POSSUM Workshop 2012- Day 1 Notes

Program for day 1 (see http://askap.org/possum/Meetings/SydneyPossumWorkshop2012)

Tuesday May 8th

Time	Topic	Speaker / Chair
09:00-09:30	Update on ASKAP	Aidan Hotan
09:30-10:00	ASKAPSOFT & Computing	Tim Cornwell
10:00-10:30	Morning Tea	
10:30-12:00	The Software Pipeline	Cormac Purcell
12:00-13:30	Lunch	
13:30-15:00	Coordination with EMU / ASKAP & EMU	Ray Norris
15:00-15:30	Afternoon Tea	
15:30-17:00	RM_Synthesis and RM-Clean	Xiaohui Sun

Attending:
Landecker (video)
Willis (video)
Robishaw (video)
GeisbÃ¼sch (video)
Kaczmarek
Stil
Gaensler
Sun
Akahori
C. Anderson
Lenc
Rudnick
Kothes
O'Sullivan
Purcell
Norris
Gheissari
S. Brown
Seymour
McClure-Griffiths (briefly)
Beck (briefly)
Cornwell
Hotan
Whiting

ASKAP update - Aidan Hotan

BETA PAF deployment in 2 stages: antennas 1/3/6, and then 8/9/15.
BETA itself ready by end of 2012.

Parallel effort taking place with Parkes Test Facility.

Next two PAFs undergoing final assembly in Marsfield.
BETA box beamformers and compute cluster fully operational

Software correlator - allows early interferometry, only 16 MHz b/w, can handle 18 beams

MRO: now has 1 GB/s link back to Geraldton (dedicated link) and to Sydney (shared link)

BETA box hardware will remain until central site array is operational.

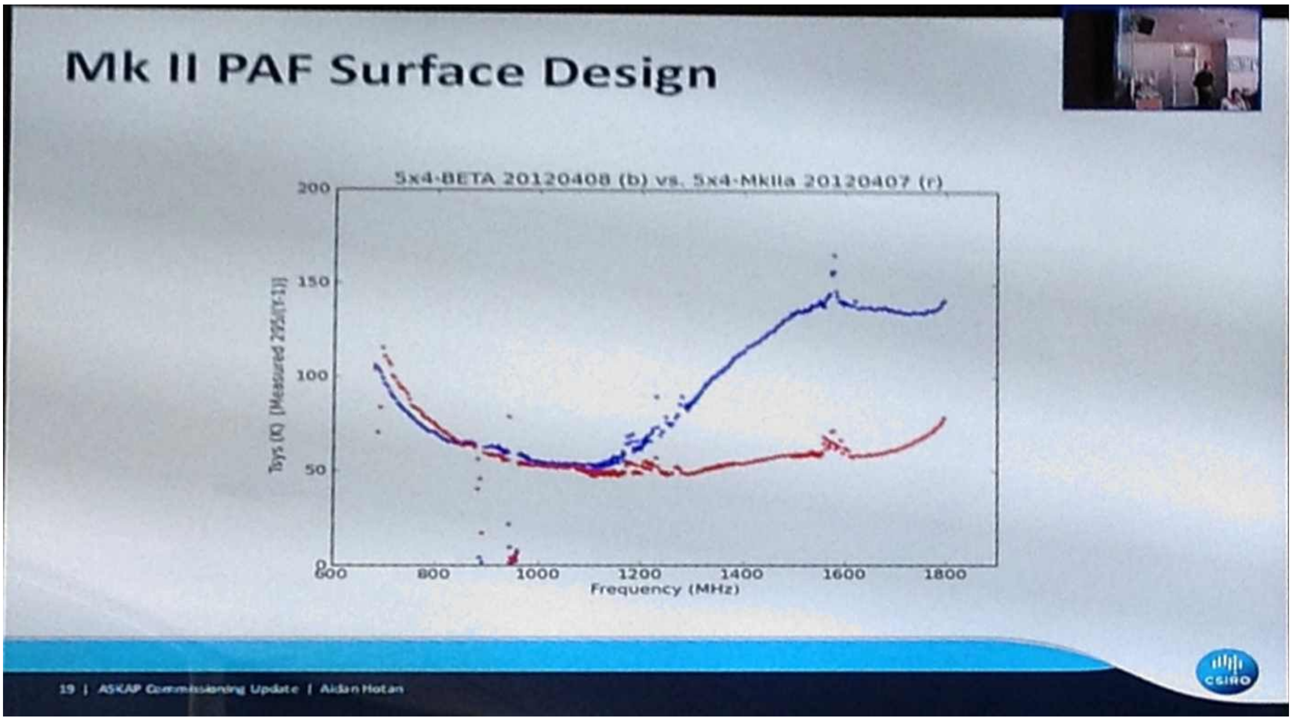
At the moment: 19 antennas installed, 15 partially installed, 2 still in crates. All 36 antennas installed by end of June 2012.

2012 commissioning plan:
- BETA is primarily an engineering test instrument - focus is understanding how to make best images with PAF arrays
- some early science may be possible once operational procedures have been established - then handover to SCOM2
- staged deployment of BETA allows for refinement of acceptance and commissioning test strategies
- detailed commissioning plan is being formulated

Steps towards BETA science
- long-term tracking of sources with sky-mount antenna
- phase closure with three phased-array feeds, bore-sight beams
- baseline vector verification, delay tracking tests
- first astronomical imaging with s/w correlator (cascading level of complexity, starting with pt src and single beam, leading up to ATCA reference fields with multiple beams)
- provision of simple measurement sets for pipeline testing
- analysis of beam shapes and stability, weighting schemes
---> **SCOM-2 involvement likely to begin during this phase (September 2012)**

Latest developments:
- Further $4M for another 6 Mk2 PAFs (12 Mk2 PAFs, 6 Mk1 PAFs total)
- ongoing discussion about configuration of 12 Mk2 antennas (conflicting baseline preferences among SSTs)
- ADE CDR scheduled for late June 201
- expect first demonstration ADE systems in August: updated PAF element design with improved T_sys in high band
- deployment of first ADE system to MRO in 2013

