LOFAR
Introduction to data reduction

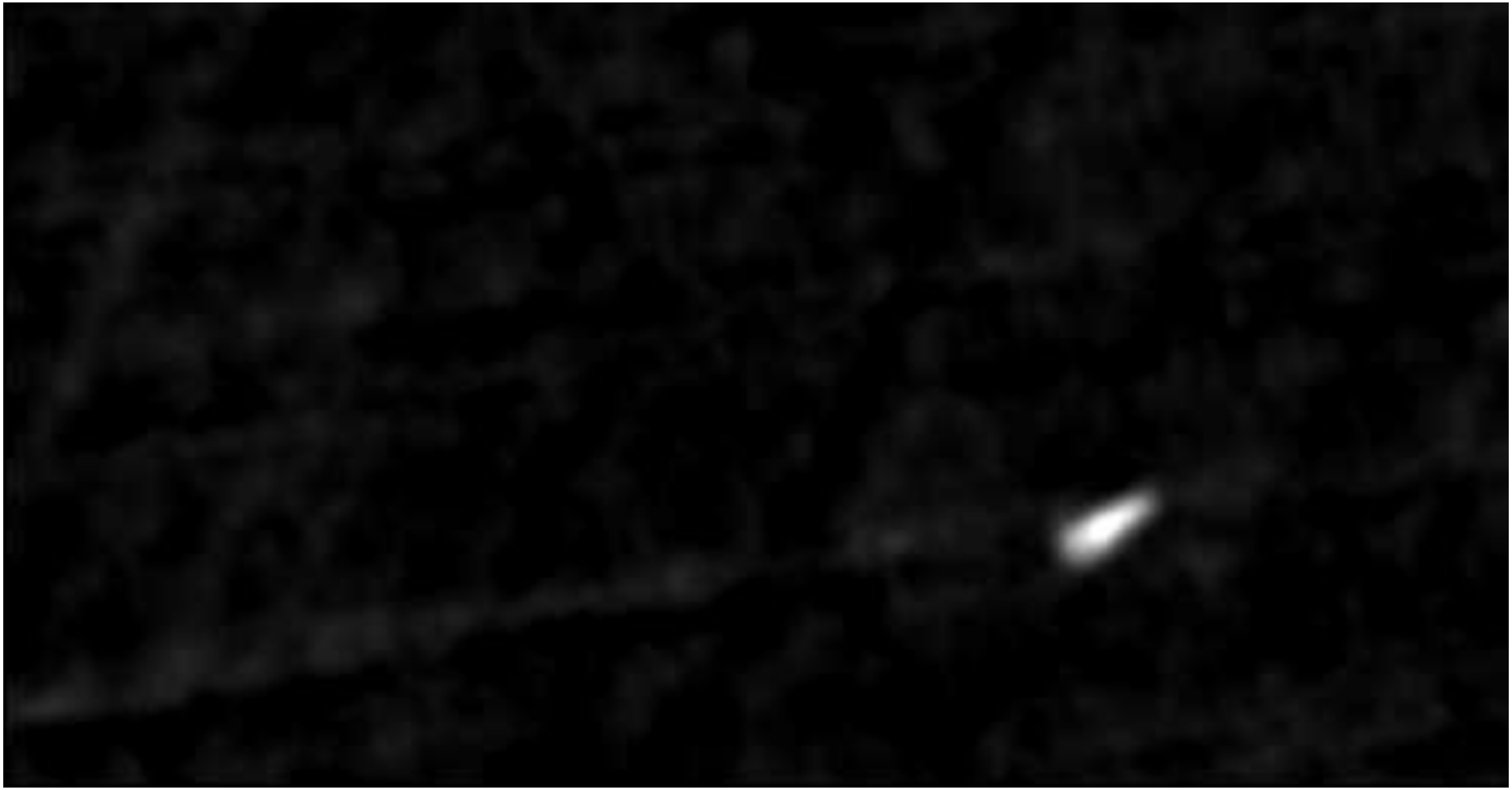
Krzysztof T. Chyży
Jagiellonian University
Kraków, Poland

Aleksandar Shulevski, Natalia Nowak, Błażej Nikiel-Wroczyński, Wojciech Jurusik, Robert Drzazga

Based on the LOFAR school November 2014:
Roberto Pizzo, John McKean, Wilfred Frieswijk
Tammo Jan Dijkema et al.

IC10 X-1,Credit: Aurore Simonnet/Sonoma State University/NASA

(Prawie) nic nie jest doskonałe… szczególnie na długich falach



74 MHz Virgo A VLA

LOFAR cookbook

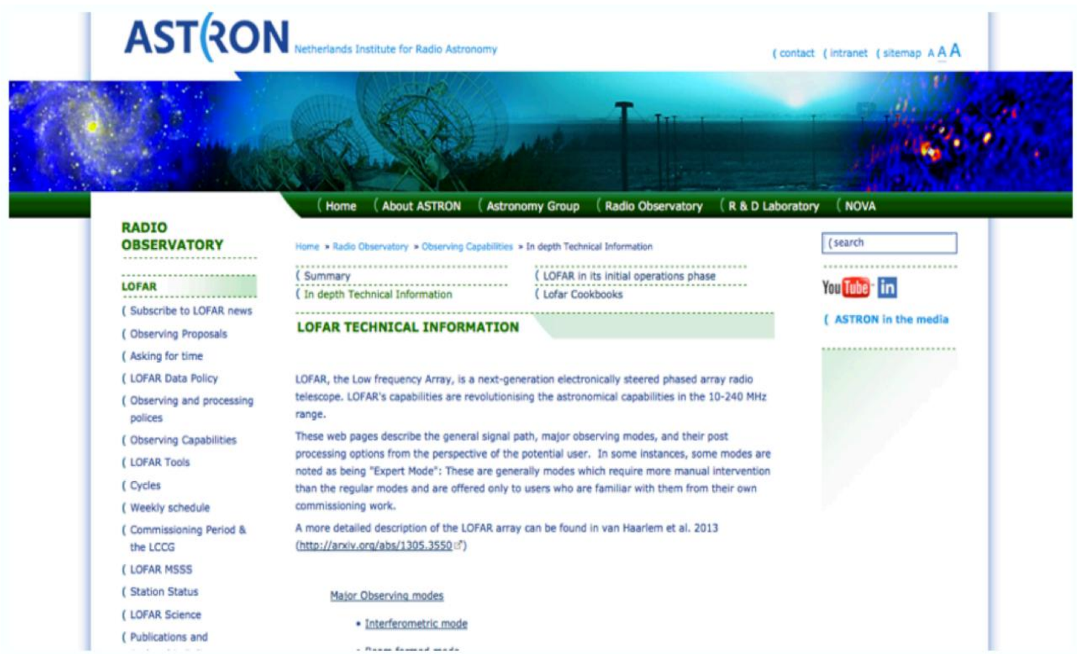
https://www.astron.nl/radio-observatory/lofar/lofar-imaging-cookbook

LOFAR wiki

http://www.lofar.org/wiki/doku.php?id=start

ASTRON website

http://www.astron.nl/radio-observatory/astronomers/technical-information/lofar-technical-information

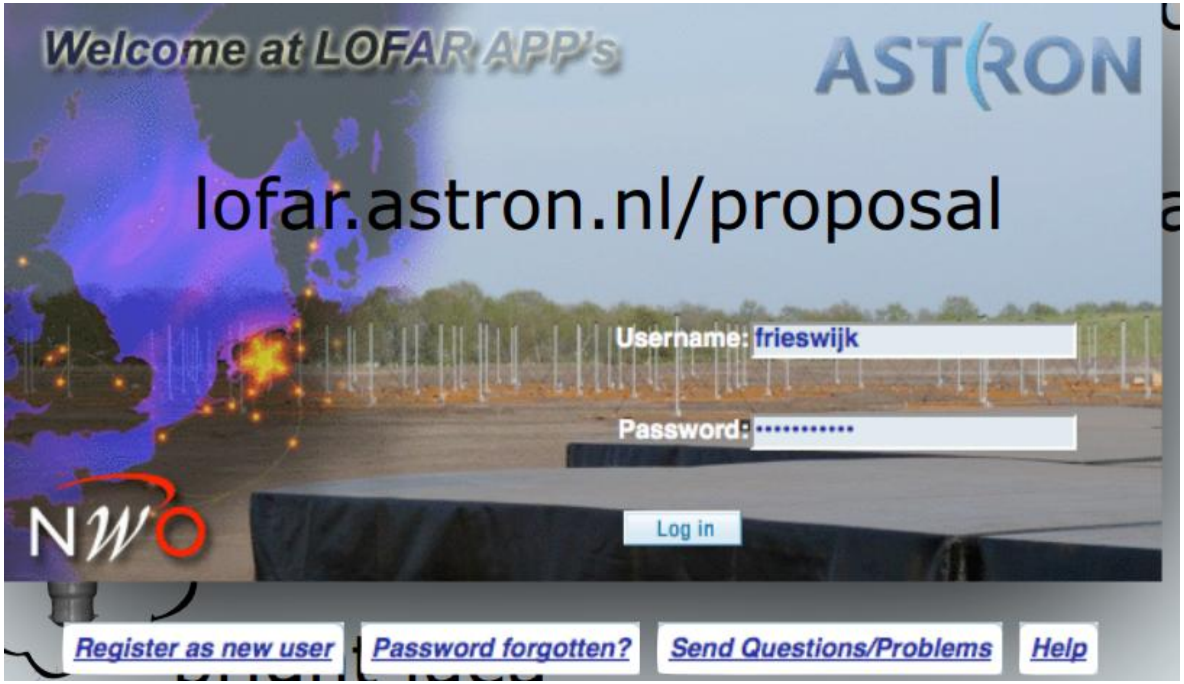


LOFAR PL

www.oa.uj.edu.pl/lofar

LOFAR wiki: register by e-mail ro_wiki@astron.nl

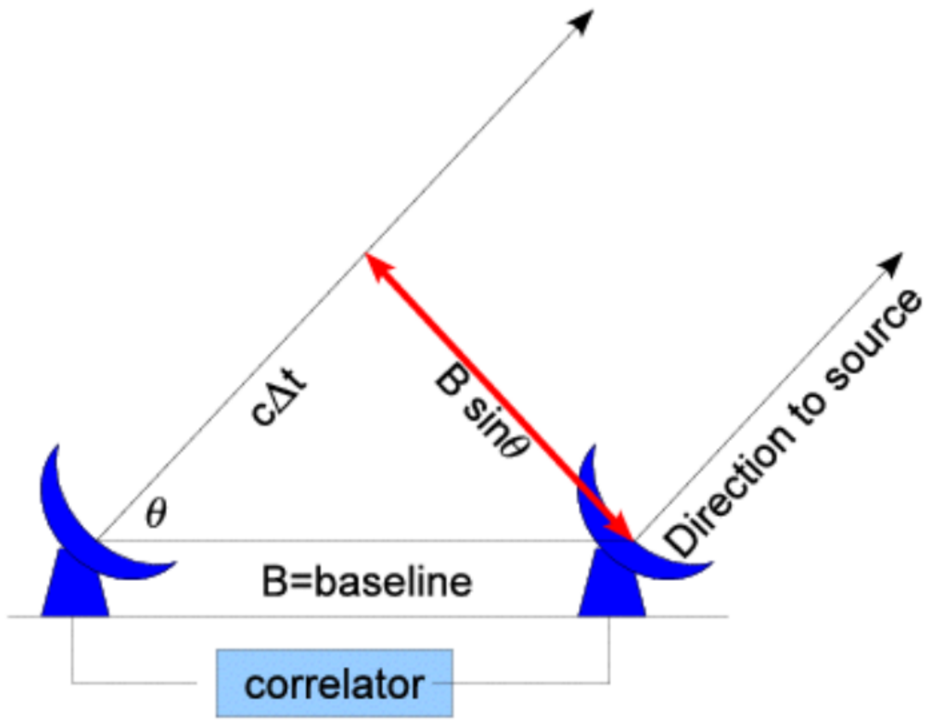
Submitting proposals: register http://lofar.astron.nl/proposal



Proposals call: Mid-September 2016
After accepted: acess created by RO to MoM, LTA, CEP clusters

Radio interferometry

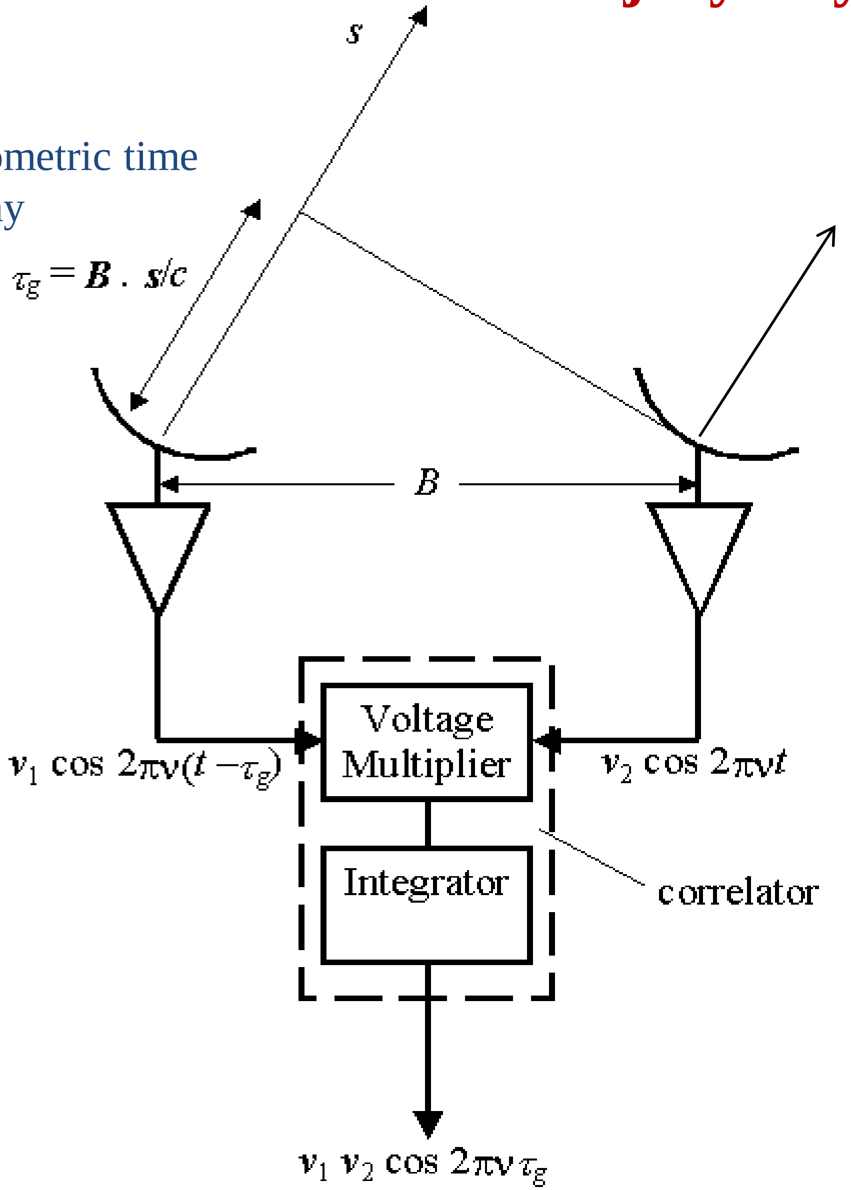
- Interferometric techniques essential for competitive low frequency radio astronomical observations



$$\text{Angular resolution } \Theta \approx \frac{\lambda}{B_{\max}}$$

Pojedynczy interferometr

Geometric time delay



$$P[\underbrace{\cos(\omega\tau_g)}_{\text{Unchanging}} + \underbrace{\cos(2\omega t - \omega\tau_g)}_{\text{Rapidly varying,}}]$$

$$R_C = P\cos(\omega\tau_g)$$

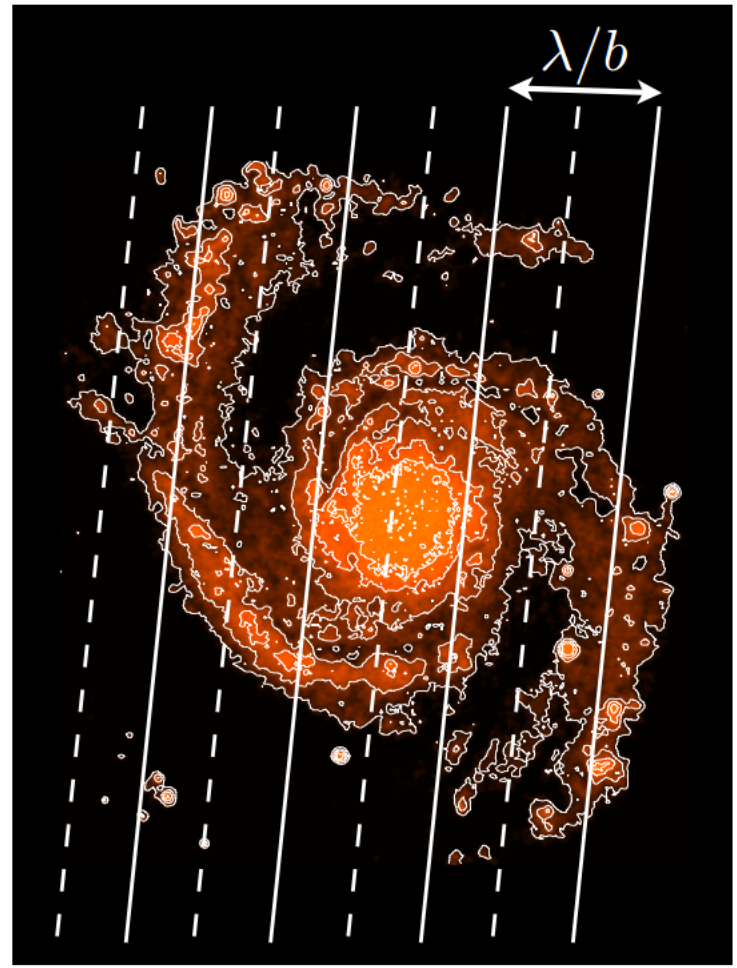
Schematic

- The cosine correlator can be thought of as casting a sinusoidal fringe pattern on the sky (of angular scale λ/b). The correlator multiplies the source intensity distribution by this fringe, and integrates the product over the sky.

$$R_c = \int\int I_\nu(\vec{s}) \cos(2\pi\nu\vec{b}\cdot\vec{s}/c)\,\mathrm{d}\Omega$$

source brightness ($I_\nu(\vec{s})$)

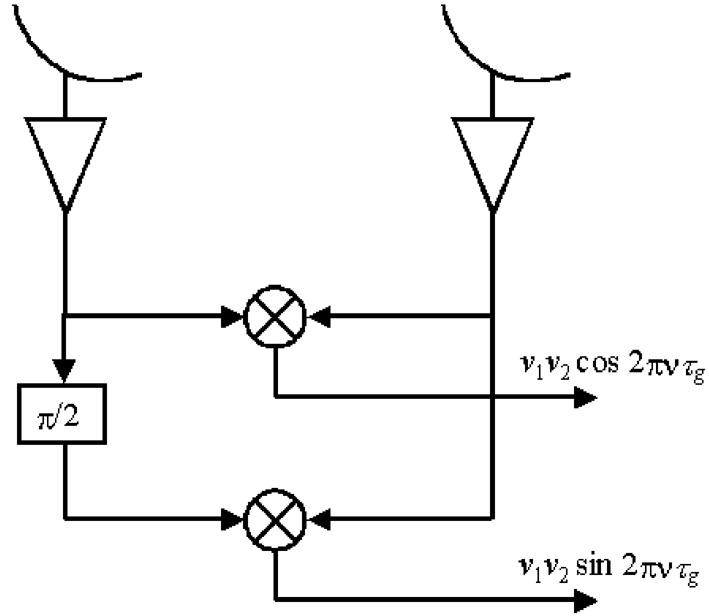
fringe pattern ($\cos(2\pi\nu\vec{b}\cdot\vec{s}/c)$)



- The orientation of the fringe is set by the baseline geometry

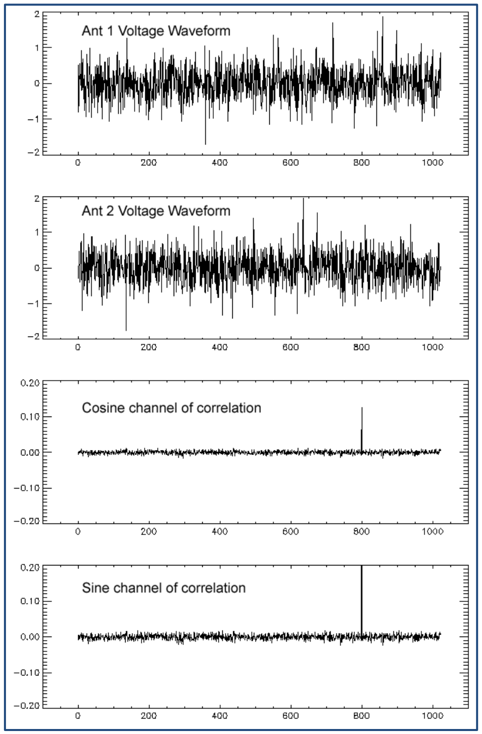
  The fringe separation is set by baseline length, and the observing wavelength

Interferometr z zespolonym korelatorem



$$Rc = \int_{zr} B(s)\,P(s)\cos\,[2\pi\overrightarrow{D_\lambda}\cdot\vec{s}]d\Omega$$

$$Rs = \int_{zr} B(s)\,P(s)\sin\,[2\pi\overrightarrow{D_\lambda}\cdot\vec{s}]d\Omega$$



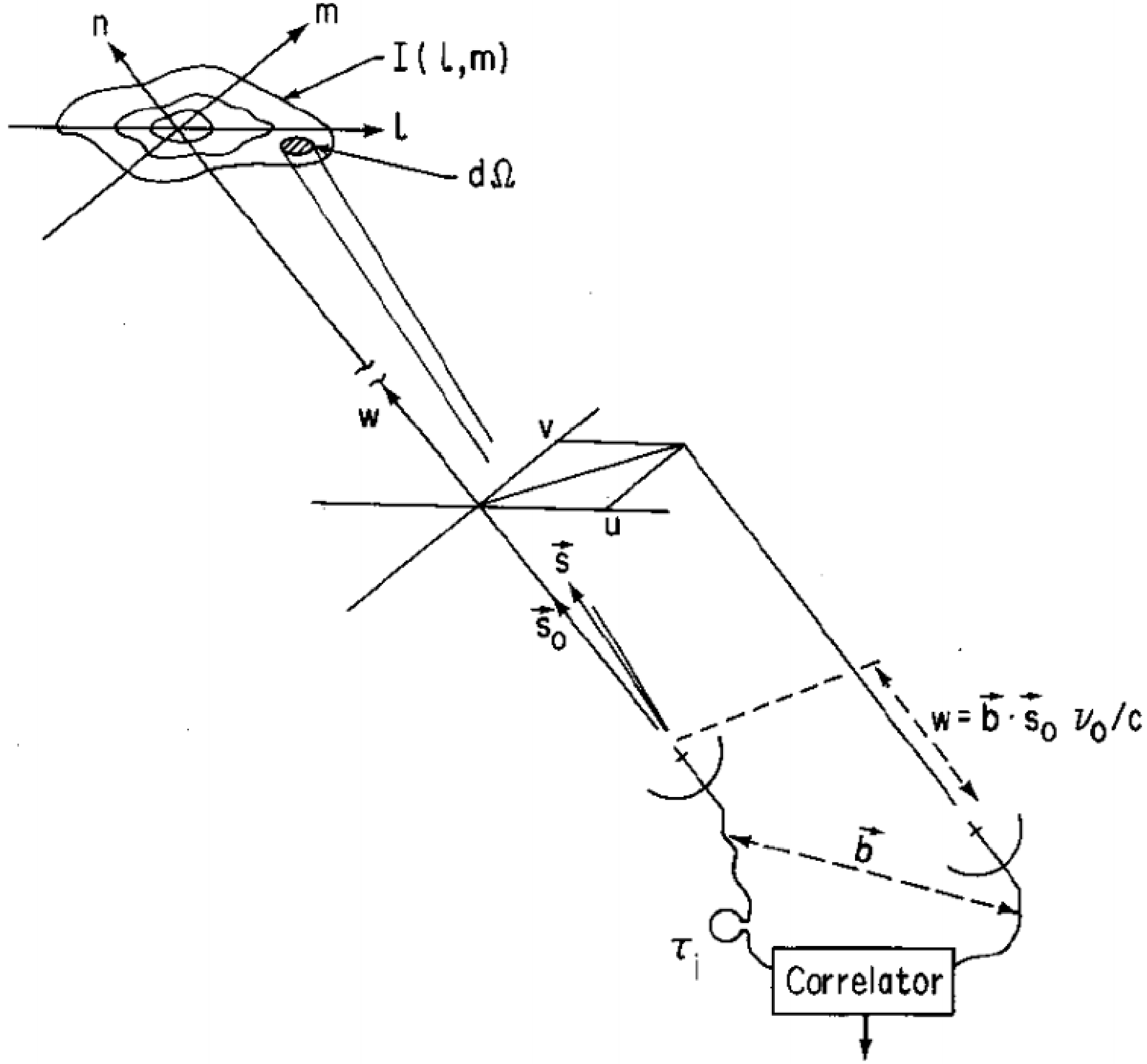
Zespolona funkcja widzialności:

$$V \equiv Rc - iRs = Ae^{i\phi}$$

$$V(\vec{s}) = \iint B(\vec{s})\,P(\vec{s})\mathrm{e}^{-2\pi i\,\overrightarrow{D_\lambda}\cdot\vec{s}}d\Omega$$

$$\overrightarrow{D_\lambda}\cdot\vec{s} = ul + vm + nw \cong ul + vm$$

Płaszczyzna UV



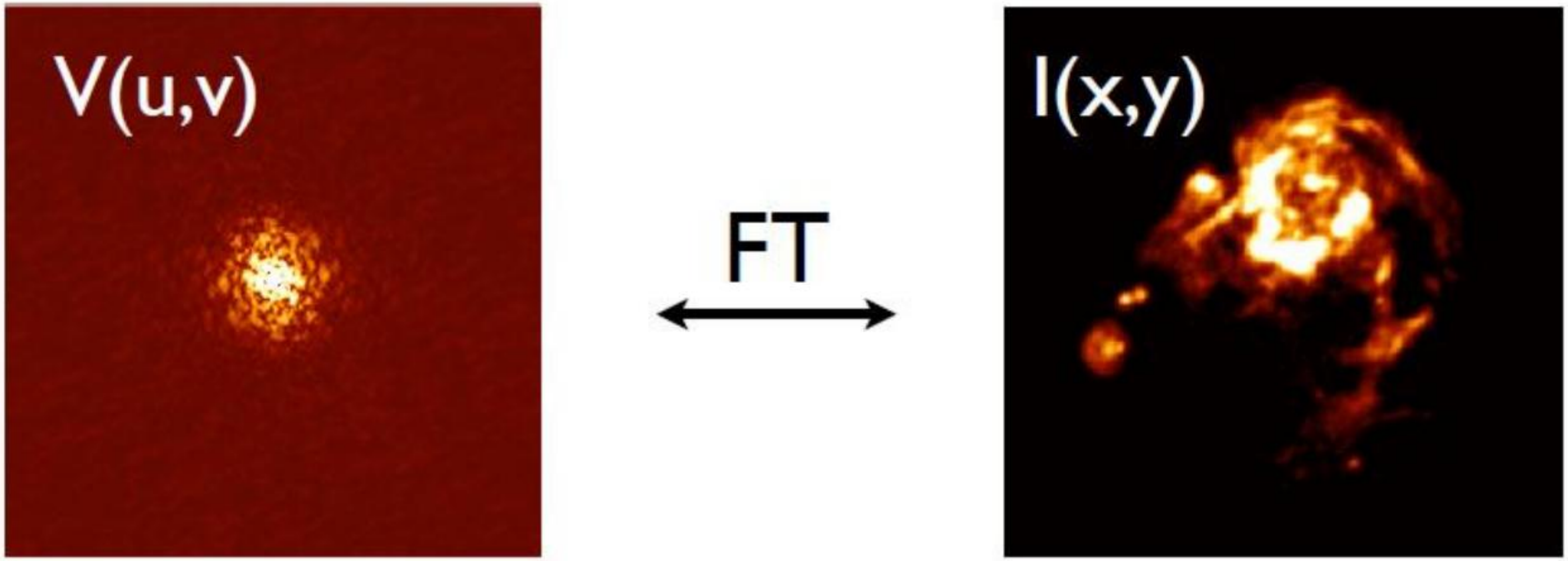
$$V(u,v) = \iint B(l,m)P(l,m)e^{-2\pi i(ul+vm)}\,dl\,dm$$

Synteza apertury

Funkcja widzialności dla pojedynczego interferometru (określone wartości u,v) ma postać:

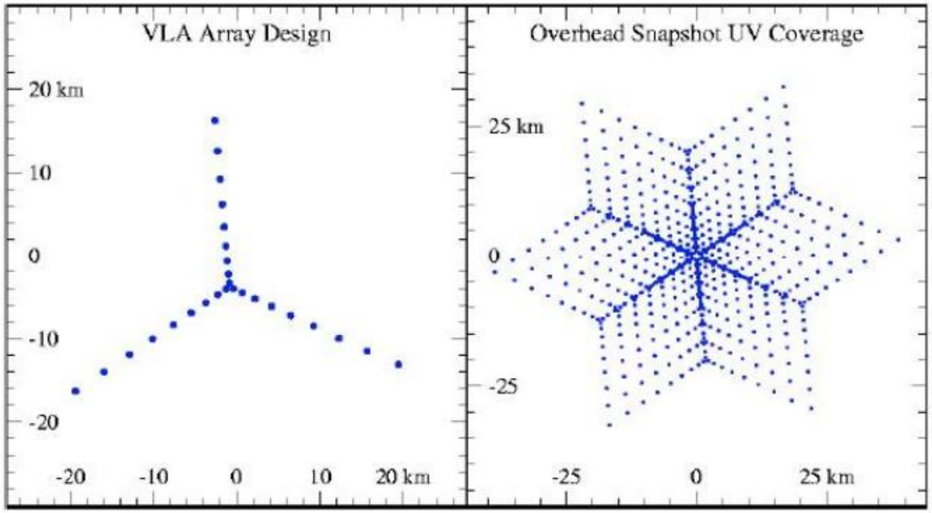
$$V(u,v) = \iint B(l,m)P(l,m)e^{-2\pi i(ul+vm)}\,dl\,dm$$

Całkowanie można rozciągnąć na $\pm\infty$ gdyż rozkład źródła na niebie (B) jest i tak skończony, zlokalizowany. Traktując (u.v) jako zmienne widzimy, że **zespolona funkcja widzialności jest transformatą Fouriera jasności źródła.**



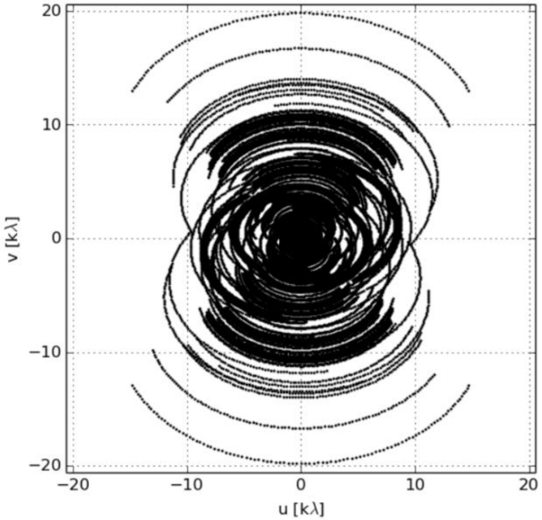
Zatem mając zbiór różnych (u,v) czyli różnych pojedynczych radiointerferometrów możemy dokonać odwrotnej **transformaty Fouriera funkcji widzialności uzyskując rozkład jasności na niebie** pomnożony przez charakterystykę pojedynczej anteny:

$$B(l,m)P(l,m) = \iint_{-\infty}^{+\infty} V(u,v)\,e^{2\pi i(ul+vm)}du\,dv$$



- Dużą aperturę syntezujemy z mniejszych
- Synteza możliwa przez przestawianie anten (w dowolnym czasie)
- Rotacja ziemi – **supersynteza (synteza rotacyjna)**
- Zasady syntezy apertury zostały sformułowane w 1959 r. przez **Martina Ryle'a** (Ryle 1960, Ryle i Hewish 1960), za co uzyskał **nagrodę Nobla**.

