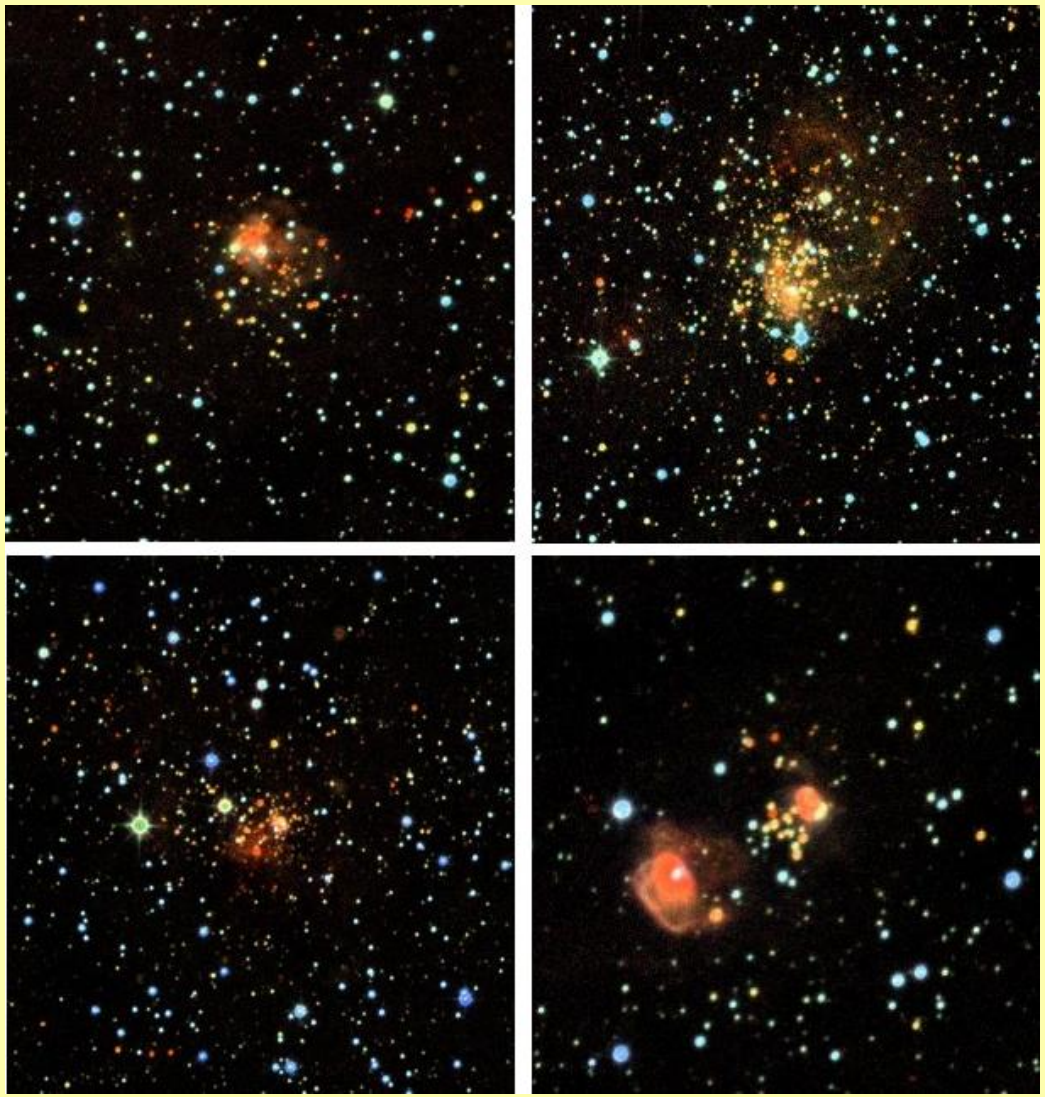
1) Start search with only class = +1 sources from the WSA or VSA. Then run the algorithm on all sources after a cluster candidate has been found.
2) Repeat search with N±1 sub-windows on each axis to deal with edge effects.

Results of Bayesian and visual searches

- We ran David Samuel's MATLAB Code on UKIDSS DR4, using a local copy of gpsSource that contains only FramesetID, RA, Dec, J, H, K, MergedClass for each of the ~555 million sources.
- Searched with K mag < 17, 16, 15, 14, 13
- 316 candidate clusters were confirmed as good candidates, based on visual inspection alone.
- A further 161 candidate clusters were identified by my visual search
  - Required at least 10 stars or extensive nebulosity in cases with no overdensity.
  - The search also detected numerous star formation regions with no cluster.
- Total is 477candidate clusters. ~50%.are new.

