MATLAB支撑的天文研究平台及其应用范例

中科院国家天文台 刘超

太原 2008-11-29

VO Philosophy

- 伽利略的望远镜的启发
- 科学驱动的VO
- VO解放科学家的创造力
  - VO的目标不是让科学研究简单化
  - VO的目标是使天文学家具有更加强大的能力来进行发现
  - VO工具应该成为科学家的画笔而不是模子
- VO要带给科学家更大的自由度

- MATLAB和Astrobox
- 应用
  - 网络访问及数据库应用
  - 并行化的模型拟合
  - 天文图像处理
  - 球面几何的计算和可视化
  - 大数据量的多种统计回归方法
  - 海量多维数据处理和可视化

MATLAB原有环境

- ToolBoxes:
  - Statistics ToolBox
  - Mapping ToolBox
  - Optimize ToolBox
  - ANN ToolBox
  - GA ToolBox
- I/O
  - FITS
  - ASCII
  - HDF5
- Database
  - JDBC
- Network
  - HTTP post/get
  - SOAP
  - FTP
- Java
- 并行处理环境

AstroBox天文工具箱

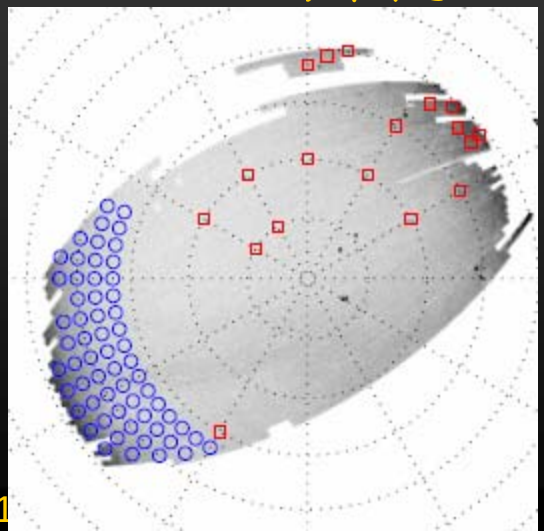
- 单位转换
  - hms2degree, dms2degree, degree2hms, degree2dms
- 坐标转换
  - eq2gal,gal2eq
  - eq2horz, horz2eq
  - UVW2Vr, Vpm2UVW
  - cylin2desc, desc2cylin
  - lambert
  - mjd2date, date2mjd
- 数据操作
  - asciiread,asciiwrite, readgadget, writegadget
  - cut, conecut, maskcut, linecut, polycut
  - bin2D, median2D, DistributeGR, normalize
  - KDensity, abinterp, polyitems
- 数据处理
  - parlsqlm, asregress, tls, lpca
  - aic, kernelreg
  - unirandsph, ngauss
  - locusPoints
  - fitisochrone
  - juricMr, juricStellarlocus
  - ultravioletexcess
  - 。。。
- 数据库应用
  - conesearch
- VO接口
  - plastic
  - readVOT

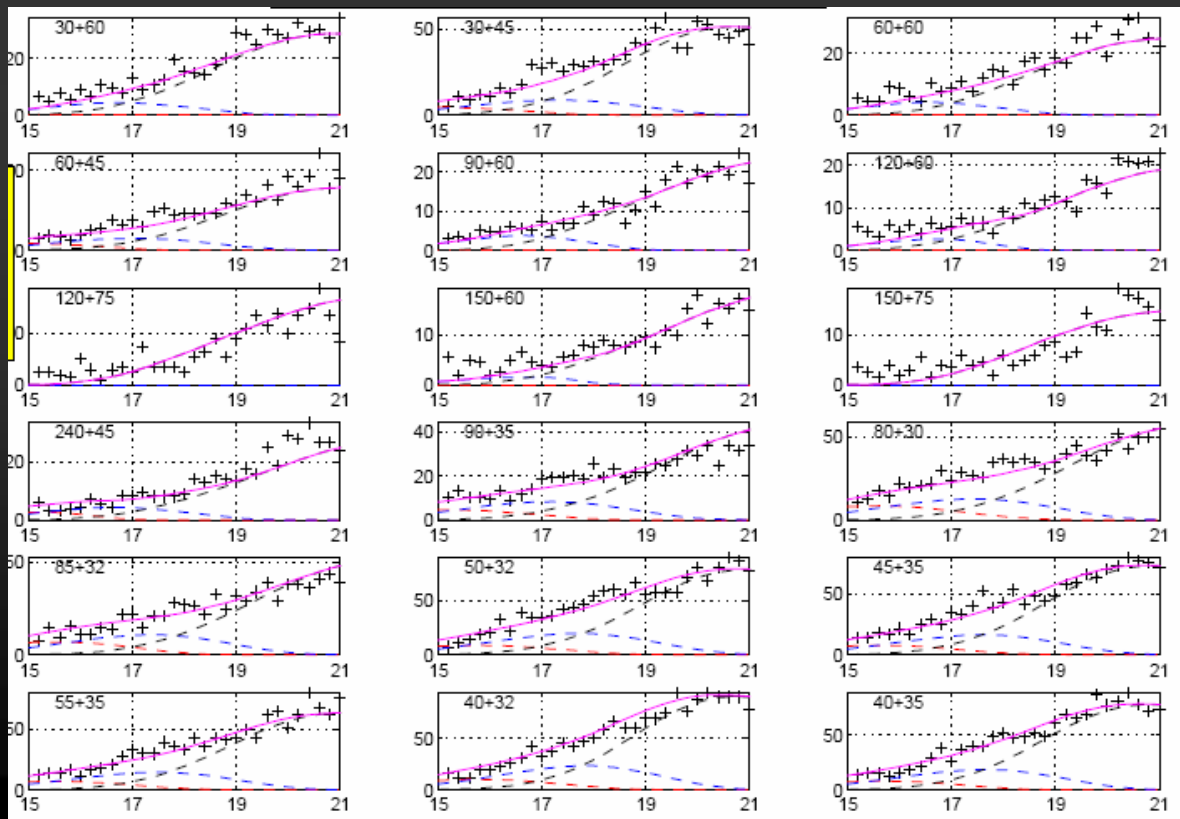
应用一：下载SuperCOSMOS

- step1:应用MATLAB的网络访问功能不断发出HTTP POST指令，通过SuperCOSMOS的网页表单下载压缩的星表数据
- step2:解压星表数据文件，增加HTM索引
- step3:利用MATLAB数据库功能将下载数据插入数据库

