Evolutionary Map of the Universe: Overview and Status

Ray Norris

CSIRO ASTRONOMY AND SPACE SCIENCE (CASS)
www.csiro.au

Ray Norris, CSIRO Astronomy & Space Science,
& University of Western Sydney

ASKAP Science

38 proposals submitted to ASKAP

2 selected as being highest priority

- EMU all-sky continuum (PI Norris)
- WALLABY all-sky HI (PI Koribalski & Staveley-Smith)

8 others supported at a lower priority

- COAST pulsars etc
- CRAFT fast variability
- DINGO deep HI
- FLASH HI absorption
- GASKAP Galactic
- POSSUM polarisation
- VAST slow variability
- VLBI

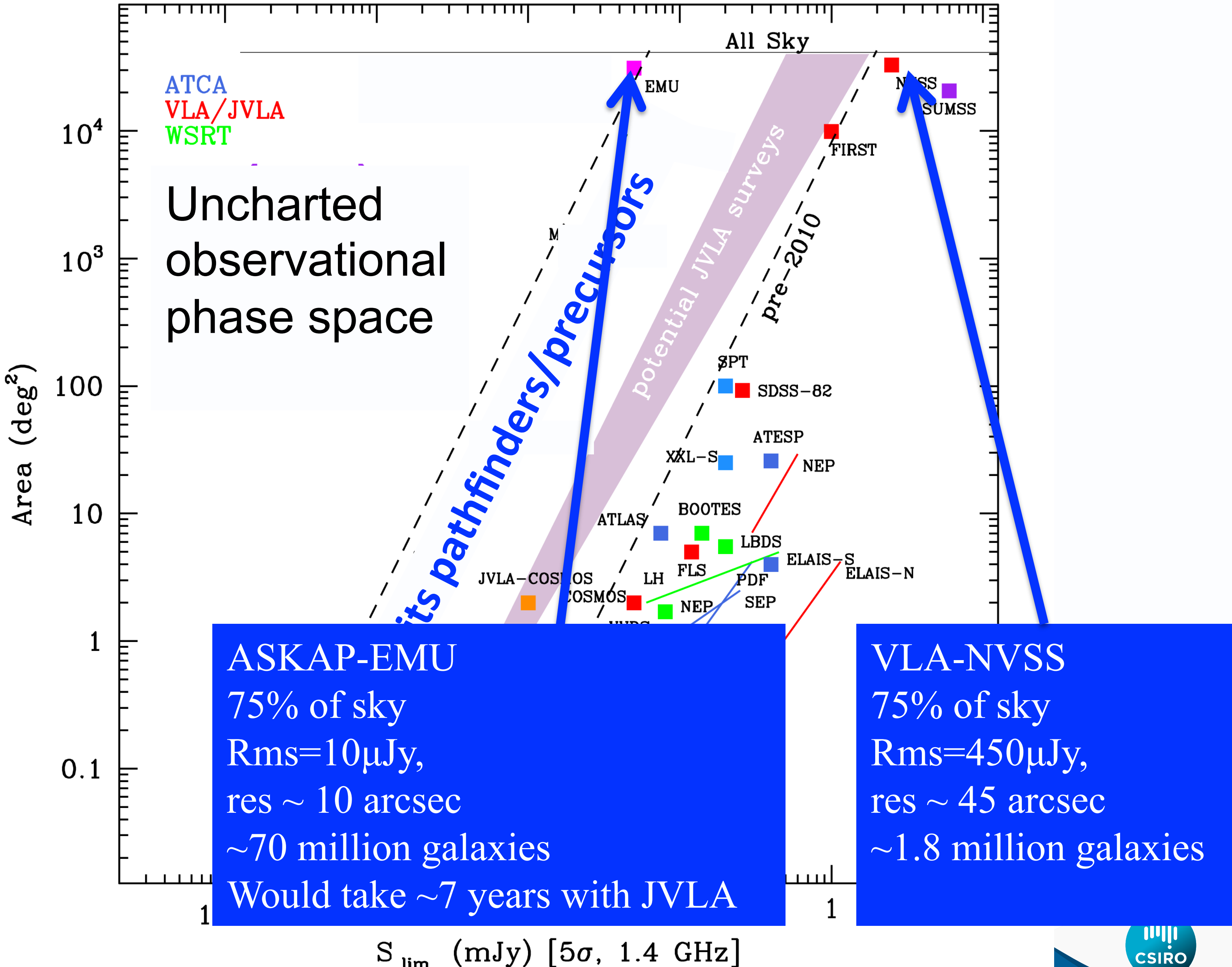


Diagram courtesy of Isabella Prandoni

EMU Overview

Deep radio image of 75% of the sky (to declination +30°)

Frequency range: 1100-1400 MHz

40 x deeper than NVSS
- 10 μJy rms across the sky

5 x better resolution than NVSS (10 arcsec)

Better sensitivity to extended structures than NVSS

Will detect and image ~70 million galaxies at 20cm

All data to be processed in pipeline

Images, catalogues, cross-IDs, to be placed in public domain

Survey starts 2016(?)

Science Goals

1) **Evolution of SF from z=2 to the present day,**

- using a wavelength unbiased by dust or molecular emission

2) **Evolution of massive black holes**

- how come they arrived so early? How do binary MB
- what is their relationship to star-formation?

3) **Explore the large-scale structure ological parameters of the Universe.**

– E.g, Independent tests of da odels

4) **Explore an uncharted observational parameter space**

- almost certainly findi ses of object.

5) **Explore Clu use low-surface-brightness radio objects**

6) Genera of the Galactic Plane

7) Cr acy for surveys at all

gths (Herschel, JWST, ALMA, etc)

How did galaxies form and evolve?

CSIRO

How does EMU differ from earlier surveys?

1. Scale – increases the number of known radio sources by a factor of ~30

2. Will not be dominated by AGN – about half the galaxies will be normal SF galaxies

3. Ambition – includes:

- Cross-identification with optical/IR catalogues
- Ancillary data (redshifts etc)
- Key science projects as an integral part of the project

4. Explicitly includes "discovering the unexpected"

The fundamental challenge to EMU

- Data processing is in real time, so we can't tweak our data reduction algorithms – we need to get it right BEFORE we take data
- No proprietary period on processed data, so need to ensure that key science papers are produced quickly, to reward teams for their work
- This also means that we get the science out quickly, and others can then build on that!