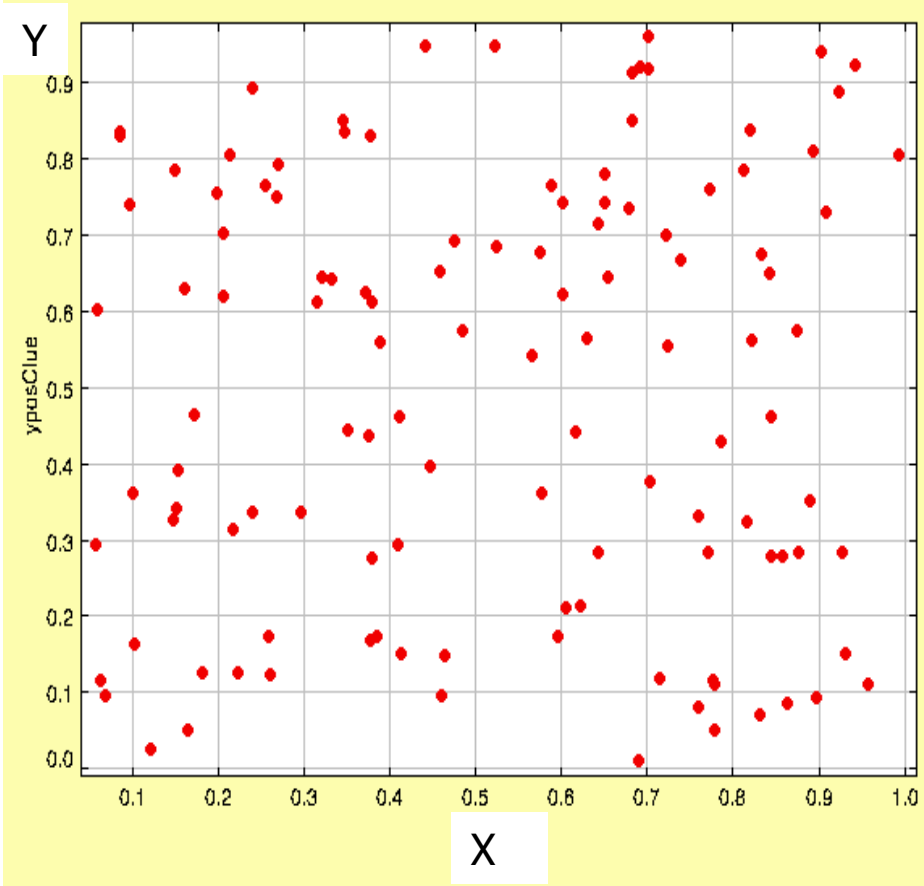
  The new cluster ones show some bias toward small angular size and location in the outer Galaxy.

Examples of new clusters (colour images generated by the WSA)

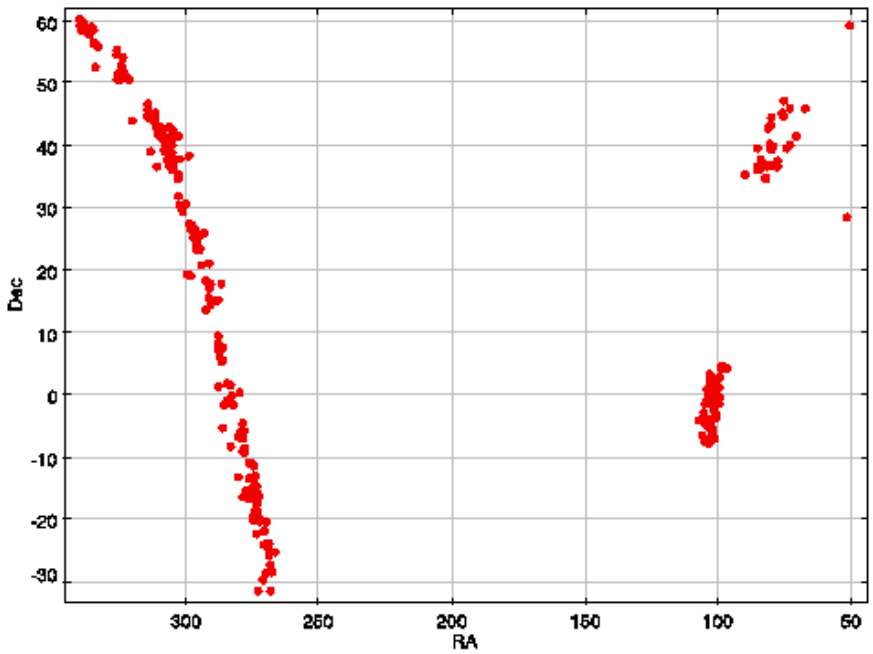


Properties of the Bayesian Cluster Sample(1)