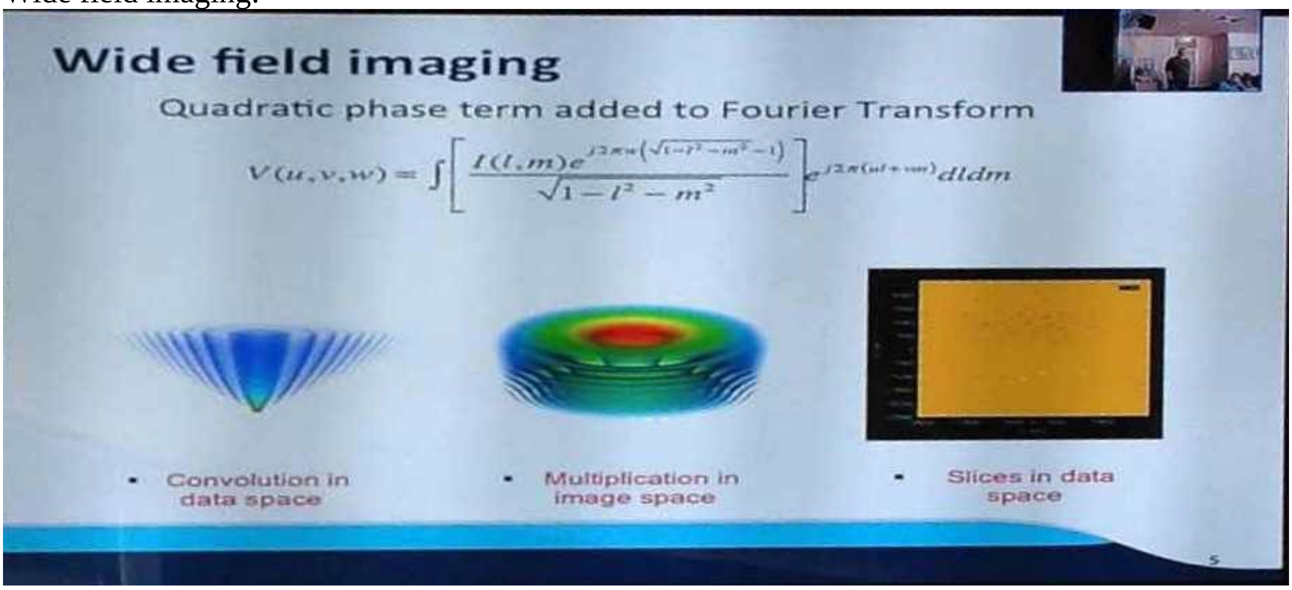
Performance of old vs new PAFs:



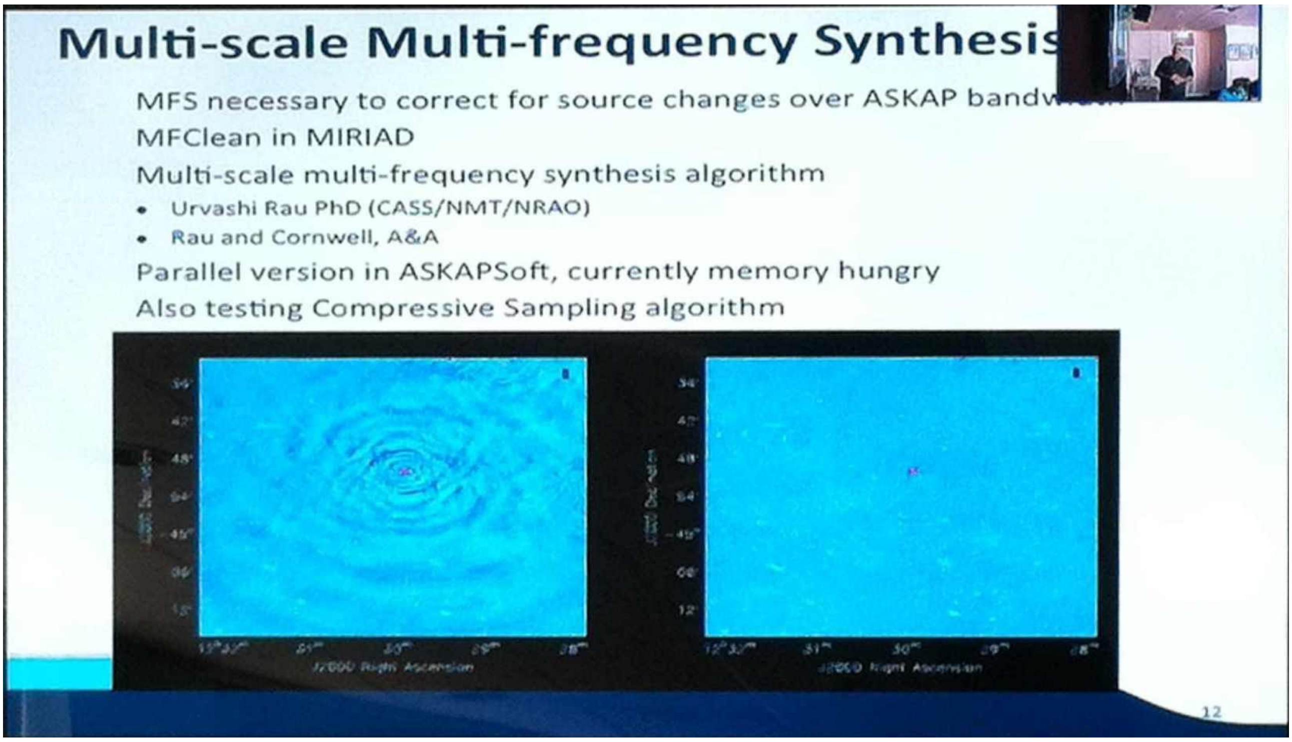
ASKAP Computing Status - Tim Cornwell

Main activities at the moment
- telescope operating system (TOS)
- software correlator
- ADE
- science team interactions (new version of ASKAP Science Processing document coming at end of year; comments welcome at any time; have sought input from science teams on improvements to source fitting code; provided access to our prototype for SSP testing)
- SKA work package preparations

Wide field imaging:



If you use slices in data space and don't update the slicing too often, you get a position error in the image, plus the source is also distorted. Combining data slicing and convolution (?) fixes this.



Note that Multi-scale Multi-frequency Synthesis is currently only Stokes I only. *It's not out of scope to include Mike Bell's Faraday synthesis into the ASKAP pipeline.*

Science Processing Data Challenges:

- imaging data challenge
-- develop simulated processing pipeline on Pawsey 1A supercomputer
-- create complete pipeline, use as regression test
-- drive improvements in s/w to be ready for BETA
-- visibilities simulated from model of sky & telescope
-- then treated as we would real observations

- Ingest data challenge
-- simulation of ingest pipeline running on Pawsey 1A supercomputer
-- control ingest process via TOS
-- move visibilities & metadata from MRO to Pawsey 1A

Full ASKAP will require 10,000 cores.

Time scale for computing roll-out is such that it will potentially be challenging to process 3-km baselines with ASKAP-12.

Summary:

- ASKAP comouting capabilities coming along well
- within a factor of a few for calibration and imaging
- new problems always pop up: cube merge is I/O limited
- still lots of work to do

- Data Challenge now being deployed onto Pawsey EPIC system

Cormac Purcell - POSSUM Software Pipeline

There will be two separate pipelines: EMU (single plane output, multi-frequency synthesis) and POSSUM (cube output, with 30-300 channels). Positions identified in EMU image, and then spectra extracted at this position in POSSUM cube

Don't take spectrum at peak position: interpolate over 5x5 box around peak.