LOFAR



Luki w pokryciu UV

$$B(l,m)P(l,m) = \iint_{-\infty}^{+\infty} V(u,v)\, e^{2\pi i(ul+vm)} du\, dv$$

Całkowanie wzoru na B jest niemożliwe gdyż nigdy nie mamy pełnego pokrycia płaszczyzny UV.

Funkcja wagowa *w* (w płaszczyźnie UV)

- *w(u,v)*=1 – jest baza interferometru
- *w(u,v)*=0 – nie ma

$$B(l,m)P(l,m) * W(l,m) = \iint_{-\infty}^{+\infty} V(u,v)\, w(u,v)\, e^{2\pi i(ul+vm)} du\, dv$$

gdzie: *W(l,m)= TF w(u,v).*

Powyższa konwolucja odpowiada skanowaniu nieba wiązką anteny mającej charakterystykę *W(l,m)*. Jest to tzw. **zsyntezowaną wiązką** układu anten (Point Spread Function), „**brudna" wiązka**

Dla dyskretnej ilości danych obserwacyjnych:

$$W(l,m) = \sum_k w_k\, e^{2\pi i(u_k l+v_k m)}$$

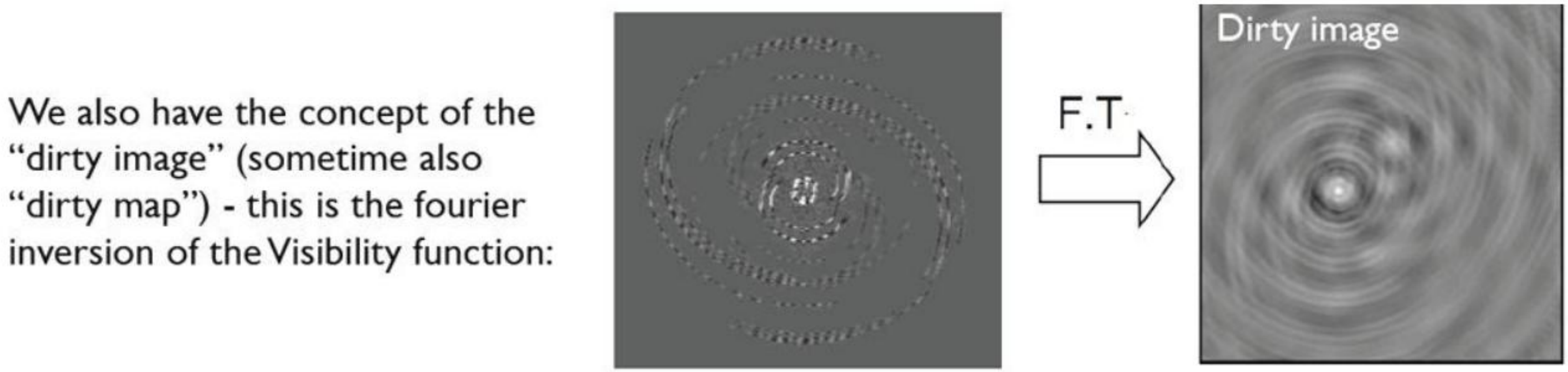
k- jest kolejnym punktem płaszczyzny UV.

Odtwarzanie rozkładów jasności metodą bezpośrednią – brudna mapa

Brudną mapą nazywamy rozkład jasności uzyskany bezpośrednio z transformaty Fouriera funkcji widzialności zniekształcony niepełnym pokryciem płaszczyzny UV:

$$B^D(l,m) = B(l,m)P(l,m) * W(l,m) = \sum_k w_k \, V(u_k, v_k) e^{2\pi i(u_k l + v_k m)}$$

k- jest kolejnym punktem płaszczyzny UV.



Szybka Transformata Fouriera.

- Zbiór punktów na płaszczyźnie UV nie tworzy regularnej siatki. Dlatego dokonuje się interpolacji danych aby uzyskać dane w oczkach siatki regularnej o rozmiarach równych 2^n gdzie n dowolne.
- W praktyce najczęściej rozmiary map to 512x512, 1024x1024, 2048x2048.
- Możemy wtedy zastosować tzw. szybką transformatę Fouriera. Przyśpiesza to wielokrotnie rachunki.

Metoda CLEAN

Najczęściej stosowaną metodą do uzyskiwania rozkładów jasności z danych z systemów interferometrycznych jest tzw. metoda CLEAN (czyszczenia). Została wprowadzona przez Högbom (1974). Doczekała się też różnych modyfikacji.

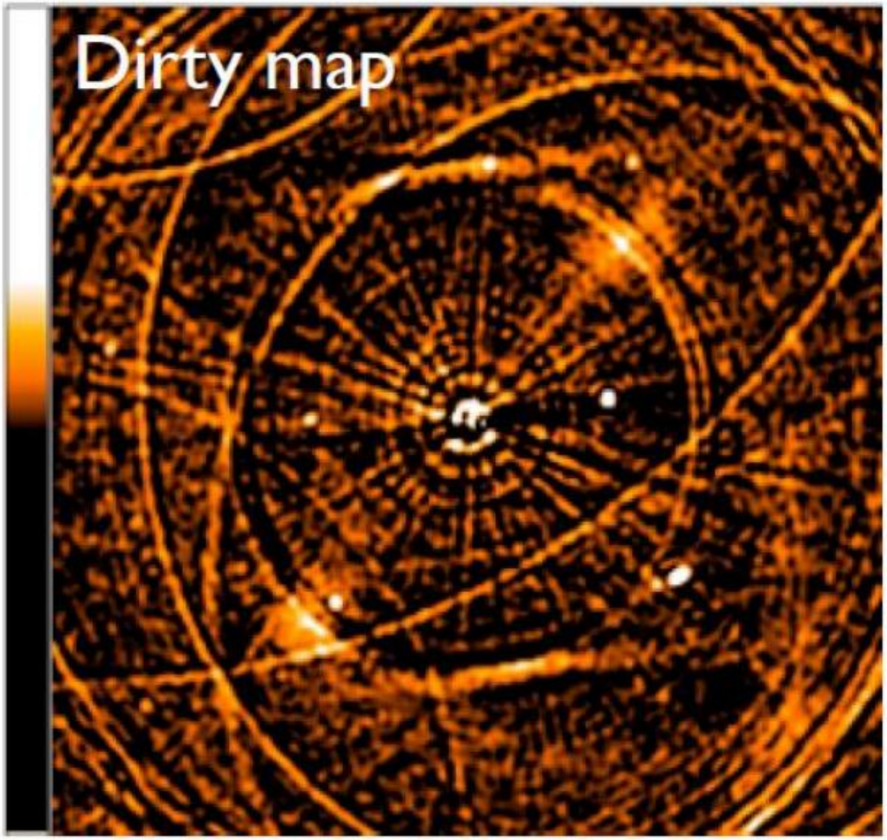
W oryginalnej metodzie CLEAN rozwiązujemy równanie na brudną mapę zakładając, że rozkład jasności można przybliżyć **sumą źródeł punktowych** o amplitudach A_k i położeniu (l_k, m_k) oraz pewnego rezidualnego rozkładu (bardzo słabe źródła, szumy):

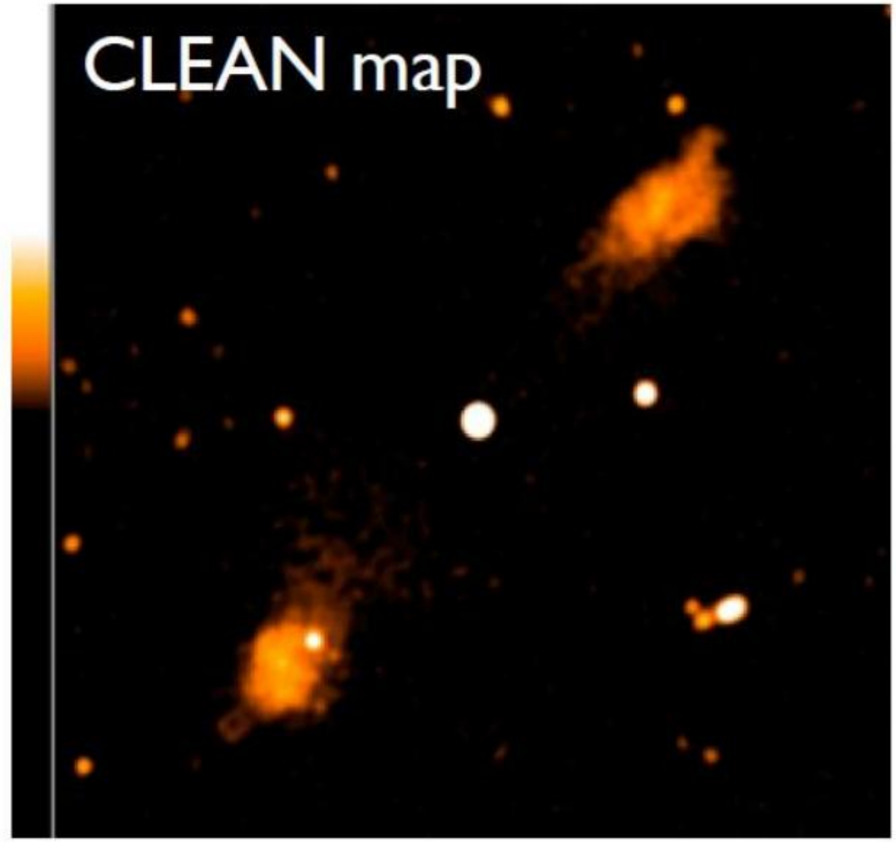
$$B^D(l,m) = \sum_k A_k(l_k, m_k) W(l - l_k, m - m_k) + B_r(l,m)$$

Źródła punktowe skonwoluowane są w powyższym wzorze z brudną wiązką.

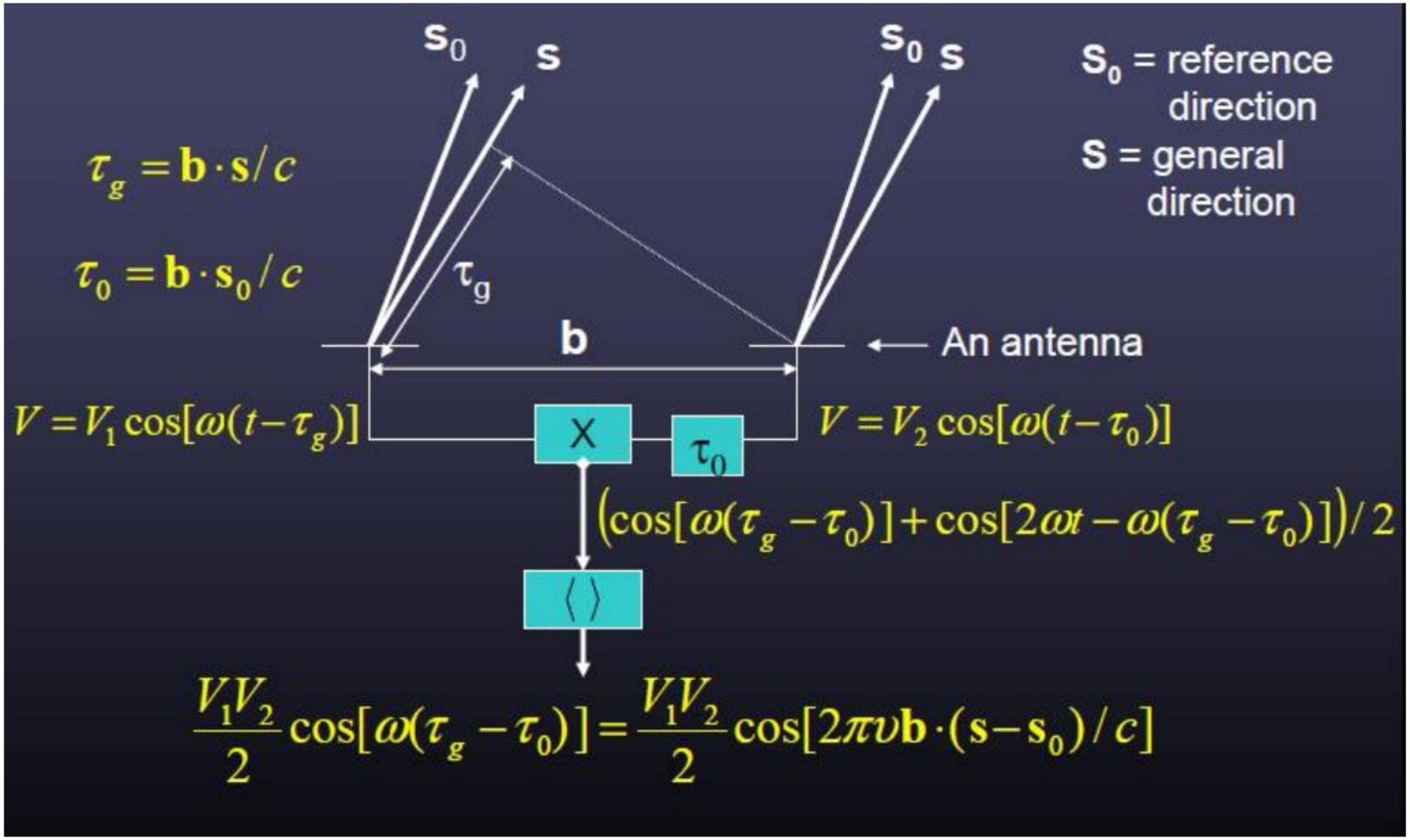
Równanie to nie da się rozwiązać analitycznie i rozwiązujemy je iteracyjnie:

1. Wyznaczyć brudną mapę i brudną wiązkę za pomocą inwersji fourierowskiej
2. Znaleźć miejsce (punkt) na brudnej mapie w którym jasność jest największa
3. Odjąć z brudnej mapy brudną wiązkę w znalezionym miejscu (odjęcie źródła punktowego), o jasności będącej ułamkiem 0<q<1 (z reguły ok. 0.8-0.9) jasności źródła w tym punkcie. Zapisać pozycję i amplitudę odjętego źródła (tzw. CLEAN component)
4. Powrócić do punktu 2 z wynikową brudną mapą (nowy rozkład jasności) albo zakończyć iterację jeśli: znajdowany kolejny punkt o największej jasności jest mniejszy od zadanej wartości (np. 3xszumy na mapie); bądź przekroczyliśmy zadaną liczbę kroków (np 1000); bądź rozkład jasności przestaje się zmieniać.
5. Dodać do brudnej mapy z ostatniego kroku (mapy rezidualnej B_r) odjęte źródła - CLEAN components - ale z tzw. **czystą wiązką** (zastąpi W powyższym równaniu). Czystą wiązkę uzyskuje sie jako wiązkę gaussowską najlepiej dopasowaną do centralnej części brudnej wiązki. Otrzymamy w ten sposób czystą (wyczyszczoną) mapę.

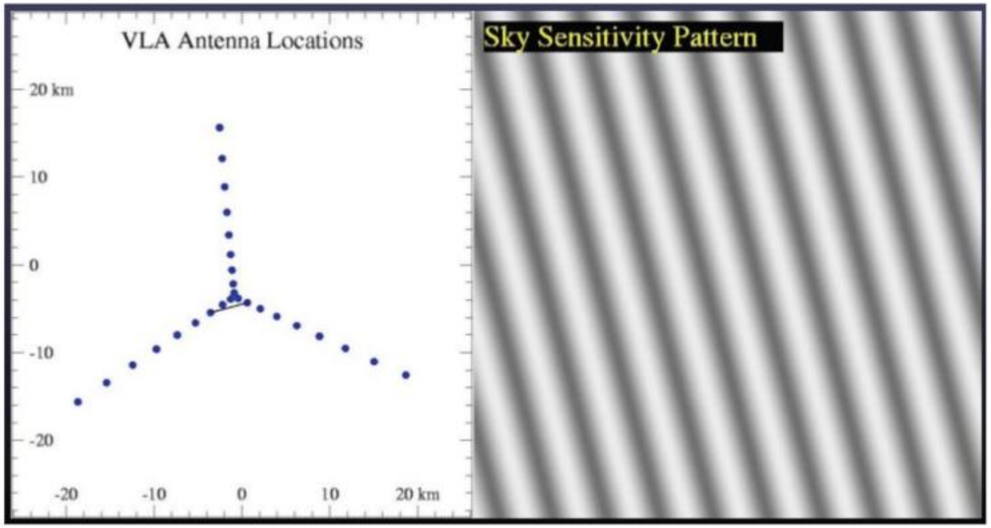

Dirty map


CLEAN map

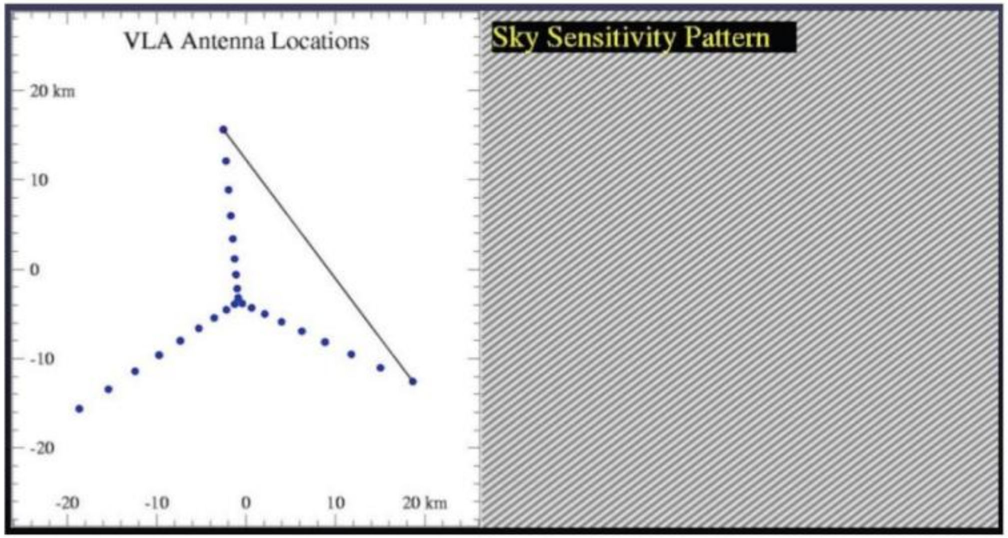
Rozpoznawanie i usuwanie błędów obrazów.



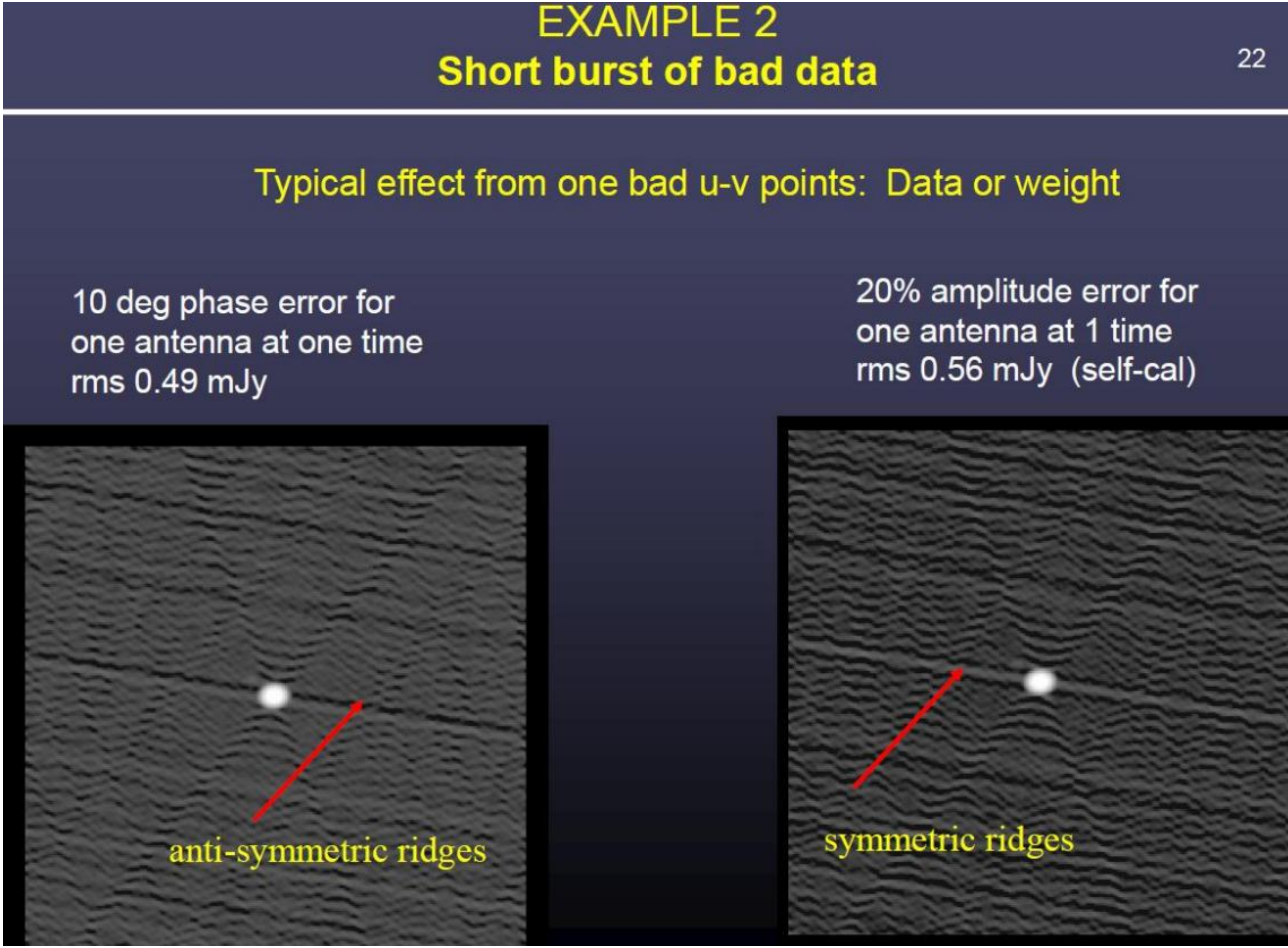
Przesunięcia w fazie równoważne są opóźnieniom czasowym, a te z kolei związane są z położeniem obiektu na niebie.



Fringes projected on to the sky for a short VLA baseline



Zatem występujące określone struktury na mapie (powtarzające się w jakimś kierunku i z jakąś częstością) zdradzają ich pochodzenie.



LOFAR – krótkie piki z powodu jonosfery i RFI

Kalibracja

Kalibracja - proces wyznaczania zespolonego zysku (gain G) pozwalającego właściwie wyskalować obserwowaną funkcję widzialności: $V_{ij}^{obs} = G_{ij} V_{ij}^{true}$

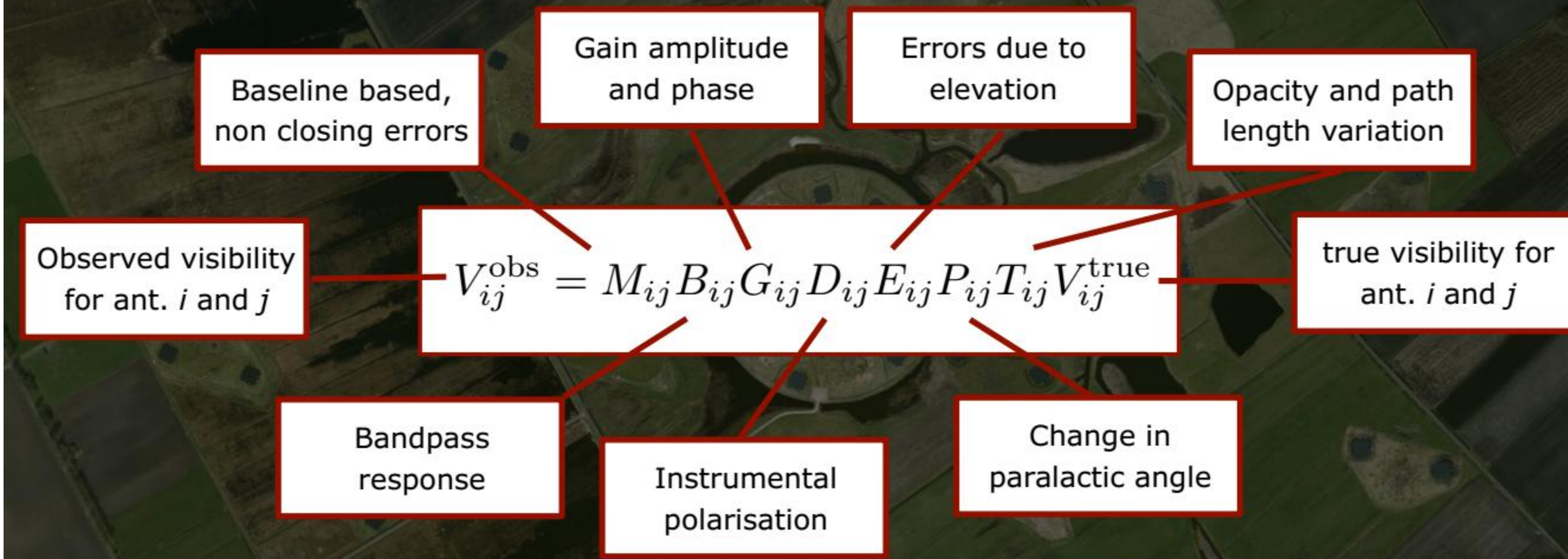
Osiągane przez minimalizację różnic z modelową widzialnością dla znanego źródła (kalibratora):

$$V_{ij}^{obs} - G_{ij} V_{ij}^{\mathrm{mod}}$$

Musimy zatem zaobserwować znane źródło (kalibrator, o znanej strukturze, najlepiej punktowej)