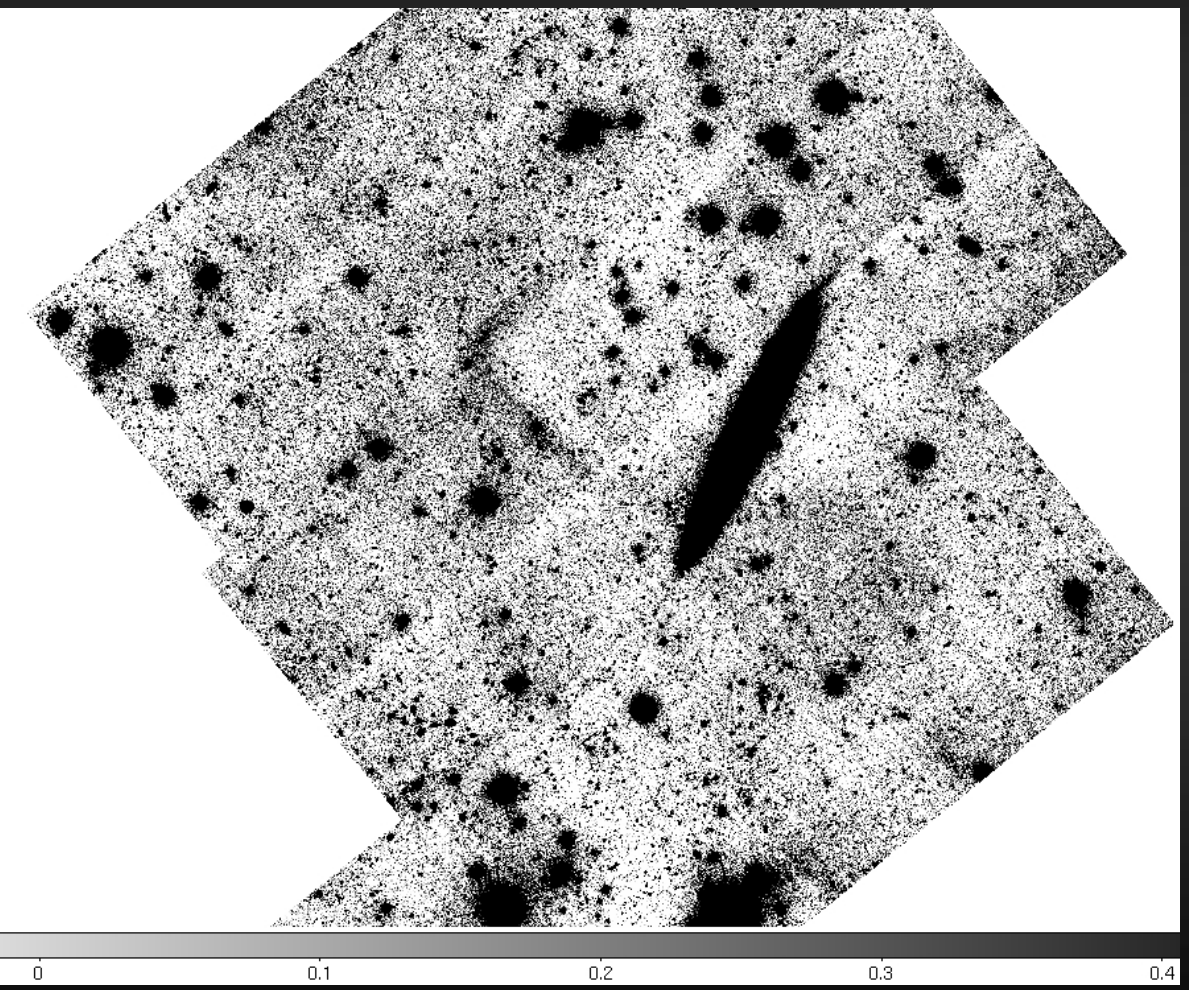
应用二：银河系计数模型

- 用若干视线方向上F型星的计数拟合薄盘、厚盘和晕的特征参数（9个参数）
- 选择天区
- 拟合方法：GA
- 8CPU并行





应用三：寻找极暗的银河系伴星系



NGC5907

合并SDSS多波段的图像（Montage）

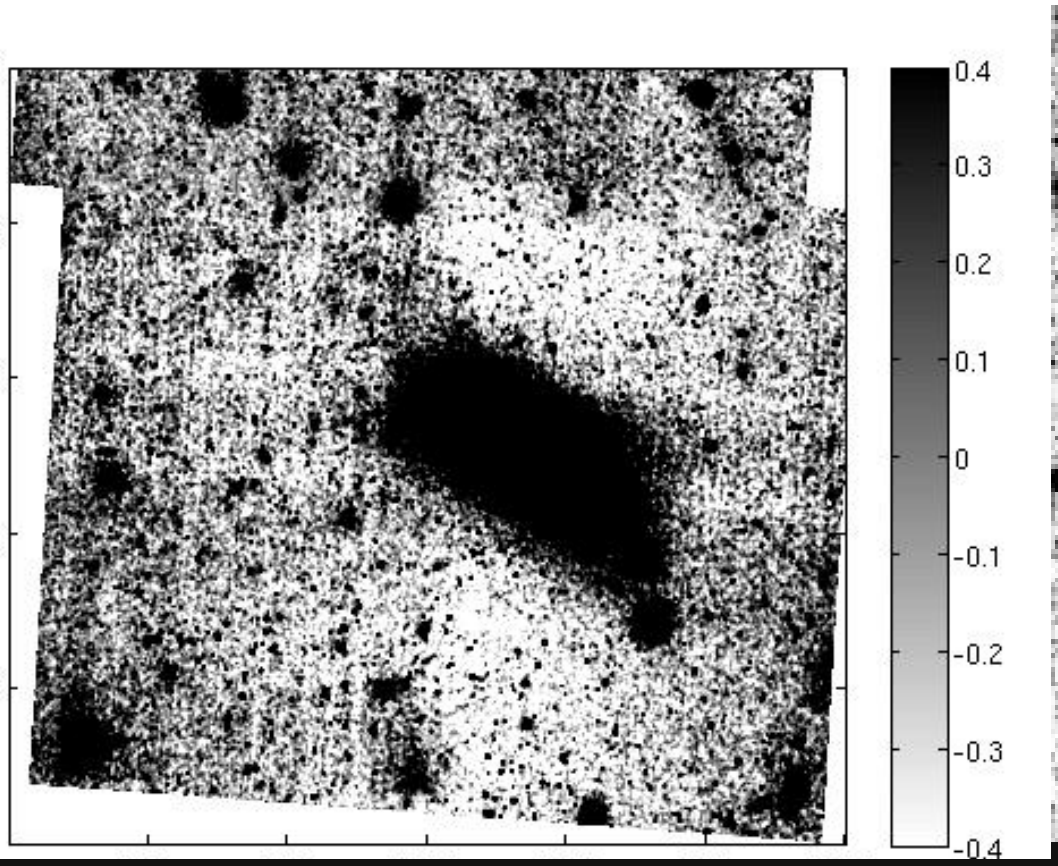
FITS导入Matlab，进行图像增强处理

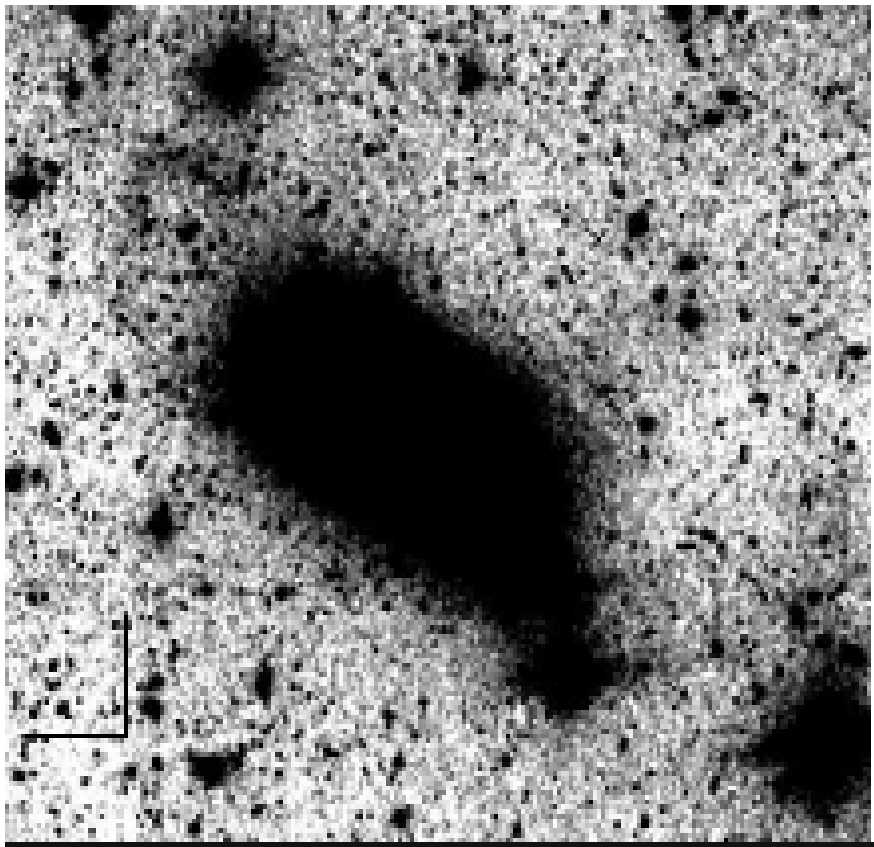
再现面亮度约27mag/sec2（R）的星流

Martinez-Delgado et al. 2008

应用三：寻找极暗的银河系伴星系

NGC4013

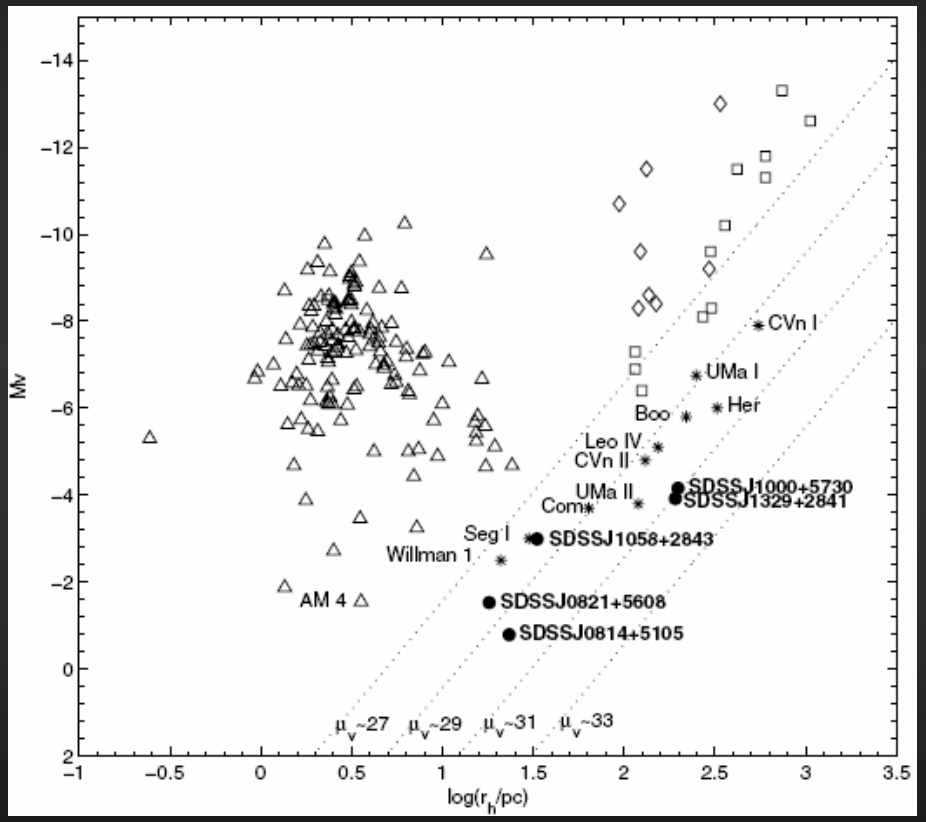




Martinez-Delgado et al. 2008

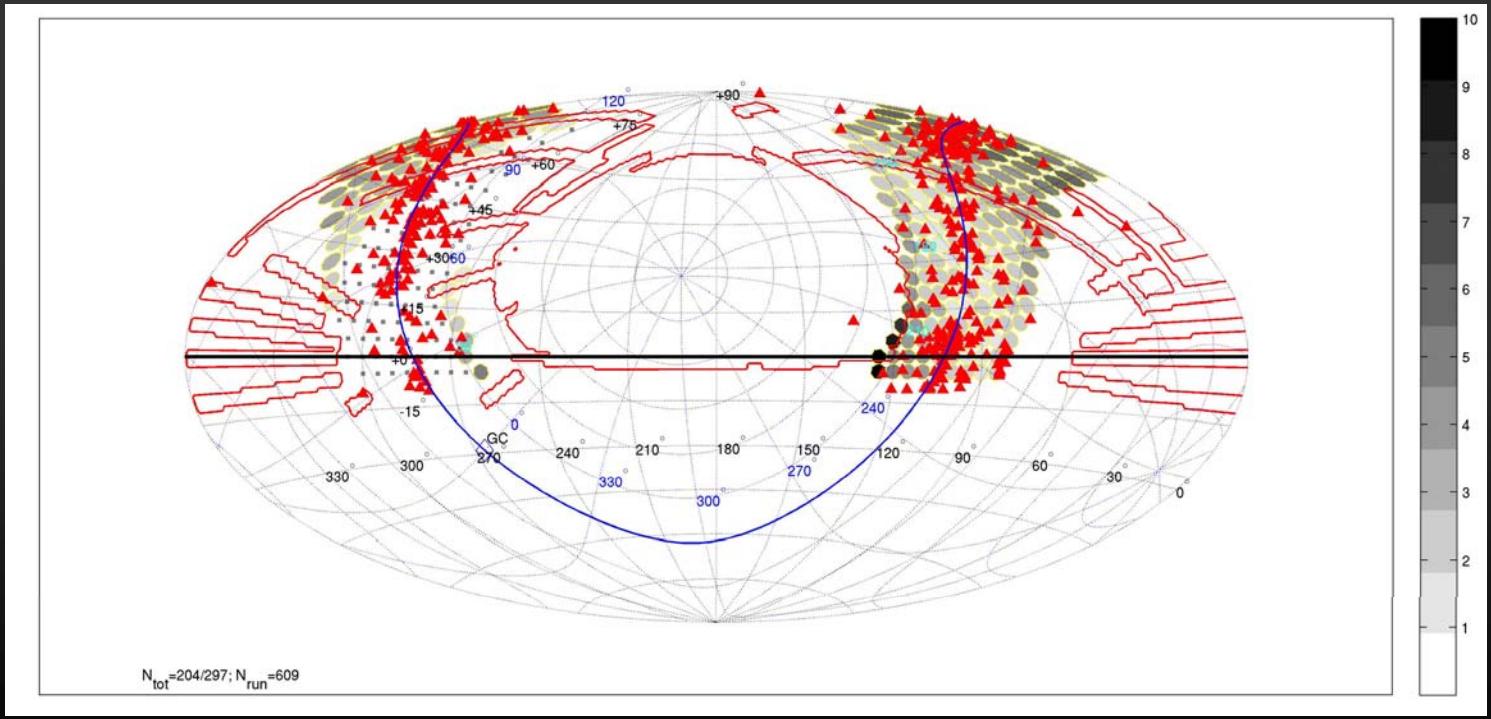
应用三：寻找极暗的银河系伴星系

- 上面两个星流的表面亮度约27mag/sec^2
- 差不多是银河系周围矮星系的亮度上限
- 搜索方案
  - 使用Stripe82的约40次重复观测图像叠加
  - 重点搜索星表中已经发现的overdensity