Bias correction: simple correction formula if we only consider sources with S/N above 8 sigma.

Larry: Faraday depth spectrum needs to be very finely binned (20-30 pixels per RMTF) if you want to use clean components in RM spectrum to determine complexity. Better approach is to look for variations in P vs lambda^2. So could reconsider Shea's criterion for complexity.

Questions:
- should we now be considering Faraday synthesis at the imaging stage and can it be done? [yes & yes?]
--> ACTION: Tim Cornwell says POSSUM should evaluate algorithm and make scientific statement of its benefit
- currently writing end-to-end pipeline written in Python and interfacting w MySQL database. Is this suitable for transfer to ASKAPsoft team? [CASS will not implement pipeline for us]Â
- algorithms & modules in well-commented procedures (based on Tim Robishaw's work)
- storage of measured parameters in MySQL database (table definitions)
- storage of extracted parameters (Q, U and RM spectra) in FITS files
- interface w EMU: how do we deal w multi-component sources? Do we just accept point sources?

Some uncertainty expressed by Whiting and Cornwell as to whether ASKAP computing will perform RM synthesis! A sentence was added to the ASKAP Science Processing Document from v1 to v2 in Section 2.11, saying:

"The standard continuum pipeline will use multi-frequency synthesis to represent the sky brightness by Taylor terms in frequency (see Section 3.8.7 for details). An alternative approach would be to deconvolve channels separately and creating a â€œcontinuum cubeâ€, most likely with all Stokes parameters. It is expected this will be possible for between âˆ≈ 30 and 300 channels, depending on the processing capability (this is still to be determined at time of writing). The output of this would be suitable for Rotation Measure Synthesis. The current plan is not to do RM synthesis in the pipeline processing."

After discussion, suggestion from Cornwell is that CASS can probably provide input needed for PPC, but for PPA we are on our own.

ACTION: Cornwell & Gaensler to establish whether CASS will perform RM synthesis or if we have to do this ourself. Fold this into discussion of whether we should implement Faraday synthesis instead.

Plans going forward for Purcell:

- demonstration of pipeline
- delivery of pipeline code to ASKAP computing group (reports 49, 62)

Gaensler: Understanding, based on Section 5 of ASKAP Processing Document (v2), is that we will specify algorithms for quality control, and that ASKAP computing will implement these. This will take data products from level 5 to level 6.

But Norris says default should be to look at outputs by eye.

ACTION: POSSUM quality control document to be developed

Larry: what are the core products/projects for POSSUM?

Â 1. RM grid: errors of <~6 rad/m^2 for sources brighter than 10 sigma; systematic errors at 1 rad/m^2 level don't matter?
Â 2. Info on intrinsic properties of compact sources
Â 3. Info on extended sources & diffuse emission

Note though that there are going to be some (many) sources in category #1 that have some level of complexity. Ask questions:
- is it good enough for the grid (even if it has some level of complexity?) i.e. is the observed RM dominated by pure rotation? Working definition could be cases where error in RM due to complexity is small compared to other sources of error
- this is basically like the old way of doing angle fitting, where we accept source with chi^2 better than some threshold

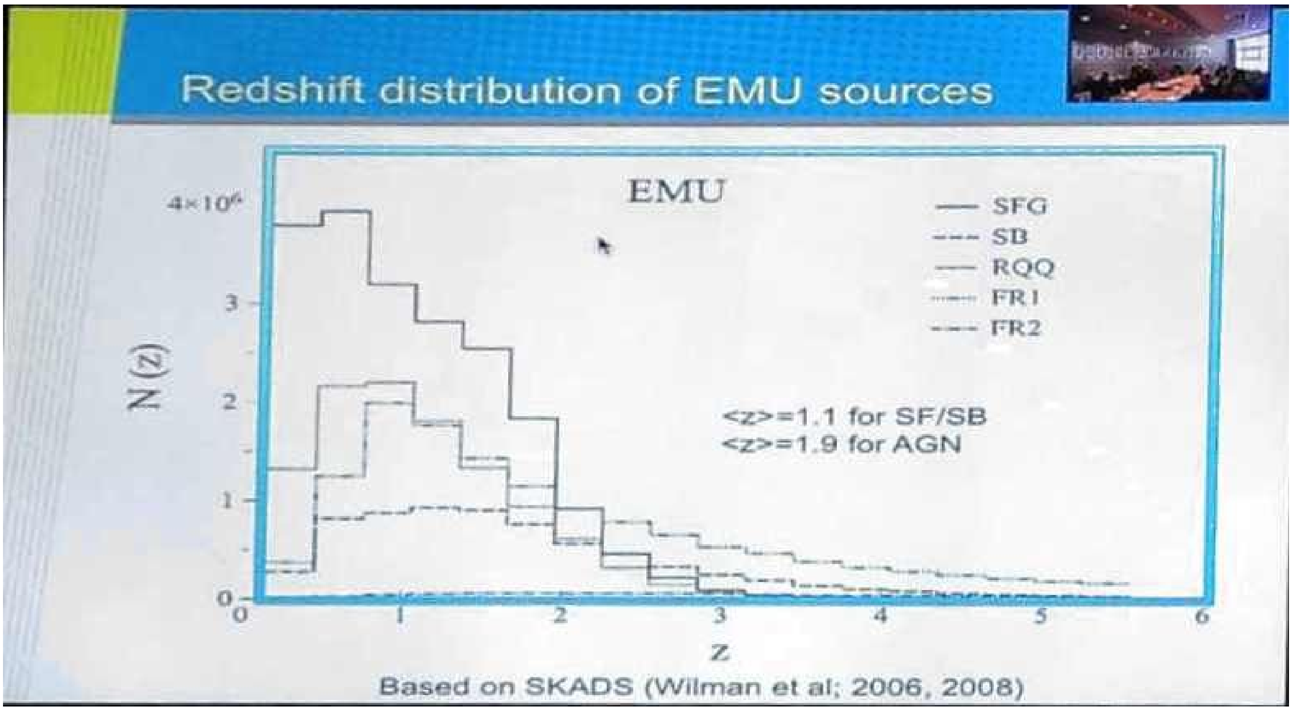
--> comes back to quality control criteria

Ray Norris - EMU & POSSUM (also with Andrew Hopkins, Nick Seymour)

EMU: 40x deeper than NVSS, 10 uJy over 75% of the sky. 70 million galaxies, 1100-1400 MHz.