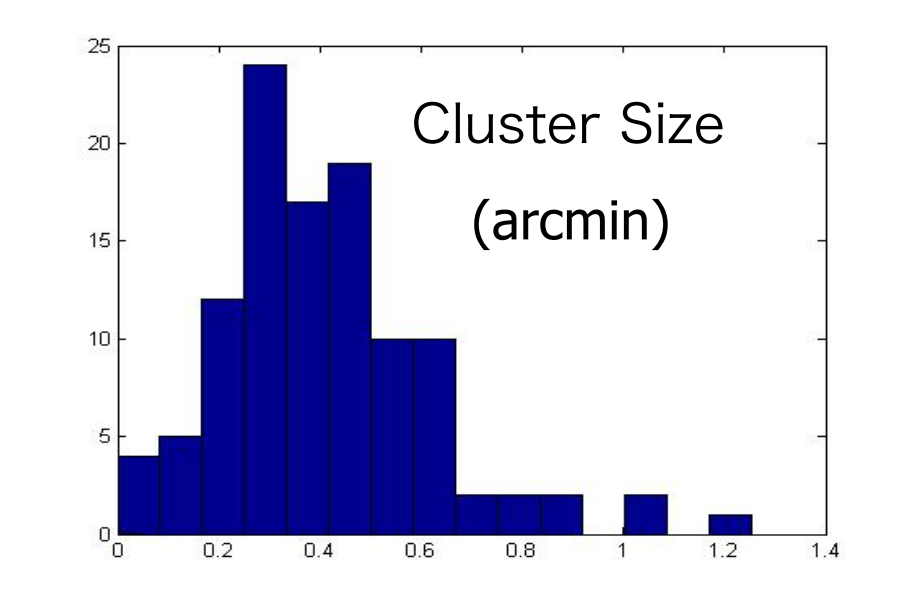
X-Y distribution on WFCAM array

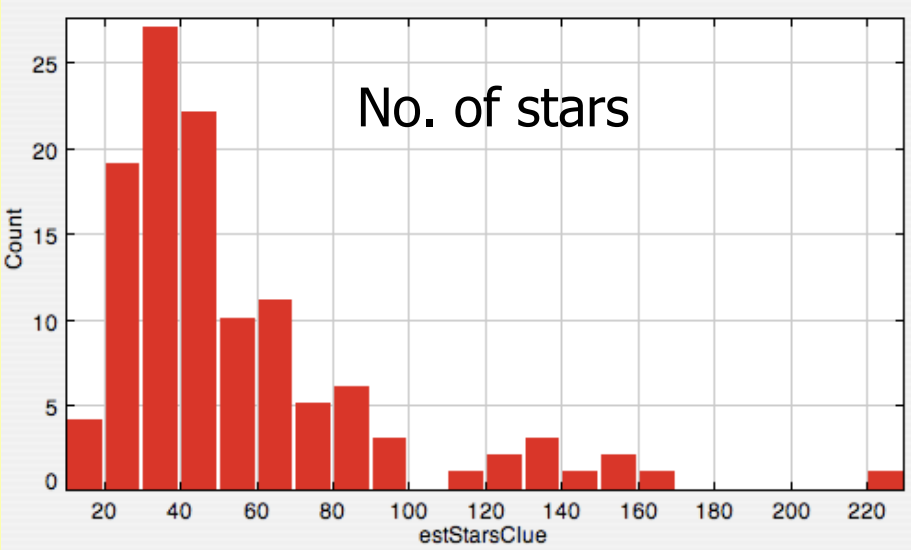


RA & Dec on sky



Properties of the Cluster Sample (2)





Caution: the cluster size given by the Bayesian code may be highly inaccurate. The size results show sensitivity to the initial randomly generated position.

Decontamination

- To confirm the reality of cluster candidates and derive cluster properties we need to remove the unrelated field stars.