Addressing the Challenge

- Ensure critical tasks (e.g. source extraction, photo-z) are being actively worked on, with the right timescale
- Ensure that the people who work on them will be rewarded for their efforts,
- Provide incentives to work on EMU tasks,
- Ensure we are able to write the key papers as quickly as possible after EMU data becomes available:
  - avoid half-baked papers
  - “hit the ground running”
  - Try out techniques/software on pathfinder data sets (ATLAS, COSMOS, ATLAS-SPT)

Important: ensure that we write papers NOW (~15 EMU journal papers so far).

EMU is not so much a large project as an archipelago of small projects

CSIRO

EMU and its pathfinders

Comparison: NVSS
3π sr
Rms=450 μJy
1.8 million galaxies

ATCA – ATLAS
(2006-2013)
6 antennas single-pixel

7 sq deg
Rms=15 μJy
6000 galaxies

ATCA – ATLAS - SPT
(2013-2015)
6 antennas single-pixel

100 sq deg
Rms=40 μJy
30,000 galaxies
300 clusters?

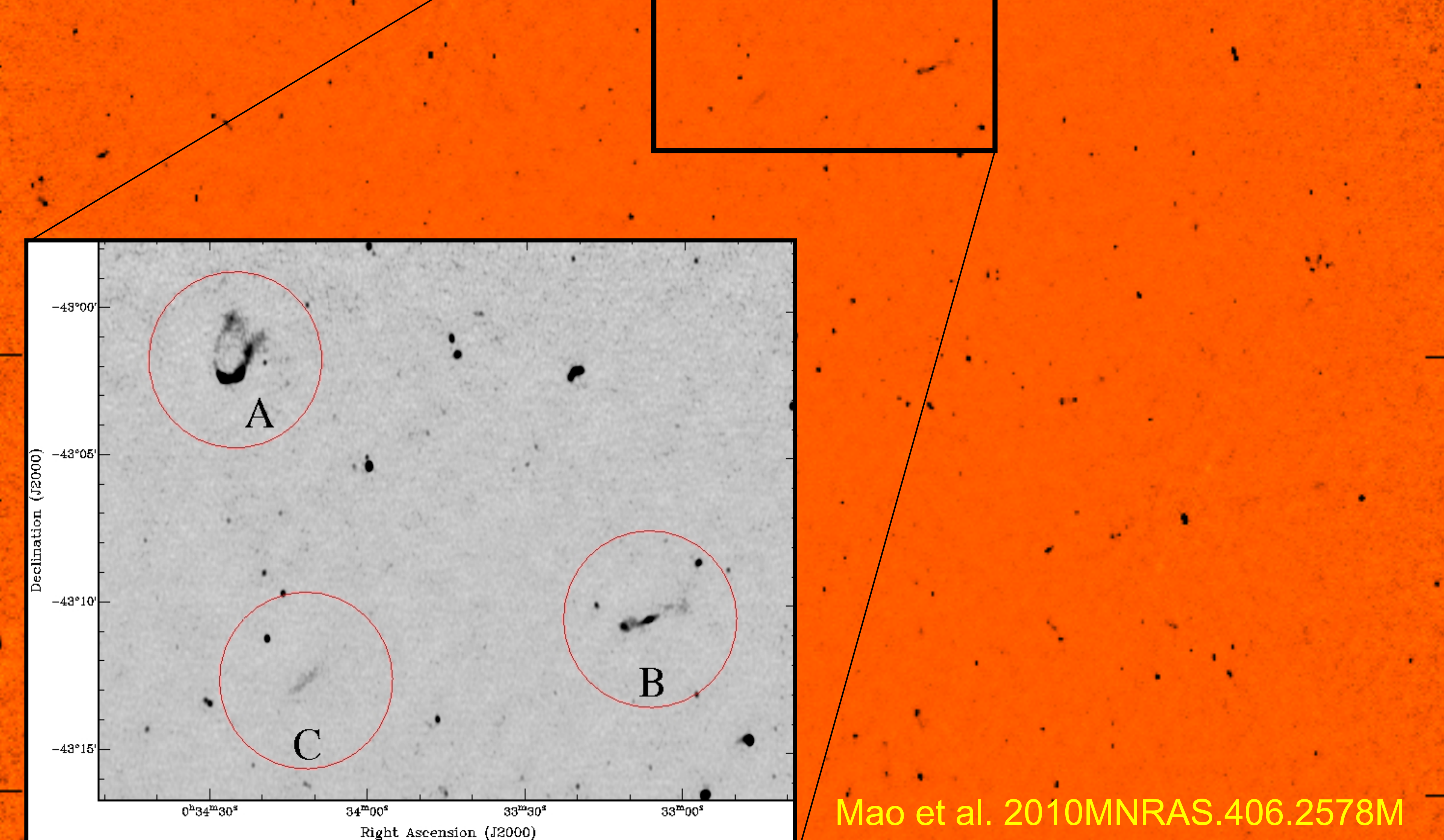
ASKAP – early science
(2016)
12 antennas MkII PAF