The Measurement Equation

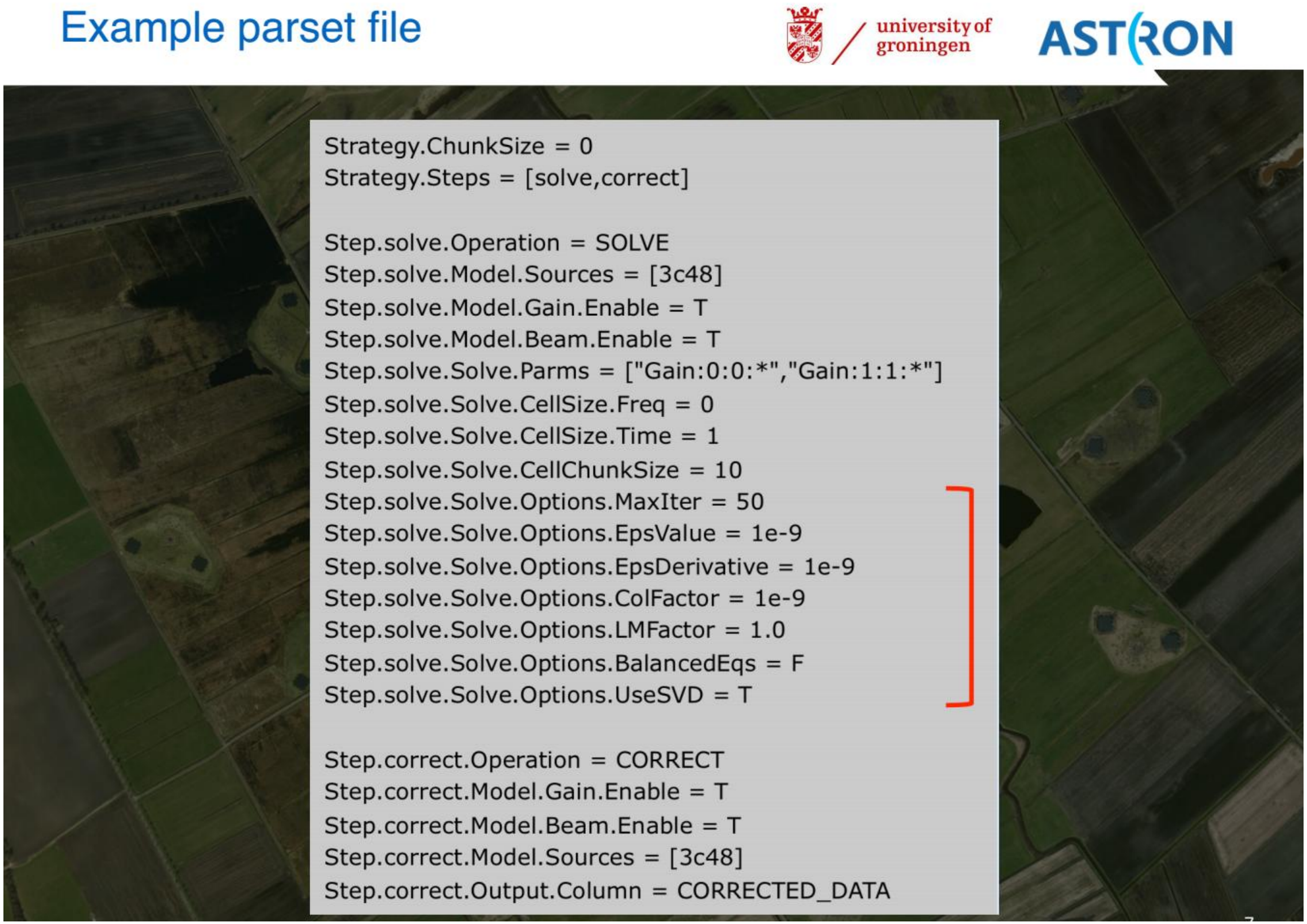
- **RIME:** The radio interferometer measurement equation, as used by CASA etc. for the calibration,



$$V_{ij}^{\mathrm{obs}} = M_{ij} B_{ij} G_{ij} D_{ij} E_{ij} P_{ij} T_{ij} V_{ij}^{\mathrm{true}}$$

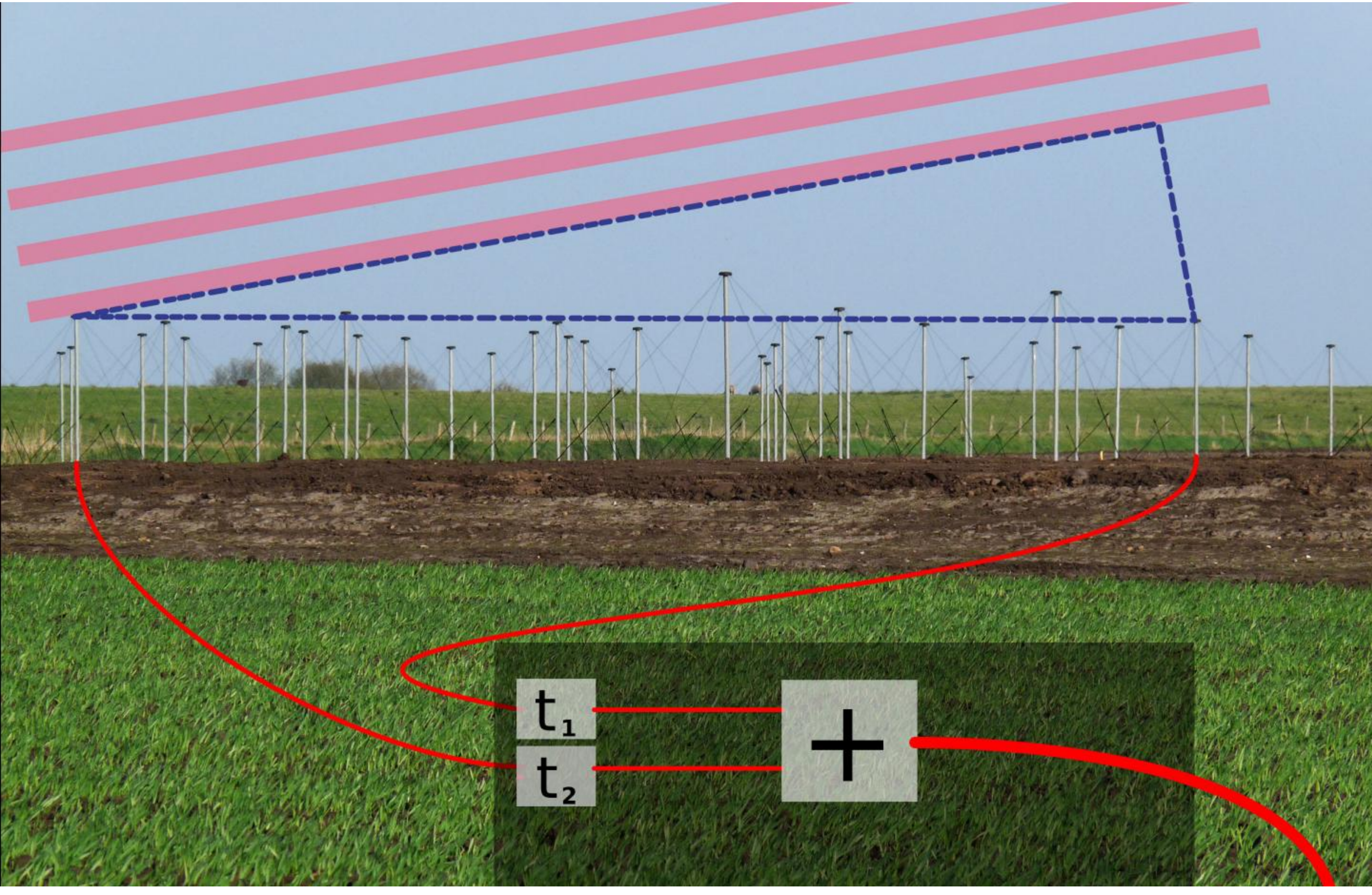
- **Calibration** involves solving this inverse problem to determine what set of parameters are needed to minimise the difference between the *observed* visibilities and the *model* visibilities (our best guess at the *true* visibilities).

Black Board Self-calibration (BBS) is a software package that is designed for the calibration and simulation of LOFAR data




```
> calibrate-stand-alone -f <MS> <parset> <source catalog>
```

LOFAR - Station beam forming



THE LOFAR SYSTEM: DATA FLOW

LOFAR

ASTRON

Station signals collected in the station cabinets

Signal sent to COBALT for correlation

CEP2

CEP3

Data sent to CEP2 for initial RO processing – products might get copied to CEP3

Products sent to the long-term archive

- Entire process is overseen by Operators, Science Support and Software Support groups

Visibilities, output of the correlator

Correlator outputs a measurement set (MS) per subband, containing visibilities: a complex number for each

- Timeslot
- Baseline (combination of two antennas)
- Channel
- Correlation (XX, XY, YX, YY)

- We *cross correlate* the signals across all *baselines*.
  - *Baseline = pair of antenna fields.*
- For each baseline, we compute:

$$\begin{bmatrix} \langle X_1 X_2^* \rangle & \langle X_1 Y_2^* \rangle \\ \langle Y_1 X_2^* \rangle & \langle Y_1 Y_2^* \rangle \end{bmatrix}$$

- Parameters:
  - Integration time (typically 1 s)
  - Frequency resolution (typically 3 kHz)

- *Subband* = 195 kHz wide.
- A Polyphase Filter (PPF) separates *channels*
- Example:
  - 64-pt PPF produces 3 kHz channels

- Stokes parameters:

$$I = |X|^2 + |Y|^2$$

← amplitude

$$Q = |X|^2 - |Y|^2$$
$$U = 2Re(XY^*)$$
$$V = 2Im(XY^*)$$

} phase aspects

Correlator – Storage Format

- Format: Measurement Sets

  SubArray Pointing

  - Example: L12345_SAP000_SB010_uv.MS/

  Observation ID   Subband Index

- Includes:
  - Obs specification, subband details
  - Participated hardware (broken antenna info)
  - Sample weights

- 1 file = 1 subband

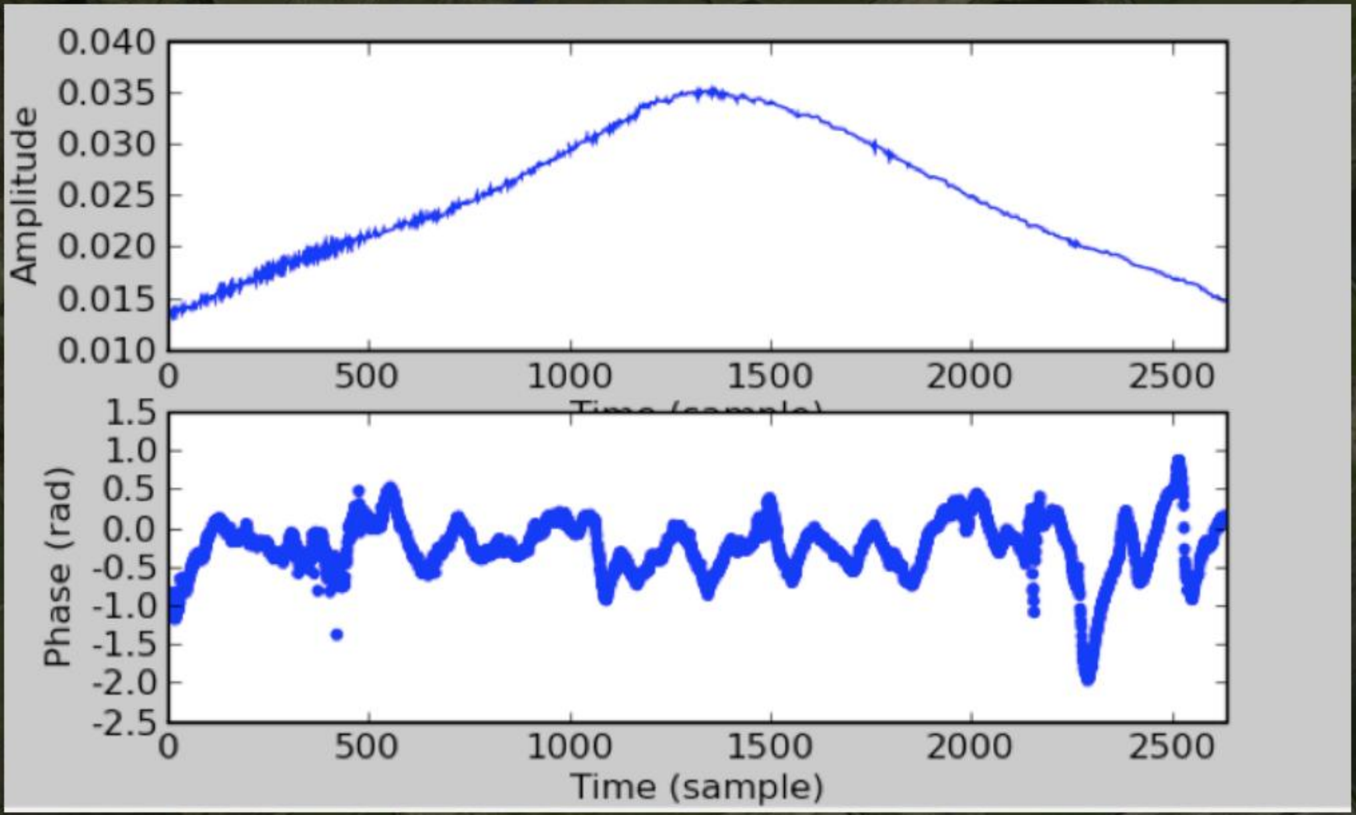
Cechy szczególne obserwacji na niskich częstotliwościach

1. Wiązka LOFAR-a
2. RFI
3. Jonosfera
4. Kalibracja zależna od częstotliwości
5. Polaryzacja - RM Synthesis (wykład Roberta Drzazgi)

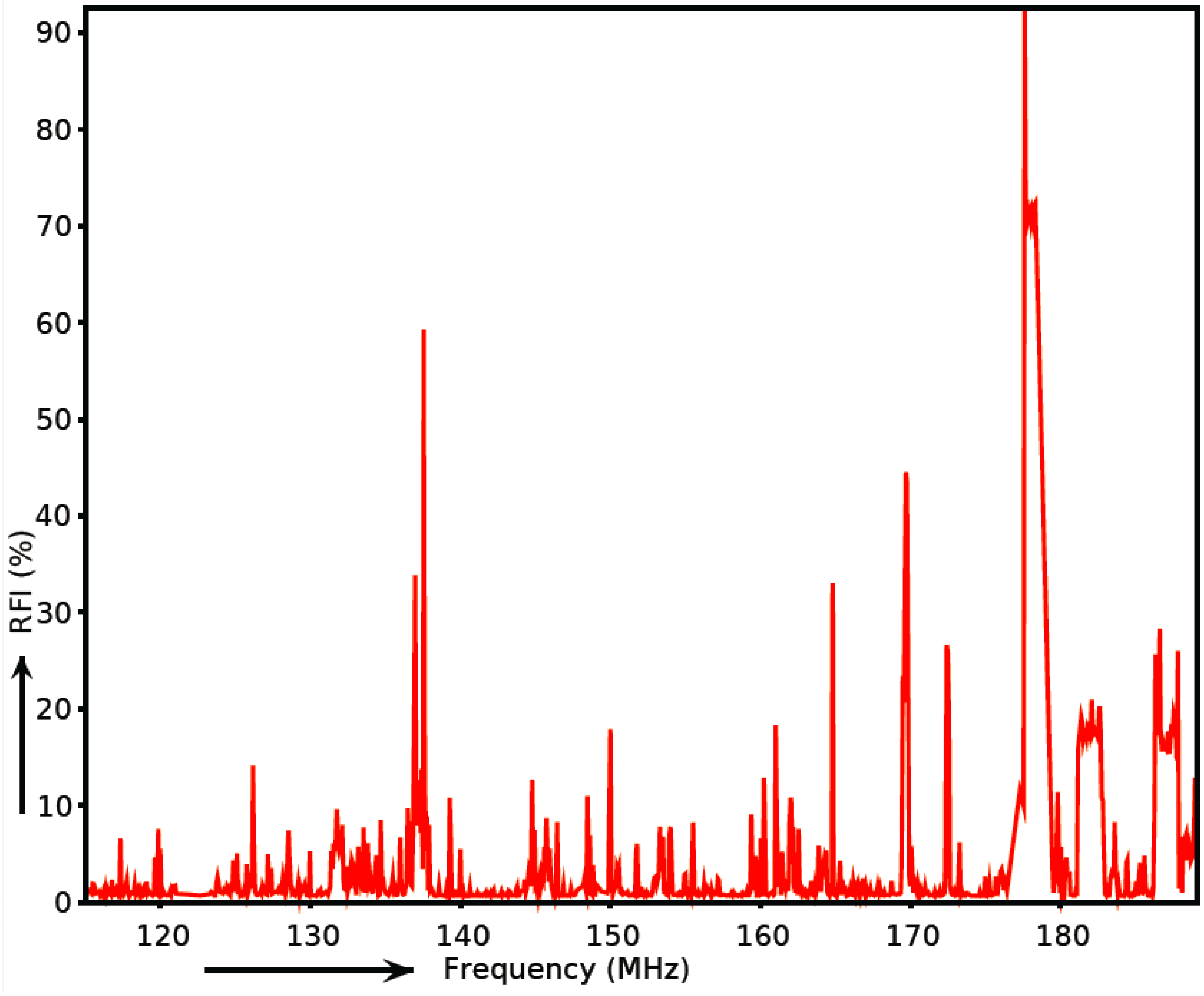
Wiązki LOFAR-a

Znajomość ograniczona

- Unlike typical dish-based interferometers (e.g. JVLA, WSRT) the gain (amplitude) of the visibilities are not constant.
- This is due to the source moving through the beam (+atmosphere+ionosphere), effectively the change in the projected area of the station. Need beam correction in the calibration.

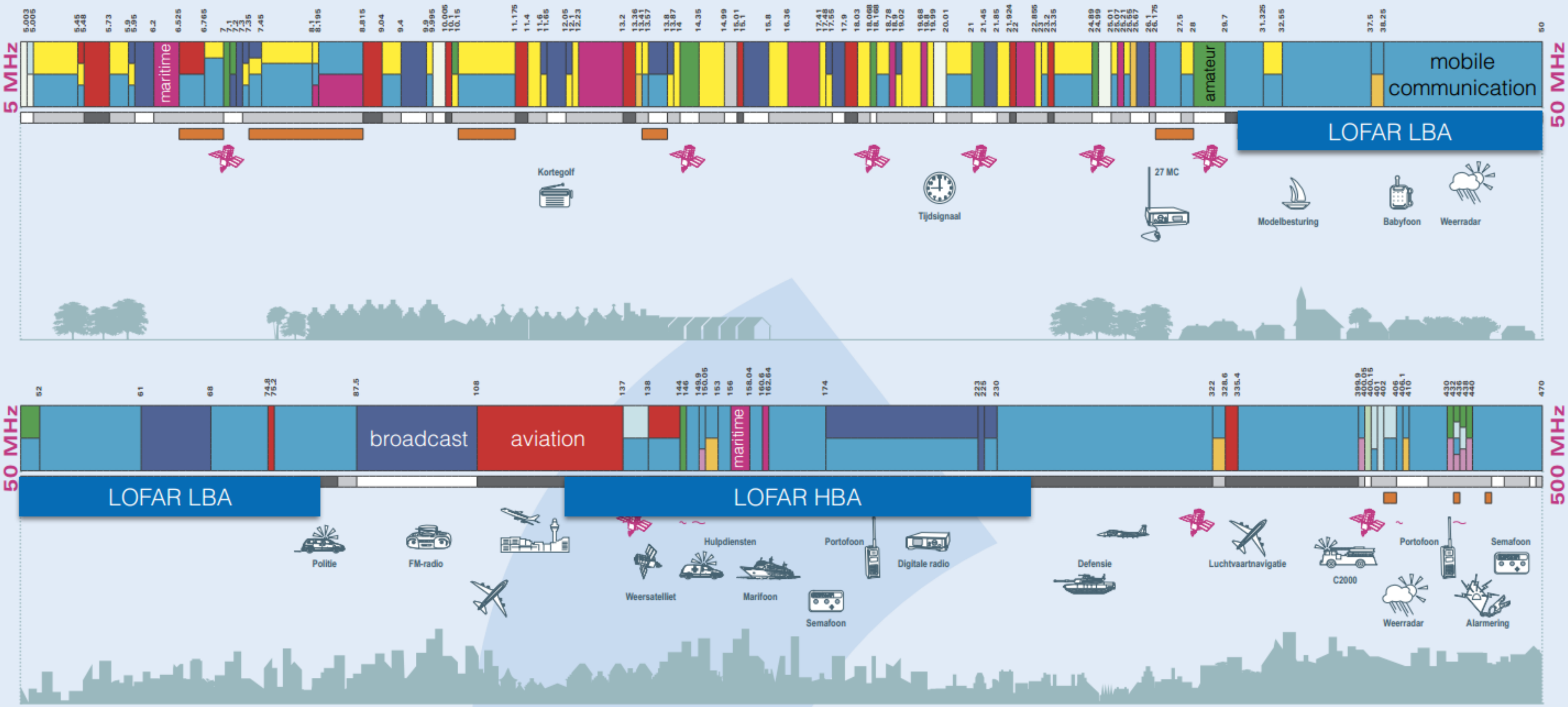


RFI – duże m.in. DAB



Frequency allocations, interfering signals

ASTRON



Source: frequentiespectrumkaart 2005

Most RFI near LOFAR is narrowband and/or short duration.

Correcting for the beam

Variable beams as a function of time mean that the contribution from each source will vary over time to the visibilities (must convolve sky model with beam model).

$$V_\nu(u,v) = \int\int A_\nu(l,m) I_\nu(l,m) e^{-2\pi i(ul+lm)} dl dm$$

More sophisticated calibration that includes the beam (a-projection is being implemented in CASA for the JVLA).

Issues:

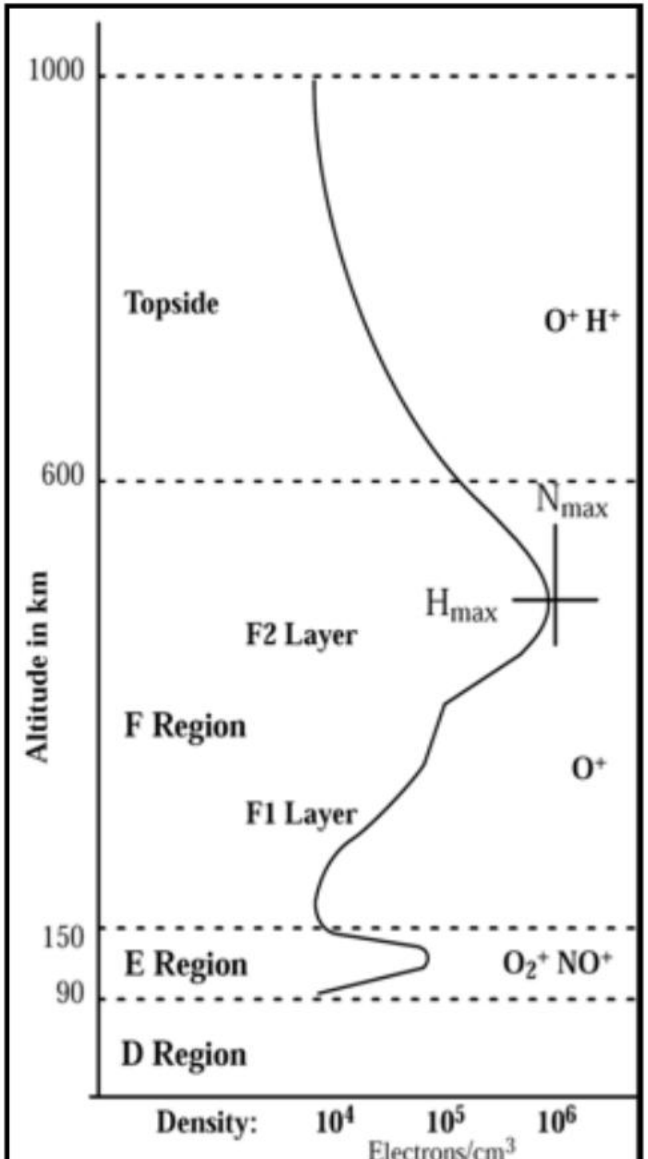
1) How well do we know the beam? Recall, the beam is the FT of the aperture. What happens if a dipole stops working?

2) The beam changes as a function of frequency (FWHM ~ λ / D).

An error in your model can be absorbed in the calibration

$$\vec{V}_{ij} = J_{ij} \vec{V}_{ij}^{\mathrm{IDEAL}}$$

The ionosphere

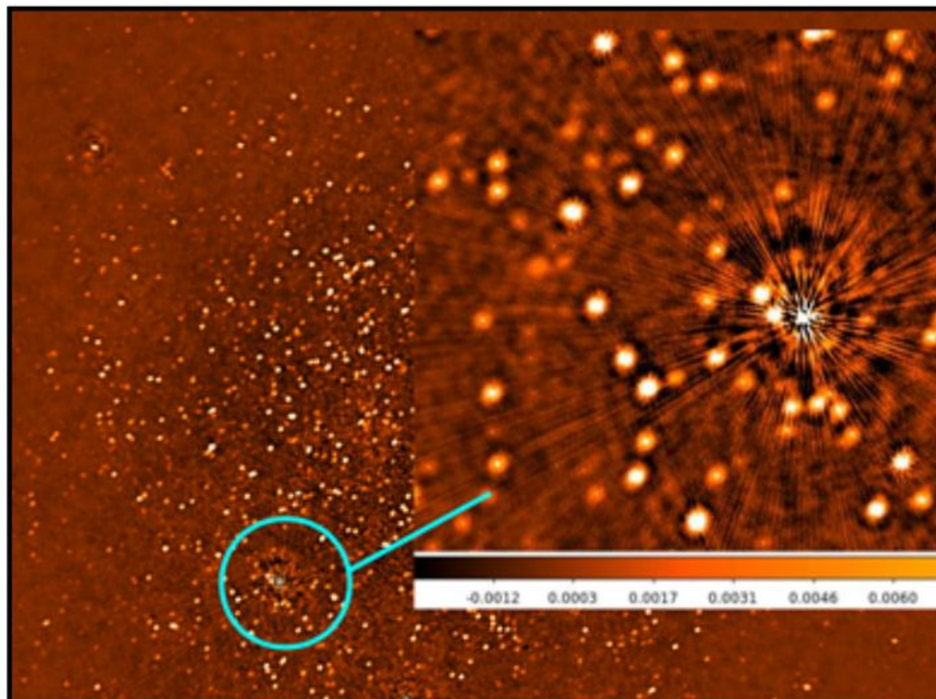


The ionosphere is a reflecting (to long wavelengths) layer of the atmosphere at ~ 125 km.

Structure and electron density changes with altitude.

Effects radio waves through:

1) Reflection (transparency)
2) Scintillation (continuum imaging)
3) Faraday rotation (polarisation)



Direction independent effects – DIE

$$G_i\,(t)$$

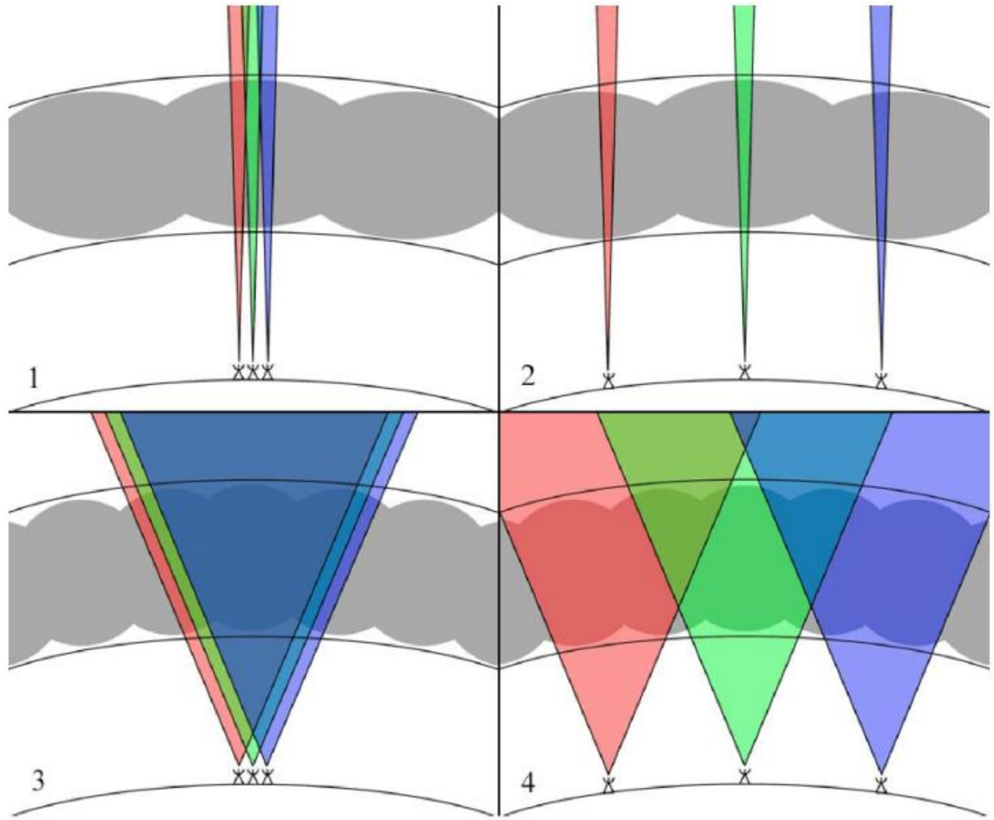
Direction dependent effects - DDE

$$G_i\,(t,\alpha,\delta)$$

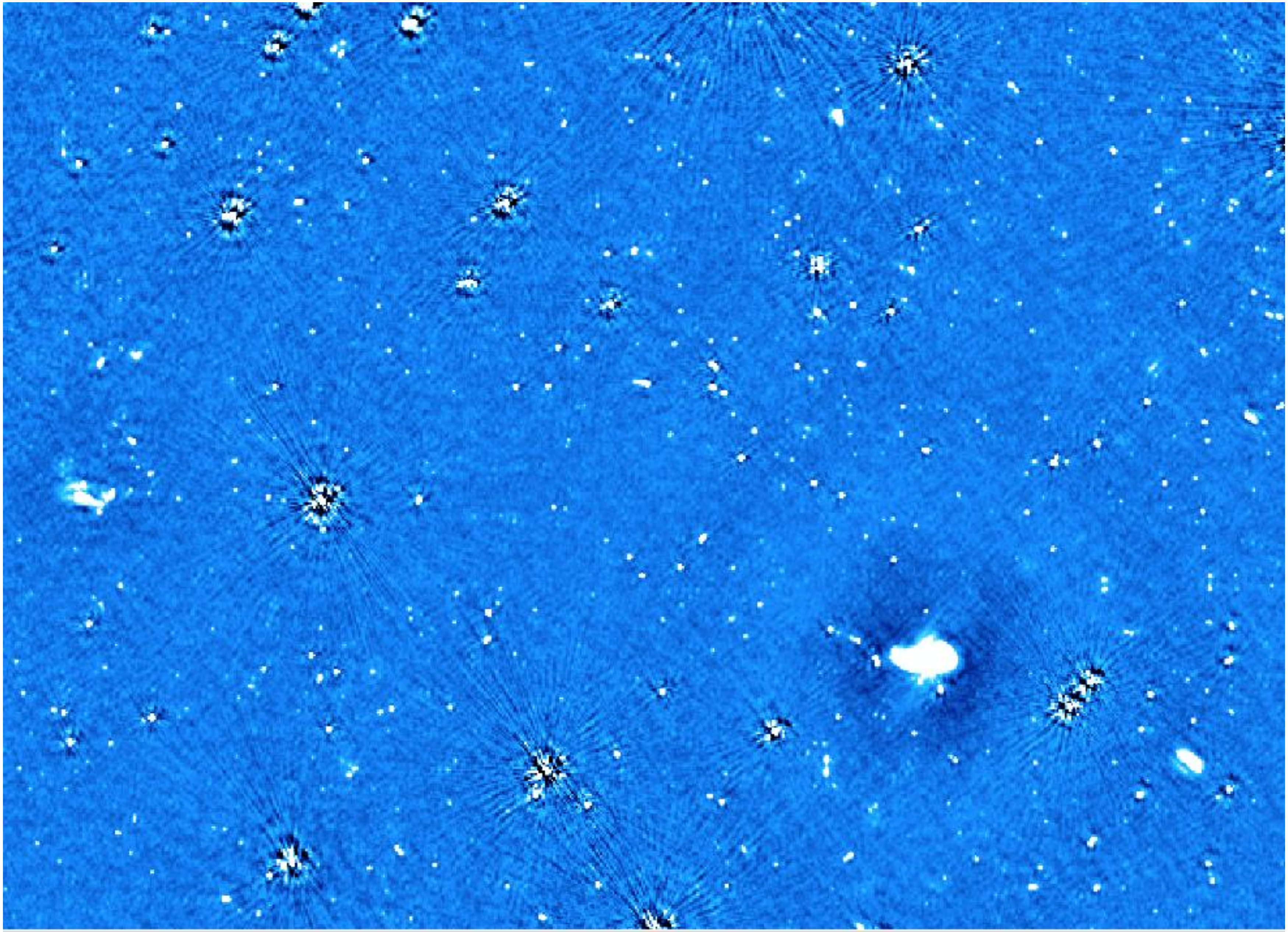
- Regimes 1 & 2 ionospheric phase error has no FoV variation – self cal OK