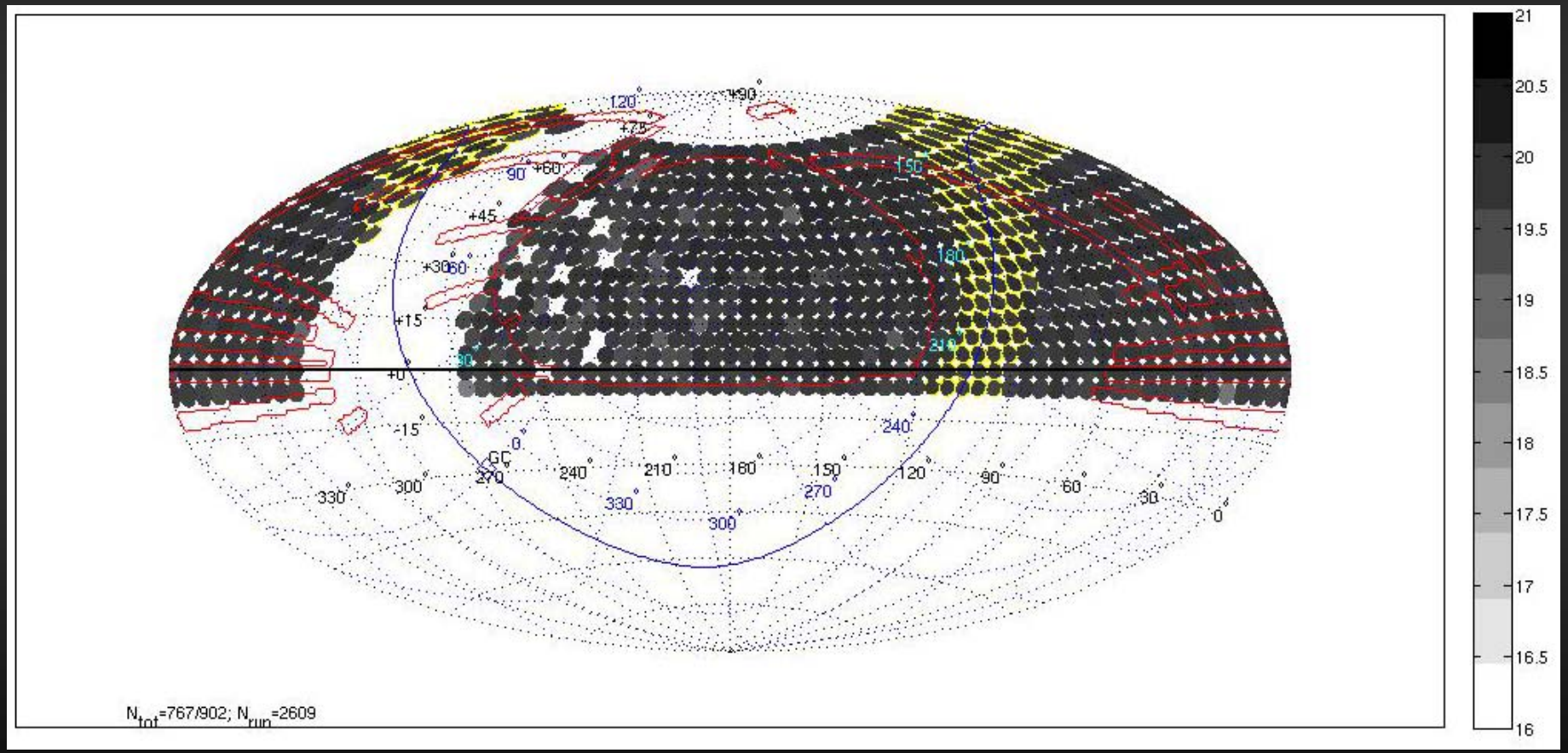


应用四：LAMOST巡天条件研究

- 利用MATLAB Mapping ToolBox提供的球面投影功能完成天气和站址对LAMOST巡天计划影响的研究

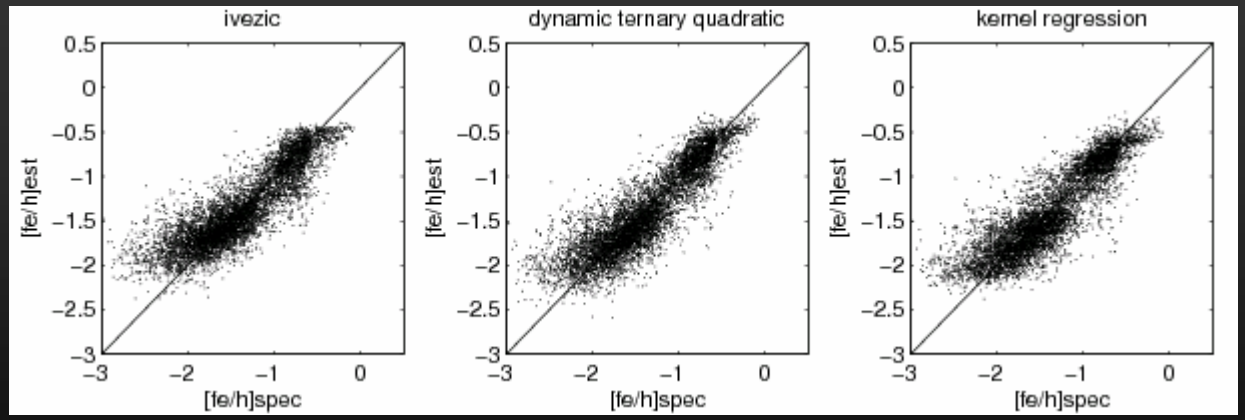


应用四：LAMOST巡天条件研究



应用五：测光金属丰度估计

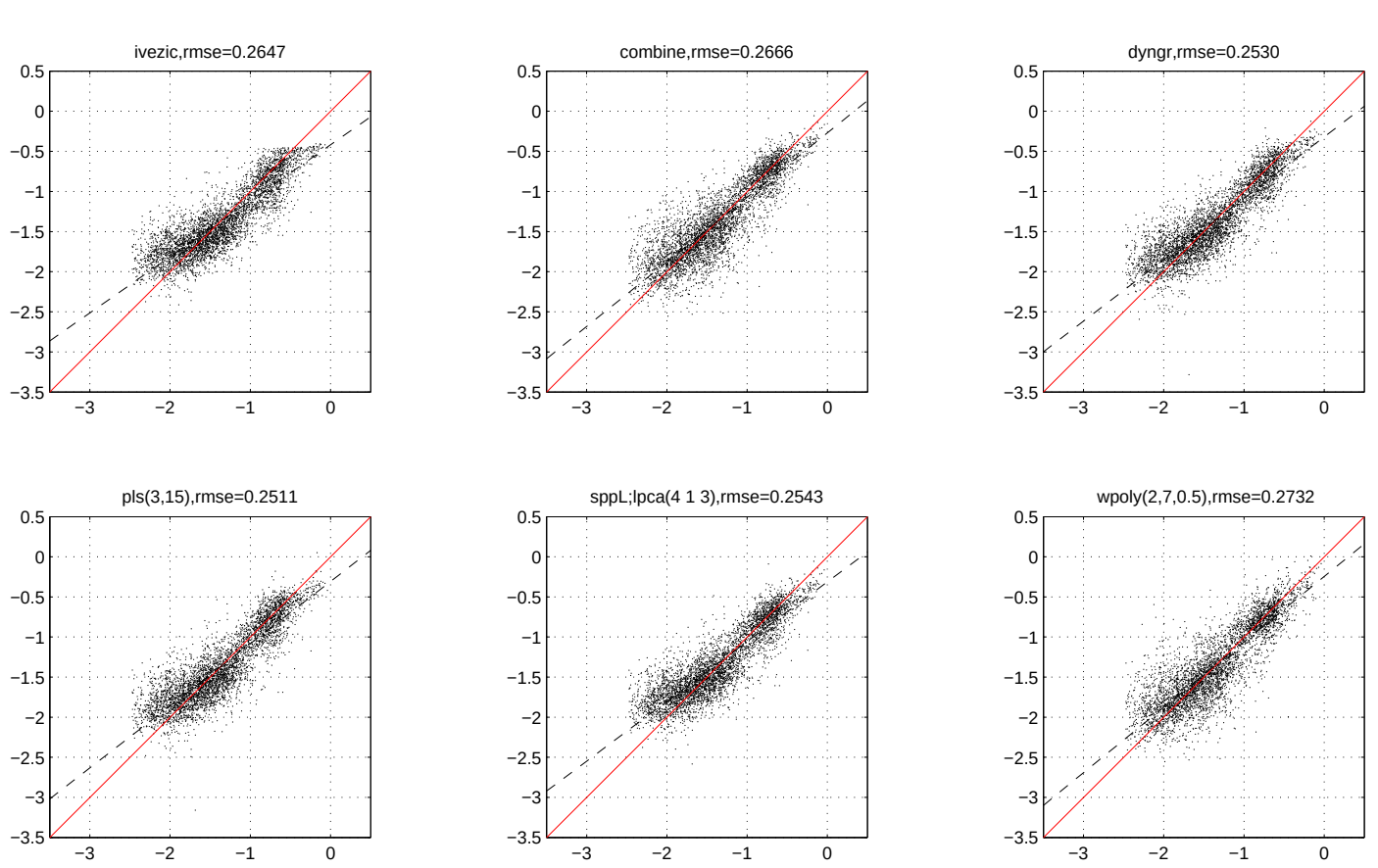
- 利用MATLAB完成恒星金属丰度的多种回归方法的测光估计
  - 多项式回归，分段动态多项式回归，核回归，LPCA，加权的多项式回归，TLS，MARS