- Decontamination programme applied to 246 of the 316 Bayesian clusters with JHK data.
  - Used Dirk Froebrich algorithm:
  - Compute rccm for every star in the cluster area

$$r_{\mathrm{ccm}} = \sqrt{\frac{1}{2}(J^i - J^j)^2 + (JK^i - JK^j)^2 + (JH^i - JH^j)^2}.$$

Calculate 10th smallest value of rccm, called r10_ccm.

Count number of stars, Nccm, in control field within r10_ccm of each star in CCM space.

Then

$$P_{\mathrm{ccm}} = 1.0 - \left(\frac{N_{\mathrm{ccm}}}{10}\right)\left(\frac{A_{\mathrm{cl}}}{A_{\mathrm{con}}}\right)$$

Decontamination (2)

Used a uniform 2 arcmin radius and a 1 arcmin wide control annulus just outside.

Many PMS clusters have high extinction so we also ran a 2D version of the algorithm using only H and K band data.

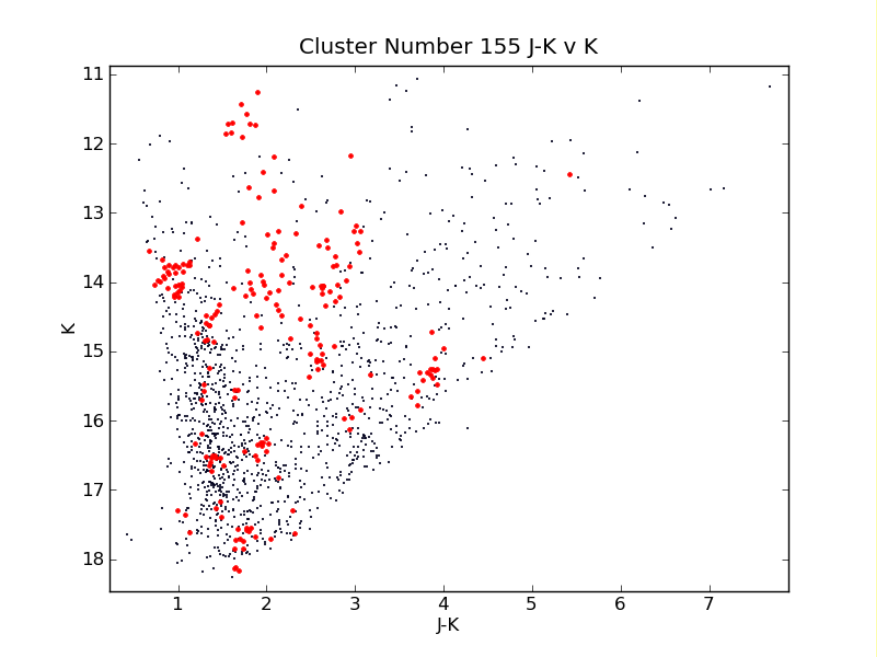
Reality of clusters is determined by inspecting decontaminating diagrams and overplotting high probability members on the FITS images.

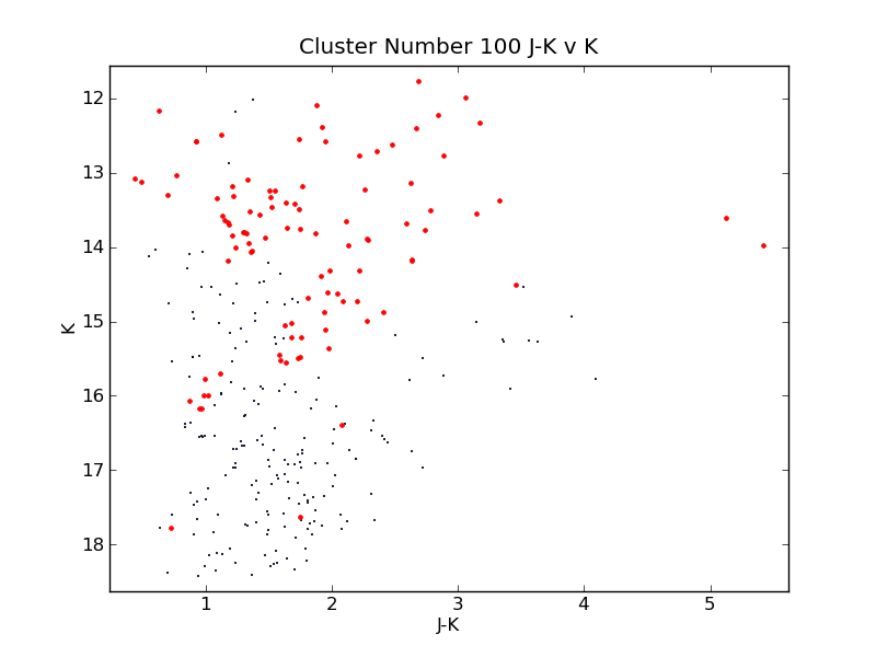
David Bush has been working on the Bayesian sample.