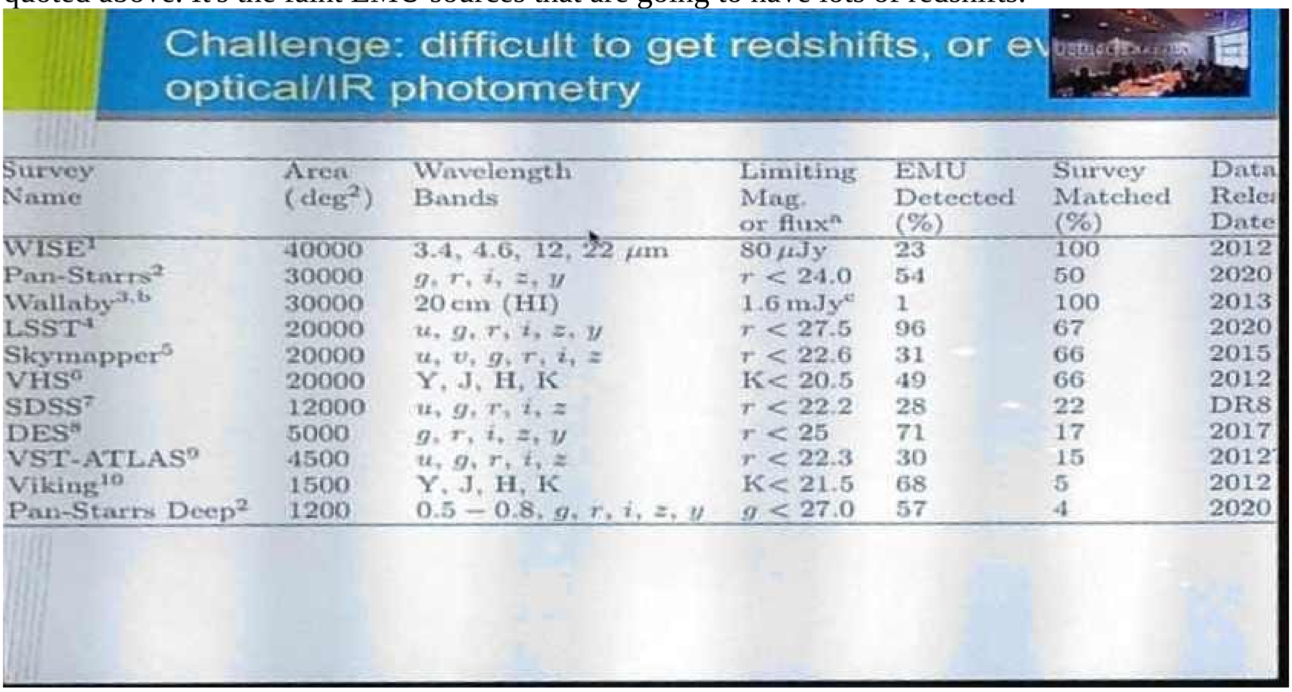
Science goals: now include cosmology

Redshift distribution:



Big challenge: optical/IR IDs and redshifts for all sources. Expect photometry for 50%-70% (at least one band), and photo-z for 30%. Spectroscopic redshift for 1% (mainly WALLABY, followed by SDSS). TAIPAN could double the number of spectroscopy redshifts.

For POSSUM: brightest sources in EMU (>5 mJy) are the most relevant ones. Most of these will be at AGNs. Strongest AGNs tend to be at high z and won't have optical counterparts or redshifts. Brightest star-forming galaxies will have redshifts. Â So if anything, fraction of bright EMU sources with redshifts (i.e. those relevant for POSSUM) will be *lower* than overall percentages quoted above. It's the faint EMU sources that are going to have lots of redshifts.



Survey Name	Area (deg^2)	Wavelength Bands	Limiting Mag. or flux[a]	EMU Detected (%)	Survey Matched (%)	Data Rele Date
WISE[1]	40000	3.4, 4.6, 12, 22 μm	80 μJy	23	100	2012
Pan-Starrs[2]	30000	g, r, i, z, y	$r < 24.0$	54	50	2020
Wallaby[3,b]	30000	20 cm (HI)	1.6 mJy[c]	1	100	2013
LSST[4]	20000	u, g, r, i, z, y	$r < 27.5$	96	67	2020
Skymapper[5]	20000	u, v, g, r, i, z	$r < 22.6$	31	66	2015
VHS[6]	20000	Y, J, H, K	K< 20.5	49	66	2012
SDSS[7]	12000	u, g, r, i, z	$r < 22.2$	28	22	DR8
DES[8]	5000	g, r, i, z, y	$r < 25$	71	17	2017
VST-ATLAS[9]	4500	u, g, r, i, z	$r < 22.3$	30	15	2012
Viking[10]	1500	Y, J, H, K	K< 21.5	68	5	2012
Pan-Starrs Deep[2]	1200	$0.5 - 0.8, g, r, i, z, y$	$g < 27.0$	57	4	2020

Pipeline will automate cross-IDs. Probably Bayesian algorithm.

Expect to cross-ID up to 70% of 70 million objects.
won't have optical/IR IDs.

Remaining sources (7 million sources) --> "Radio Zoo" - citizen identification

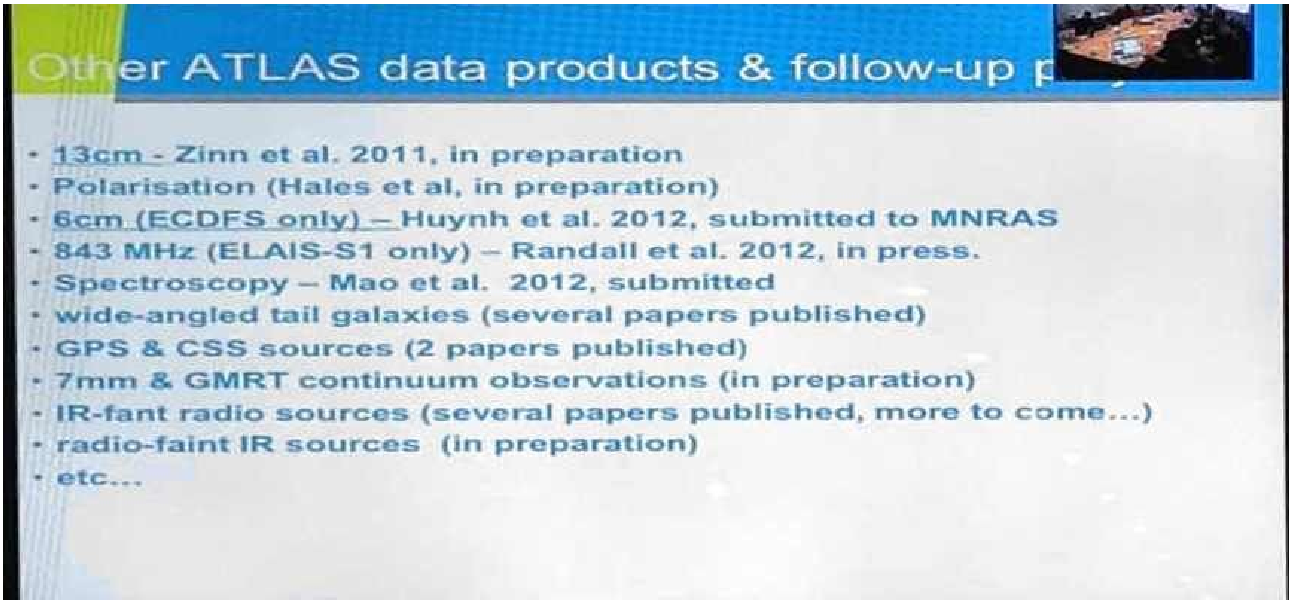
Phases of EMU:
1. Design Study 2009-2012
2. Commissioning 2012-2013
3. Early science with ASKAP-12 2013-2014?
4. EMU full survey 2014-2015?