Huang, Liu & Yan submitted

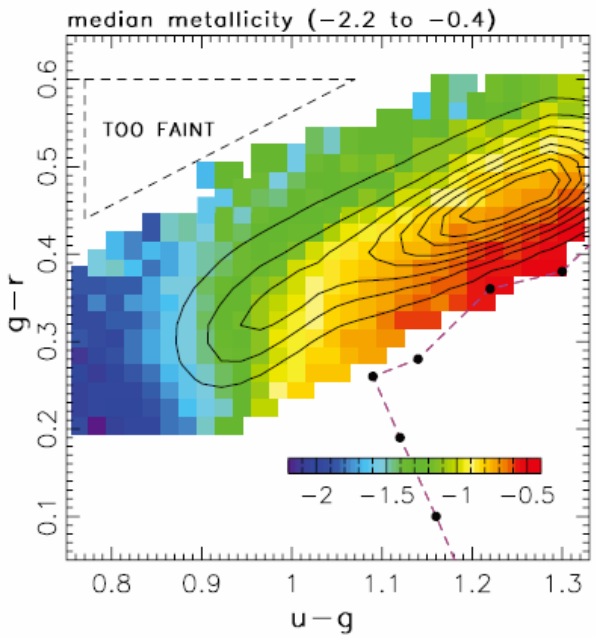
应用五：测光金属丰度估计

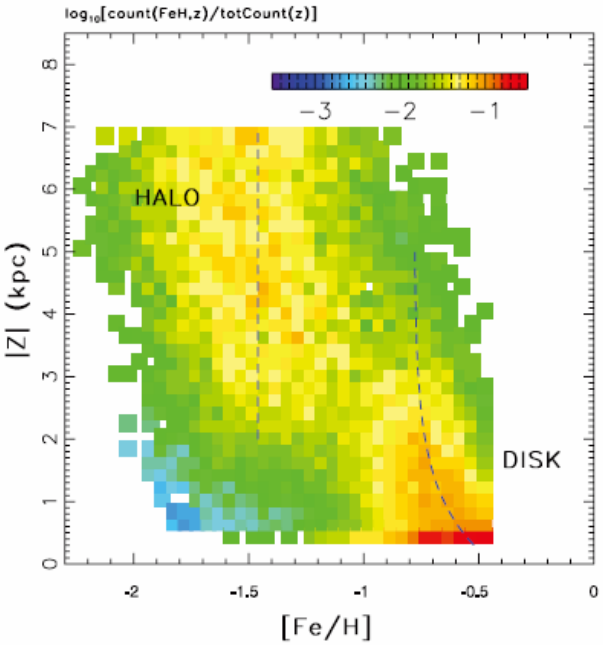
– 并行化参数拟合

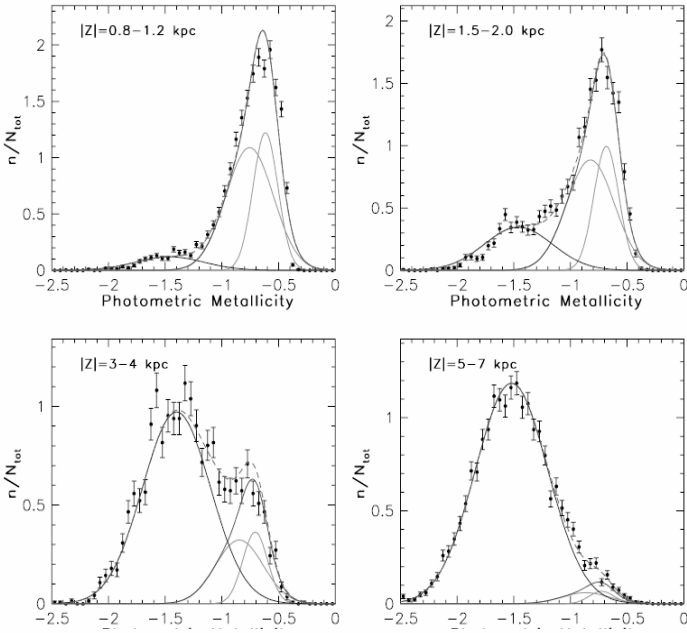


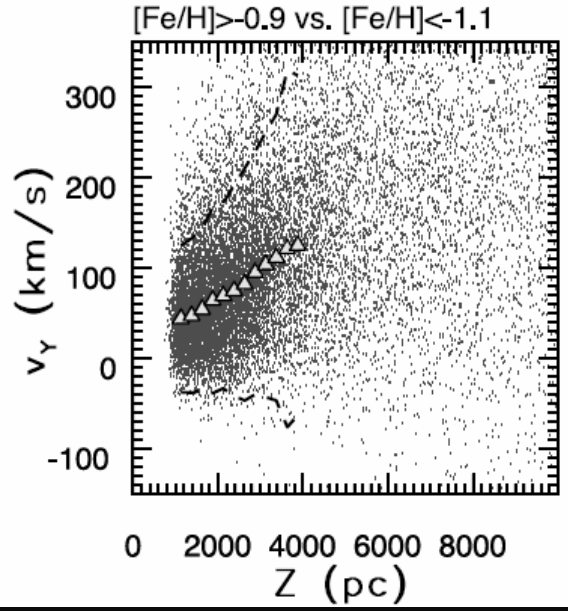
应用六：恒星金属丰度分布和银河系结构

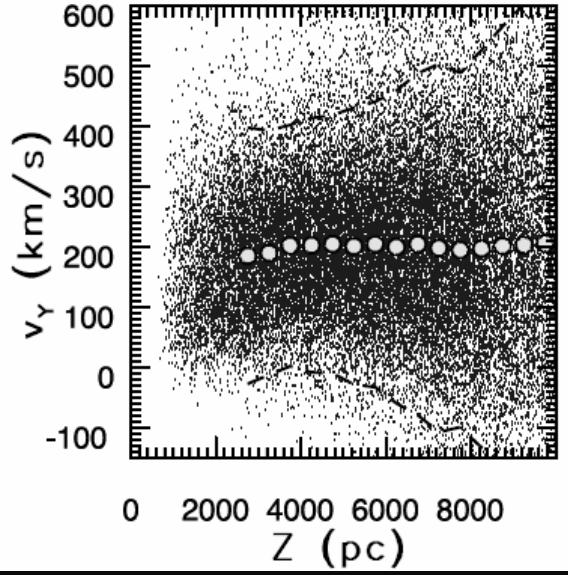
- 背景
  - 测光金属丰度的估计在[Fe/H]=-1.5~-0.5是相对可靠的
  - Ivezic et al. 2008给出了三次多项式回归的估计方法，并用来研究银河系的整体结构
  - Monaco et al. 2007, Chou et al. 2007等发现Sagittarius Tidal Stream上存在金属丰度分布的差异
- 目的
  - 用单位体积恒星的金属丰度分布（MD）来研究银河系的子结构

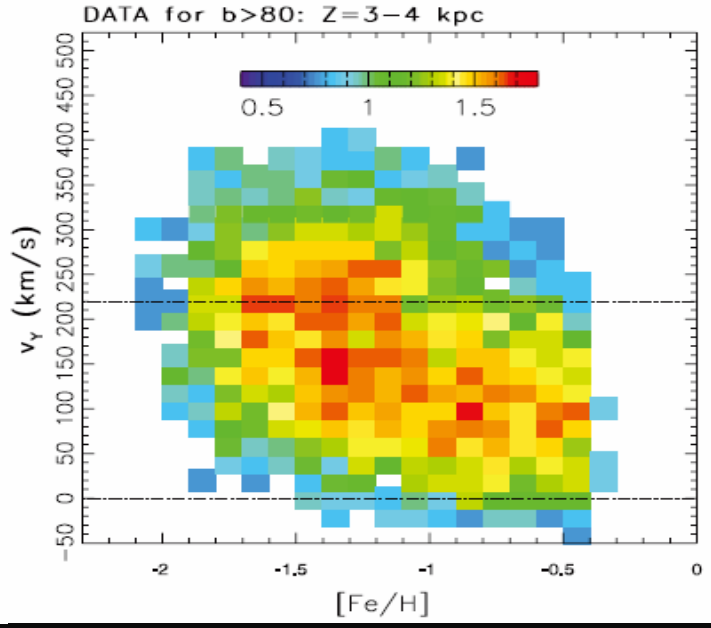

median metallicity (−2.2 to −0.4)
TOO FAINT
−2 −1.5 −1 −0.5
g − r
u − g


log10[count(FeH,z)/totCount(z)]
−3 −2 −1
HALO
DISK
|Z| (kpc)
[Fe/H]

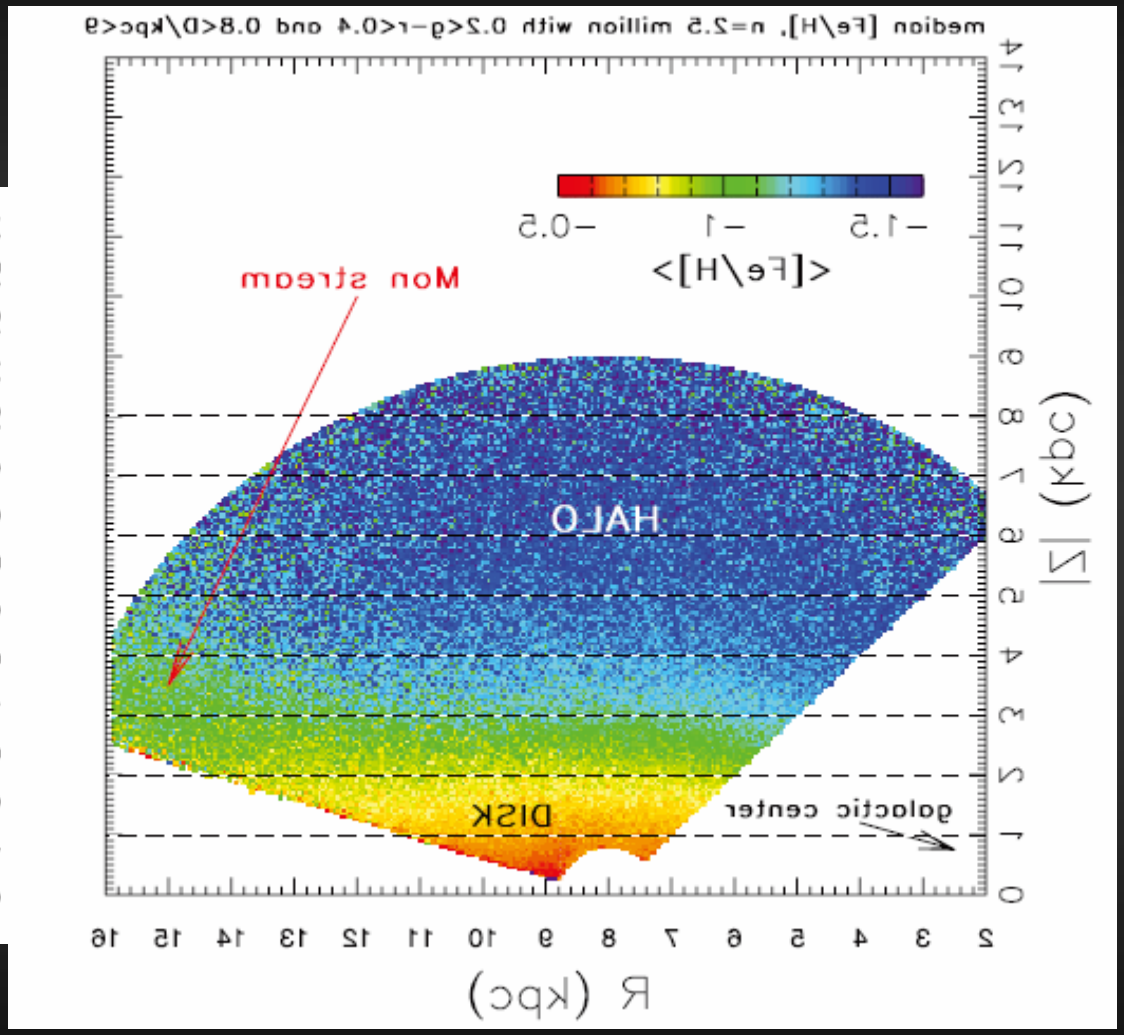

|Z|=0.8−1.2 kpc
|Z|=1.5−2.0 kpc
|Z|=3−4 kpc
|Z|=5−7 kpc
n/Ntot
Photometric Metallicity

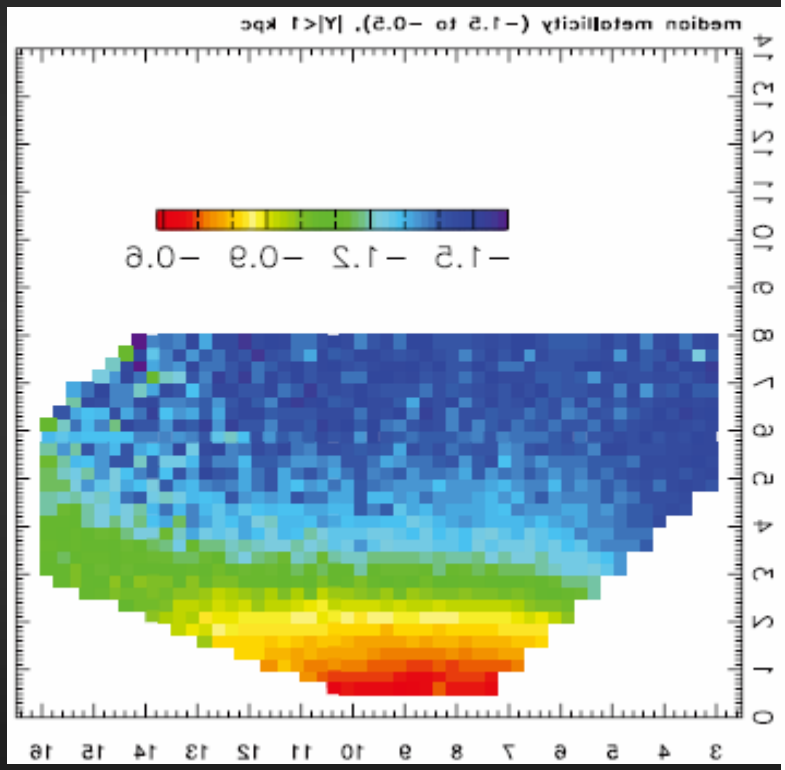

[Fe/H]>-0.9 vs. [Fe/H]<-1.1
vY (km/s)
Z (pc)


vY (km/s)
Z (pc)