- Regimes 3 & 4 have varying phase over FoV – need direction dependent algorithms

Significant effort underway: field-based, source peeling, global model, multiple scale height models, …



Intema et al. (2009)



Tim Shimwell – Tier-1

Jonosfera

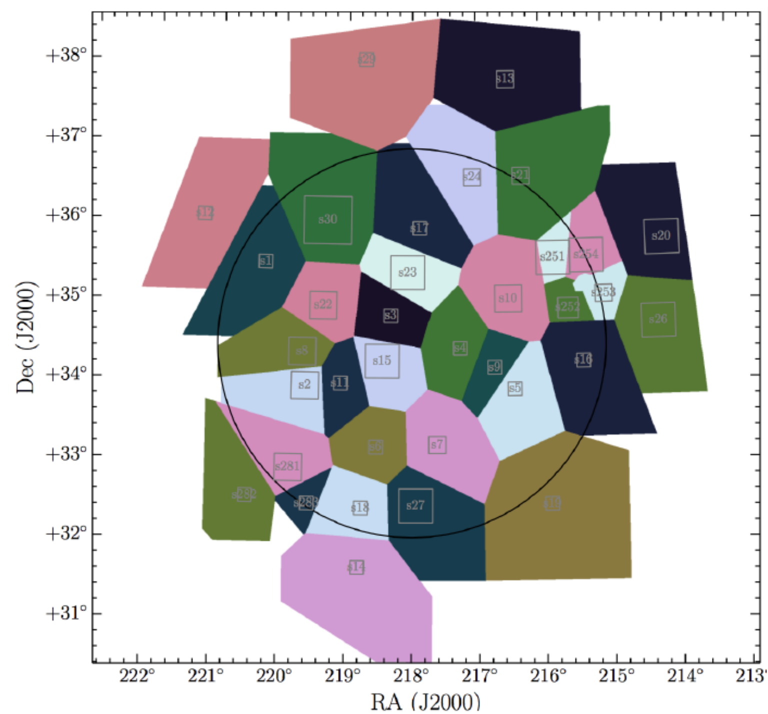
- Konieczne zawansowane techniki flagowania danych
- Pomoc: dane satelitarne GPS – modelowanie phase screens w czasie rzeczywistym
- Clock-TEC separation

$$\Delta\theta = 2\pi f \Delta t + 8.44797245 \times 10^9 \Delta TEC / f + \Delta\theta_0$$

  - ❖ BBS
  - ❖ losotto – PREFACTOR

- Korekta na jonosferyczną rotacje Faradaya (Rotation measure) (LBA – efekty wyższych rzędów)

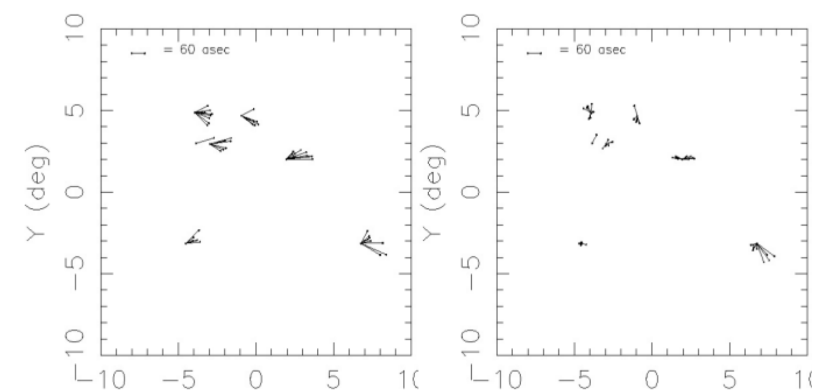
- DDE uwzględnić np. podczas obrazowania



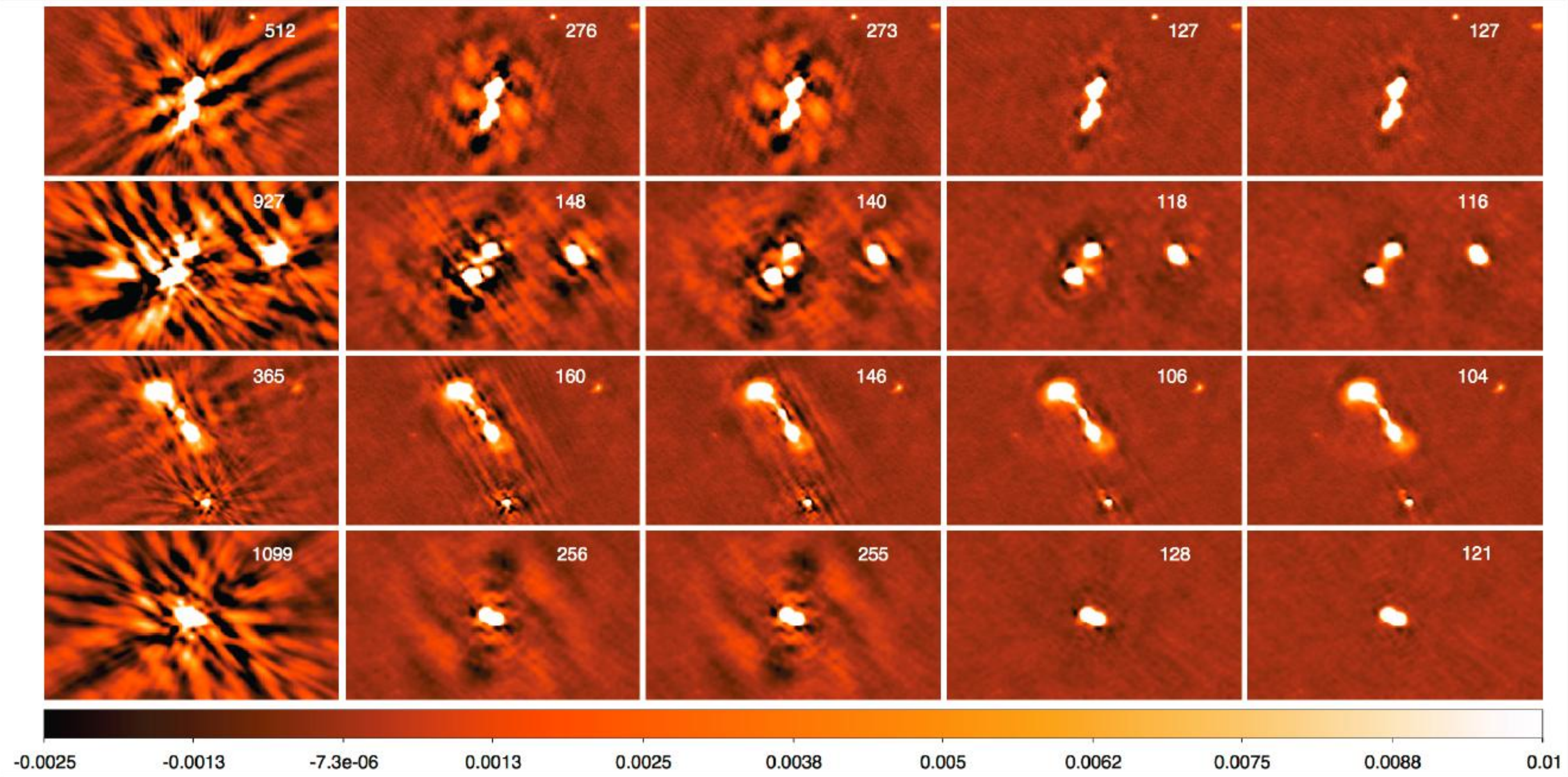
Ionospheric calibration methods

- Calibration schemes including ionosphere models are rare
- Field-based calibration (Cotton+ 2004),
  - Tracks apparent movement of multiple radio sources, fits de-distorting Zernike screen
- SPAM method (Intema+ 2009)
  - Multi-directional phase calibration, fits turbulent multi-layer phase screen
- Other DD calibration schemes exist, but without constraint on the ionosphere
  - E.g., facet-calibration (van Weeren+ 2015), SAGECal (Yattawata+ 2008, arXiv:0810.5751)

field-based calibration
(Cotton+ 2004)



Facet calibration

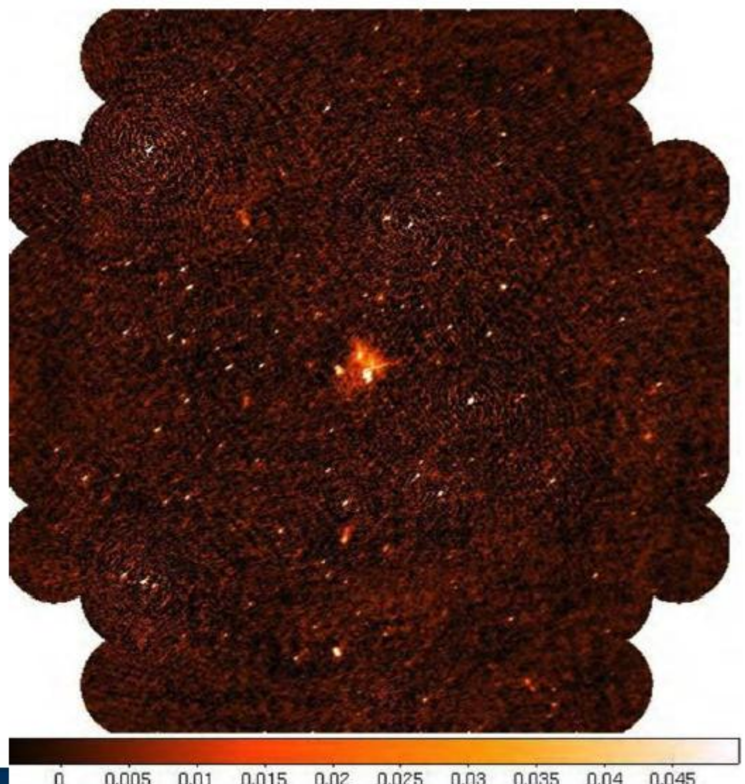


Demonstrating direction dependent calibration (van Weeren R. J., et al., 2016, ApJS, 223, 2)

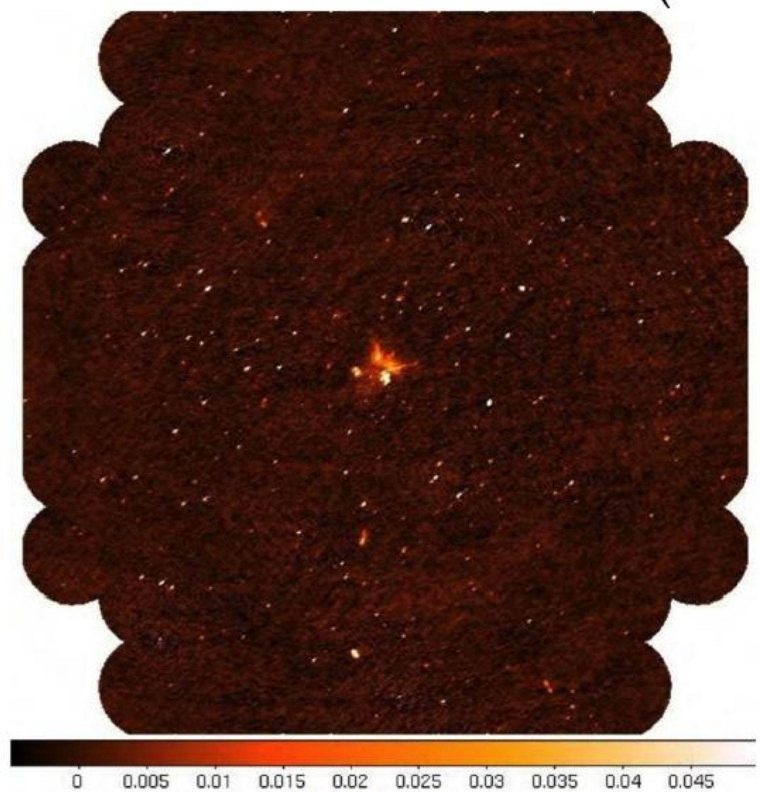
Ionosphere: SPAM

- Source Peeling and Ionospheric Modeling: SPAM (python + AIPS)
- Constrain ionospheric phase model based on calibration phases from 'peeling' (sequential self-calibration) of bright sources
- Fit a phase screen to pierce point solutions and apply to imaging

Intema et al. (2009)



Self-calibration



SPAM calibration

Multi-Frequency Synthesis (MFS)

ASTRON

We can represent the sky in emission interms of a Taylor expansion about some reference frequency (see Rau & Cornwell 2011).

Taylor co-efficient images

MS model image

Build *I(v)* model:

$$\boldsymbol{I}^{\mathrm{m}}_{\nu} = \sum_{t=0}^{N_{\mathrm{t}}-1} w^{t}_{\nu} \boldsymbol{I}^{\mathrm{sky}}_{t} \quad \text{where} \quad w^{t}_{\nu} = \left(\frac{\nu - \nu_0}{\nu_0}\right)^{t}$$

A power-law model is used to describe the spectral dependence of the sky emission.

Parameterise:

$$\boldsymbol{I}^{\mathrm{sky}}_{\nu} = \boldsymbol{I}^{\mathrm{sky}}_{\nu_0} \left(\frac{\nu}{\nu_0}\right)^{\boldsymbol{I}^{\mathrm{sky}}_{\alpha} + \boldsymbol{I}^{\mathrm{sky}}_{\beta} \log\left(\frac{\nu}{\nu_0}\right)}$$

Sky images:

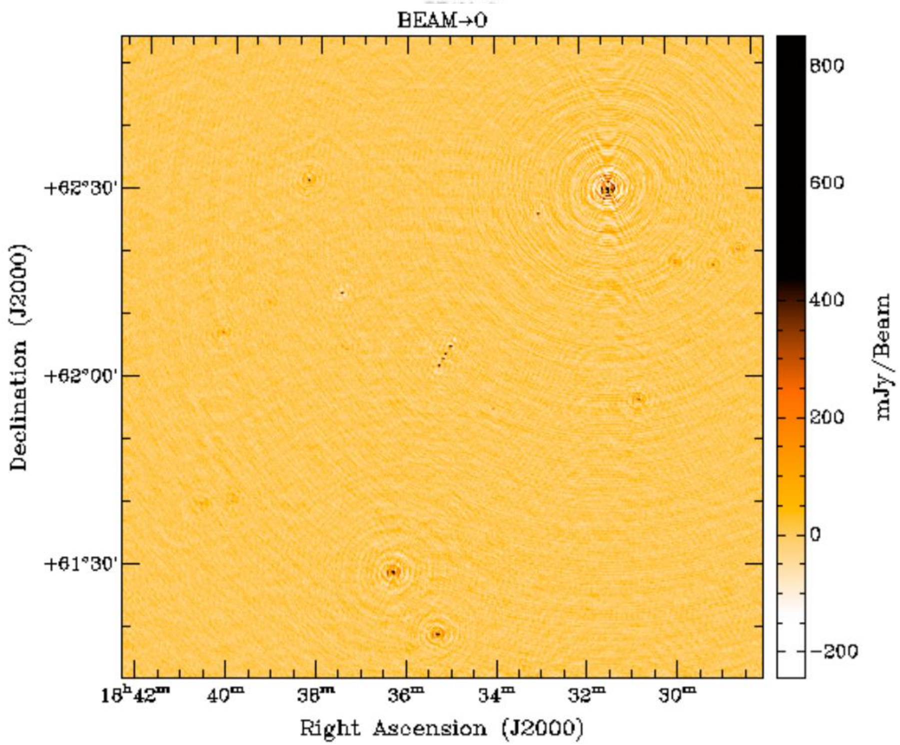
$$\boldsymbol{I}^{m}_{0} = \boldsymbol{I}^{\mathrm{sky}}_{\nu_0} \;;\; \boldsymbol{I}^{m}_{1} = \boldsymbol{I}^{\mathrm{sky}}_{\alpha} \boldsymbol{I}^{\mathrm{sky}}_{\nu_0} \;;\; \boldsymbol{I}^{m}_{2} = \left(\frac{\boldsymbol{I}^{\mathrm{sky}}_{\alpha}(\boldsymbol{I}^{\mathrm{sky}}_{\alpha} - 1)}{2} + \boldsymbol{I}^{\mathrm{sky}}_{\beta}\right) \boldsymbol{I}^{\mathrm{sky}}_{\nu_0}$$

Direction dependent calibration

The solution to these issues is to calibrate the gains, not in a single position, but over several positions (10s to 100s) across the sky.

$$\vec{V}_{ij} = \sum_s J_{ij,s} \vec{V}_{ij}^{\mathrm{IDEAL}}$$

Sarod Yattawatta



Computationally expensive and the robustness is a matter of (current) debate.

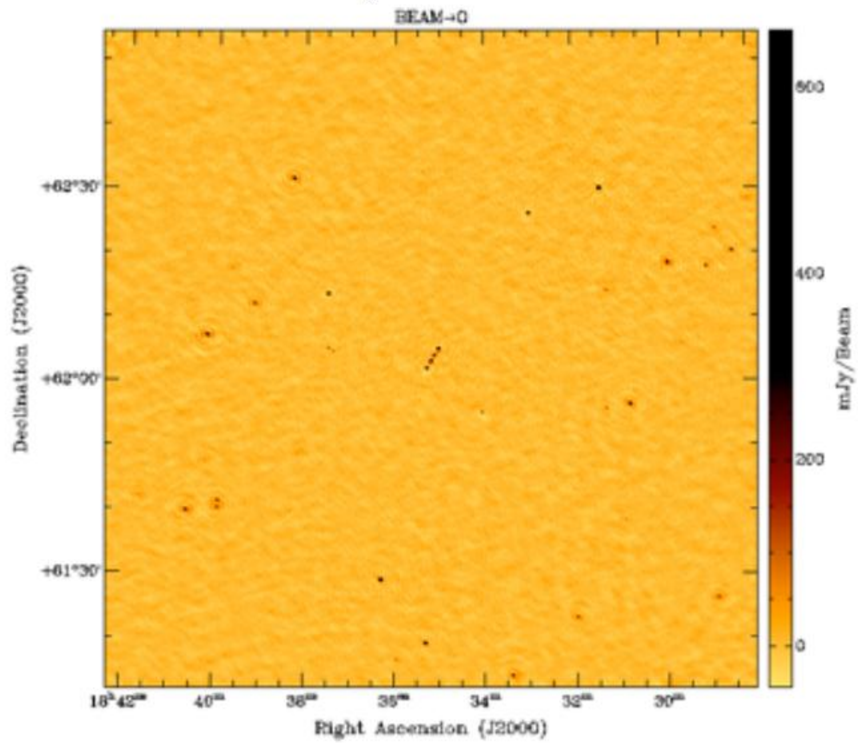
Direction dependent calibration

ASTRON

The solution to these issues is to calibrate the gains, not in a single position, but over several positions (10s to 100s) across the sky.

$$\vec{V}_{ij} = \sum_s J_{ij,s} \vec{V}_{ij}^{\mathrm{IDEAL}}$$

Sarod Yattawatta

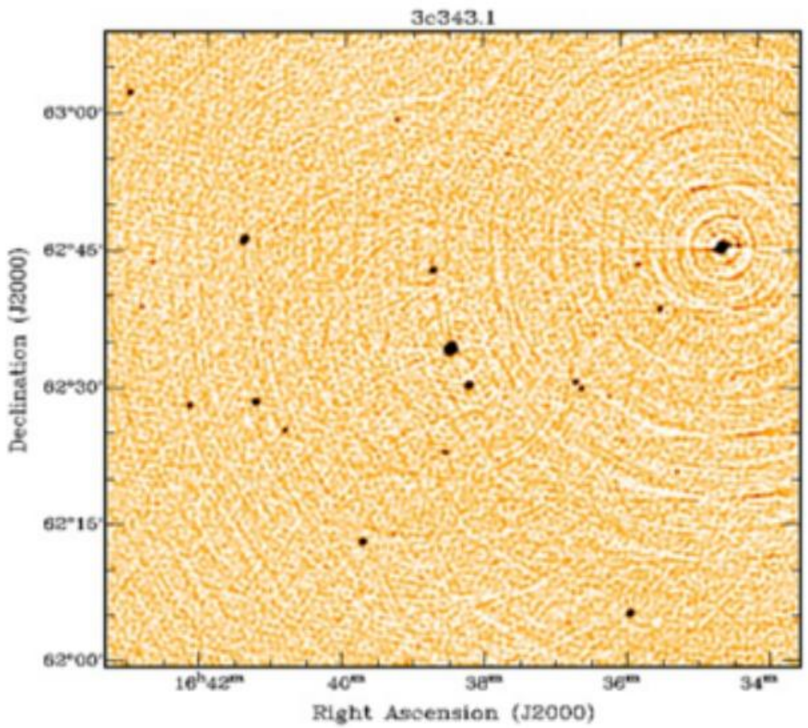


Computationally expensive and the robustness is a matter of (current) debate.

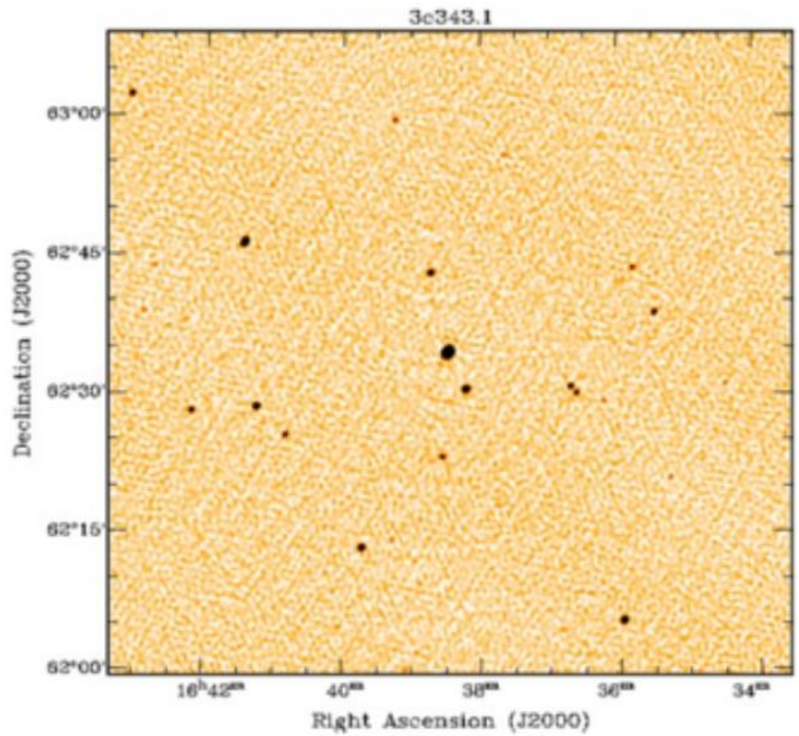
Direction dependent calibration

Alternatively, calibrate in one direction at a time and remove the troublesome sources (called peeling)

Tom Oosterloo



Before.



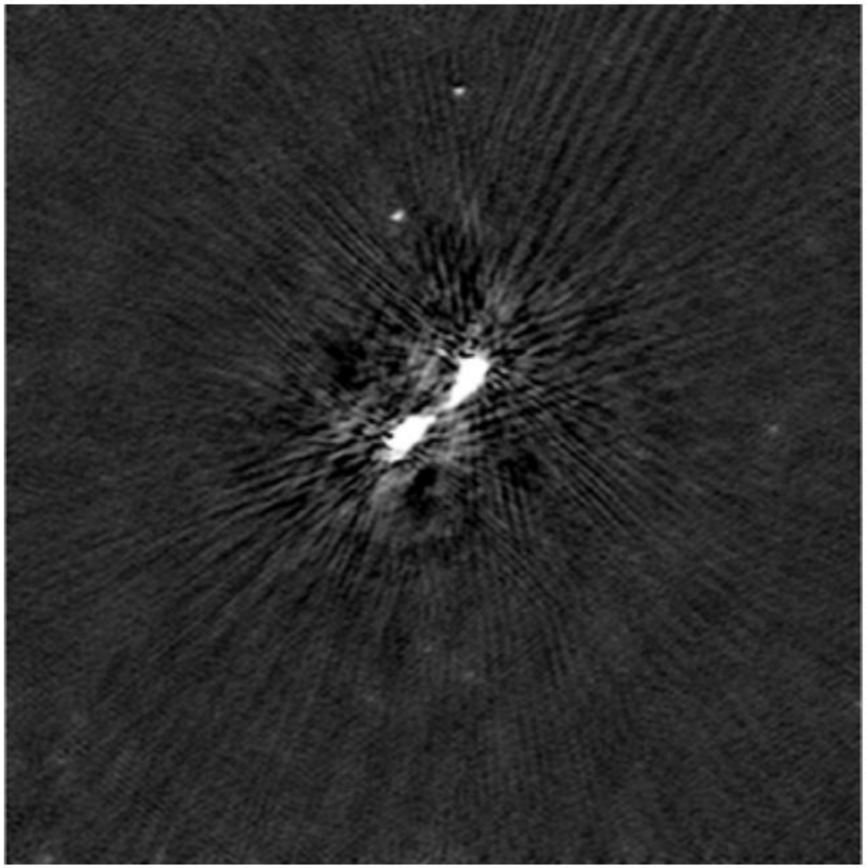
After peeling.