ATLAS project: 2000 galaxies so far (rms ~ 30 uJy). Expect 10,000 galaxies at completion of project.

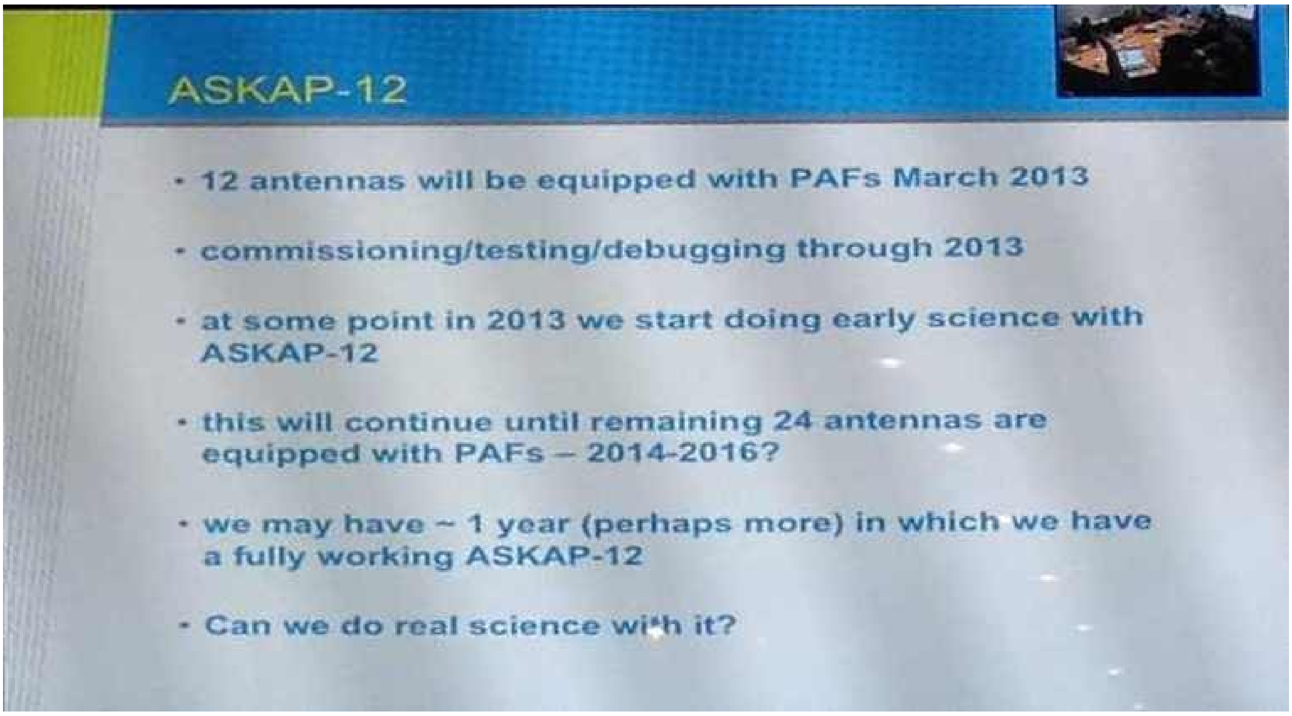
ATLAS DR2 (Hales): all pre-CABB data, focus on polarisation; slightly depeer in CDFS than DR1

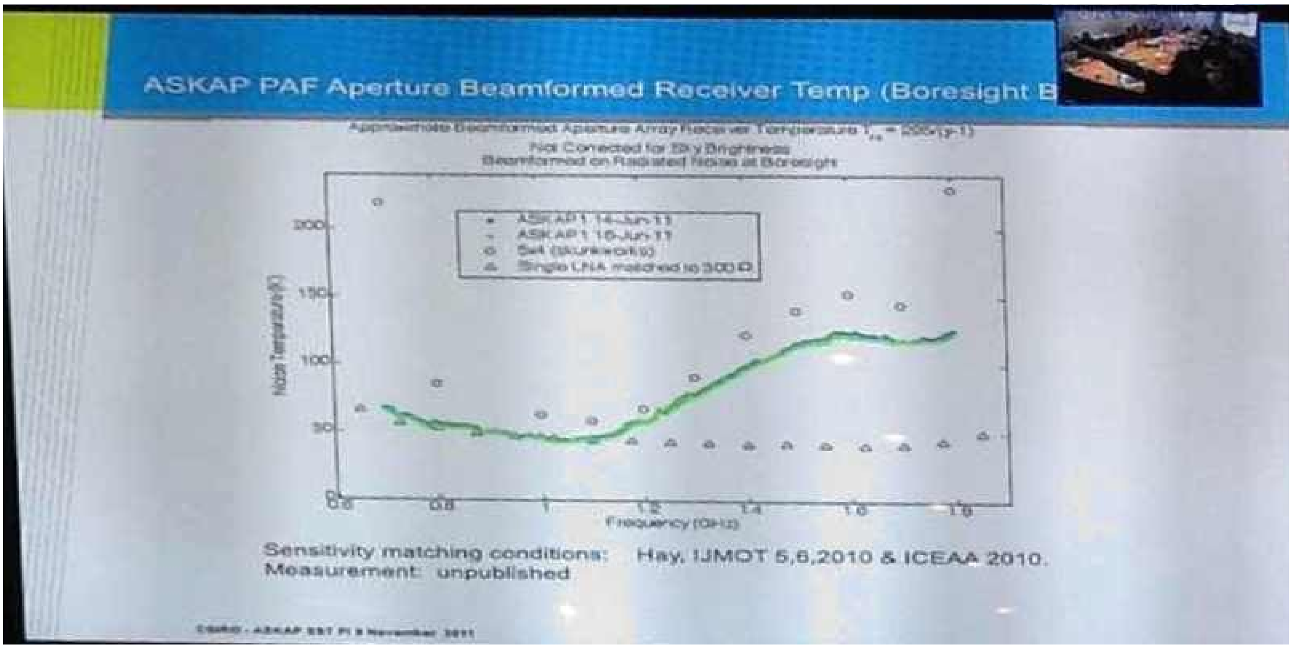
ATLAS DR3 (Banfield): final definitive ATLAS release, including pol; all 20cm data (pre- and post-CABB).