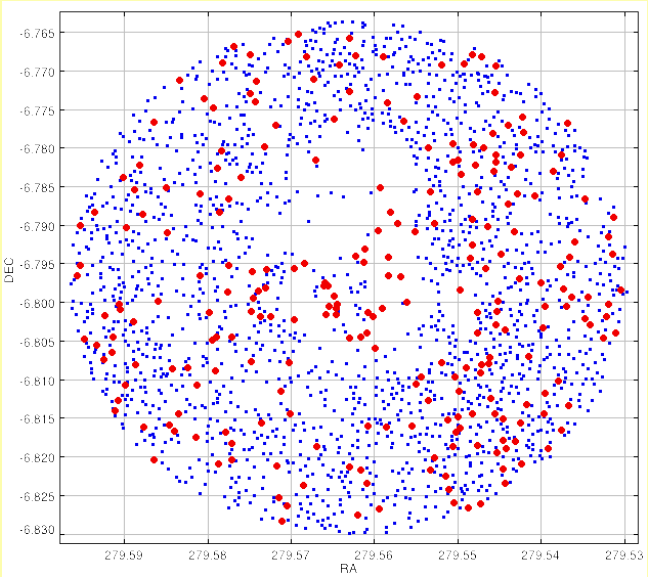
Result: 151/167 new cluster candidates checked so far are real. Some of the rest are still candidates.

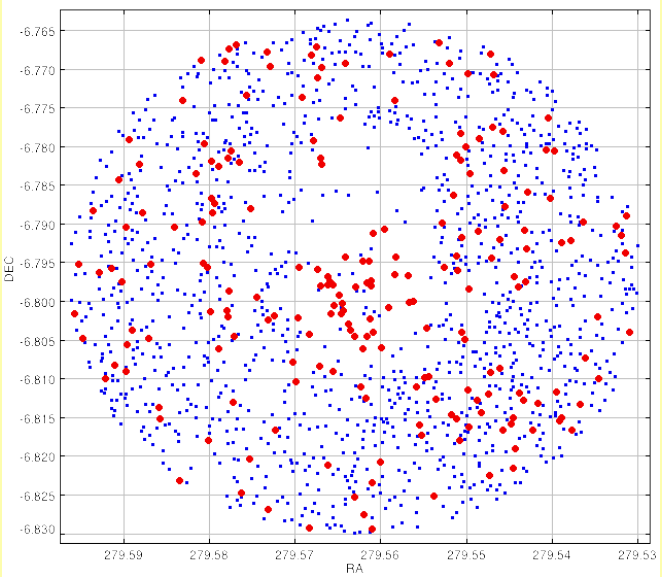
117/167 new cluster candidates are classified as pre-main sequence.