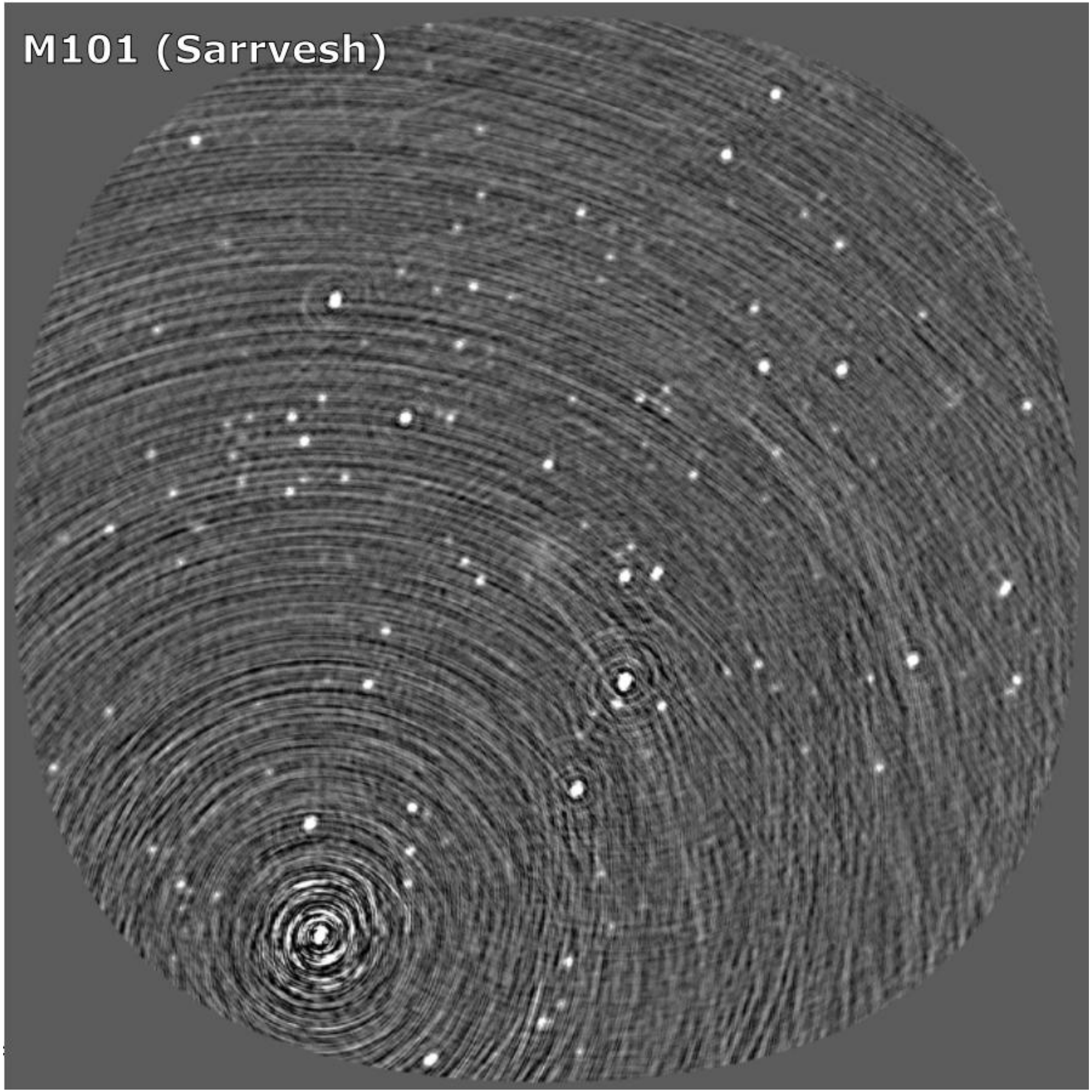
```
Step.solve.Solve.Parms = ["DirectionalGain:0:0:*","DirectionalGain:1:1:*"]
```

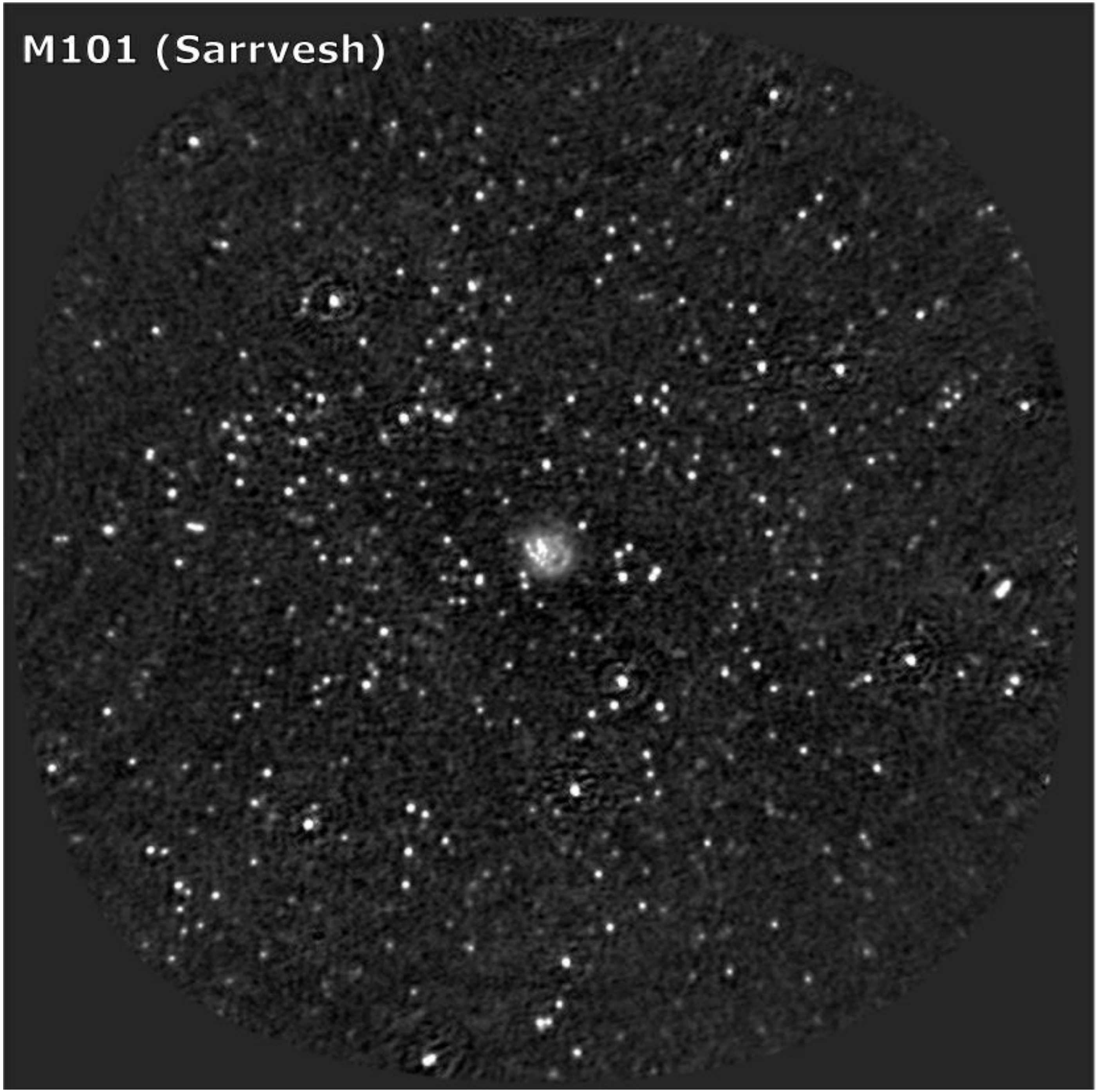
Direction dependent calibration

Alternatively, calibrate in one direction at a time and remove the troublesome sources (called peeling)

Wendy Williams

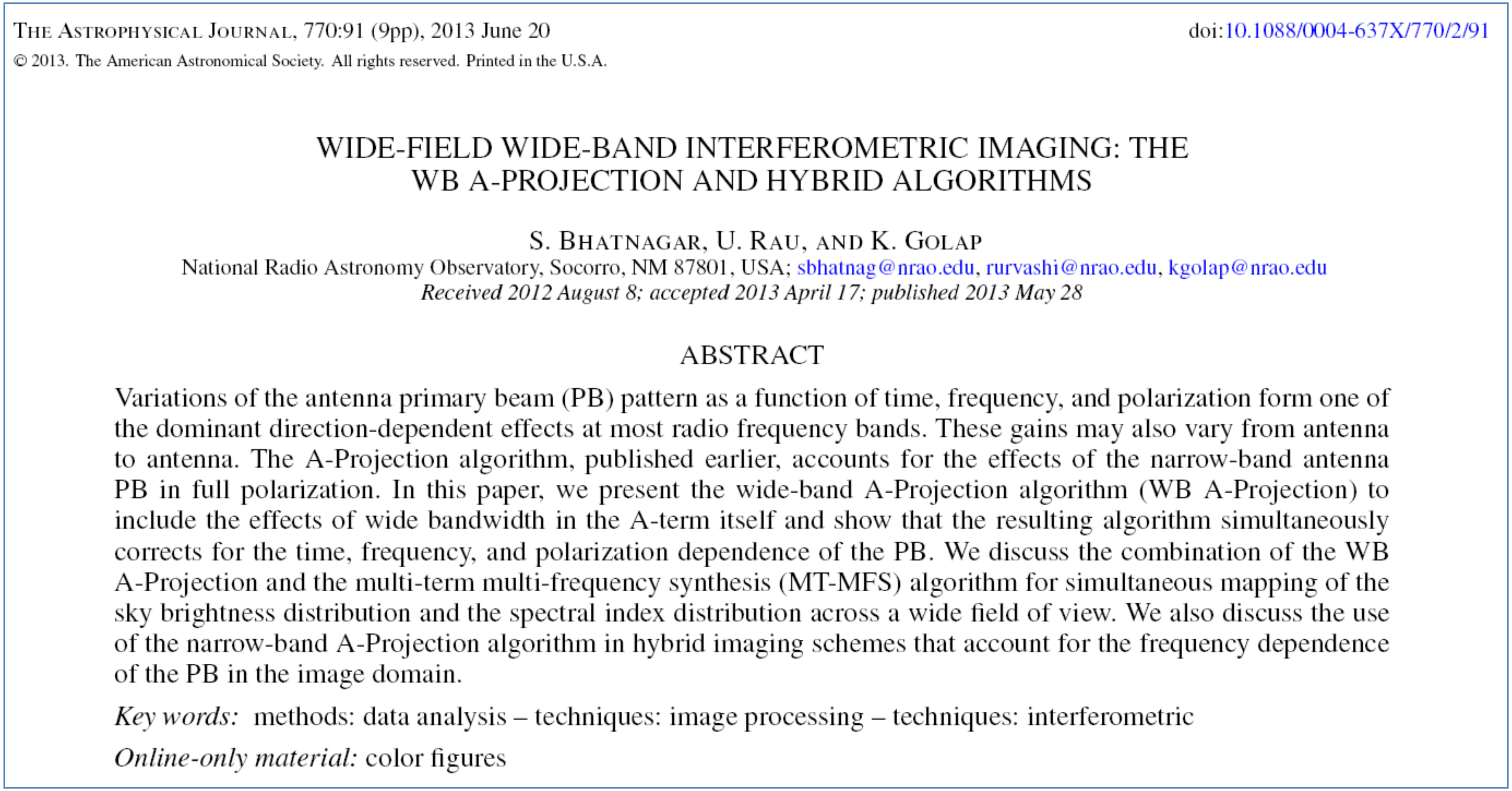



M101 (Sarrvesh)


M101 (Sarrvesh)

Intensywne prace teoretyczne i modelowe

Szybko rozwijane metody i implementacje



THE ASTROPHYSICAL JOURNAL, 770:91 (9pp), 2013 June 20 doi:10.1088/0004-637X/770/2/91

WIDE-FIELD WIDE-BAND INTERFEROMETRIC IMAGING: THE WB A-PROJECTION AND HYBRID ALGORITHMS

S. BHATNAGAR, U. RAU, AND K. GOLAP
National Radio Astronomy Observatory, Socorro, NM 87801, USA; sbhatnag@nrao.edu, rurvashi@nrao.edu, kgolap@nrao.edu
Received 2012 August 8; accepted 2013 April 17; published 2013 May 28

ABSTRACT

Variations of the antenna primary beam (PB) pattern as a function of time, frequency, and polarization form one of the dominant direction-dependent effects at most radio frequency bands. These gains may also vary from antenna to antenna. The A-Projection algorithm, published earlier, accounts for the effects of the narrow-band antenna PB in full polarization. In this paper, we present the wide-band A-Projection algorithm (WB A-Projection) to include the effects of wide bandwidth in the A-term itself and show that the resulting algorithm simultaneously corrects for the time, frequency, and polarization dependence of the PB. We discuss the combination of the WB A-Projection and the multi-term multi-frequency synthesis (MT-MFS) algorithm for simultaneous mapping of the sky brightness distribution and the spectral index distribution across a wide field of view. We also discuss the use of the narrow-band A-Projection algorithm in hybrid imaging schemes that account for the frequency dependence of the PB in the image domain.

Key words: methods: data analysis – techniques: image processing – techniques: interferometric

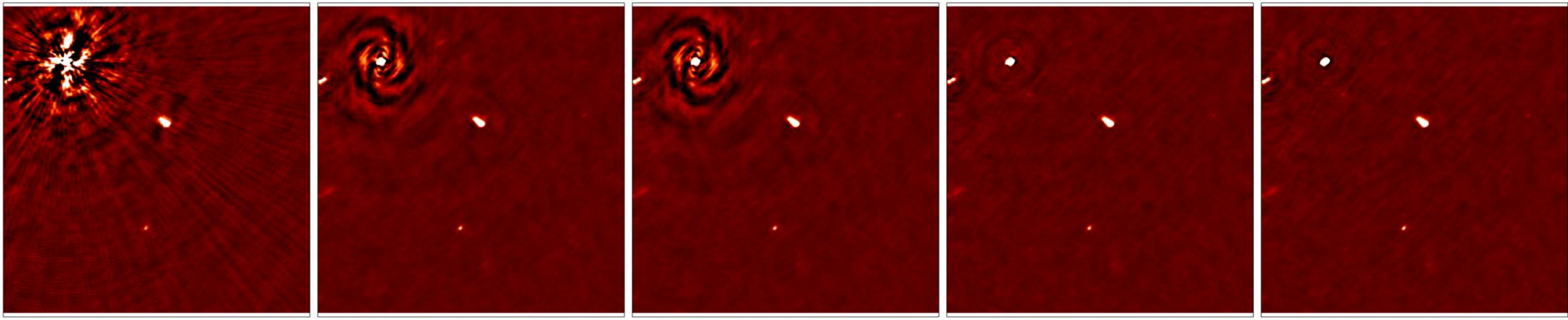
Online-only material: color figures

Obecnie najbardziej popularne podejście do DDE

- PREFACTOR (DIE)
- FACTOR (DDE)

https://github.com/lofar-astron/factor

New cookbook (v.19)
Next school!



Redukcja danych LOFAR-a

1. **Wstępna ocena jakościowa danych V_{ij}**
   - Przypisanie wag do Vij
   - Usunięcie zaburzeń Vij - flagowanie
   - Demixing
   - Średniowanie

2. **Kalibracja i poprawa V_{ij}**
   - Modelowe $\mathbf{V_{ij}}$ (kalibrator) => wyznaczenie Gij
   - Przeniesienie rozwiązań na źródło
   - Kalibracja źródła

3. **Dekonwolucja i obrazowanie (imaging)**
   - I_{lm} = FFT V_{ij} (dirty image)
   - Cleaning
   - Selfcalibration

1. Wstępna ocena jakościowa danych V_{ij}

1. Utworzenie wag
   Inspekcja jakości danych – casaplotms
   Usunięcie (flagging) błędnych danych i RFI (radio frequency interference)
2. "Demixing" – usunięcie wpływu odległych od centrum b. silnych źródeł (A-team)
3. Średniowanie (kompresja) danych w czasie i częstotliwości
4. Automatyzacja zadań (skrypty), dobór parametrów

Compression is critical to decreasing processing time, which is necessary given the data volumes.

LOFAR Cookbook and DPPP documentation

DPPP – Default Preprocessing Pipeline (also NDPPP)

Command-line tool, input as a parset, output as feedback on screen.

```
> DPPP myreduction.parset
```

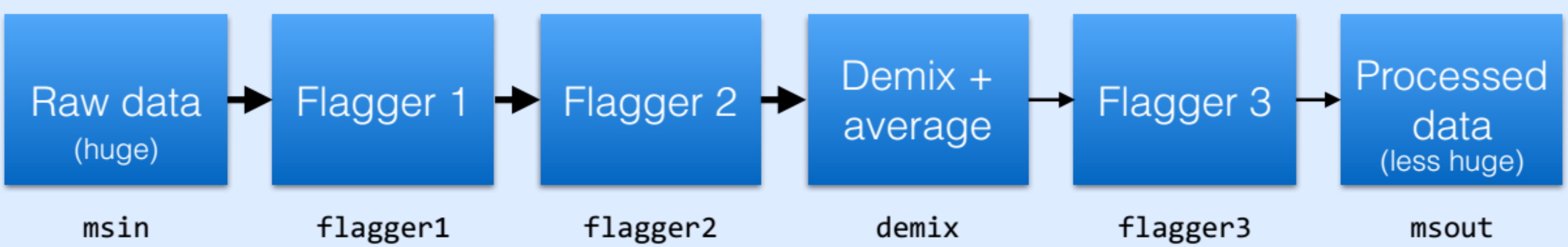
Overriding some parameters from the command line:

```
> DPPP myreduction.parset msin=L91.MS
```

Documentation:
http://www.lofar.org/operations/doku.php?id=engineering:software:tools:ndppp

```
msin=L123.MS
msout=L123_dppp.MS
steps=[flagger1, flagger2, demix, flagger3]

flagger1.type=aoflagger

flagger2.type=preflagger
flagger2.baseline=*&&& # autocorrelations

demix.type=demixer
demix.subractsources=[CygA,CasA]
demix.skymodel=Ateam.sourcedb

flagger3.type=aoflagger
```

Typical pipeline:



Weights, autoweight

ASTRON

Associated to each visibility $v_{i,j} = \hat{v}_{i,j} + n_{i,j}$ (between station i and j) is a weight $w_{i,j}$.

To exploit the data as much as possible, weights should be set such that noisy visibilities get down-weighted.

$$w_{i,j} = \frac{N_{\text{samples}}}{\sigma_i^2 \sigma_j^2}$$

Variance of noise of one station σ_i estimated from autocorrelation $v_{i,i}$.

Weights are computed when DPPP reads raw data and `msin.autoweight=true`.
Autoweight should be performed only once on the data.

Weights are stored in the column `WEIGHT_SPECTRUM`.

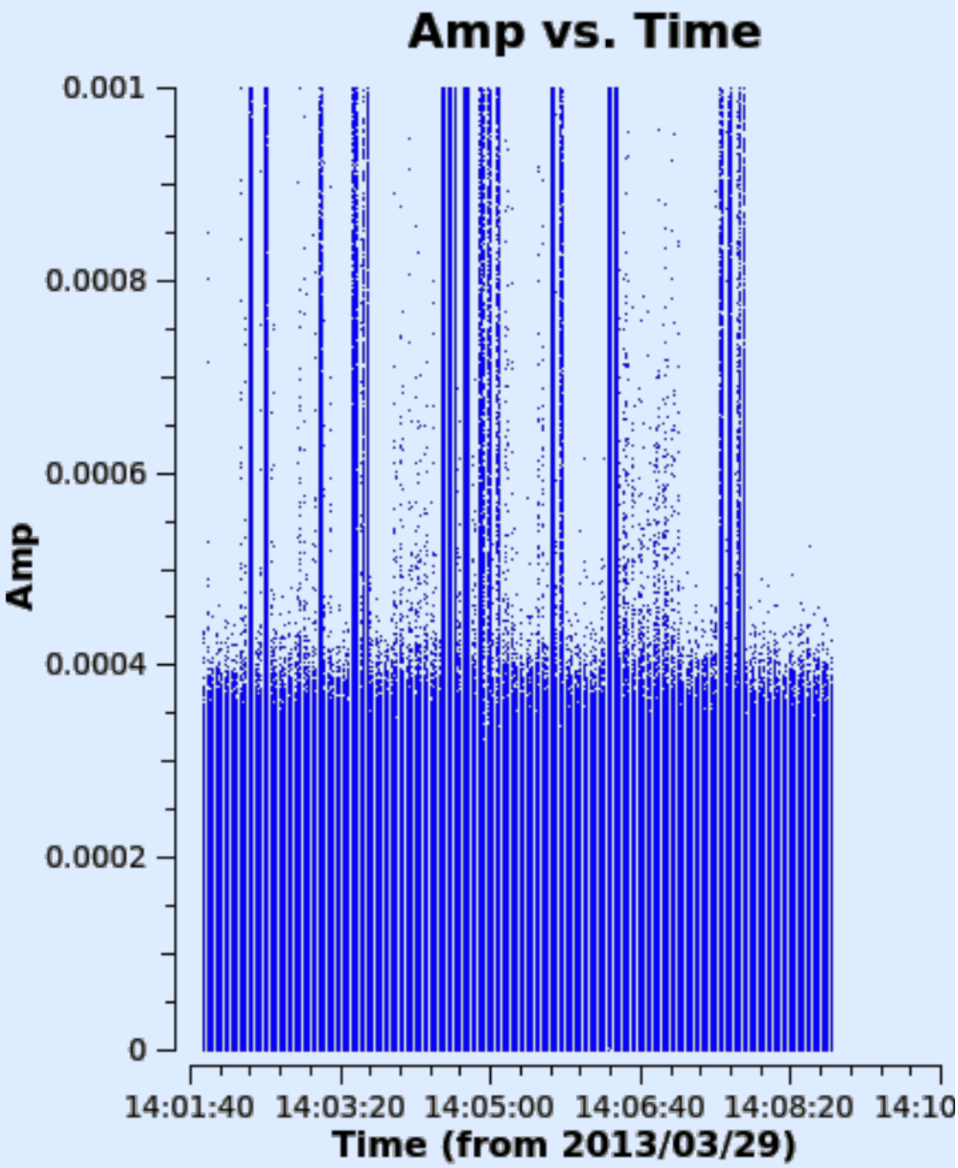
Flagging

Some samples affected by radio frequency interference (RFI).

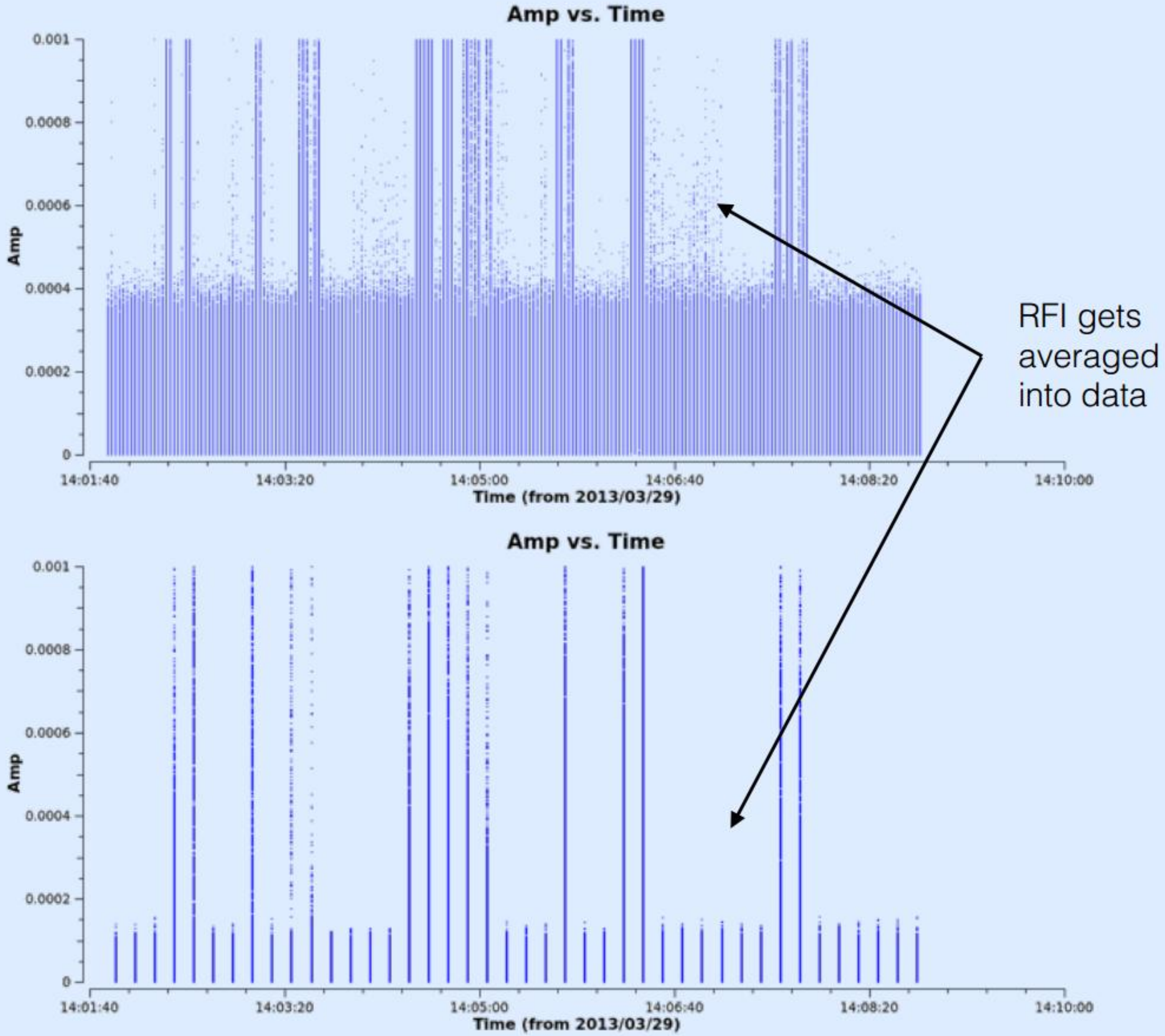
This makes these samples unsuitable for further processing.

We will flag them, and pretend they were never there.

Data is not deleted, just a check is put in `FLAG` column in MS.



Data should be flagged at high resolution



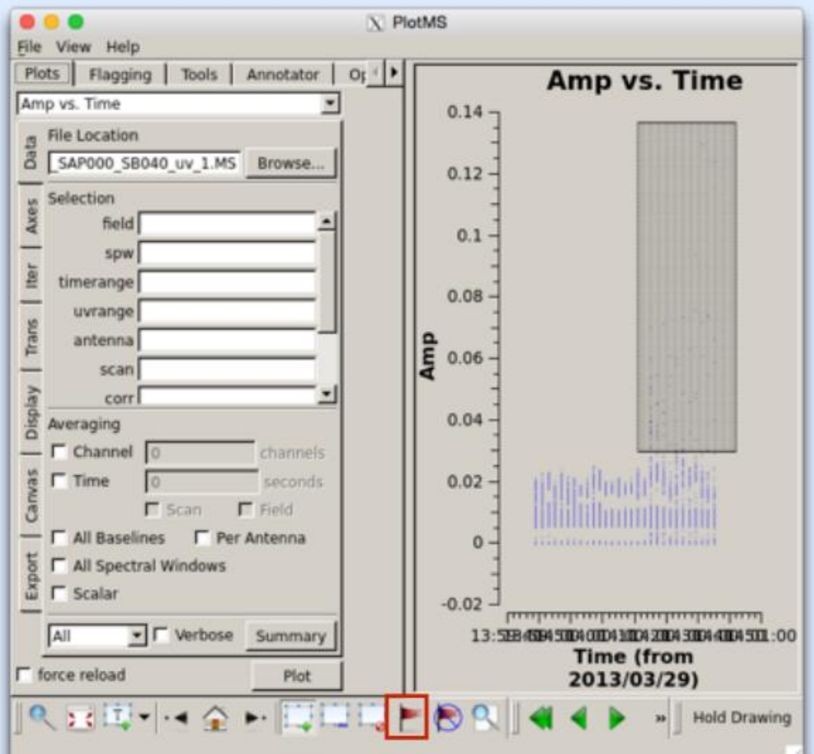
Three methods of flagging

ASTRON

1. Manual Flagging

Inspect data, select visibilities to flag

Can be done with `casaplotms`



2. Semi-automatic flagging

For example:

- Flag all autocorrelations
- Flag all signal stronger that 100 Jy
- Flag the first channel

Can be done with `DPPP`, step `preflagger`

3. Automatic flagging

For example using AOFlagger

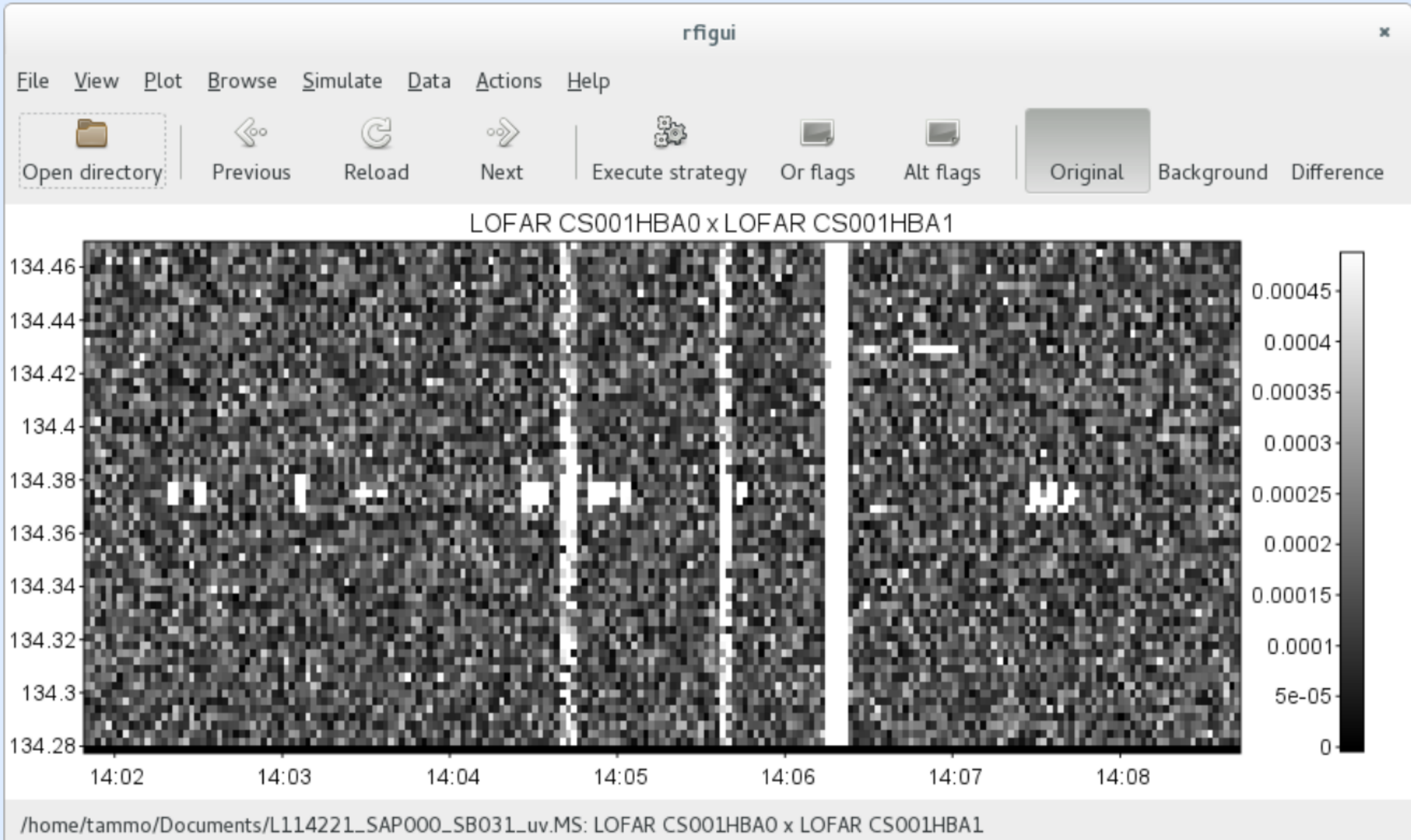
Flags based on time-freq statistics (per bl).
Performs best on long time ranges!