DATA for b>80: Z=3−4 kpc
0.5 1 1.5
vY (km/s)
[Fe/H]

~2.5million F/G stars,
(Ivezic et al. 2008)

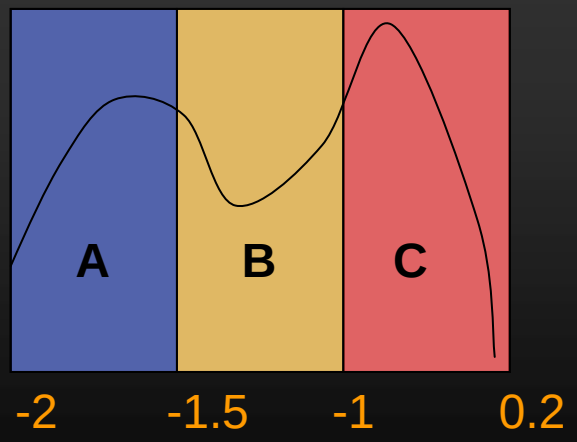


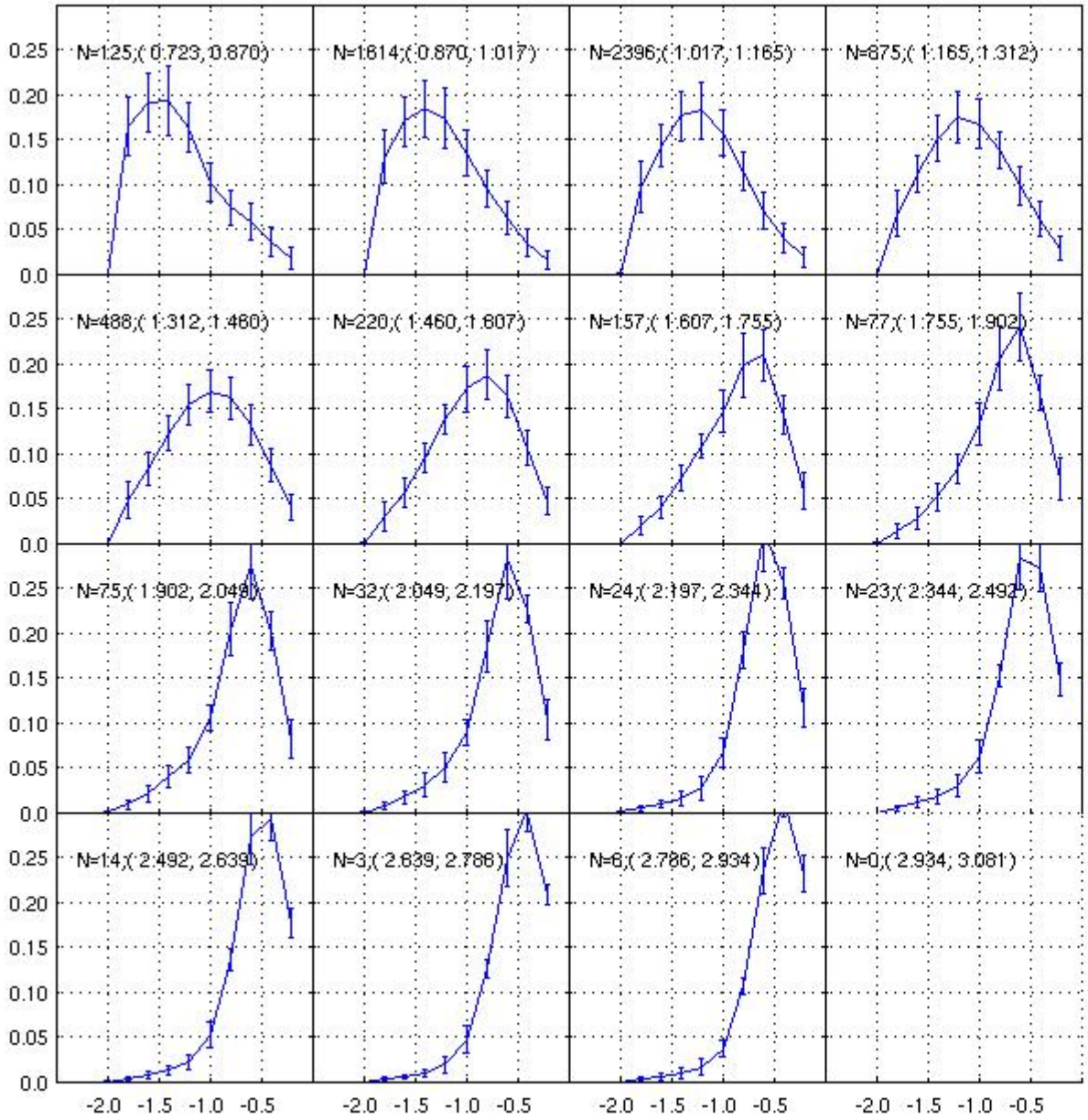


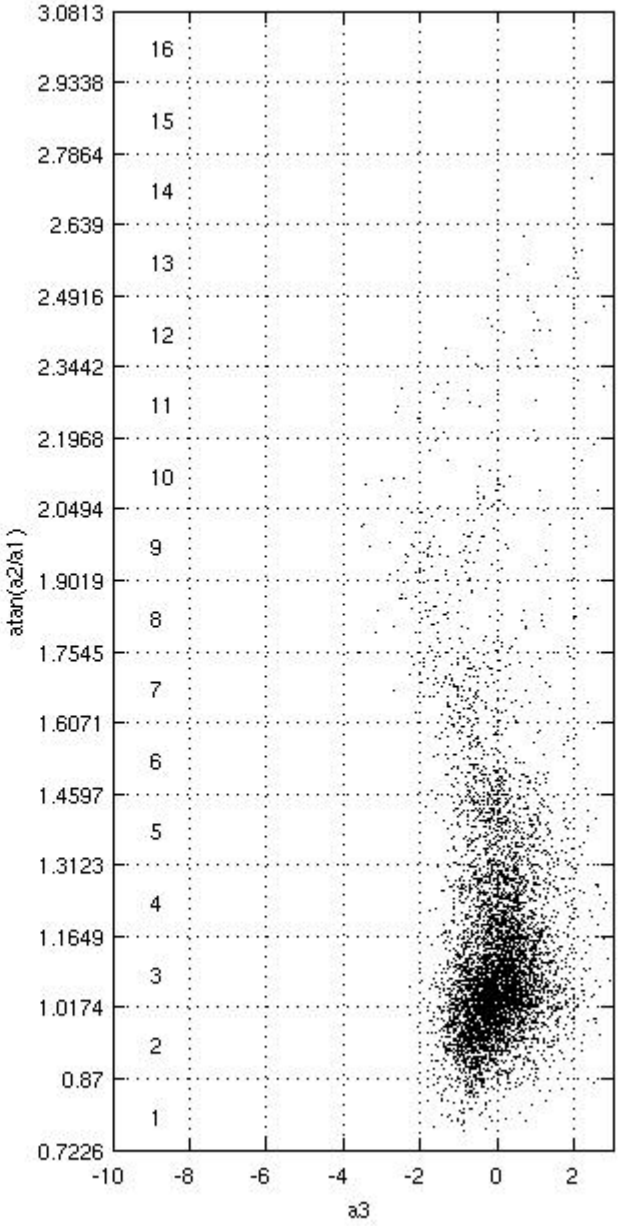
1.04million, |Y|<1kpc

应用六：恒星金属丰度分布和银河系结构

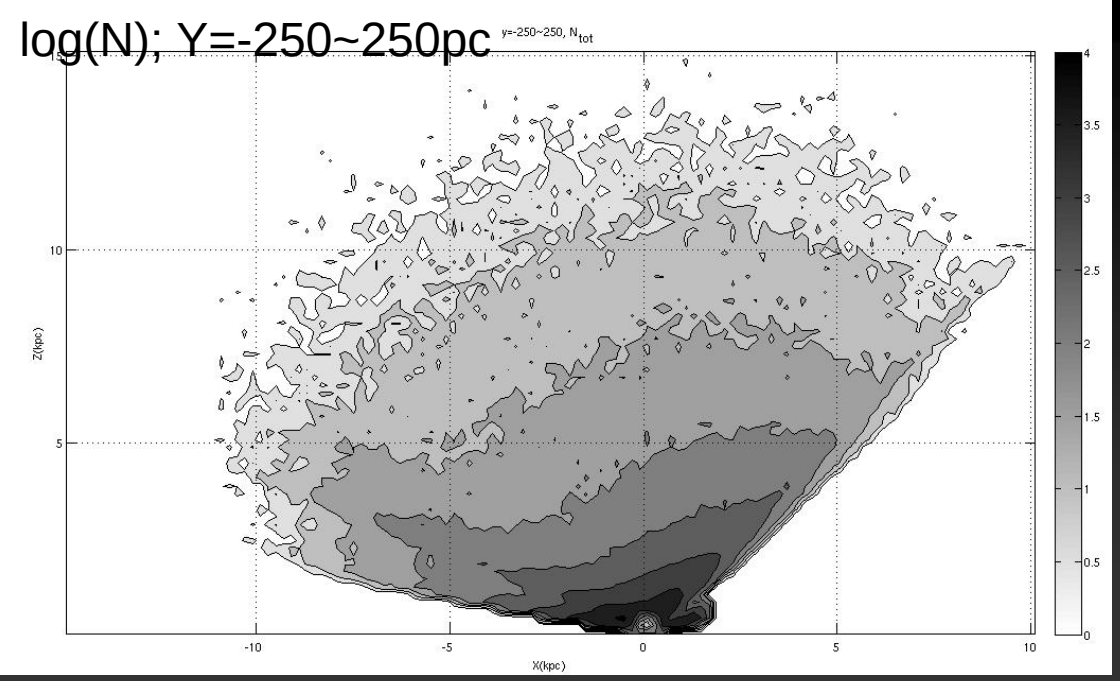
- 数据：
  - SDSS DR7的恒星
    - g<=21, g-r=[0.2,0.4]
    - 测光距离（Juric et al. 2008）
    - 测光金属丰度（Ivezic et al. 2008）
- 数据处理：
  - A: MD(-2，-1.5)；B: MD(-1.5，-1)；C: MD(-1，-0.2)
    - (A-B)/(A+B+C); (B-C)/(A+B+C)
  - 对MD进行PCA变换
    - phai=atan((a2+a20)/(a1+a10))
  - 用halo和disk的MD模板进行匹配
    - 得到halo或disk在每个MD中的比例