1000 sq deg
Rms=30 μJy
0.5 million galaxies

ASKAP – EMU
(2017-2018)
30-36 antennas MkII PAF

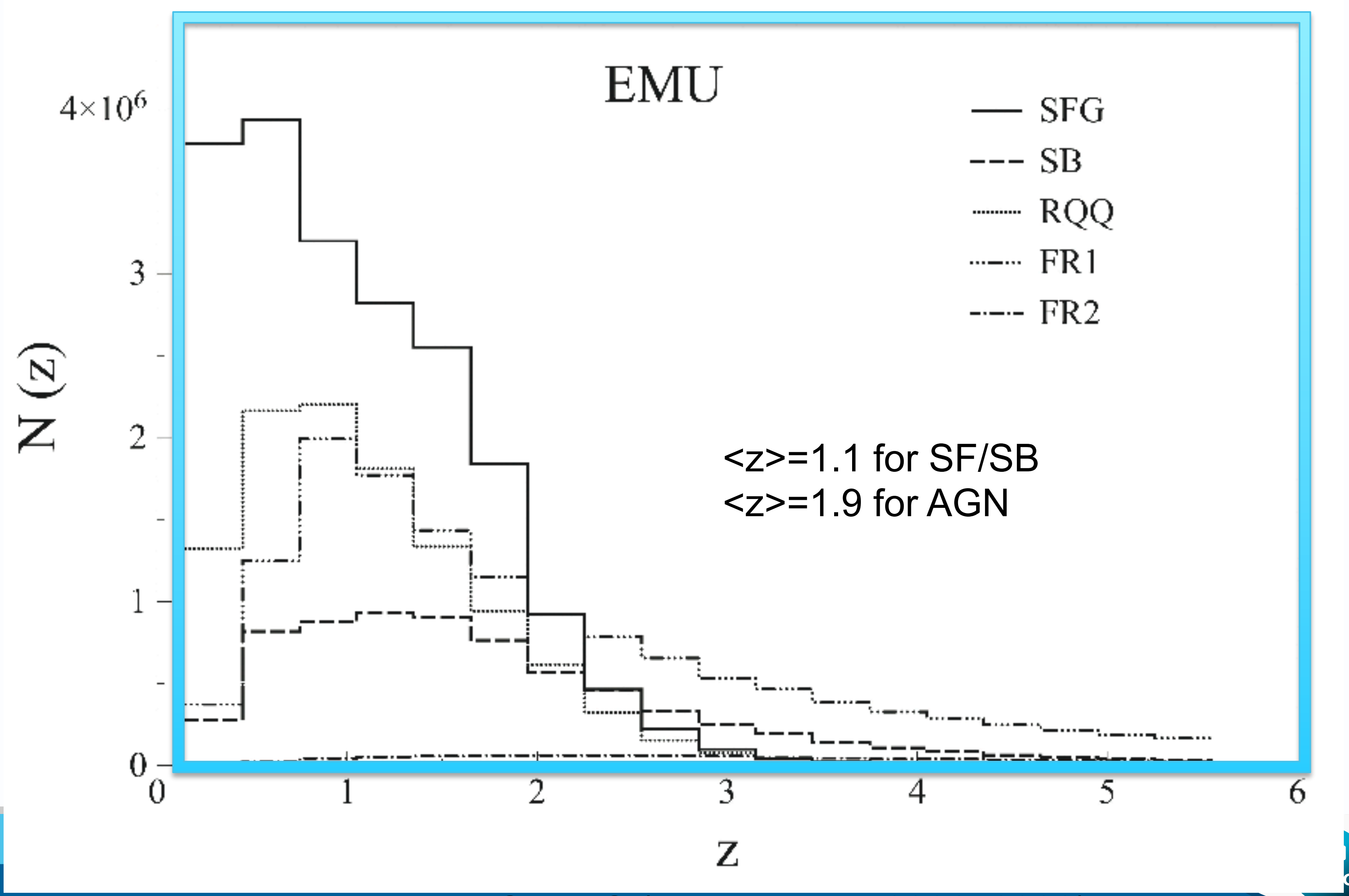
3π sr
Rms=10 μJy
70 million galaxies

The EMU Pathfinder:
ATLAS=Australia Telescope Large Area Survey
7 sq deg to rms=15 μJy



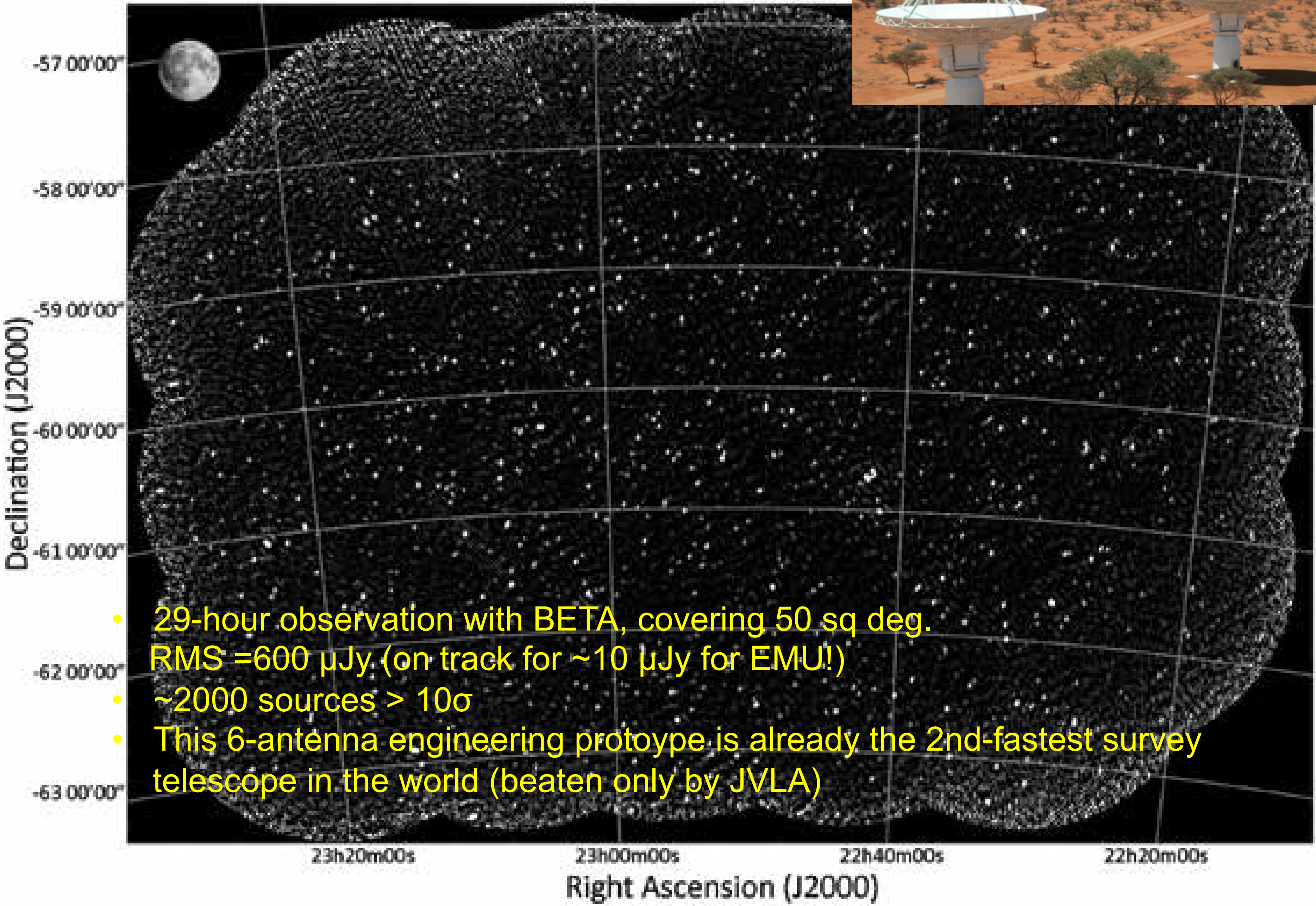
Mao et al. 2010MNRAS.406.2578M

Redshift distribution of EMU sources



Based on SKADS (Wilman et al; 2006, 2008)

PAFs work (even with only 6 antennas)!