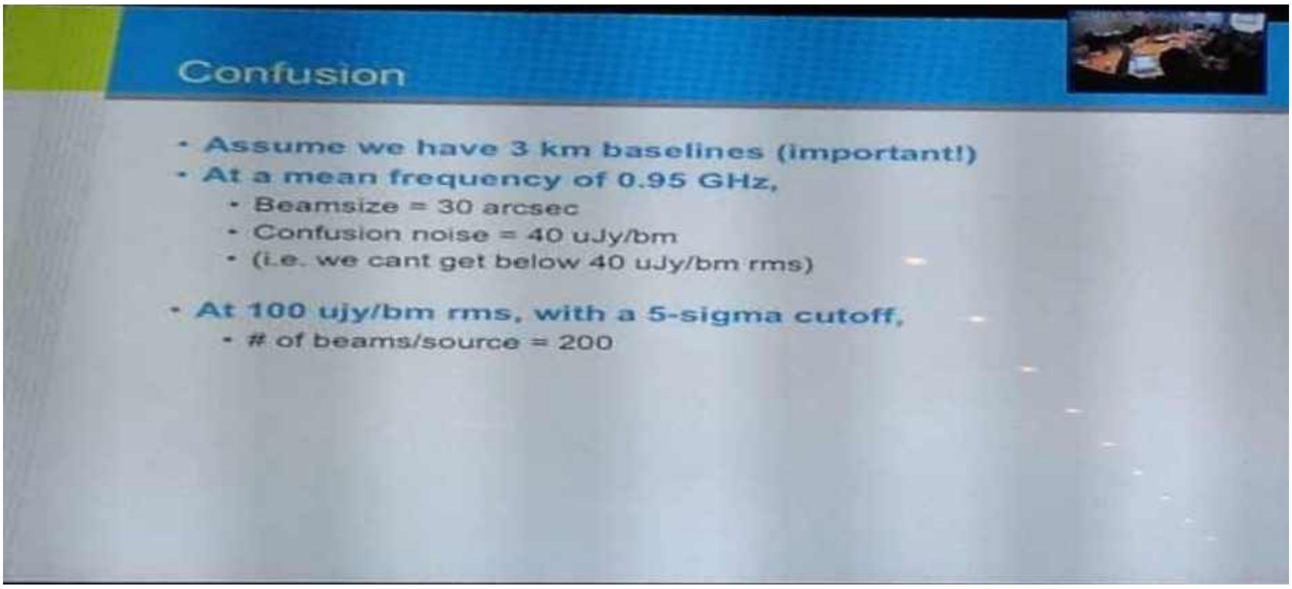
ATLAS fields go to same depth as EMU & POSSUM, over smaller FoV --> we know what EMU & POSSUM will look like!

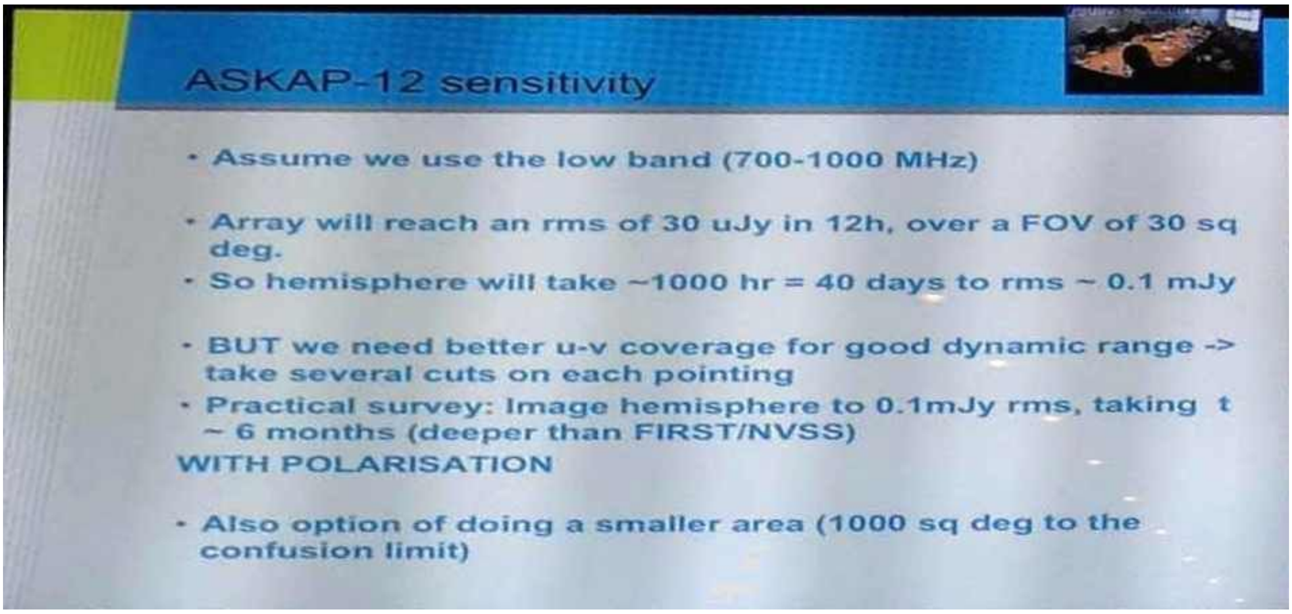


EMU science with ASKAP-12:


ASKAP-12
12 antennas will be equipped with PAFs March 2013
commissioning/testing/debugging through 2013
at some point in 2013 we start doing early science with ASKAP-12
this will continue until remaining 24 antennas are equipped with PAFs – 2014-2016?
we may have ~ 1 year (perhaps more) in which we have a fully working ASKAP-12
Can we do real science with it?