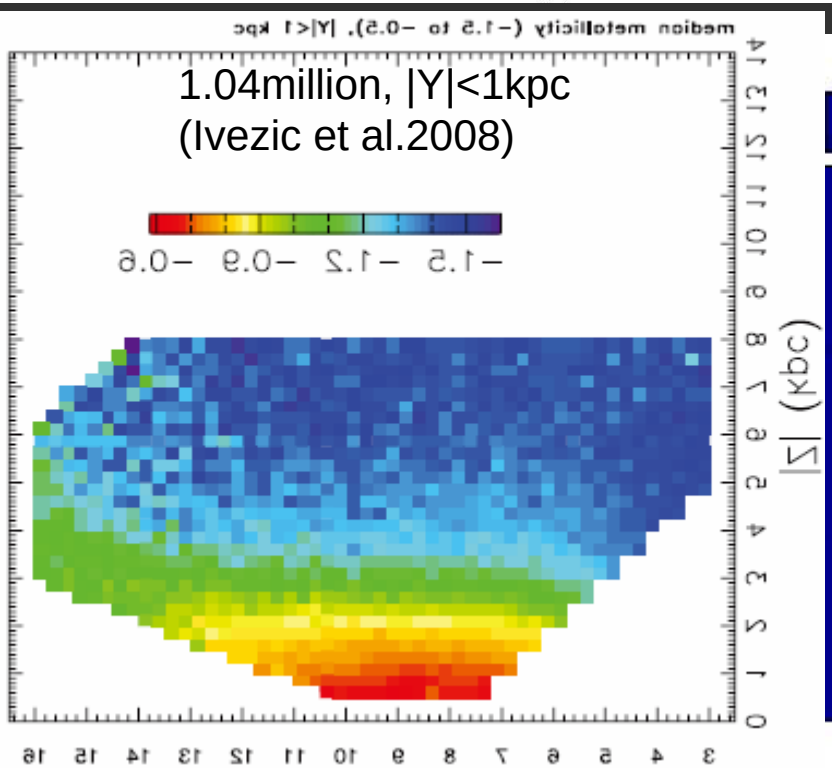


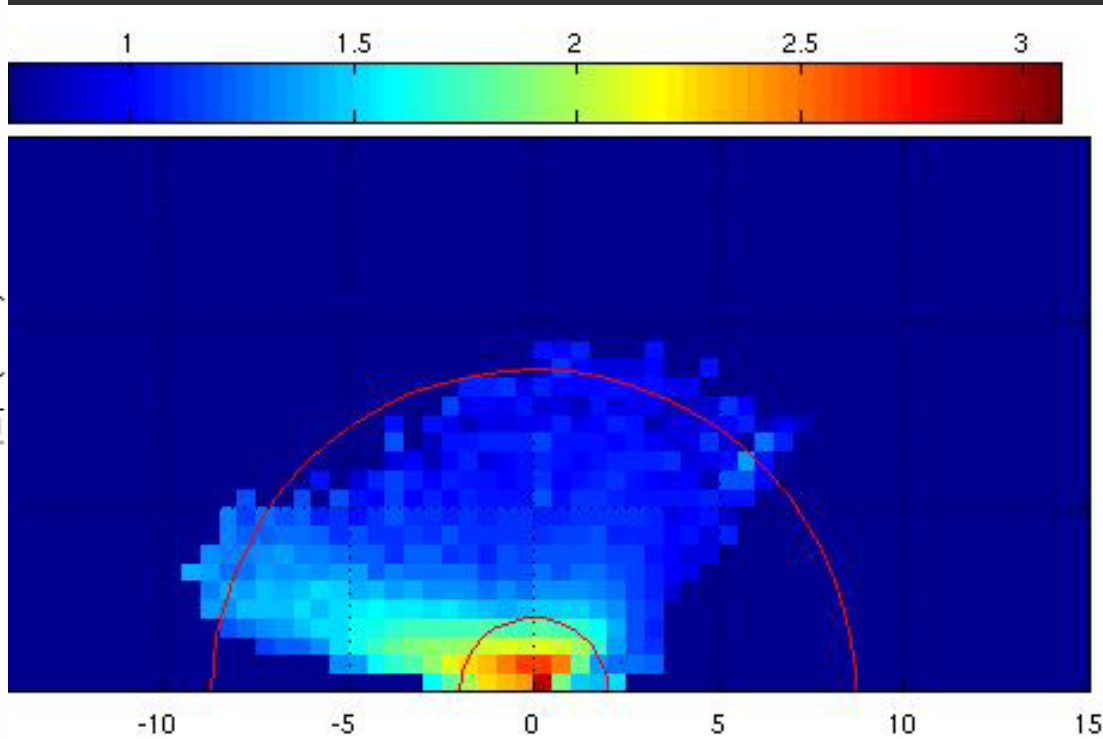

N=125;(0.723, 0.870)
N=1814;(0.870, 1.017)
N=2396;(1.017, 1.165)
N=875;(1.165, 1.312)
N=488;(1.312, 1.460)
N=220;(1.460, 1.607)
N=157;(1.607, 1.755)
N=77;(1.755, 1.902)
N=75;(1.902, 2.049)
N=32;(2.049, 2.197)
N=24;(2.197, 2.344)
N=23;(2.344, 2.492)
N=14;(2.492, 2.639)
N=3;(2.639, 2.786)
N=6;(2.786, 2.934)
N=0;(2.934, 3.081)


atan(a2/a1)
a3

log(N); Y=-250~250pc

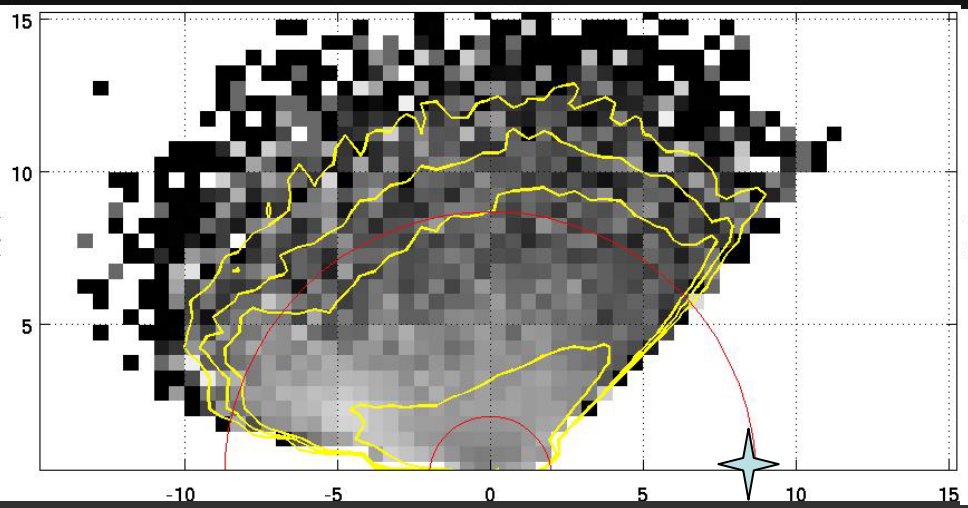




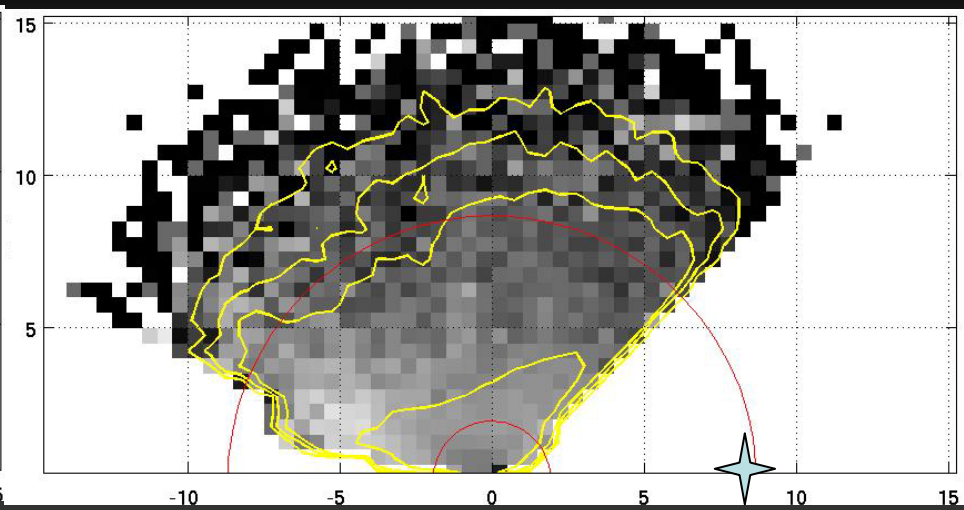


(A-B)/(A+B+C)