Examples of decontamination: red points have membership probability P > 0.5





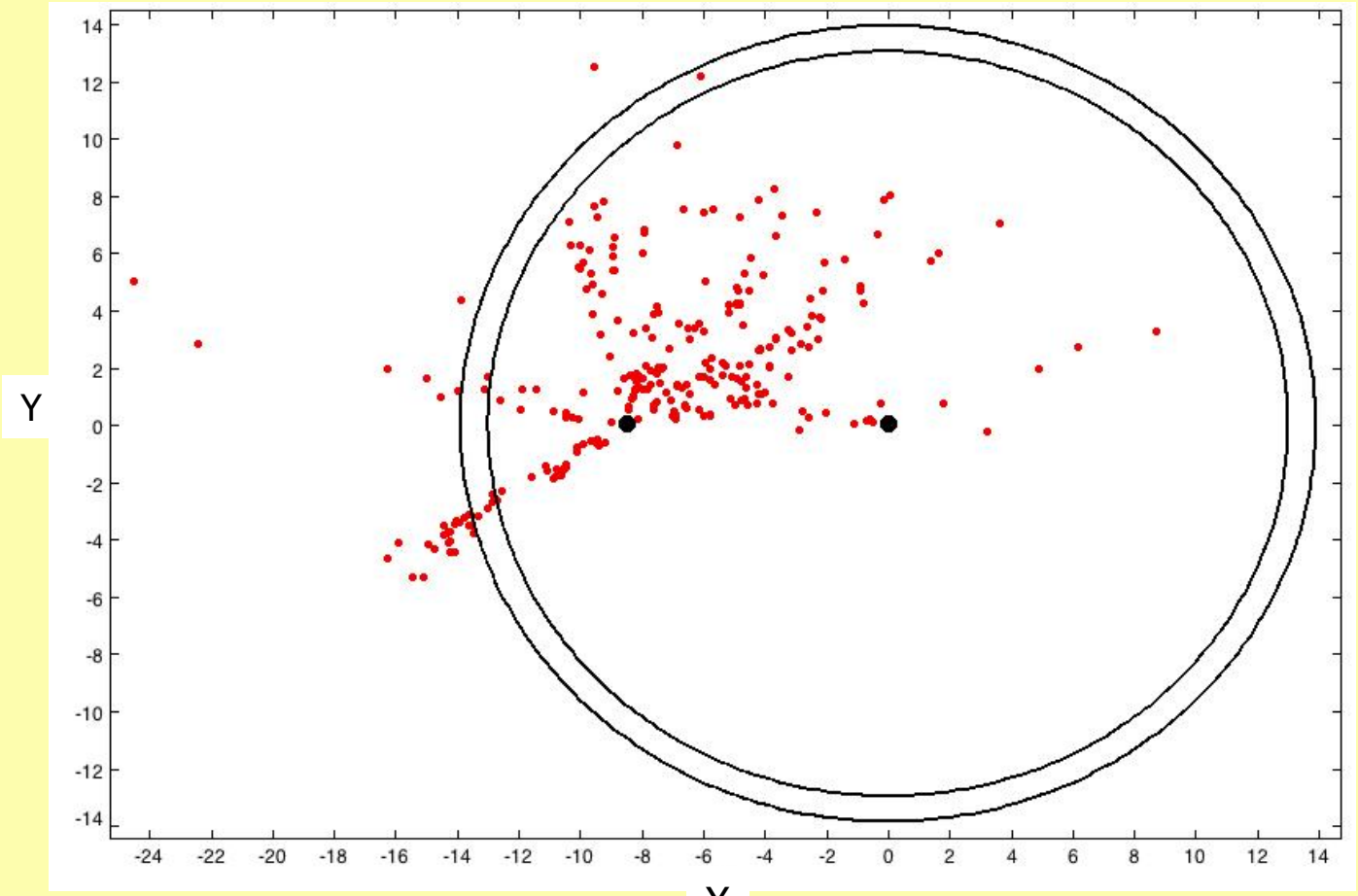




Distance determination

- Distance determination is clearly essential.
- Most clusters have distances in the literature – even new clusters often have kinematic distances to associated molecular clouds.
- 265/477 distances have been taken from “trusted sources”
  - Wouterloot & Brand 1993 & Wouterloot & Brand 1989 (kinematic distances for IRAS-selected molecular clouds in the outer Galaxy) 60”
  - Roman-Duval et al.2009 (kinematic - CO clouds in the BU-FCRAO GRS region) 400”
  - Anderson et al.2009 (kinematic - HII regions in the BU-FCRAO GRS region) 60”
  - Red MSX Survey (RMS) website at Leeds (used a variety of methods) 75”
  - Globular cluster distances from the Hayes catalogue 45”
  - Russeil 2003 (mix of stellar and kinematic distances) 500”
- Of the 212 other clusters, the literature has been checked with Simbad for 106 of them. 54/106 of these now have distances. (Thanks Dirk, Basmah Jura, Richard)
- So it looks like ~373/477 will have distances from the literature.

Galactic location (319 clusters)



More distances

- We have tried using red clump giants as standard candles
  - Giants with the same extinction as the cluster should be at the same distance.
  - It doesn't work well with the K vs J-K CMD.
  - It might work better in VVV with the Ks vs. Z-Ks diagram.

- Other approaches
  - Cluster isochrone fitting
  - Spectophotometric distances to the brightest members
  - A star distances from IPHAS

Conclusions and next steps

The GPS (and VVV) detect large numbers of clusters to distances of ~8 kpc

Most clusters can be readily confirmed as real using decontamination procedures

-- but isochrone fitting may be more problematic

-- Spitzer and WISE data should help to estimate cluster ages from the ratio of Class I and Class II sources.

-- GPS 2 epoch proper motions may also help for decontamination

Incompleteness of the search needs to be checked using well studied GMCs.

First paper will be published with cluster locations and distance estimates. More detailed analysis to follow in subsequent papers.