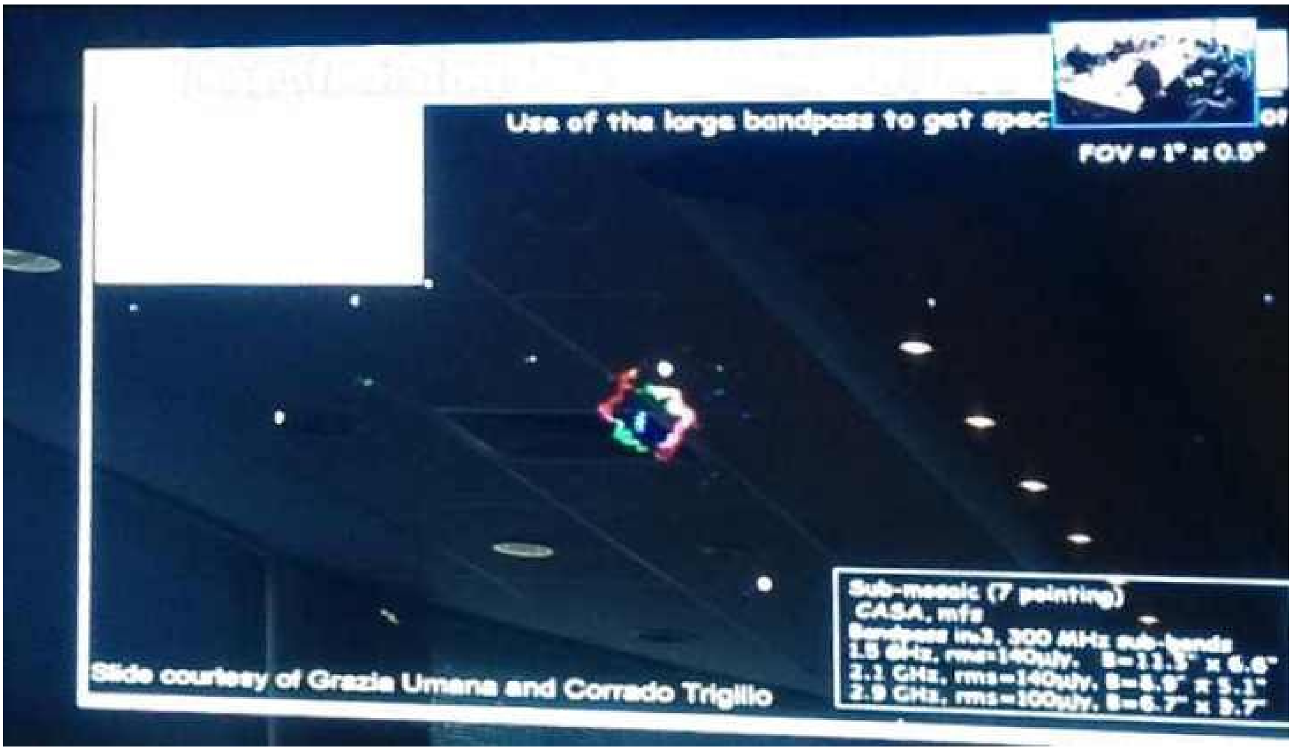

ASKAP PAF Aperture Beamformed Receiver Temp (Boresight B
Not Corrected for Sky Brightness
Beamformed on Radiated Noise at Boresight
ASKAP1 14-Jun-11
ASKAP1 16-Jun-11
Single LNA matched to 300 Ω
Noise Temperature (K)
Frequency (GHz)
Sensitivity matching conditions: Hay, IJMOT 5,6,2010 & ICEAA 2010.
Measurement: unpublished





* New project **SCORPIO**: blind survey (like ATLAS) but in Galactic Plane. 4 deg^2 to 10 uJy RMS and 10" resolution (same as EMU). ATCA project C2515 to observe 1 deg^2 as pilot - observing in Apr 2011 and Jul 2012.

Pilot results: new SNR, only ~40 arcsec across? Coords l = ???, b = 0



* Areas of collaboration w POSSUM:
- calibration and imaging
- source extraction
- value-addd data / VO
- quality control
- science (also need pol in real-time for cross-IDs and statistical redshifts)
- short spacing data (is single dish needed for EMU?)

* Source extraction
- effort has been split into compact sources vs extended sources
- extended: deferred into value-added products
- compact: Hancock (AEGEAN); Huynh; Hales (BLOBCAT)

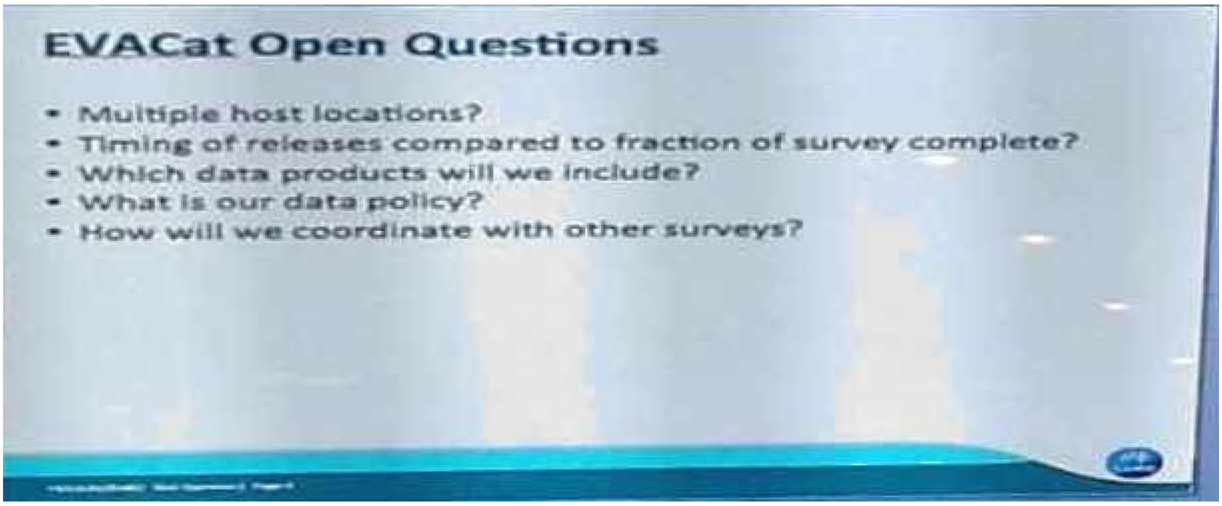
* EMU data products

- "EMU in 3D" - statistical redshifts for *all* 70 million EMU sources

... discussion of data rights, publication policy, student policy, etc etc.

Larry: wants to call the POSSUM value-added catalogue "PVACat" instead of PPA, to be completely parallel with EMU's EVACAT.

Also wants to rename the level 6 products: PPC (POSSUM pipeline catalogue), EPC (EMU pipeline catalogue)