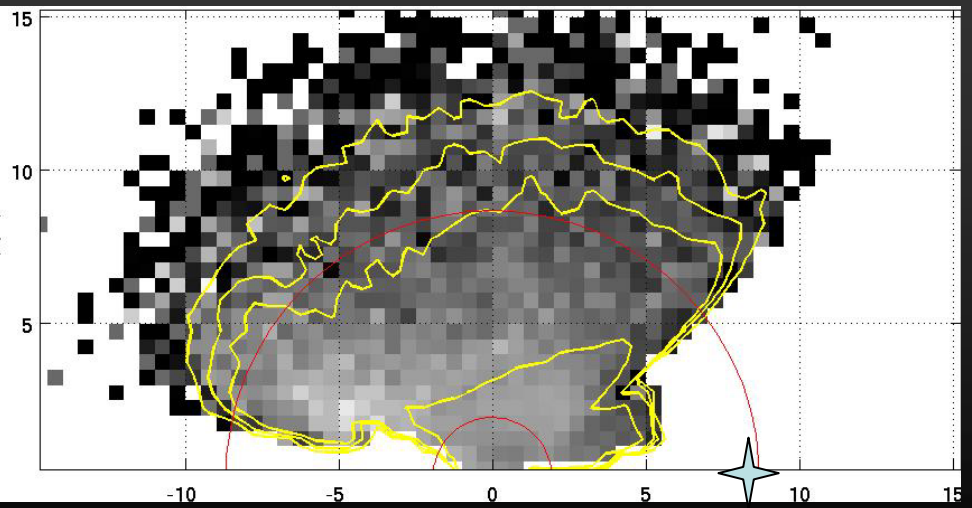
Y=-250~250pc

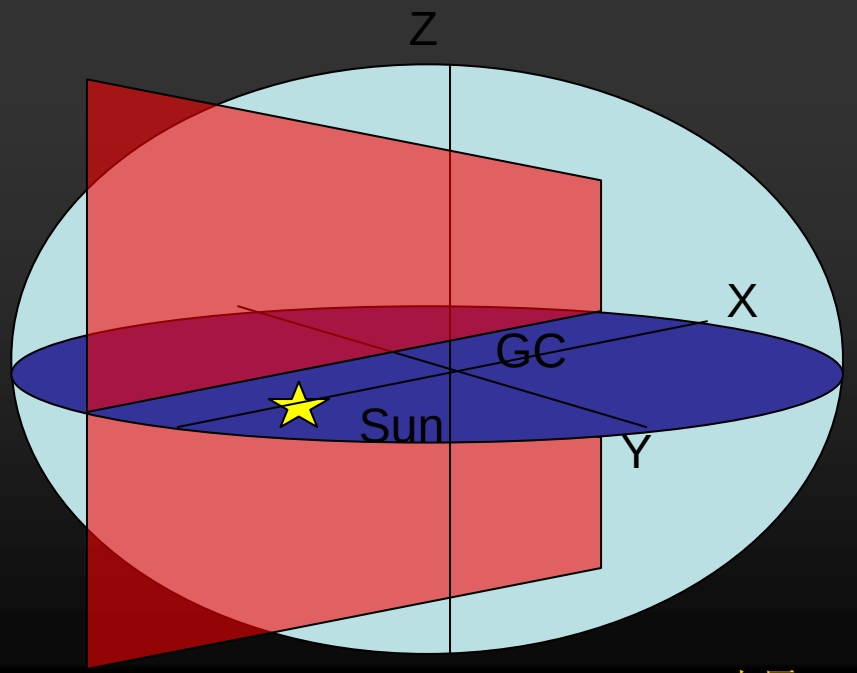


Y=-750~-250pc



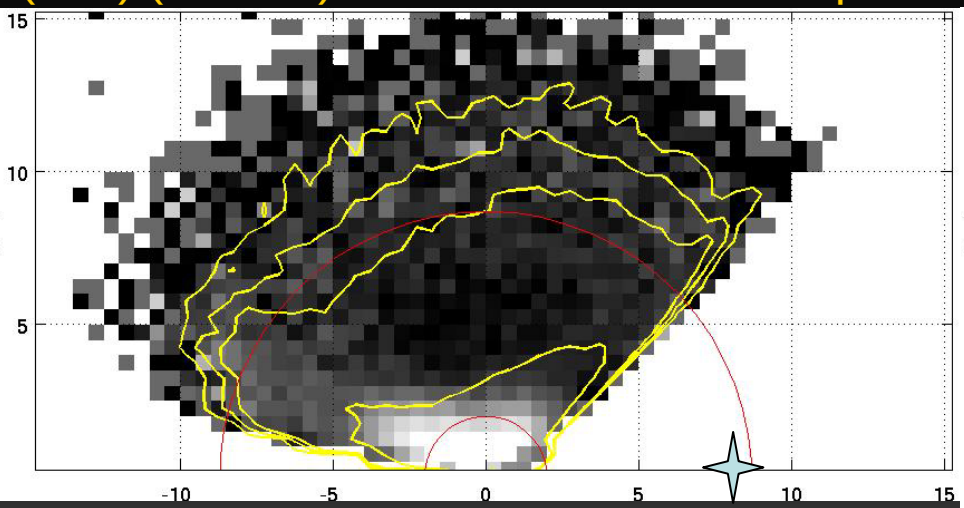
Y=250~750pc



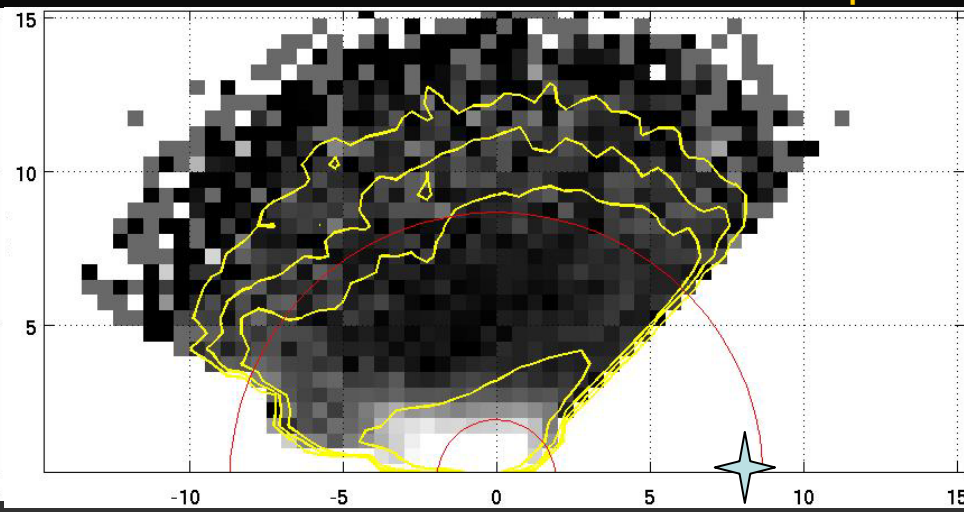


(B-C)/(A+B+C)