- 29-hour observation with BETA, covering 50 sq deg.
  RMS =600 μJy (on track for ~10 μJy for EMU!)
- ~2000 sources > 10σ
- This 6-antenna engineering protoype is already the 2nd-fastest survey telescope in the world (beaten only by JVLA)

EMU Project Structure

The EMU Management Team

- **Ray Norris (Project Leader)**
- **Andrew Hopkins & Nick Seymour** (Project Scientists)
- **Anna Kapinska** (Project Manager)
- **Josh Marvil** (ASKAP Early Science Team Leader)
- **Ian Heywood** (Guru in residence)
- Coordinator and editor of a twice-yearly EMU newsletter (vacant)
- Editor of the EMU public web page (vacant)

The EMU's EGG

- The consultative group of ~30 active, involved, EMU members

The EMU team

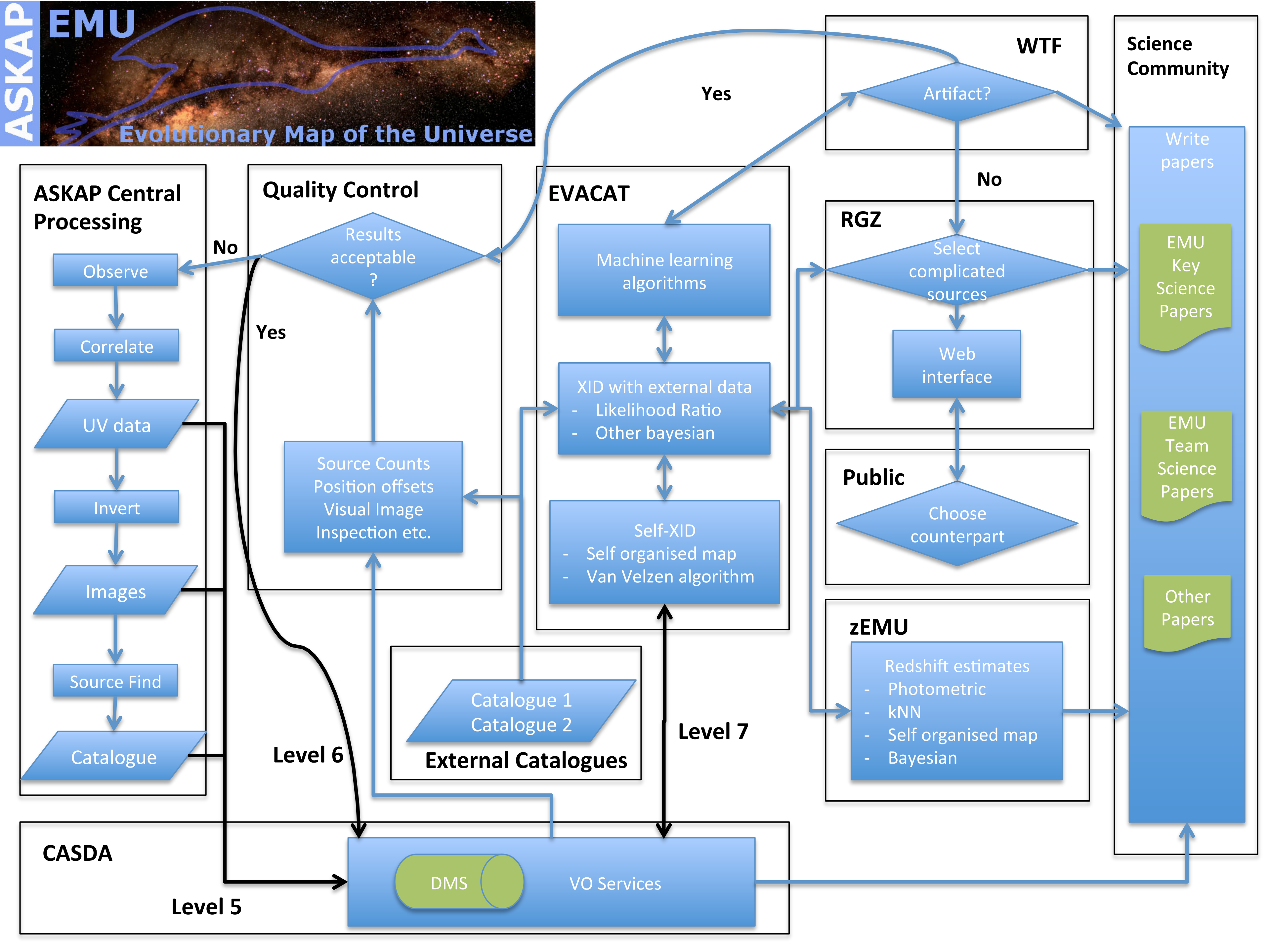
- ~300 team members from 17 countries

What is EMU doing right now?

- **ASKAP-BETA**
- **ASKAP Early Science**
- **Key science projects**
  - e.g. cosmology, AGN evolution, etc
- **Development Projects**
  - e.g. source extraction, photo-z, etc
- **Collaboration Projects**
  - building links to other wide-field surveys
- **User science projects**

Actively participating in any of these activities will result in your leading or co-authoring an EMU paper

CSIRO



ASKAP
EMU
Evolutionary Map of the Universe
ASKAP Central Processing
Observe
Correlate
UV data
Invert
Images
Source Find
Catalogue
Quality Control
Results acceptable ?
No
Yes
Source Counts
Position offsets
Visual Image
Inspection etc.
Level 6
EVACAT
Machine learning algorithms
XID with external data
- Likelihood Ratio
- Other bayesian
Self-XID
- Self organised map
- Van Velzen algorithm
Catalogue 1
Catalogue 2
External Catalogues
Level 7
WTF
Artifact?
Yes
No
RGZ
Select complicated sources
Web interface
Public
Choose counterpart
zEMU
Redshift estimates
- Photometric
- kNN
- Self organised map
- Bayesian
Science Community
Write papers
EMU Key Science Papers
EMU Team Science Papers
Other Papers
CASDA
DMS
VO Services
Level 5

EMU Key Science Projects

EMU Key Science Projects	Project Leader
EMU Value-Added Catalogue	Nick Seymour
Characterising the Radio Sky	Ian Heywood
EMU Cosmology	David Parkinson
Cosmic Web	Shea Brown
Clusters of Galaxies	Melanie Johnston-Hollitt
Cosmic star formation	
Radio-loud AGN	
Radio AGN in the EoR	
Radio-quiet AGN	
Local Universe	
The Galactic Plane	
SCORPIO: Radio Stars	Grazia Umana
WTF: Mining Data for the Unexpected	Ray Norris

Each of these will generate a number of papers, starting now, and culminating in at least one key science paper. Each project has a team working on it. You can volunteer to join a team.