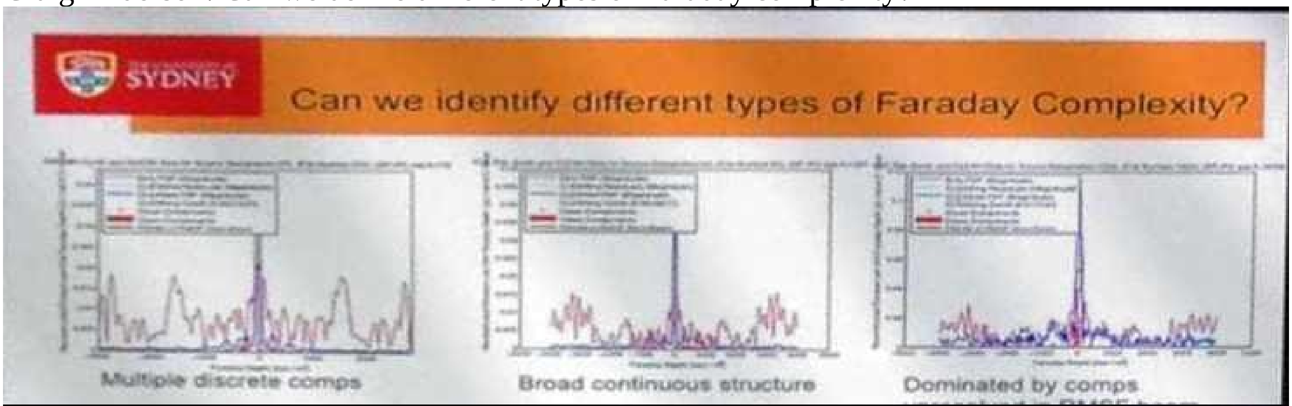
RM Synthesis discussion

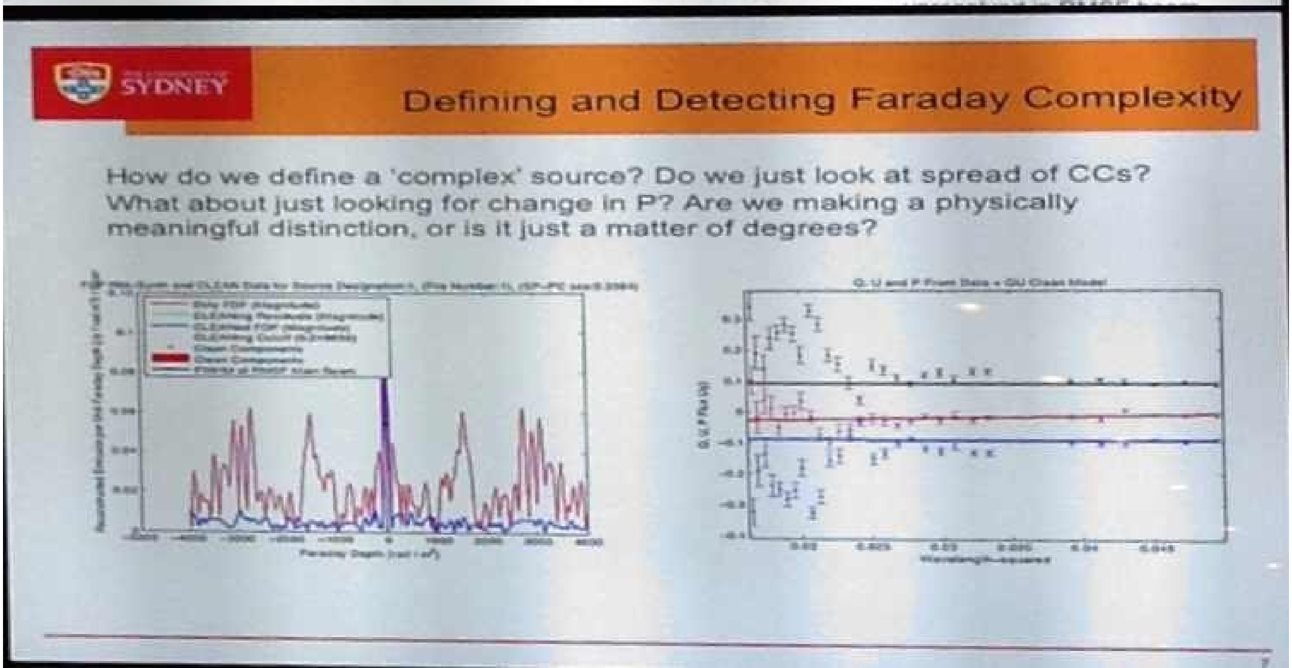
Larry Rudnick: Real vs perfect world experimentsÂ

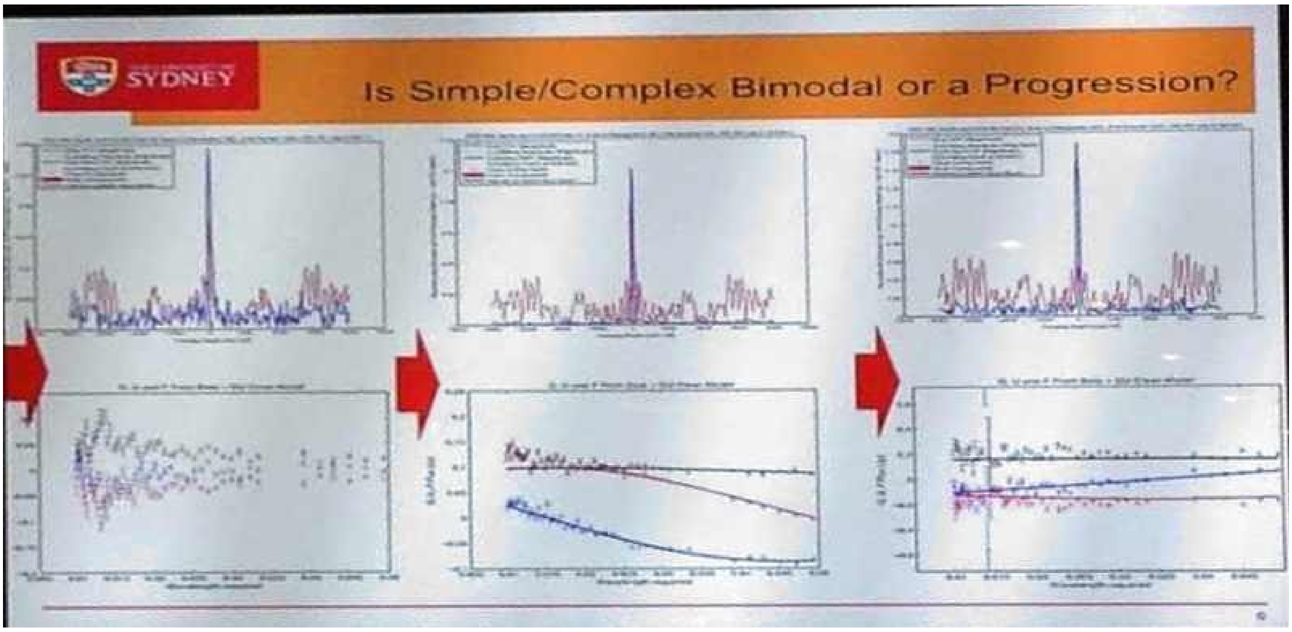
Jeroen Stil: sparse RM synthesis ([Andrecut et al. 2012])

Rainer Beck: Wavelet-based RM synthesis ([Frick et al. 2010]; [Frick et al. 2011]; [Beck et al. 2012]).

Craig Anderson: Can we define different types of Faraday complexity?






SYDNEY
Is Simple/Complex Bimodal or a Progression?