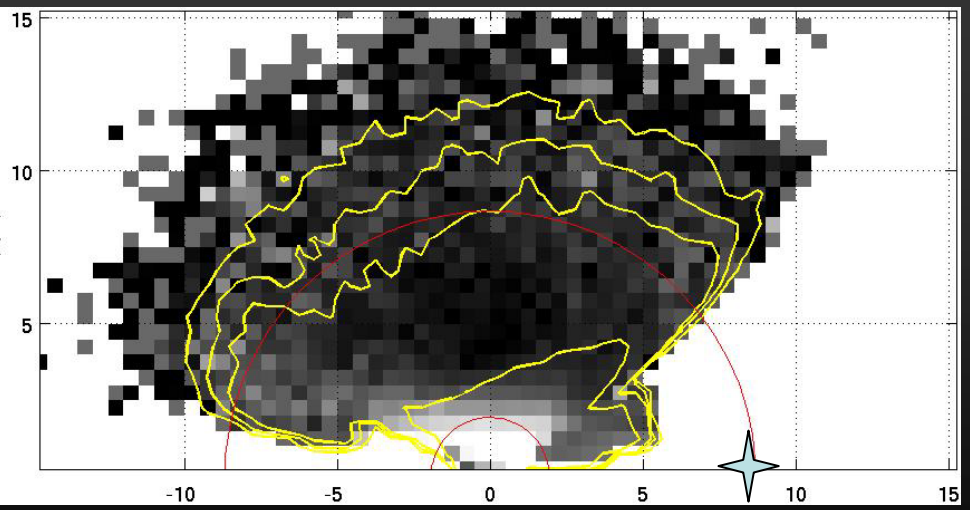
Y=-250~250pc

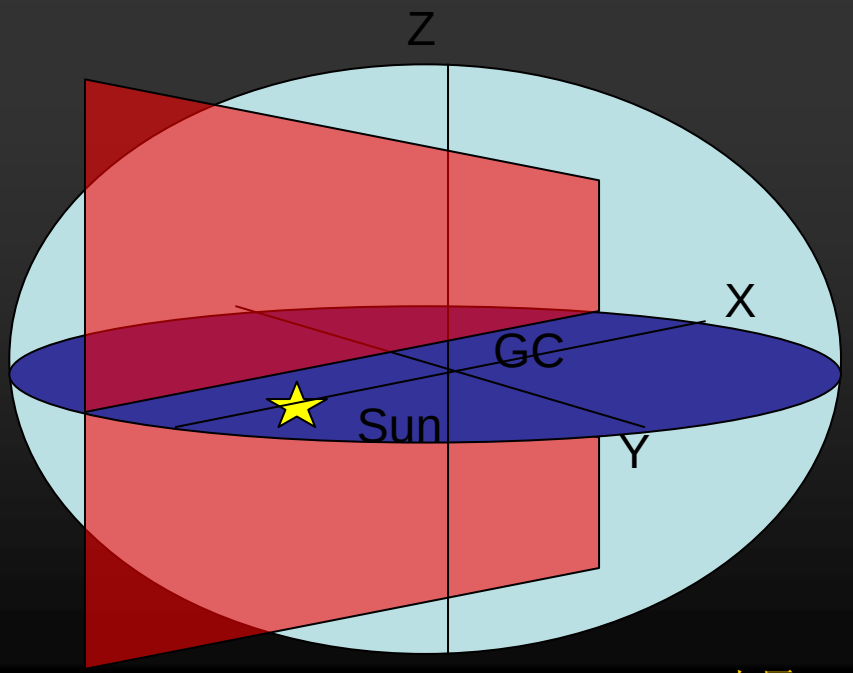


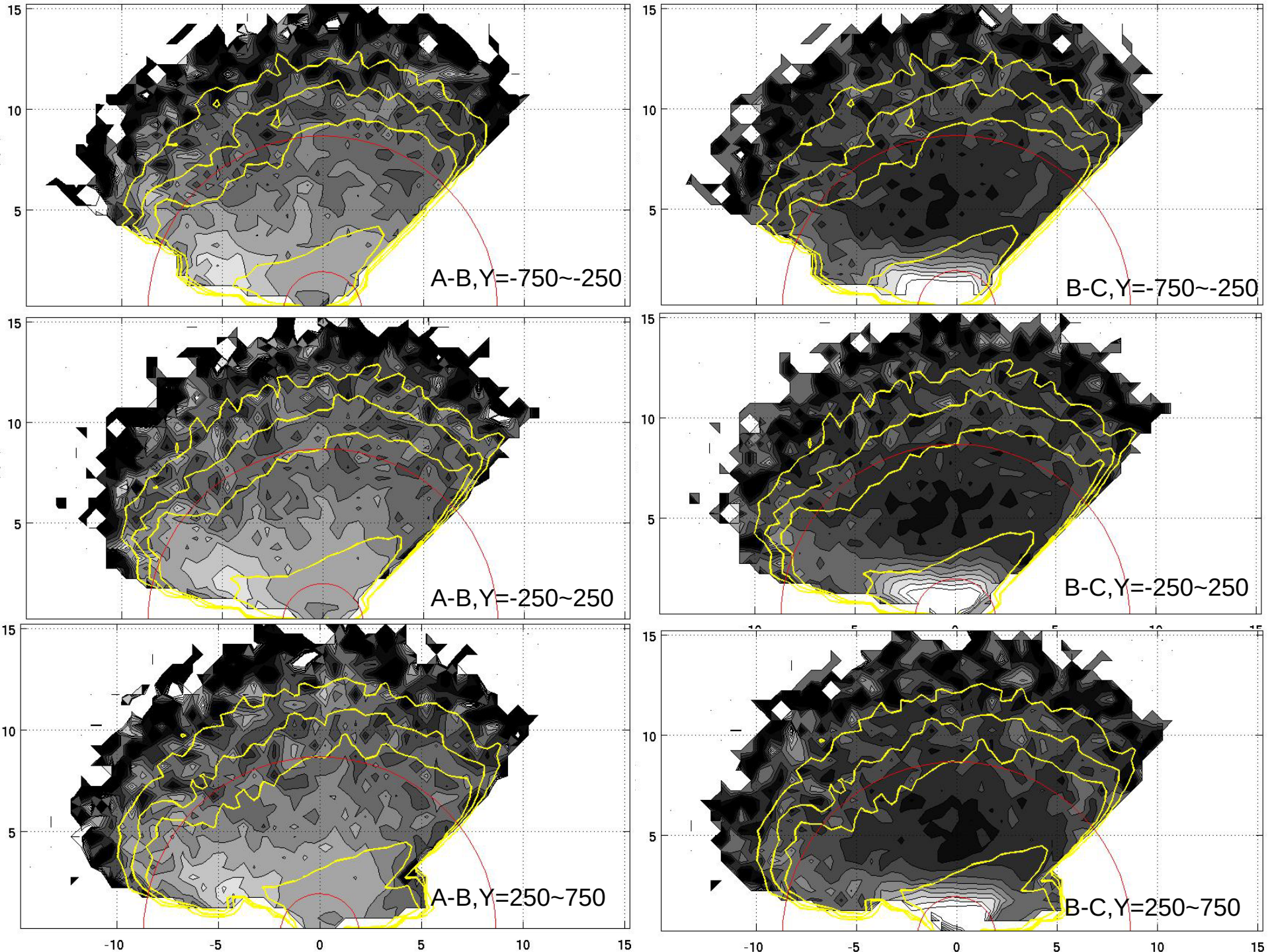
Y=-750~-250pc



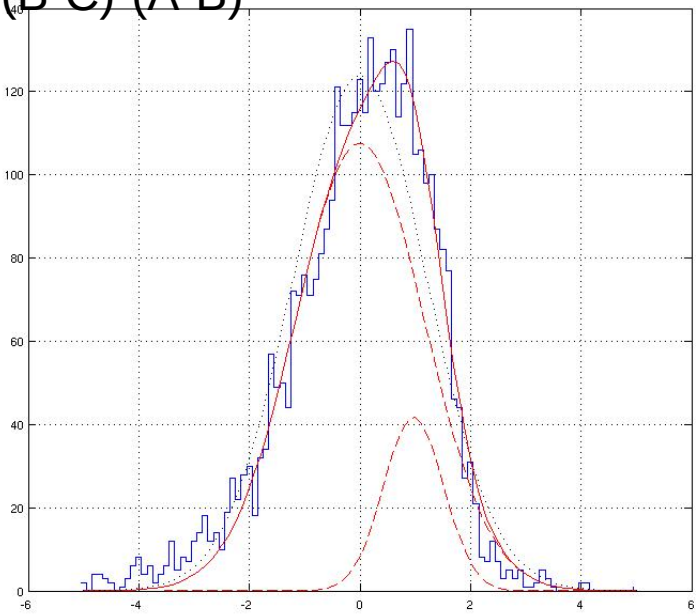
Y=250~750pc

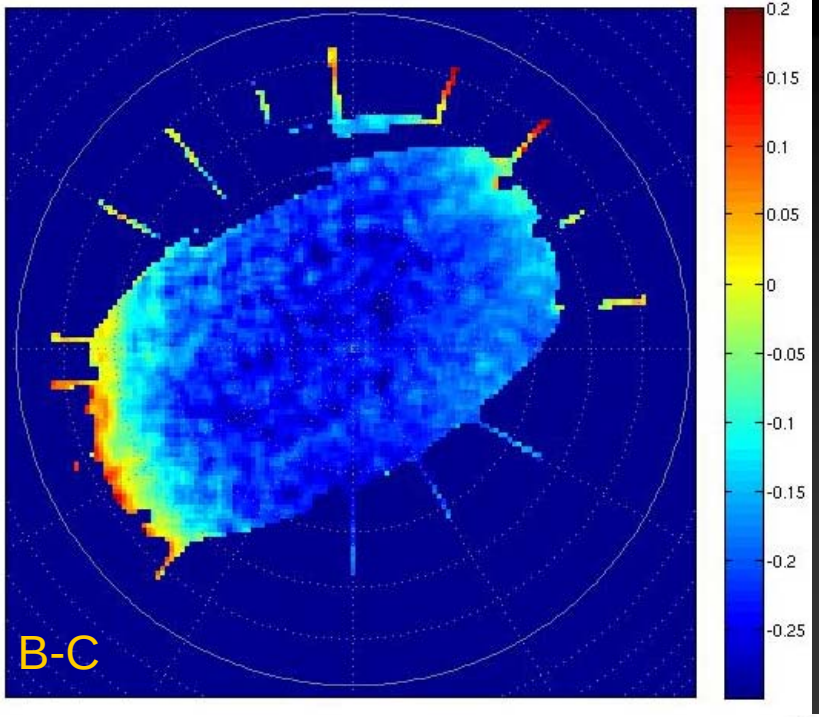


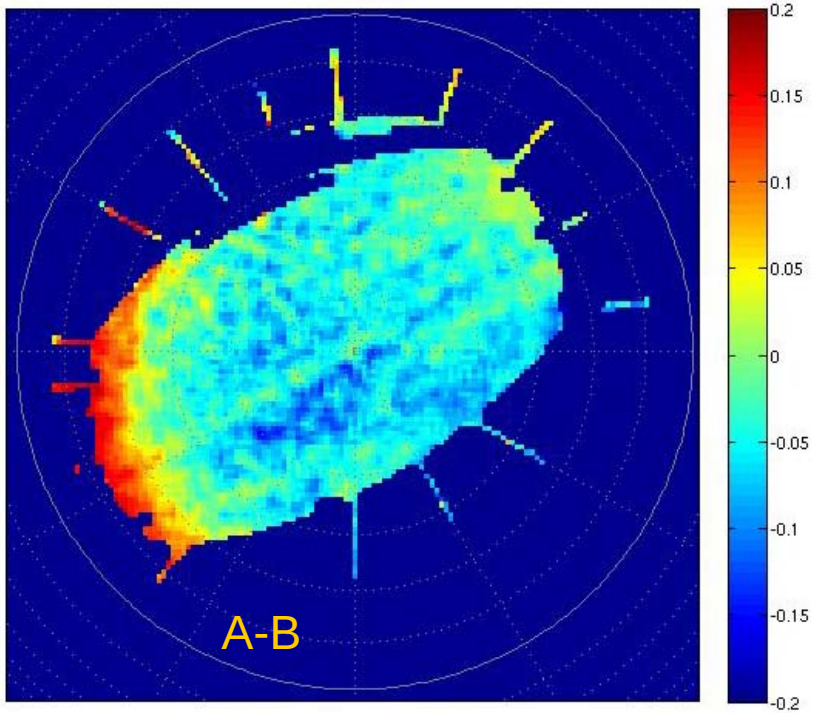


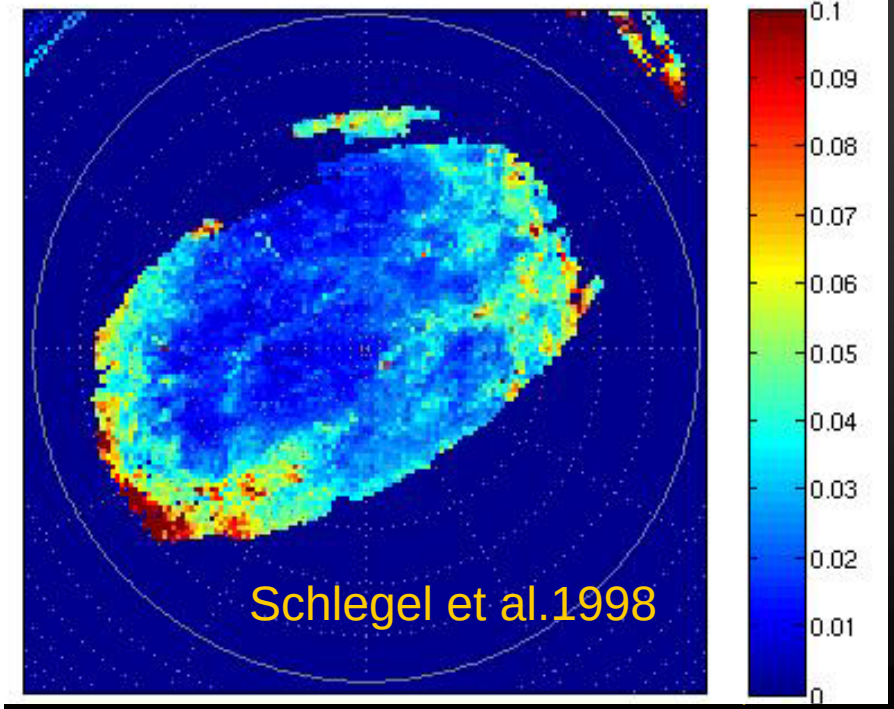

A-B,Y=-750~-250
B-C,Y=-750~-250
A-B,Y=-250~250
B-C,Y=-250~250
A-B,Y=250~750
B-C,Y=250~750
15
10
5
-10
-5
0
5
10
15

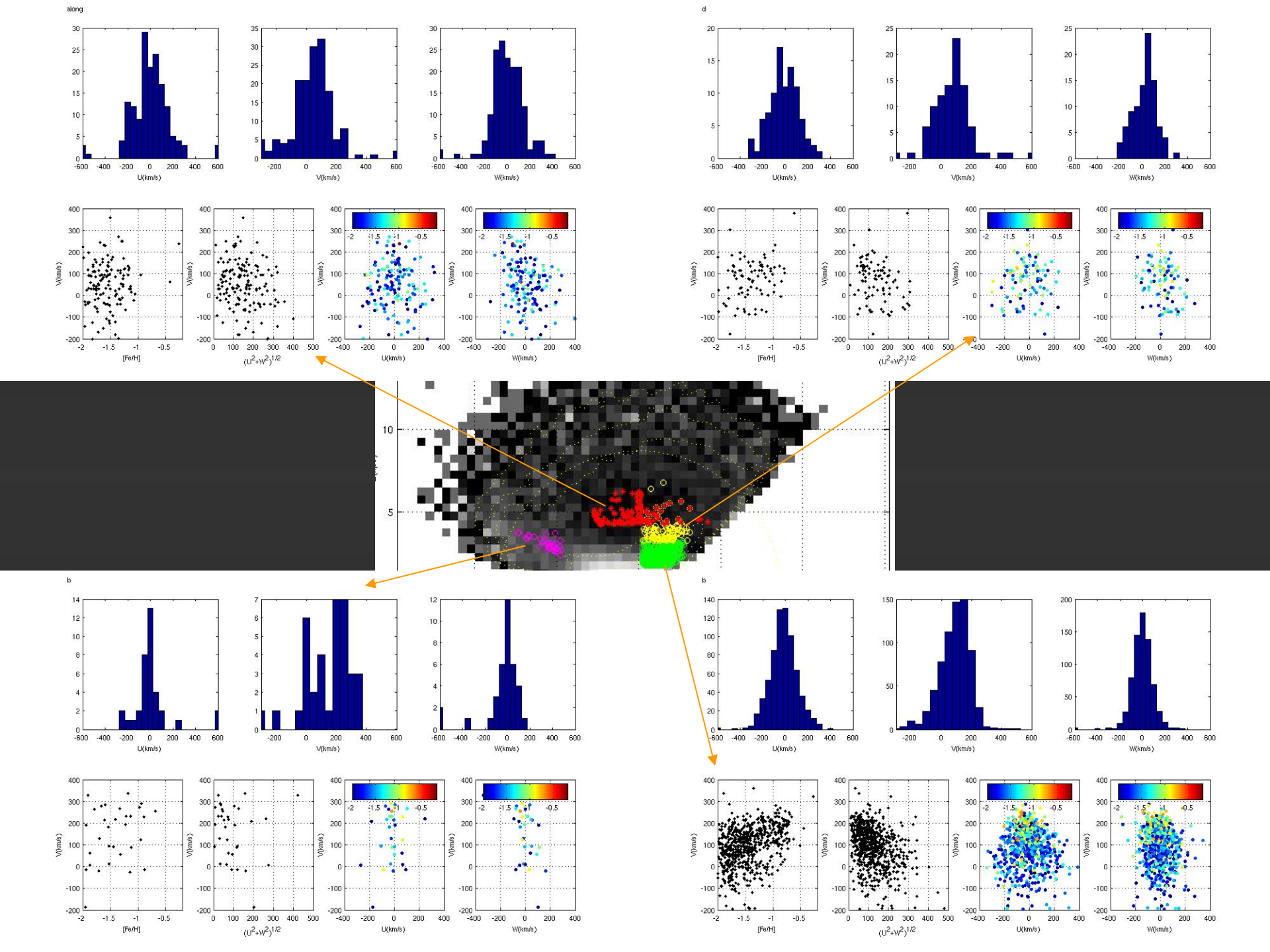
(B-C)-(A-B)


120
100
80
60
40
20
0
-6
-4
-2
0
2
4
6


B-C
0.2
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
-0.2
-0.25


A-B
0.2
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
-0.2


Schlegel et al.1998
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0


along
U(km/s)
V(km/s)
W(km/s)
[Fe/H]
$(U^2+W^2)^{1/2}$
d
b

应用六：恒星金属丰度分布和银河系结构

- 小结
  - Monoceros Ring(Mon)的MD同盘的类似，且连续变化，因而更像是盘的一个部分（flare）；Mon旋转速度和速度弥散类似于盘
  - 厚盘和晕的边界（Z~4-5kpc）非常复杂（A-B的分布），并非如Ivezic et al. 2008中所显示的平静
  - A-B和B-C的空间分布不一致，B-C在以(2kpc, 5kpc, 0kpc)为中心的一个棒状区域内出现很高的值，说明该区域在[Fe/H]~(-1.5,-1)处存在极大值，这是既不同于晕，也不同于厚盘的MD特征；运动学特征更像晕

总结

- MATLAB优势
  - 可以对海量数据进行处理（10million@laptop）
  - 支持多种形式的数据
  - 提高数据处理速度
  - 可以完成多种统计计算
  - 支持并行处理，提高运算效率
- 不足
  - Astrobox尚未产品化
  - 改变传统变成习惯

Q & A