See http://askap.pbworks.com/KeyProjects for details

CSIRO

Examples of Proposed ASKAP Early Science Projects

Ian Heywood	Characterising the Radio Sky
Glen Rees	Radio cosmolo
Andrew Hopkins	Radio-luminos
Minh Huynh	RLFs of radio-l ... ds
Anna Kapinska	RLFs of FRII ra
Ray Norris	Cluster source
Jose M. Diego	Radio-IR corre
Nick Seymour	Studies of Bro
Chiara Ferrari	Dynamically a
Tiziana Venturi	Radio emission in the Shapley Concentration:

You can join these, or propose a new one. Observations will start ~early 2016 using the early 12-antenna ASKAP. Each project will produce at least one paper.

Examples of Proposed EMU User Science Projects

Rossella Cassano	Insight on the origin of giant radio halo from a mass-selected sample of galaxy clusters.
Nick Seymour	SFR as functi environment
Ray Norris	EMU Deep Fi
Andrew Hopkins	Evolving IMF

You can join these, or propose a new one. Observations will start ~ 2017. Each project will produce at least one paper.

CSIRO

Examples of EMU Collaboration Projects

Meerkat-MIGHTEE	Ray Norris
eRosita	Nicolas
SkyMapper	Julie Ba
WISE	Tom Ja
VHS	(TBD)
LSST	Amy Ki
MWA	(TBD)
Taipan	Ray No
DES/OZ-DES	Nick Se
POSSUM	(TBD)
XXL	Vernes
Radio Galaxy Zoo	Julie Banfield
FP7-HELP	(TBD)

You can volunteer to lead one of these, or propose a new one. Each project will probably produce at least one paper, and working on one will also earn you the right to be on other EMU papers.

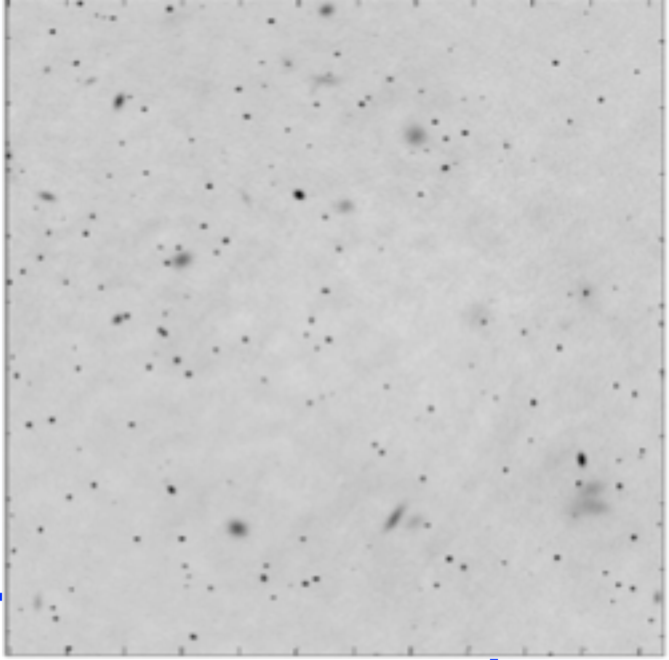
EMU Development Projects

Developers earn co-authorship on key science papers

- Ensure the EMU database satisfies our st
  and value- added.
- Develop, set up, and implement the data
- Ensure ASKAPSOFT imaging satisfies EMU
- See what special imaging is needed for th
- Ensure ASKAPSOFT source extraction sati
- Develop algorithms for extraction of diffu
- Develop the self-ID and cross-ID algorith
- Develop an "optimum photo-z algorithm
  strategy for those smaller areas of EMU c
- Develop techniques for Statistical redshif
- Explore other EMU applications for Mach

You can volunteer to lead or join one of these, or propose a new one. Each project will probably produce at least one paper, and <u>working</u> on one will also earn you the right to be on other EMU papers.

EMU Data Science Challenges: 1. Compact Source Extraction



Initial data challenge study showed existing source extraction algorithms not up to the job.

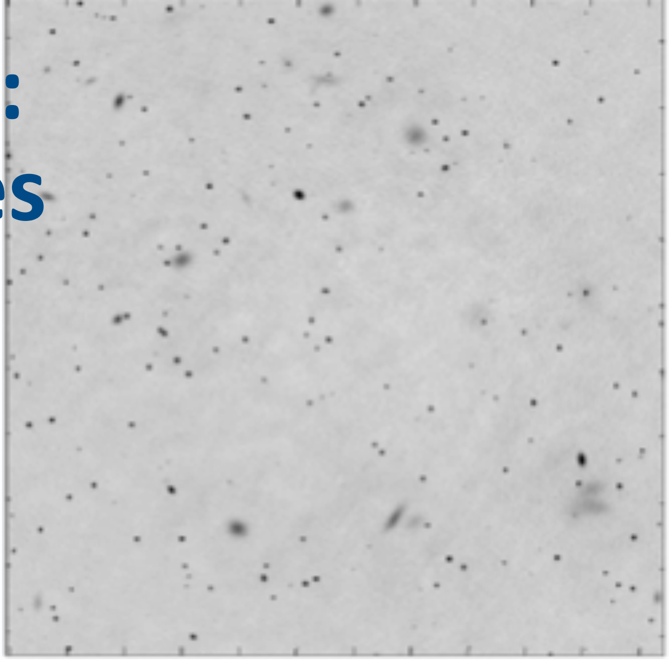
Need to build on this, identify problems, using fix existing algorithms or develop a better one.

Publications of the Astronomical Society of Australia (PASA)
© Astronomical Society of Australia 2015; published by Cambridge University Press.
doi: 10.1017/pas.2015.xxx.