Can be called from `DPPP`, step `aoflagger`

Interactive counterpart: `rfigui`

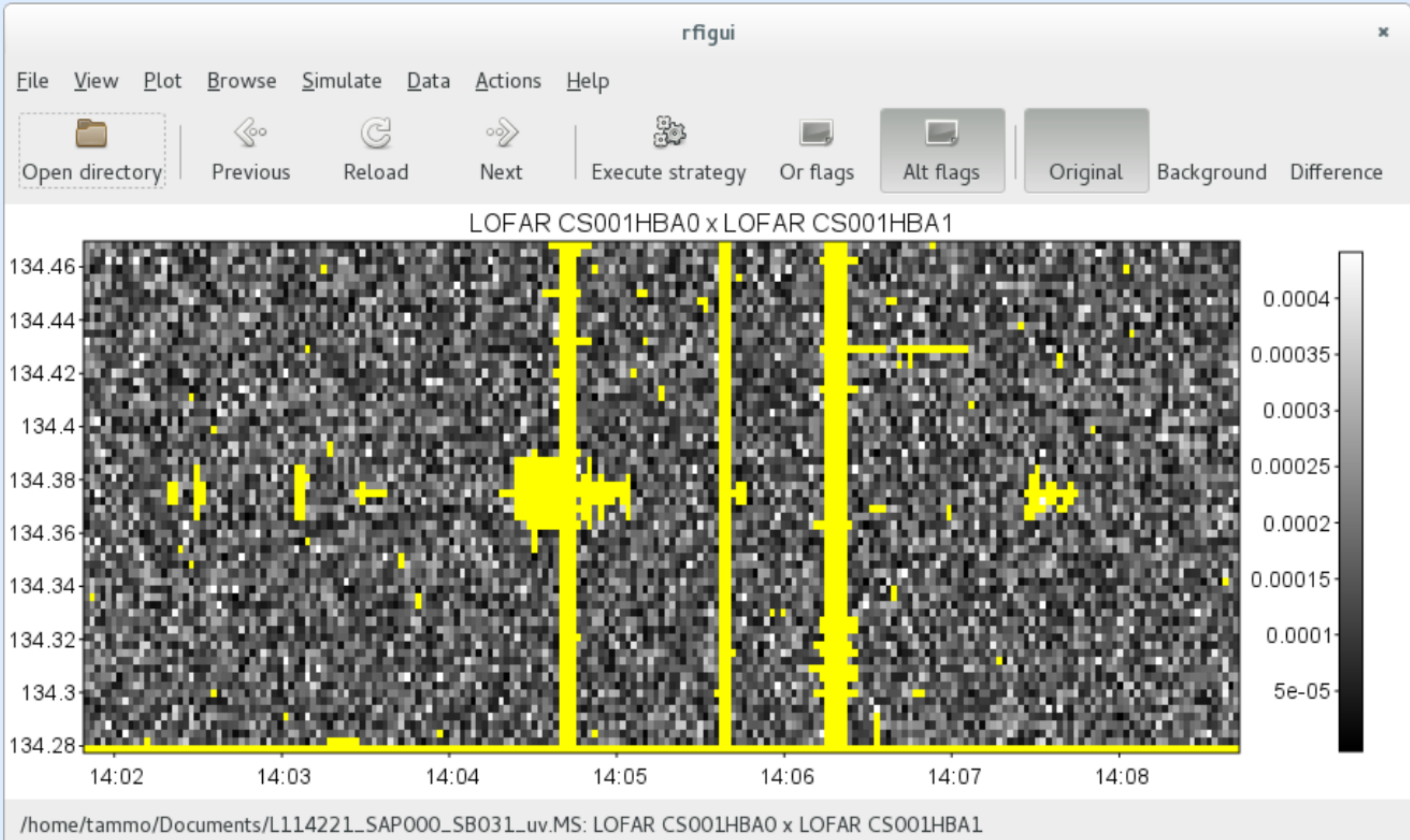
AOFlagger, rfigui

AOFlagger (André Offringa) flags data based on statistics:



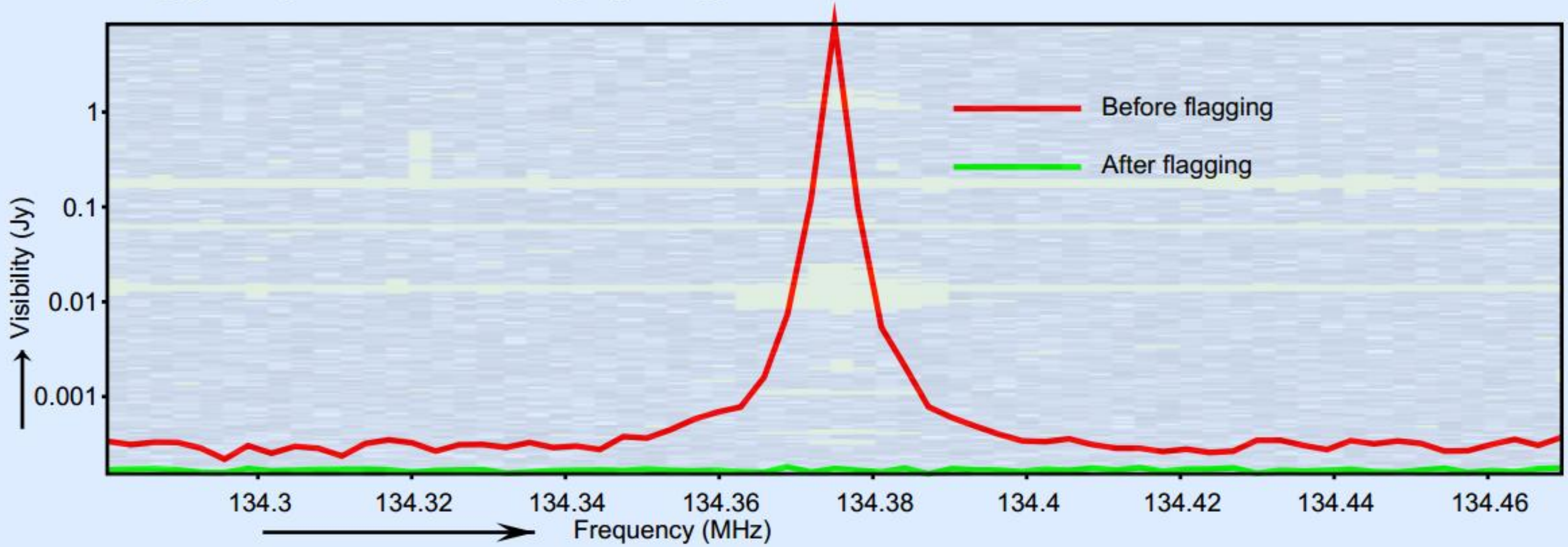
AOFlagger, rfigui

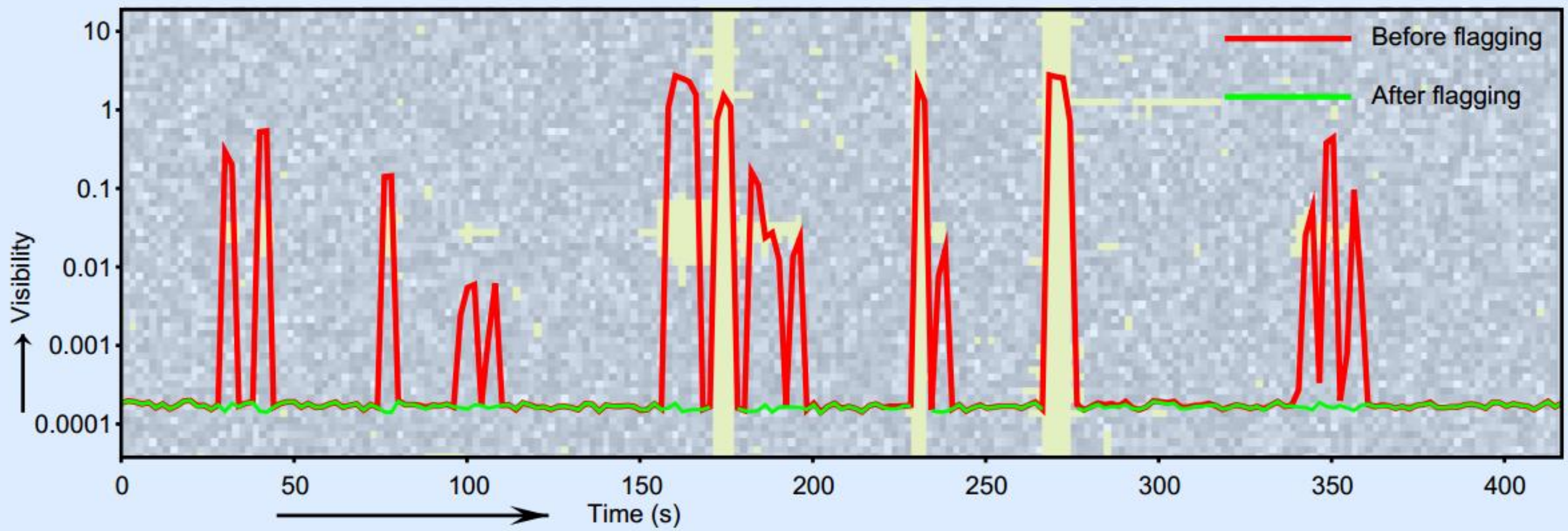
AOFlagger (André Offringa) flags data based on statistics:



AOFlagger, rfigui

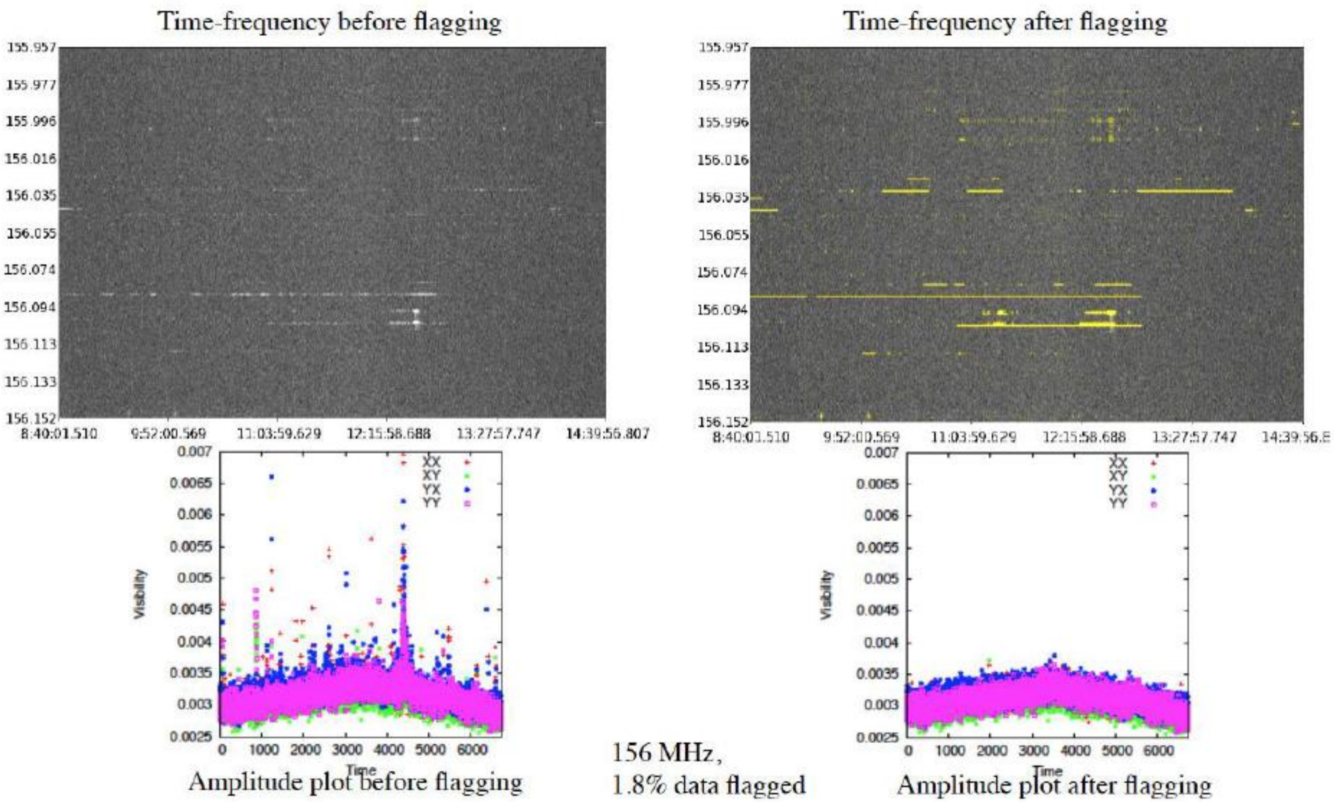
AOFlagger (André Offringa) flags data based on statistics:





RFI Flagging: AOFlagger

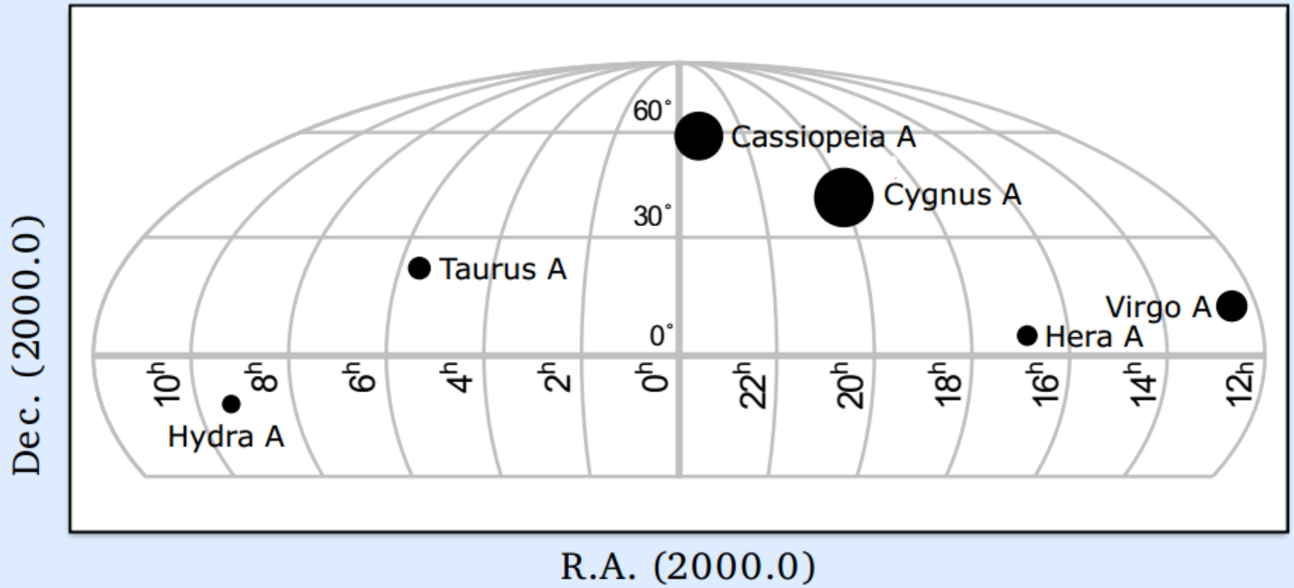
- '*SumThreshold*' method to detect series of samples with high values
- Iterative analysis for entire obs., per subband and spectral channel
- Used for LOFAR but available as stand-alone package



Offringa et al. (2010, 2012)

Demixing with BBS

Sky at low frequencies is dominated by a few sources, together called **A-team** sources.



If A-team source is affecting signal, its signal needs to be subtracted.
To subtract, the data must be calibrated against a model of the A-team source.
Time and frequency resolution needs to be such that signal from A-team sources is not too much affected by time and frequency smearing.

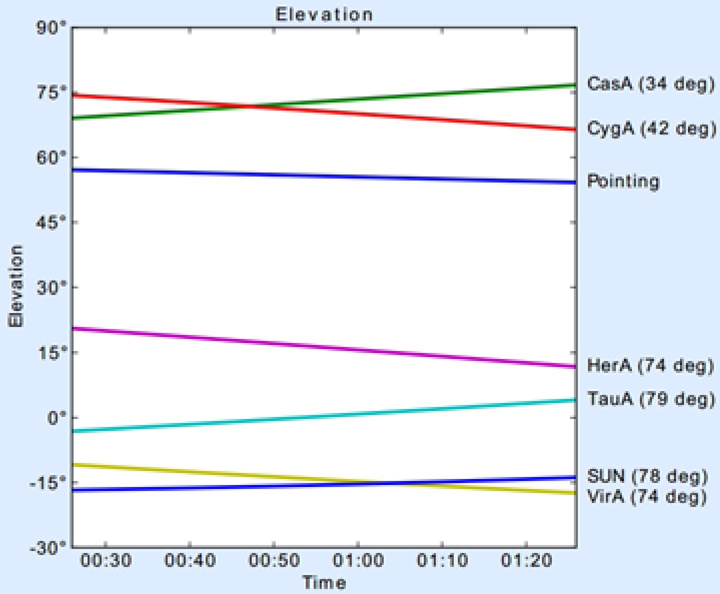
Demixing: is your data affected by A-team?

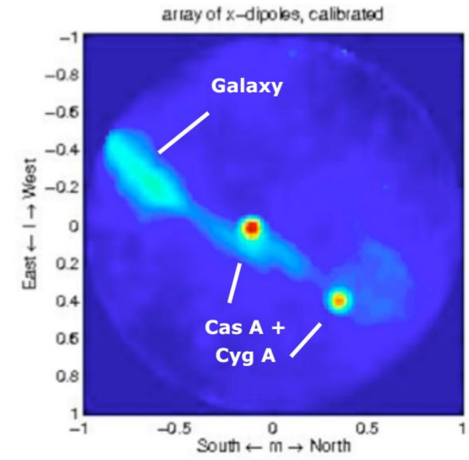
ASTRON

LBA: yes, your data is affected by CygA and CasA and perhaps more

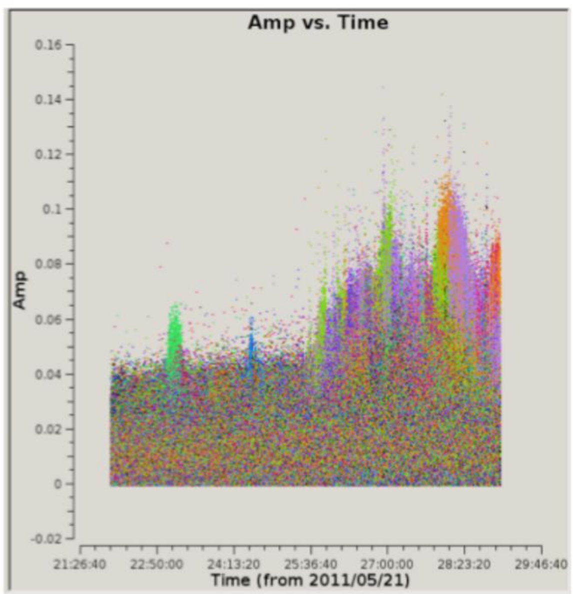
HBA: your data might be affected by A-team:

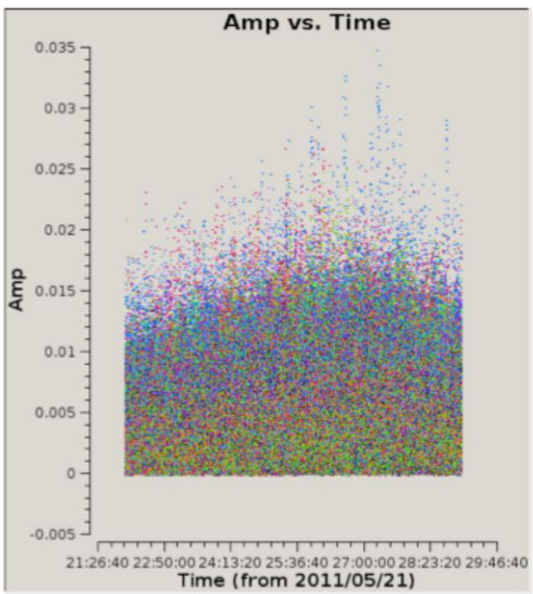
- If target is within 30° separation of A-team source
- If A-team elevation is high during observation





Model interferometer response and subtract or flag affected data



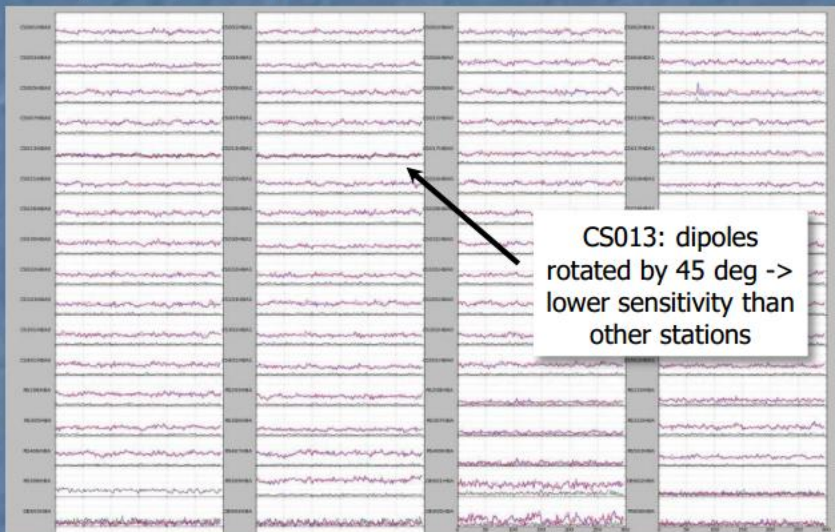


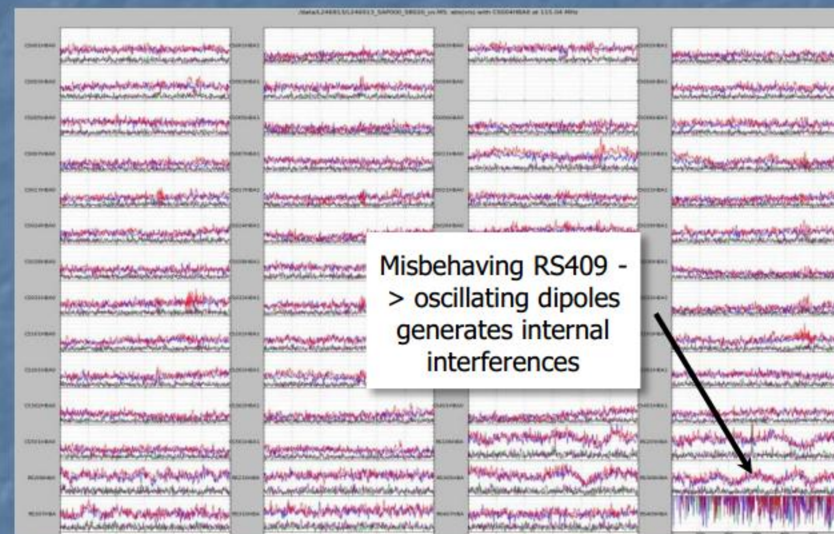
INSPECTION PLOTS: TIME SERIES

LOFAR

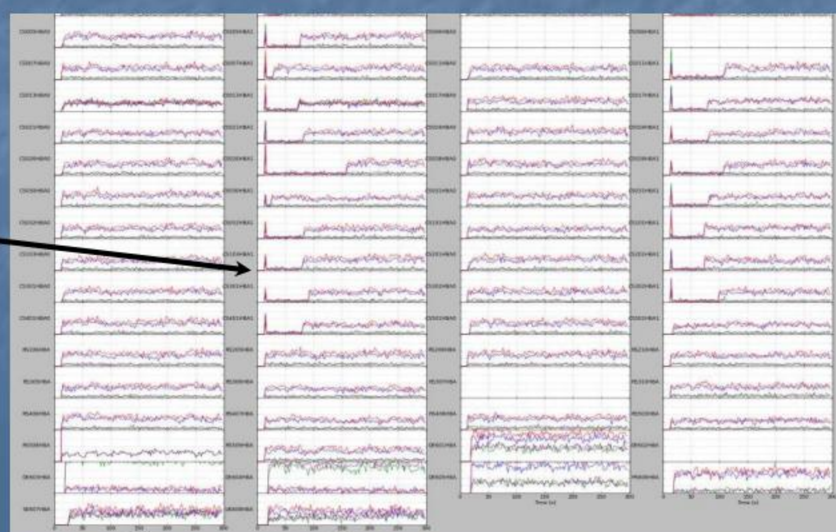
ASTRON

- See tutorial at http://www.astron.nl/radio-observatory/observing-capabilities/depth-technical-information/data-quality-inspection/data-qu





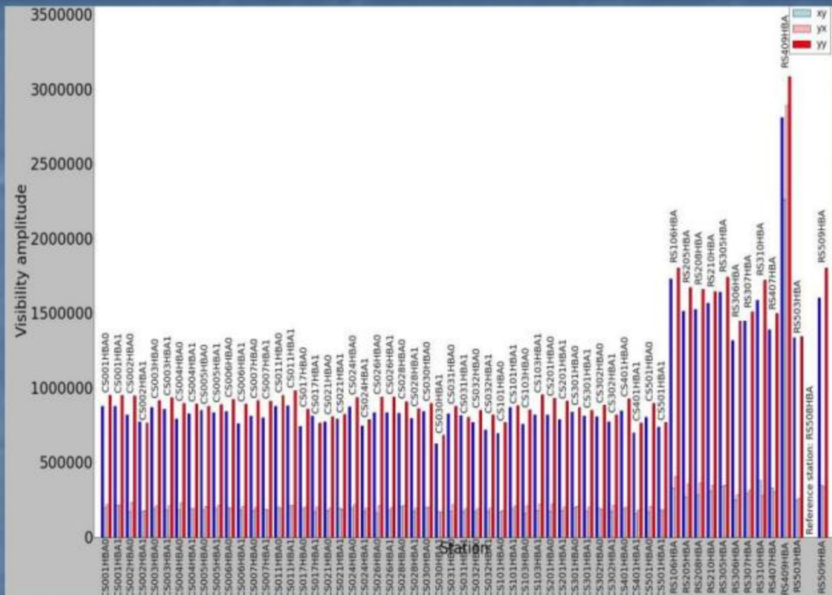
Connection problems at the beginning of the observation for all stations

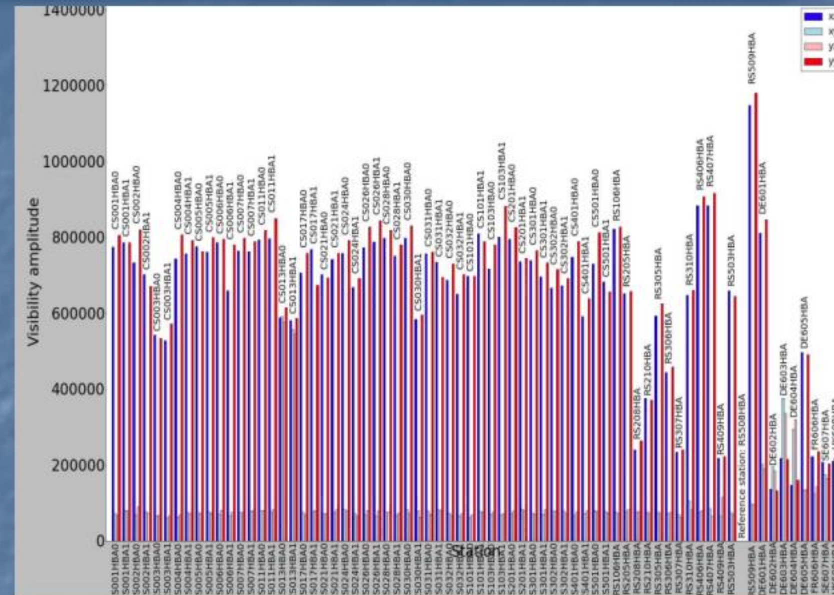


INSPECTION PLOTS: VISIBILITY AMPLITUDE

LOFAR

ASTRON





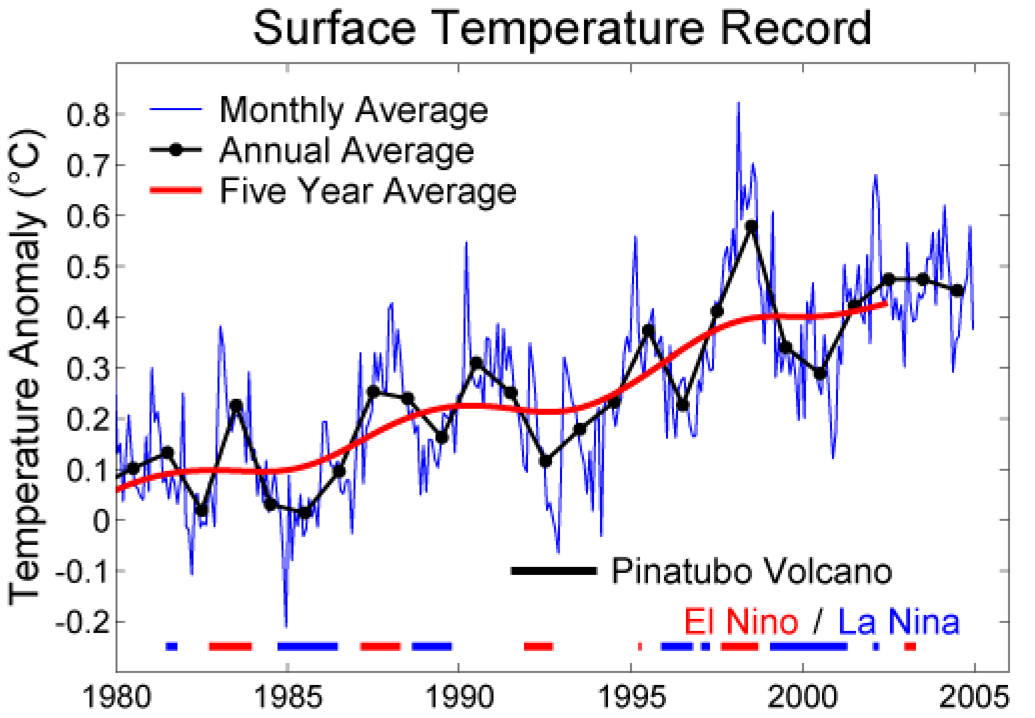
- Stations with the same characteristics, e.g. all CS should have more or less the same amplitudes; when their values differ too much the sensitivity of the station is not good (in this case RS409)
- It is good to check on a long baseline a frequency with high S/N which is relatively clean of interferences (e.g. correlator SB 77 in HBA and 301-302 in LBA)
- A way to identify the presence of Solar bursts is to check if the amplitude visibilities scales of CS have much higher values than the remote station or CS in quiet conditions.
- Visibility amplitude 3C196 (resolved at long baselines)
- Different amplitudes values among RS and IS are due to the fact that some baselines detect and resolve source structure and some other do not.
- Knowing the layout of the telescope and the characteristics of the source are the key to interpret these plots.

See tutorial at http://www.astron.nl/radio-observatory/observing-capabilities/depth-technical-information/data-quality-inspection/data-qu

Po inspekcji jakościowej, flagowaniu, demixingu, średniujemy dane w czasie i częstotliwości np.