The ASKAP/EMU Source Finding Data Challenge

A. M. Hopkins[1,*], M. T. Whiting[2], N. Seymour[3], K. E. Chow[2], R. P. Norris[2], L. Bonavera[4], R. Breton[5], D. Carbone[6], C. Ferrari[7], T. M. O. Franzen[3], H. Garsden[8], J. Gonzalez-Nuevo[4], C. A. Hales[9], P. J. Hancock[3,10,11], G. Heald[12,13], D. Herranz[4], M. Huynh[14], R. J. Jurek[2], M. Lopez-Caniego[15,4], M. Massardi[16], N. Mohan[17], S. Molinari[18], E. Orrù[12], R. Paladino[19,16], M. Pestalozzi[18], R. Pizzo[12], D. Rafferty[20], H. J. A. Röttgering[20], L. Rudnick[21], E. Schisano[18], A. Shulevski[12,13], J. Swinbank[22,6], R. Taylor[23,24], A. J. van der Horst[25,6]

EMU Data Science Challenges: 2. Extraction of diffuse sources



- No existing algorithm can routinely extract diffuse sources
- A number of algorithms in development (and have been for years!)
- We need this soon for ATLAS-SPT!

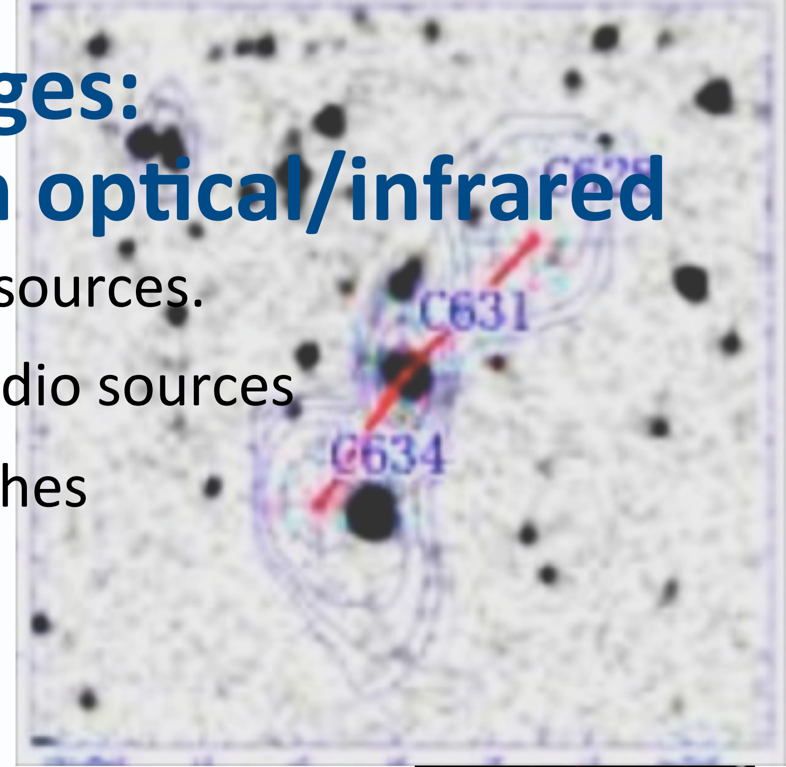
Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY
MNRAS **447**, 2243–2260 (2015)

The deep diffuse extragalactic radio sky at 1.75 GHz

T. Vernstrom,[1★] Ray P. Norris,[2] Douglas Scott[1] and J. V. Wall[1]

EMU Data Science Challenges:
3. Cross-identification with optical/infrared

- Existing algorithms work well with point sources.
- They fail badly on typical core-jet-lobe radio sources
- Currently exploring a number of approaches

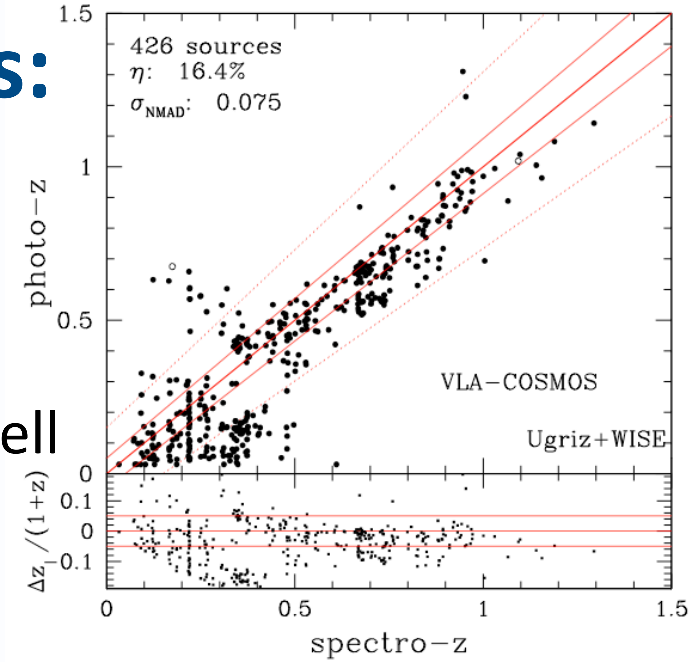


Mon. Not. R. Astron. Soc. 000, 1–8 (2002) Printed 5 May 2015 (MN LATEX style file v2.2)

Matching Radio Catalogs with Realistic Geometry: Application to SWIRE and ATLAS

Dongwei Fan[1⋆], Tamás Budavári[2,3†], Ray P. Norris[4], Andrew M. Hopkins[5]

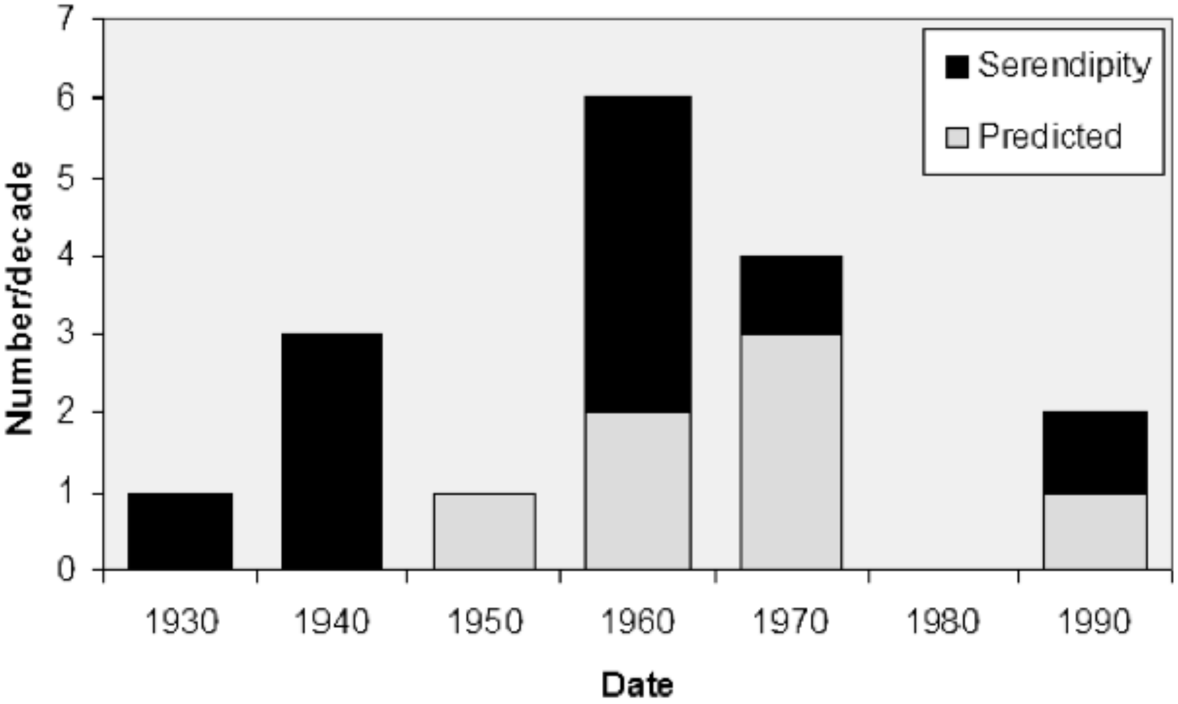
EMU Data Science Challenges: 4. Statistical redshifts



- We will measure spectroscopic redshifts for only ~1% of EMU sources
- Even photometric redshifts are hard to do well
- But for many science goals, the redshift distribution of a sample is more important than individual redshifts
- Several machine-learning algorithms are being tried.

An Example of a Key Science project: Discovering the Unexpected

What fraction of discoveries in astronomy were "Popperian"?



(b) Predicted v Serendipity

+1 for dark energy (2012)

Serendipity:11
Predicted: 7

From Ekers (2009) *PoS(sps5)007*

Discoveries with HST

Project	Key project	Planned?	Nat. Geo. top ten?	Highly cited?	Nobel prize?
Use Cepheids to improve value of H0	✔	✔	✔	✔	
study intergalactic medium with uv spectroscopy	✔	✔			
Medium-deep survey	✔	✔			
Image quasar host galaxies		✔	✔		
Measure SMBH masses		✔	✔		
Exoplanet atmospheres		✔	✔		
Planetary Nebulae		✔	✔		
Discover Dark Energy			✔	✔	✔
Comet Shoemaker-Levy			✔		
Deep fields (HDF, HDFS, UDF, FF, etc)			✔	✔	
Proplyds in Orion			✔		
GRB Hosts			✔		

Discoveries with HST (see e.g. Lallo: *arXiv:1203.0002)*

Project	Key	Planned?	Nat.	Highly	Nobel prize?
Use Cepheids to					
study intergalac uv spectroscopy					
Medium-deep s					
Image quasar hc					
Measure SMBH					
Exoplanet atmo:					
Planetary Nebul					
Discover Dark Er					✔
Comet Shoemak					
Deep fields (HDI					
Proplyds in Orion			✔		
GRB Hosts			✔		

Summary:

Of the "top ten" HST discoveries:

- **1 was a key project**
- **4 were planned by astronomers but were not key projects**
- **5 were totally unexpected (e.g. dark energy)**

The discovery of pulsars

Jocelyn Bell:

- explored a new area of observational phase space
- knew the instrument well enough to distinguish interference from signal
- observant enough to recognise a sidereal signature
- open minded – prepared for discovery
- within a supportive environment
- persistent

See Bell-Burnell (2009) PoS(sps5)014 for a personal perspective

Could Jocelyn Bell Discover the Unexpected in ASKAP data?

- Data volumes are huge – cannot sift by eye
- Instrument is complex – no single individual will be familiar with all possible artifacts
- ASKAP will be superb at answering well-defined questions (the "known unknowns")
- Humans won't be able to find the "unknown unknowns"
- Can we mine data for the unexpected, by rejecting the expected?

If not, ASKAP will not reach its full potential
i.e. it will not deliver value for money

What does ASKAP need to do to discover the unexpected?

- **Maximise the volume of new phase space**
  - A good surrogate is to use # of known objects
  - Maximised by an all-sky survey

- **Retain flexibility**
  - don't optimise the telescope ONLY for your science goals

- **Develop data mining software to search for the unexpected**

- This will be an important part of data-intensive research

mining radio survey data for the unexpected

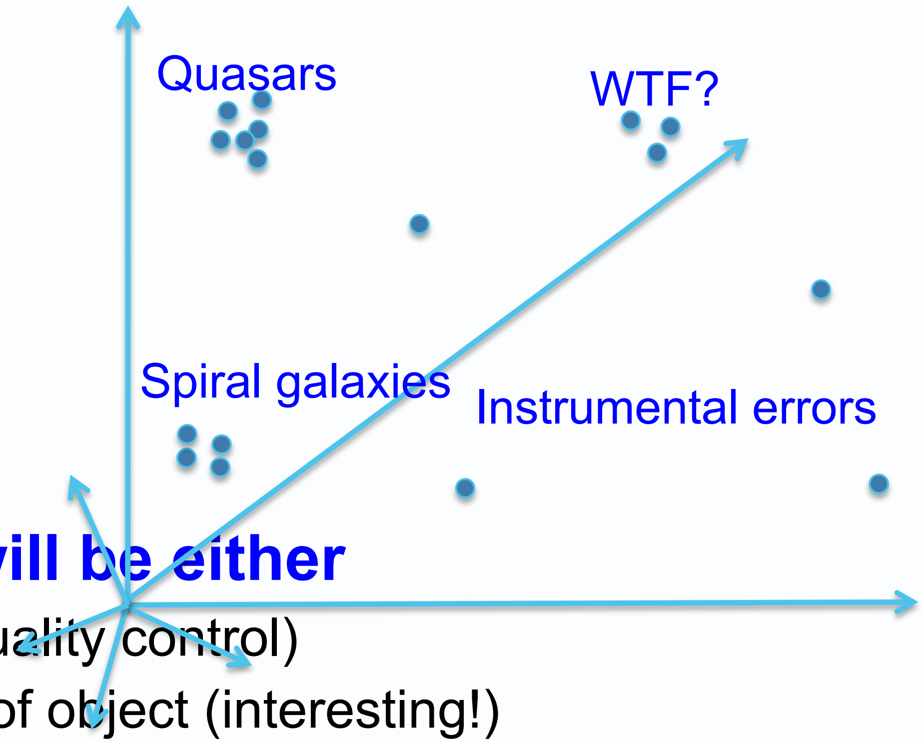
WTF?