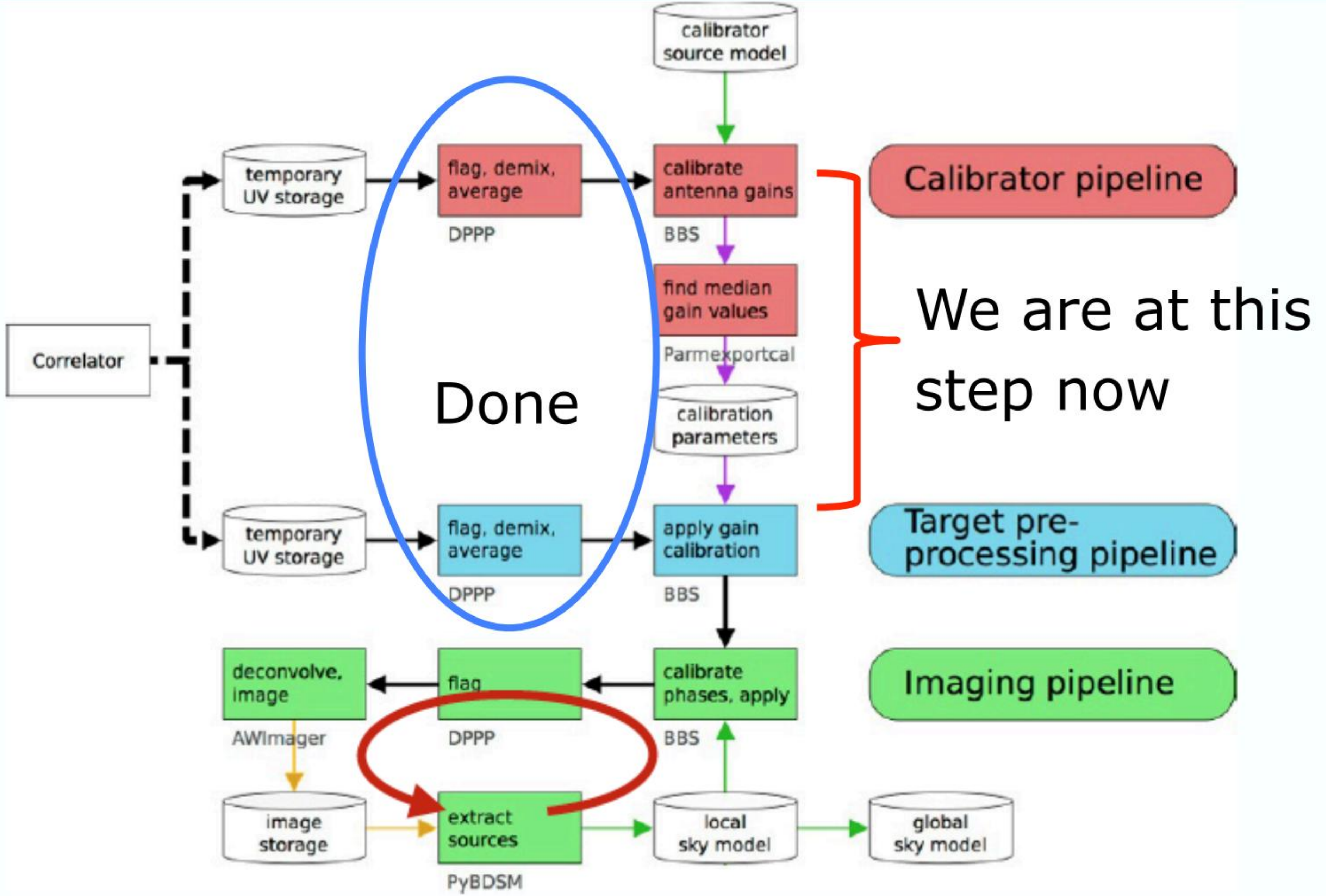
1s => 4 s

64 => 4 ch/sb



Zautomatyzuj cały proces (ułóż skrypt na wszystkie sbb) – pipeline

2. Kalibracja G_{ij}, poprawa V_{ij}



Kalibracja

1) Kalibracja „antena gains" (amplituda+faza) z danych kalibratora – model źródła znany (np. źródło punktowe) - Black Board Self-calibration (BBS)
   Ocena rozwiązań – casaplotms
2) Przeniesienie rozwiązań na target (BBS)
3) Kalibracja źródła (faza) – model źródła np. z Global Sky Model
   Ocena rozwiązań – casaplotms
4) Flagowanie błędnych rozwiązań (DPPP)
5) Automatyzacja zadań, dobór parametrów

Calibration:
Calibrate in amplitude your Target

ASTRON

- Required:
  - 1 flagged/demixed/averaged subband from Calibrator
  - 1 flagged/demixed/averaged subband from Target
  - 1 parset file
  - 1 Skymodel of the Calibrator

- N.B: frequency must be identical, quick check with msoverview

```
msoverview: Version 20110407GvD
================================================================================
           MeasurementSet Name:  /data/scratch/mahony/tutorial_t2/cal_averaged.MS      MS Version 2
================================================================================
   Observer: unknown     Project: MSSS_HBA_2013
Observation: LOFAR
Antenna-set: HBA_DUAL_INNER

Data records: 10620       Total integration time = 60.0834 seconds
   Observed from   29-Mar-2013/13:59:48.0   to   29-Mar-2013/14:00:48.1 (UTC)


Fields: 1
  ID   Code Name                RA               Decl           Epoch        nRows
  0         BEAM_0              01:37:41.299440 +33.09.35.13240 J2000        10620

Spectral Windows:  (1 unique spectral windows and 1 unique polarization setups)
  SpwID  Name   #Chans   FrameCh1(MHz)     ChanWid(kHz)  TotBW(kHz) CtrFreq(MHz)  Corrs
  0      SB-31      4    TOPO      134.300          48.828       195.3    134.3735   XX  XY  YX  YY
```

Calibration:
Write your parset file

ASTRON

solvecal.parset

Strategy.ChunkSize = 0	read entire MS into memory (By default BBS reads in the DATA column)
Strategy.Steps = [solve,correct]	
Step.solve.Operation = SOLVE	
Step.solve.Model.Sources = [3c48]	The source you want to solve for
Step.solve.Model.Gain.Enable = T	
Step.solve.Model.Beam.Enable = T	Correct for the beam
Step.solve.Solve.Parms = ["Gain:0:0:*","Gain:1:1:*"]	Solve for xx,yy (not xy,yx), * means amplitude and phase
Step.solve.Solve.CellSize.Freq = 0	Number of channels to solve for (0=all)
Step.solve.Solve.CellSize.Time = 1	Number of timeslots to solve for (1=solve for every timestamp)
Step.solve.Solve.CellChunkSize = 10	Number of solution cells to simultaneously process
Step.solve.Solve.Options.MaxIter = 50	These are the default values that determine the stop criteria – see wiki for details (applies to Options.MaxIter … Options.UseSVD)
Step.solve.Solve.Options.EpsValue = 1e-9	
Step.solve.Solve.Options.EpsDerivative = 1e-9	
Step.solve.Solve.Options.ColFactor = 1e-9	
Step.solve.Solve.Options.LMFactor = 1.0	
Step.solve.Solve.Options.BalancedEqs = F	
Step.solve.Solve.Options.UseSVD = T	
Step.correct.Operation = CORRECT	
Step.correct.Model.Gain.Enable = T	
Step.correct.Model.Beam.Enable = T	Apply beam
Step.correct.Model.Sources = [3c48]	
Step.correct.Output.Column = CORRECTED_DATA	

Zamiast BBS możliwe też użycie DPPP

Calibration:
Run BBS: extract solution from the Calibrator

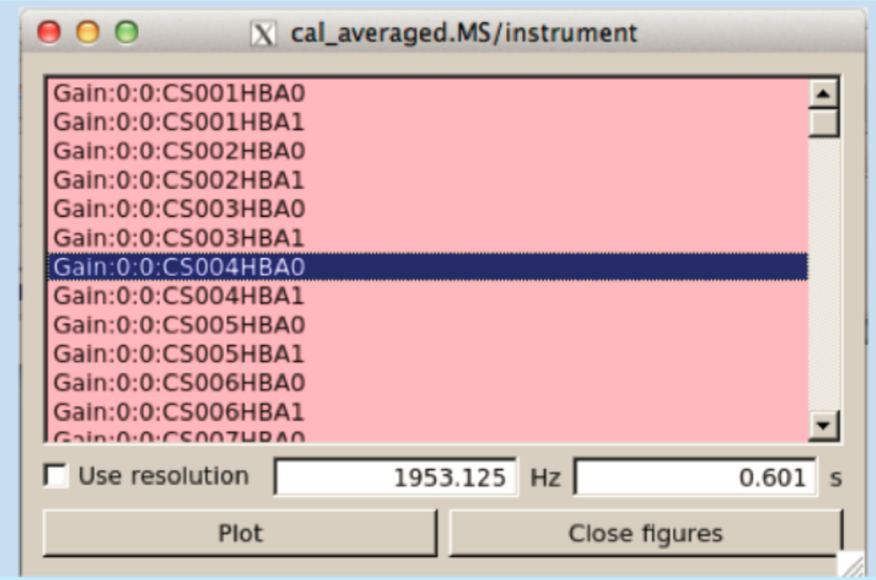
- Type:

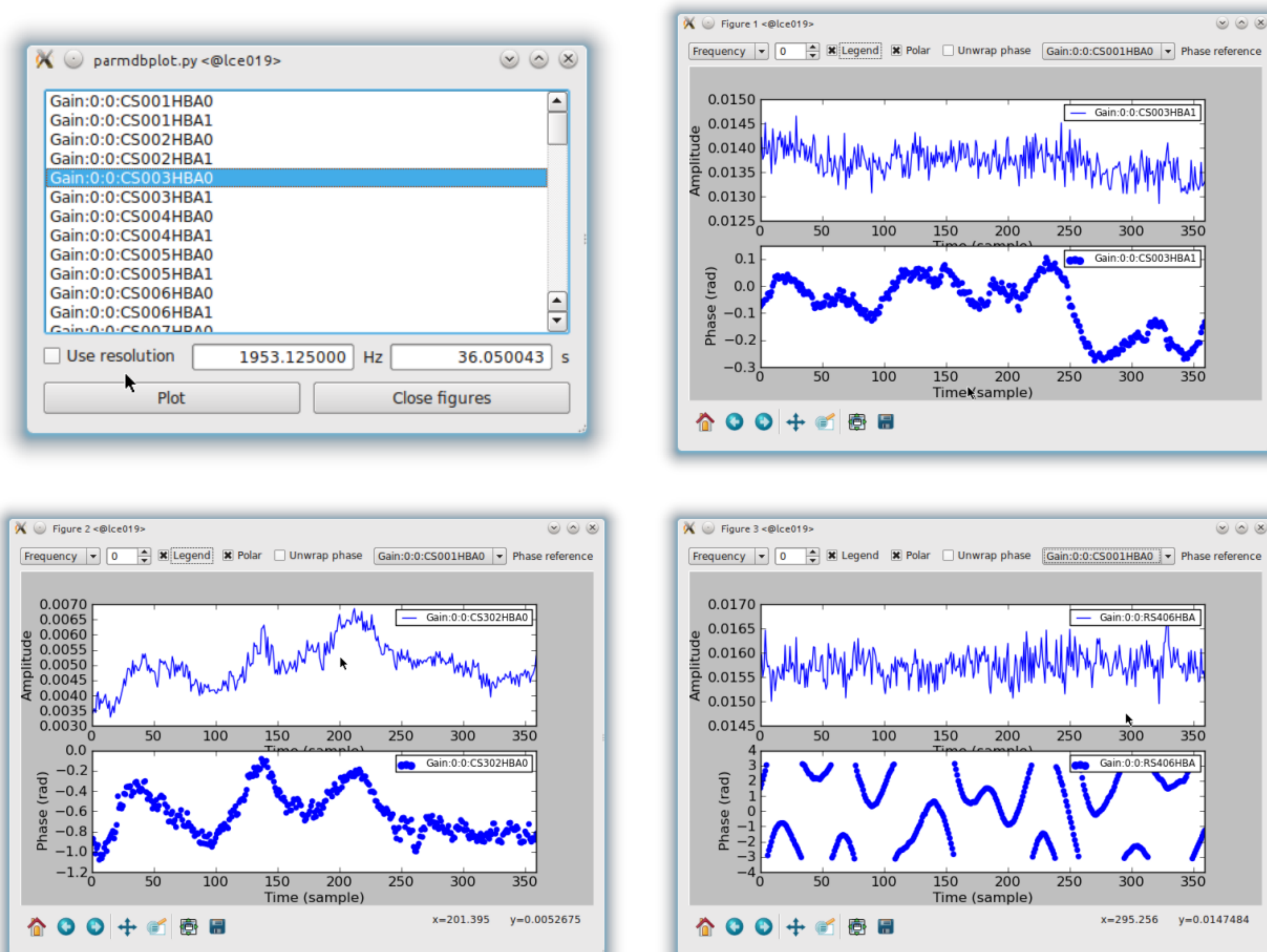
“ calibrate-stand-alone mycalibrator.MS myparset mycalibrator_Skymodel ”

- The solution is written in mycalibrator.MS/instrument

- It could (must) be visualized and investigating with parmdbplot by typing:

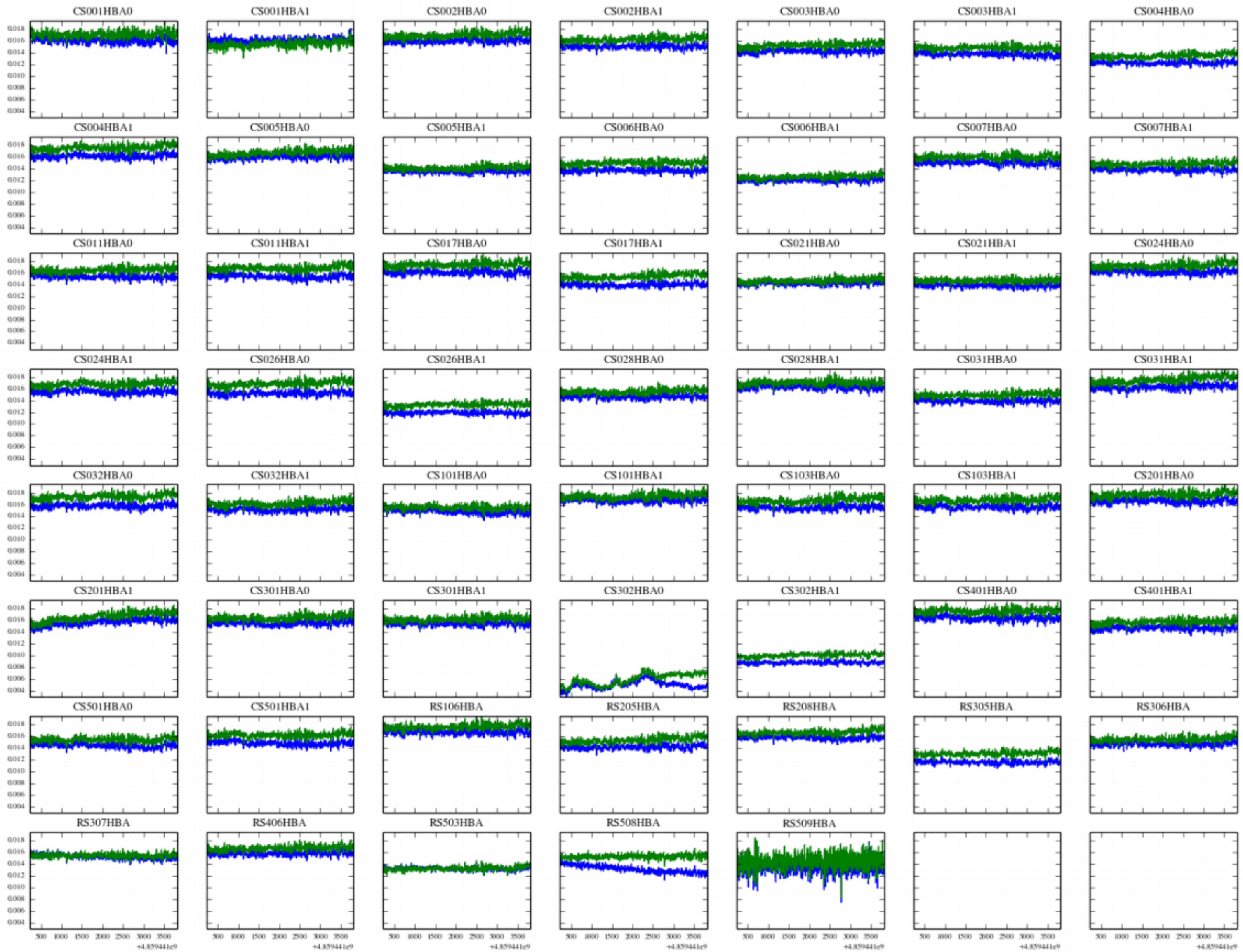
“parmdbplot.py mycalibrator.MS/instrument”





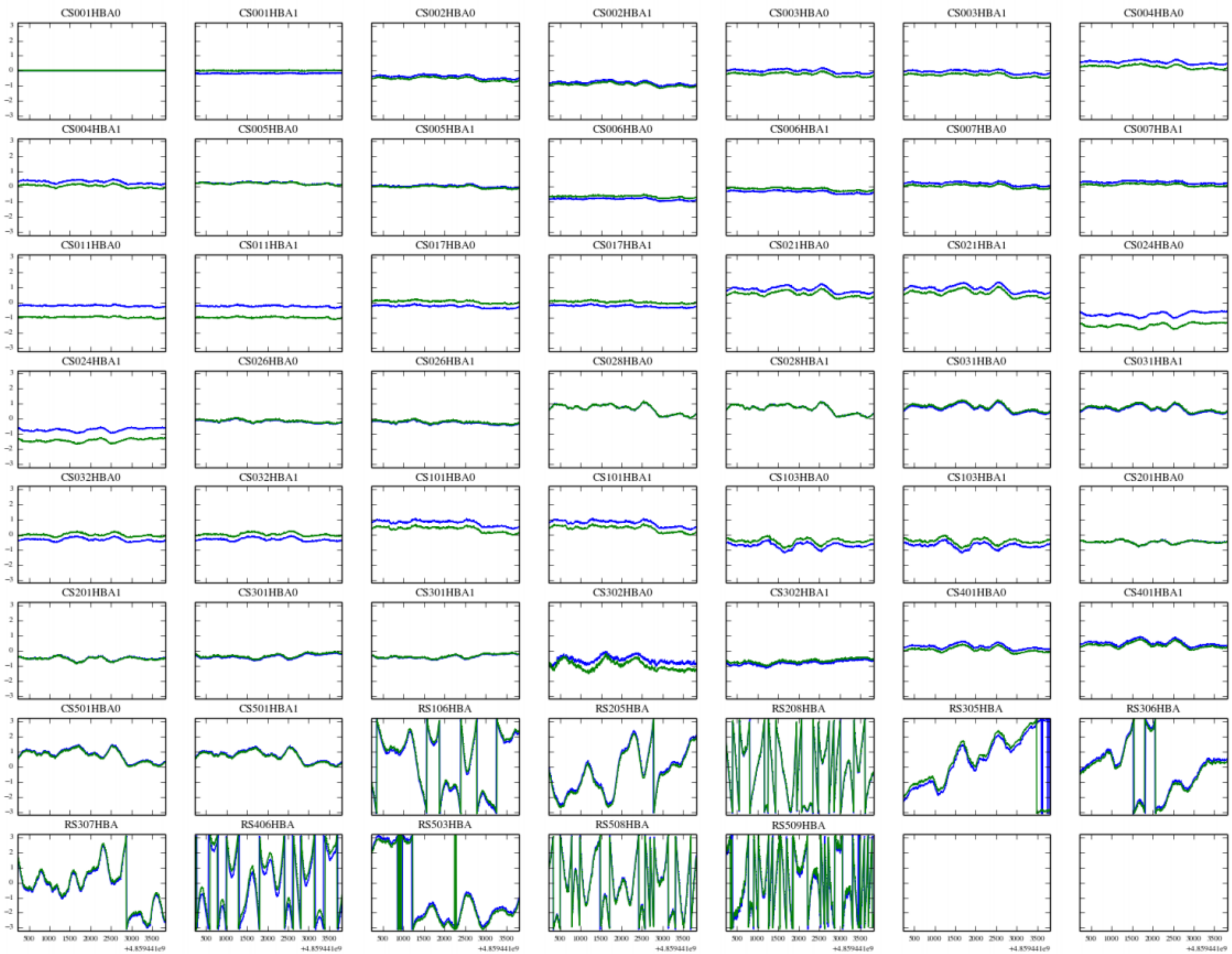
parmdbplot - like uvplot in AIPS

Amplitude solutions – should be ~constant



Phase solutions

typically worse solutions for remote stations



Calibration:
Run BBS: Transfert the solution to the Target

- We need to make the solutions time-independent to transfer the gains to the target field. So type:

"Parmexportcal in=mycalibrator.MS/instrument out=mysol"

- Apply gain solutions to target field by doing a correct step in BBS:

"calibrate-stand-alone --parmdb mysol mytarget.MS transfersolns.parset"

transfersolns.parset:

```
Strategy.ChunkSize = 0
Strategy.Steps = [correct]

Step.correct.Operation = CORRECT
Step.correct.Model.Sources = []
Step.correct.Model.Gain.Enable = T
Step.correct.Model.Beam.Enable = F
Step.correct.Output.Column = CORRECTED_AMP
```

NOTE: do NOT apply the beam in this correct step. We only want to apply the beam at the last correct step before imaging!

Calibration:
Run BBS: Phase calibration on the Target

ASTRON

- Get a skymodel for the target field
  - Run msoverview to get the co-ordinates of the pointing centre (RA=01:02:21.73, Dec=+31:27:36.0)
- Get the GSM skymodel for this field using gsm.py

```
> gsm.py –h
> gsm gsm.py targetfield.skymodel 15.59 31.46 3 1
Sky model stored in source table: targetfield.skymodel

> more targetfield.skymodel
```

```
mahony@lof013:/data/scratch/mahony/tutorial_t2$ more targetfield.skymodel
FORMAT = Name, Type, Ra, Dec, I, Q, U, V, ReferenceFrequency='60e6', SpectralIndex='[0.0]', MajorAxis, MinorAxis, Orientation

# the next lines define the sources
0049.0+3220, POINT, 00:49:01.94880000, +32.20.23.20800000, 2.8587, , , , , [-0.5724, -0.1103]
0050.2+3229, POINT, 00:50:17.52960000, +32.29.14.38800000, 2.7945, , , , , [-0.657, -0.1036]
0050.9+3050, POINT, 00:50:56.46000000, +30.50.03.58800000, 1.2646, , , , , [-0.7373, -0.1638]
0053.7+2925, GAUSSIAN, 00:53:44.60880000, +29.25.10.88400000, 5.0451, , , , , [-0.7525, 0.0181], 49.7, 35.0, 165.7
0053.8+3114, GAUSSIAN, 00:53:49.51920000, +31.14.48.91200000, 8.7384, , , , , [-0.8641], 43.4, 40.3, 32.8
0054.1+3203, POINT, 00:54:09.52080000, +32.03.43.99200000, 1.3, , , , , [-0.7]
0054.1+3101, POINT, 00:54:09.95040000, +31.01.59.41200000, 1.2422, , , , , [-0.5142, -0.2463]
0054.2+3201, POINT, 00:54:17.53920000, +32.01.06.88800000, 3.4466, , , , , [-0.6143]
0054.3+3353, POINT, 00:54:22.03920000, +33.53.36.09600000, 2.0607, , , , , [-0.4791, -0.1043]
0054.6+3219, POINT, 00:54:41.88000000, +32.19.04.58400000, 1.1784, , , , , [-0.6299, -0.1301]
0057.7+3021, GAUSSIAN, 00:57:46.60080000, +30.21.34.59600000, 4.25, , , , , [-0.7], 105.0, 52.7, 136.7
0058.0+3121, POINT, 00:58:05.69040000, +31.21.13.60800000, 2.8847, , , , , [-0.0435, -0.3402]
```

Calibration:
Run BBS: Phase calibration on the Target

ASTRON

solve_phaseonly.parset

```
Strategy.InputColumn = CORRECTED_AMP # define input column
Strategy.ChunkSize = 500
Strategy.Steps = [solve, correct]

Step.solve.Operation = SOLVE
Step.solve.Model.Sources = []  #solves for all sources in skymodel
Step.solve.Model.Cache.Enable = T
Step.solve.Model.Phasors.Enable = T
Step.solve.Model.Gain.Enable = T
Step.solve.Model.Beam.Enable = T
Step.solve.Model.Beam.UseChannelFreq = F
Step.solve.Solve.Mode = COMPLEX #use COMPLEX not PHASE
Step.solve.Solve.Parms = ["Gain:0:0:Phase:*", "Gain:1:1:Phase:*"]
Step.solve.Solve.CellSize.Freq = 0
Step.solve.Solve.CellSize.Time = 1
Step.solve.Solve.CellChunkSize = 40
Step.solve.Solve.PropagateSolutions = F   #don't use previous
Step.solve.Solve.Options.MaxIter = 50      solution as starting guess
Step.solve.Solve.Options.EpsValue = 1e-9
Step.solve.Solve.Options.EpsDerivative = 1e-9
Step.solve.Solve.Options.ColFactor = 1e-9
Step.solve.Solve.Options.LMFactor = 1.0
Step.solve.Solve.Options.BalancedEqs = F
Step.solve.Solve.Options.UseSVD = T
```

```
Step.correct.Operation = CORRECT
Step.correct.Model.Sources = []
Step.correct.Model.Phasors.Enable  = T
Step.correct.Model.Gain.Enable  = T
Step.correct.Model.Beam.Enable  = T
Step.correct.Model.Beam.UseChannelFreq = F
Step.correct.Output.Column = CORRECTED_DATA
```

UseChannelFreq – this option needs to be set to True when using datasets where multiple subbands have been combined. (this corrects for how the beam changes with frequency).