WTF = Widefield ouTlier Finder

Mining large data sets for the unexpected

WTF will work by searching the n-dimensional (large n) phase space of observables, using techniques such as

- Decision tree approach
- Zoo approach
- Cluster analysis
- k-nearest-neighbours
- self-organised maps
- Bayesian approach to combine all the above

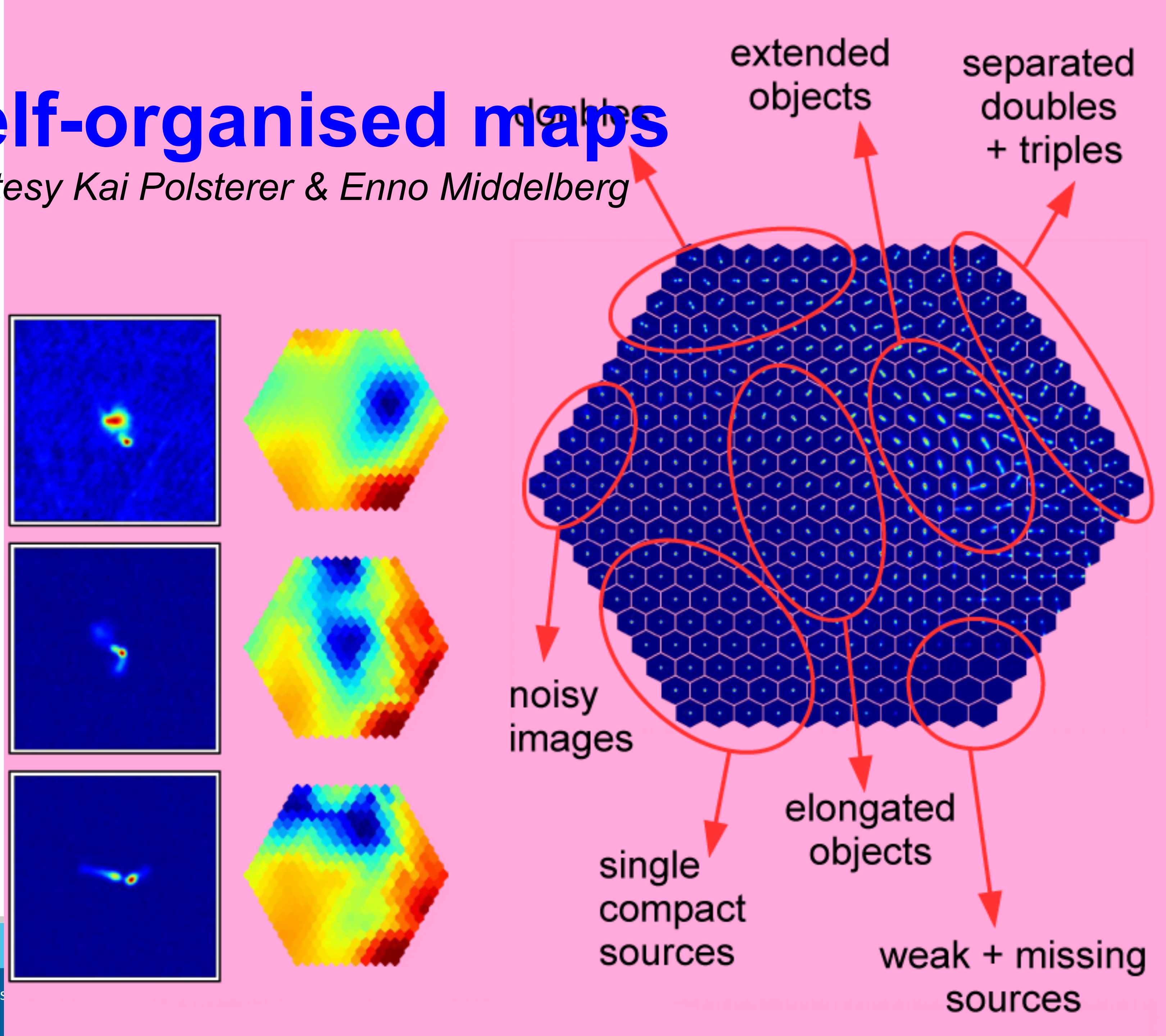


Identified objects/regions will be either

- processing artifacts (important for quality control)
- statistical outliers of known classes of object (interesting!)
- New classes of object (WTF)

Self-organised maps

courtesy Kai Polsterer & Enno Middelberg



WTF Proposal to Amazon

Challenge approaches for catalog data

- Decision tree approach
- Cluster analysis
- K-nearest-neighbours
- Self-organised maps
- Bayesian approach
- Searches in the spatial structure domain

Identified objects/regions may be either

- Processing artefacts (important for quality control)
- Statistical outliers of known classes of object (interesting!)
- New classes of object or phenomena (WTF?)

WTF Proposal to Amazon

Challenge approaches for image data

- Low-surface brightness emission or absorption
- Structure (such as rings) in the noise which are statistically significant in an appropriate domain (e.g. Hough transform and Hu moments)
- Noise properties which depart from the distribution predicted from thermal noise and confusion
- Structures with unusual spectral shapes.

Discussion

See http://askap.pbworks.com/EMU_reacceleration for more details

- Comments?
- Reactions?
- Suggestions?
- Questions?
- Suggestions for tasks?
- Suggestions for science papers?
- How can we generate papers now?
  - NB: 15 EMU journal papers so far!
- Any volunteers for (a) newsletter editor (b) web editor


YOU ARE NOW LEAVING THE
MURCHISON RADIO-ASTRONOMY
OBSERVATORY
THANK YOU FOR BEING RADIO QUIET