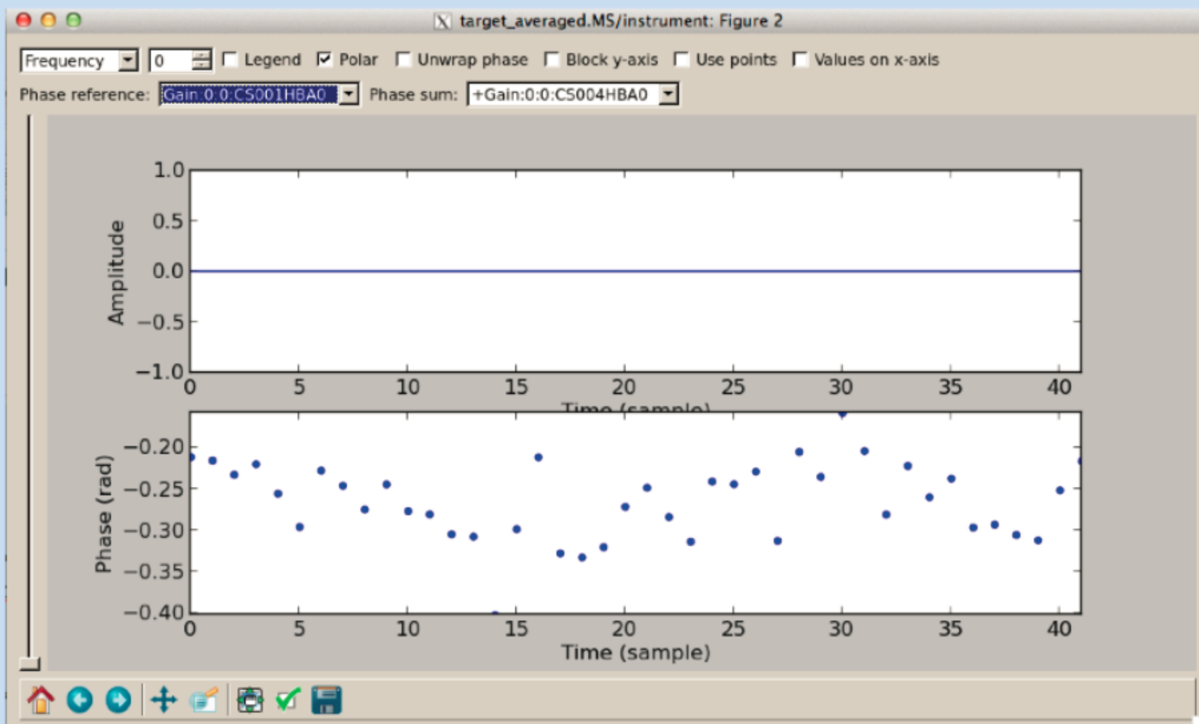
Calibration:
Run BBS: Phase calibration on the Target

- **Run BBS:**

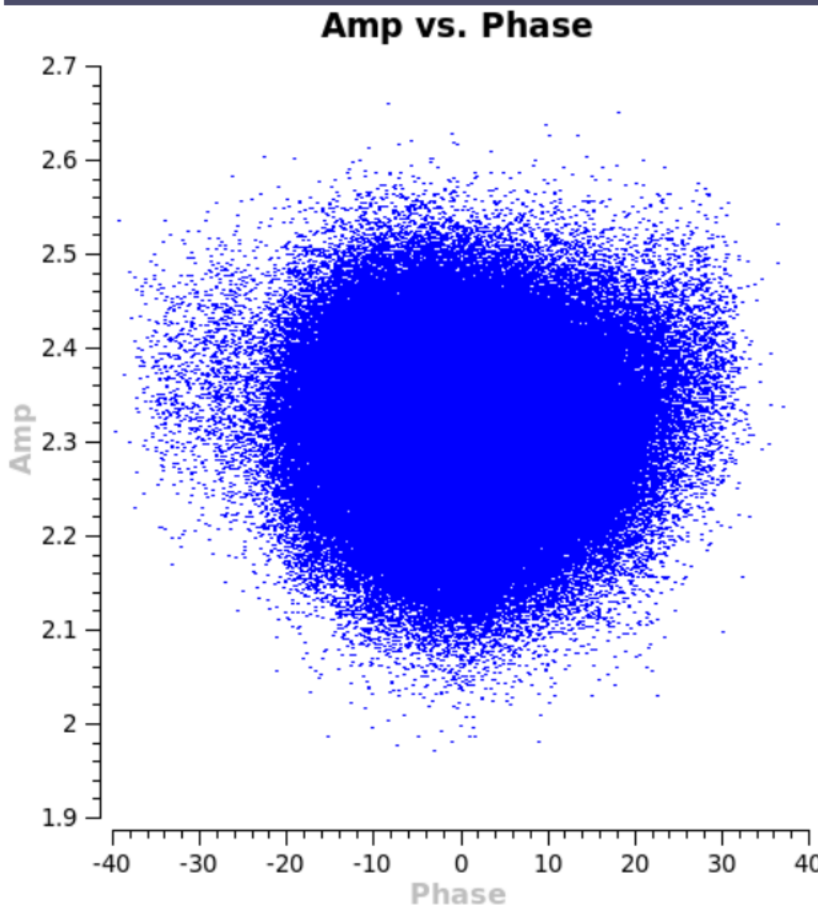
```
> calibrate-stand-alone -f target_averaged.MS/ solve_phaseonly.parset
targetfield.skymodel > phasecal.log &
```

- **Inspect solutions:**

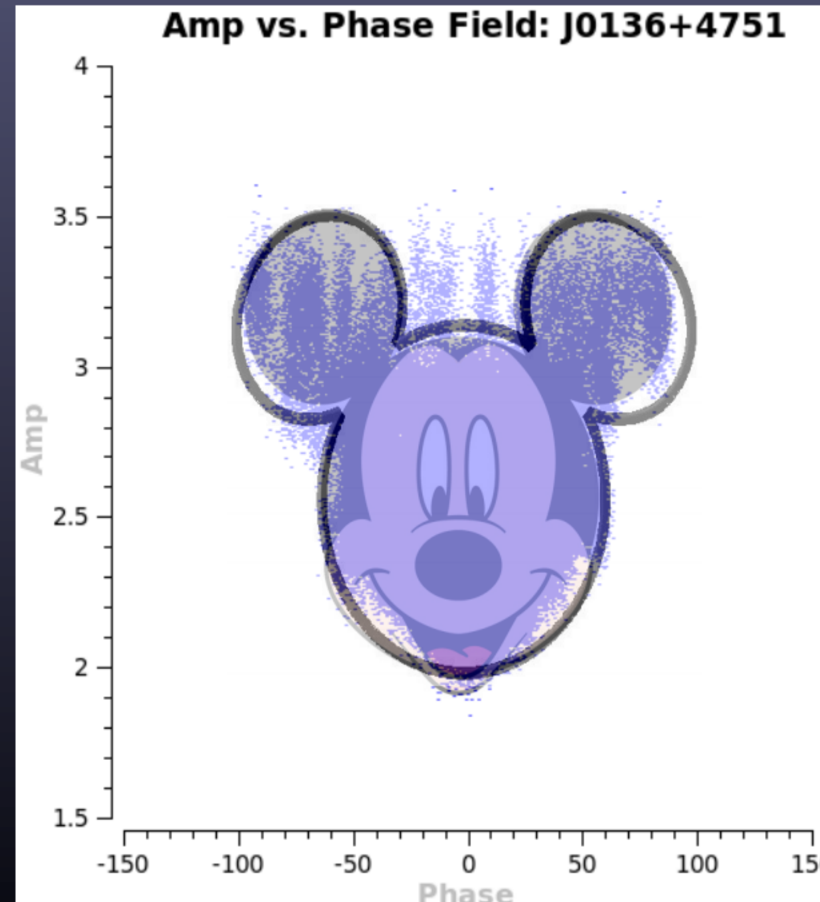
```
> parmdbplot.py target_averaged.MS/instrument/ &
```



DATA DISPLAYS(5) – Amplitude vs Phase



Good

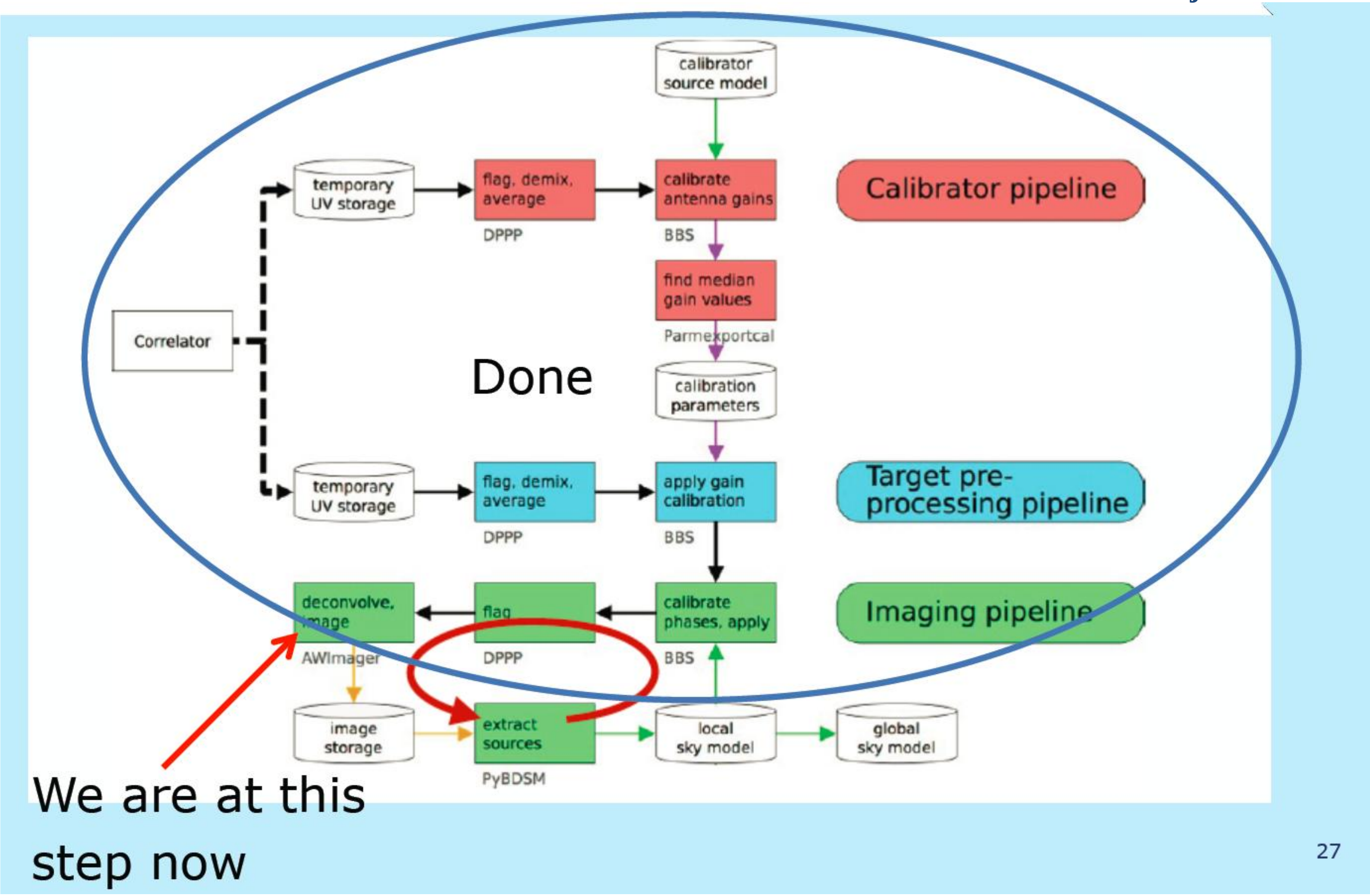


Bad

Po opracowaniu poprawnej strategii kalibracji zautomatyzuj zadania

3. Dekonwolucja i obrazowanie

$I_{lm} = FFT\ V_{ij}$



Brudna i czysta mapa – AWIMAGER (obecnie częściej wsclean)

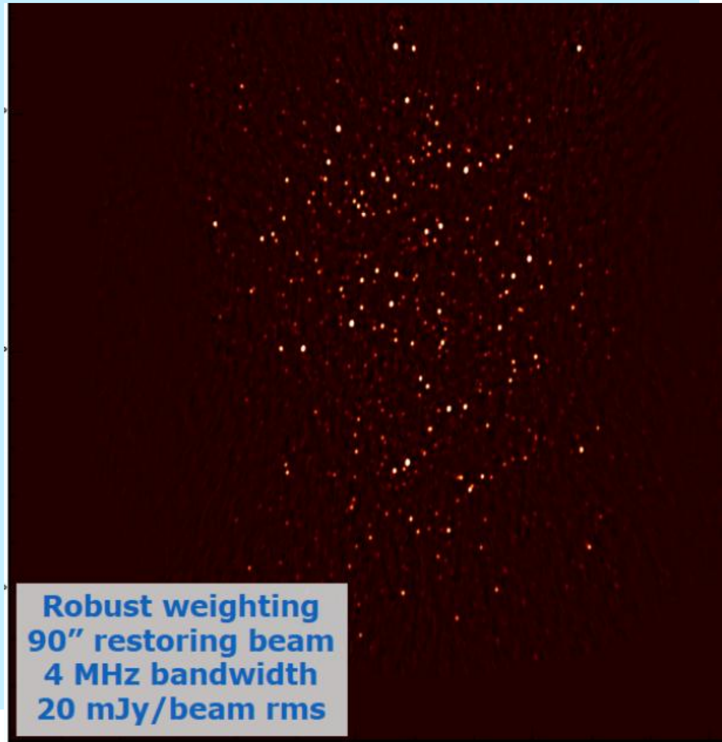
- Required:
  - 1 subband amplitude and phase calibrated from Target
  - 1 parset like:

```
ms=mytarget.MS
image=myImage
weight=briggs
robust=0
npix=3600
cellsize=5arcsec
data=CORRECTED_DATA
padding=1.18
niter=2000
stokes=I
operation=mfclark
timewindow=300
UVmin=0.1
UVmax=10
wmax=20000
fits=""
threshold=0.01Jy
```

define the size of the image:
in this case FOV=5°

define cleaning operation

define used baselines

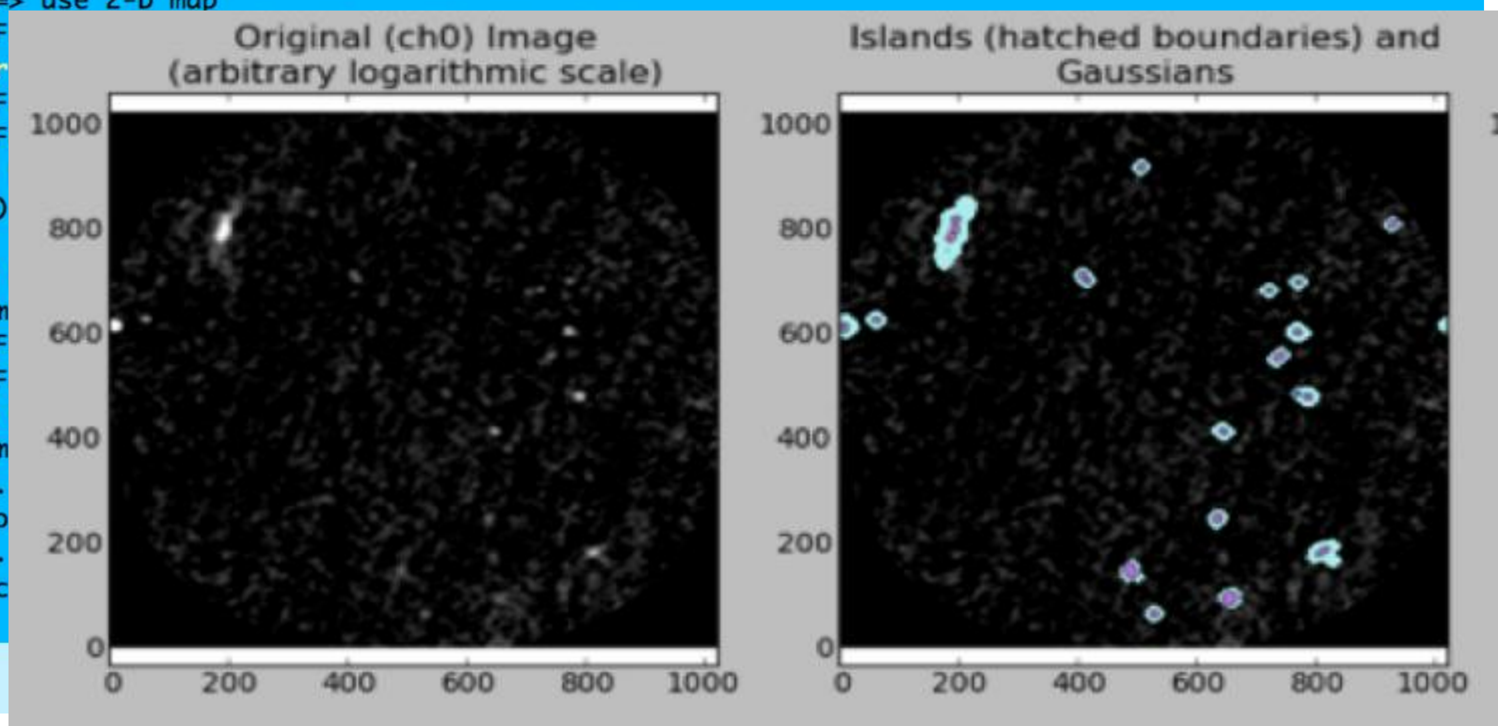


Użycie „maski” – pybdsm+awimager
(Manipulacje na modelach LSMTool)

Type: pybdsm: the inp process_image

```
filename .. 'target_clean_mask.img.restored': Input image file name
adaptive_rms_box ..... False : Use adaptive rms_box when determining rms and mean maps
advanced_opts ........ False : Show advanced options
atrous_do ............ False : Decompose Gaussian residual image into multiple scales
beam .................. None : FWHM of restoring beam. Specify as (maj, min, pos ang E of N) in degrees. E.g.,
beam = (0.06, 0.02, 13.3). None => get from header
flagging_opts ........ False : Show options for Gaussian flagging
frequency ............. None : Frequency in Hz of input image. E.g., frequency = 74e6. None => get from header.
interactive ........... True : Use interactive mode
mean_map .......... 'default': Background mean map: 'default' => calc whether to use or not, 'zero' => 0, 'const'
=> clipped mean, 'map' => use 2-D map
multichan_opts ....... F
output_opts ..........Tr
polarisation_do ...... F
psf_vary_do .......... F
rms_box ...............
E.g., rms_box = (40, 10)

rms_map ...............
calculate inside program
shapelet_do .......... F
spectralindex_do ..... F
thresh ................
detection rate algorithm
thresh_isl ............
extent of island used fo
thresh_pix ............
the mean. If false detec
```

Then GO ....



Flag bad solutions with DPPP

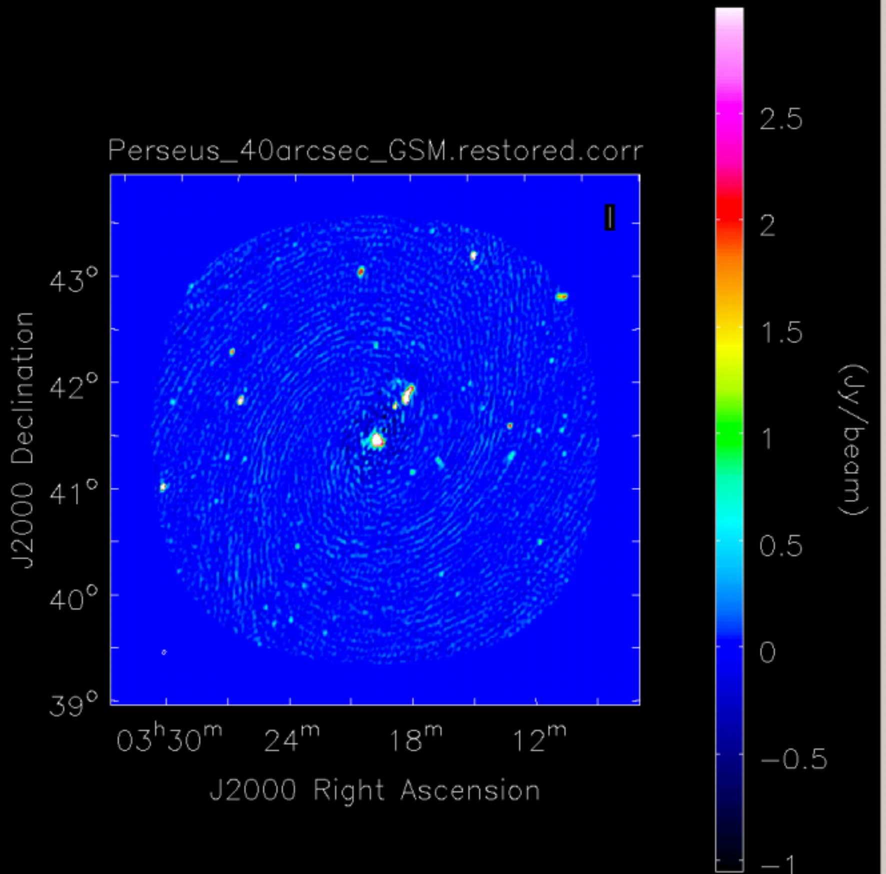
Selfcalibration

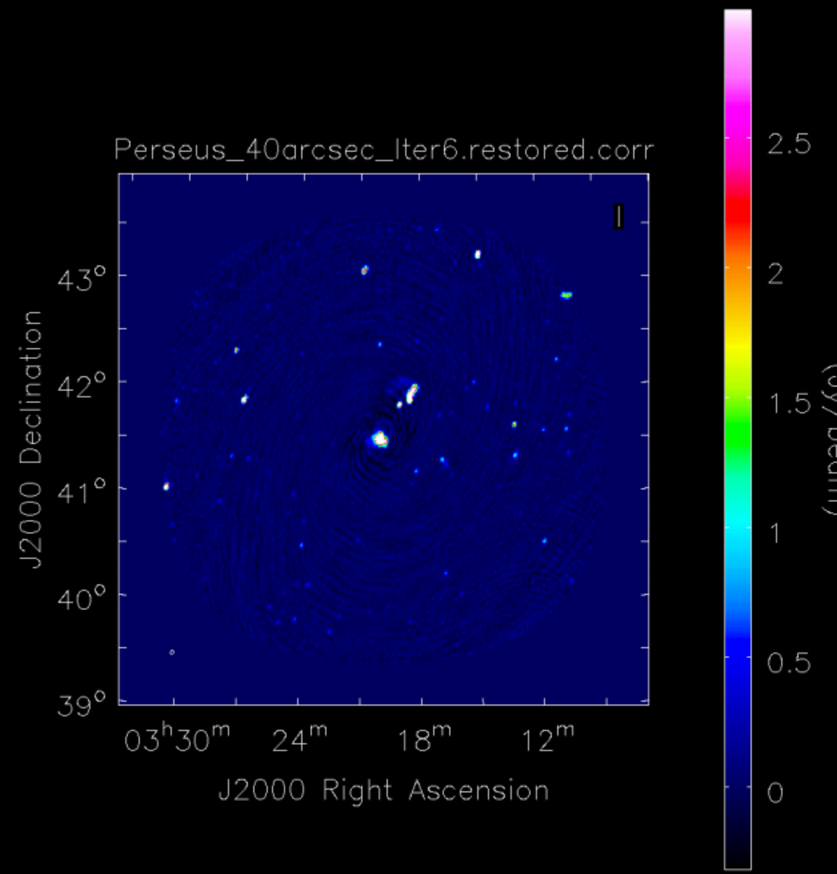
- **How to do:**
  - Define cycle like:
    - Phase Calibration/flagging
    - Imaging
    - Source extraction (sky model generation)
  - Increasing at each cycle the image resolution

Try different cleaning methods

Results on Perseus cluster:

GSM calibration	Self-calibration



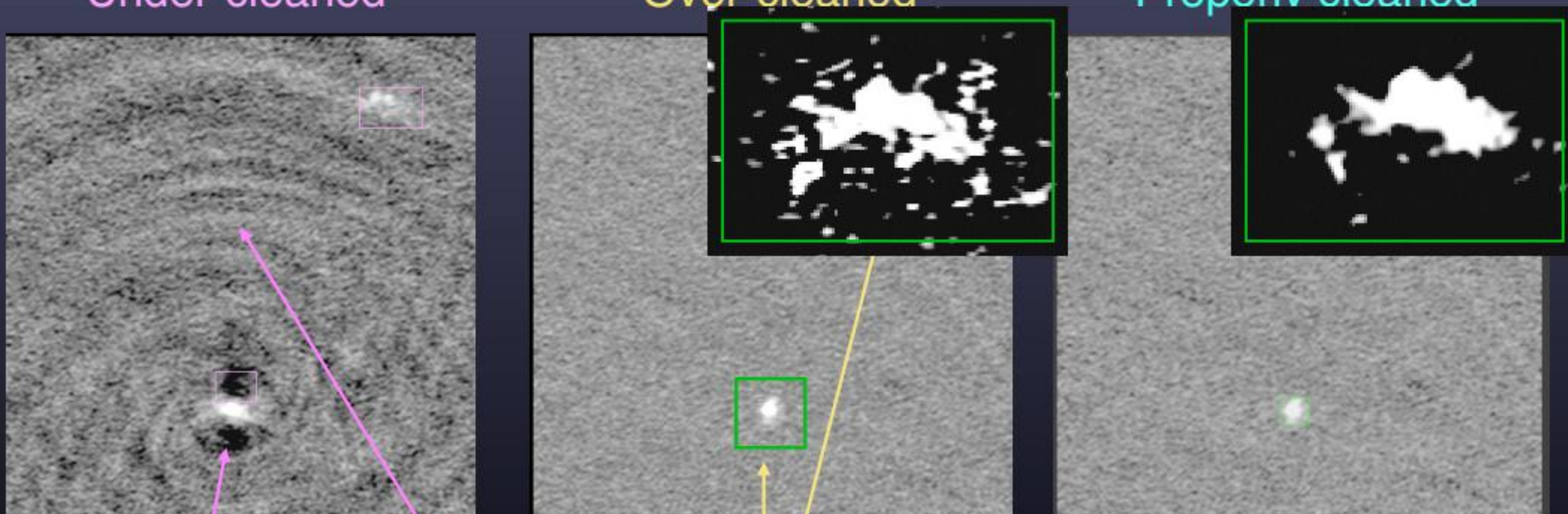


How Deep to Clean?

Under-cleaned

Over-cleaned

Properly cleaned



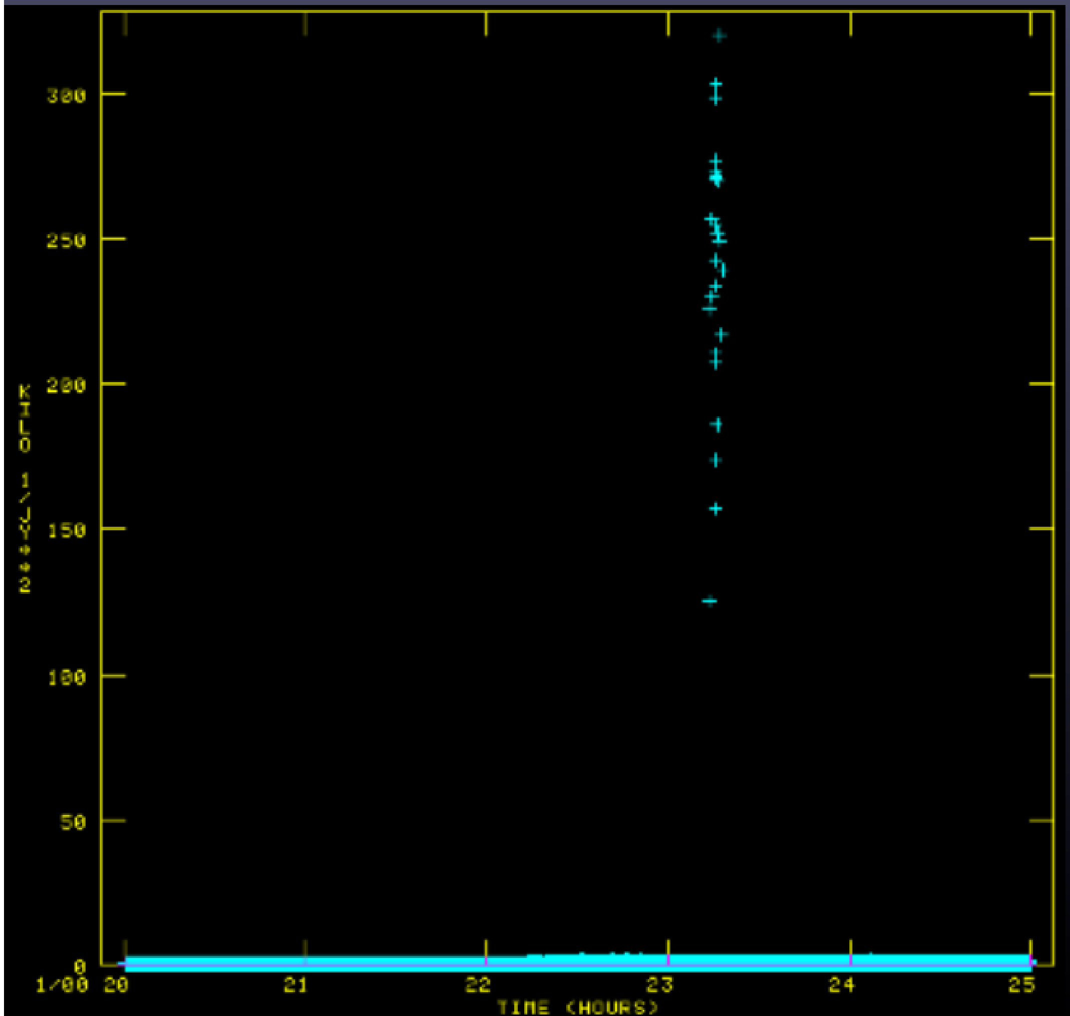
Residual sidelobes dominate the noise

Emission from second source sits atop a negative "bowl"

Regions within clean boxes appear "mottled"

Background is thermal noise-dominated; no "bowls" around sources.

Bad weighting of a few (u,v) points



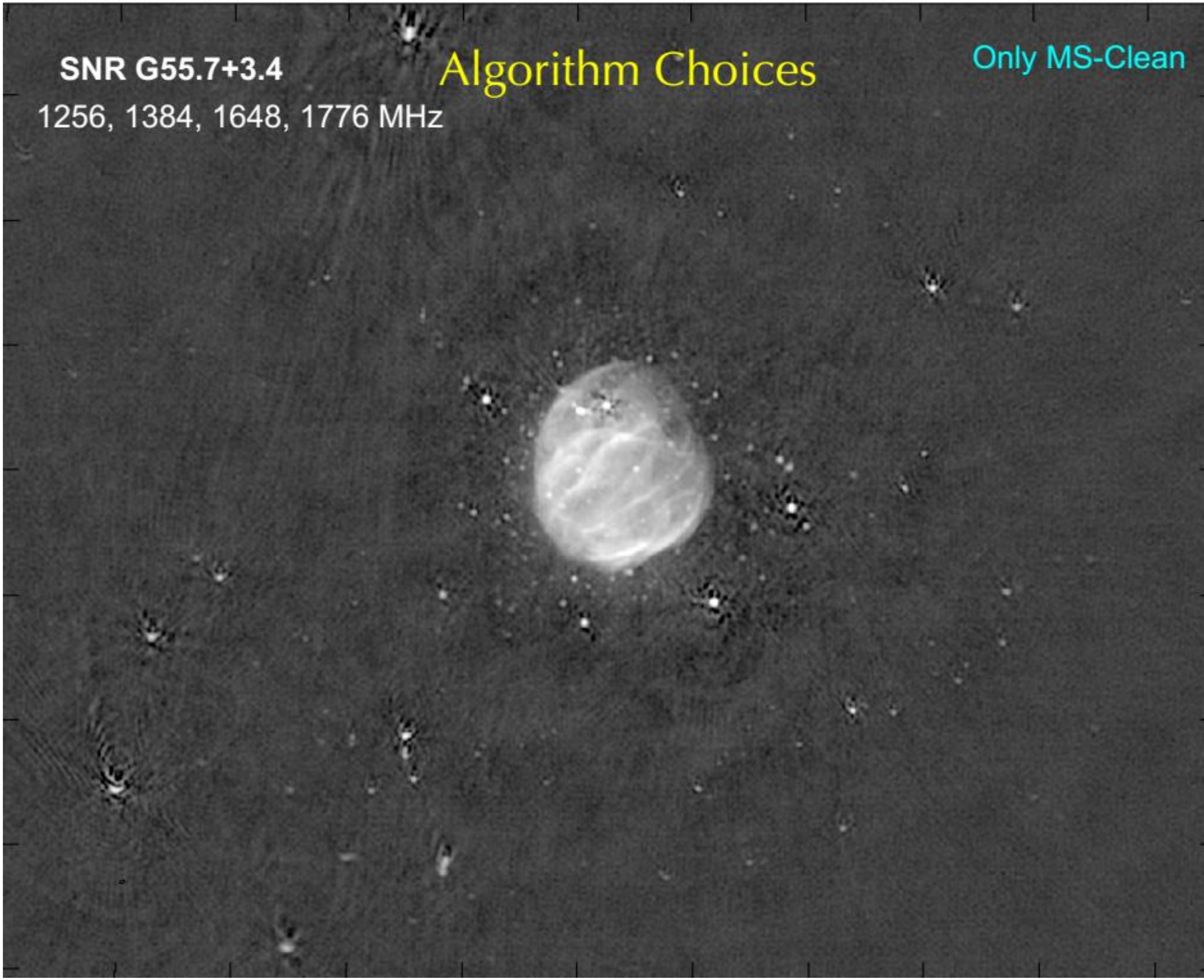
After a long search through the data, about 30 points out of 300,000 points were found to have too high of a weight by a factor of 100.

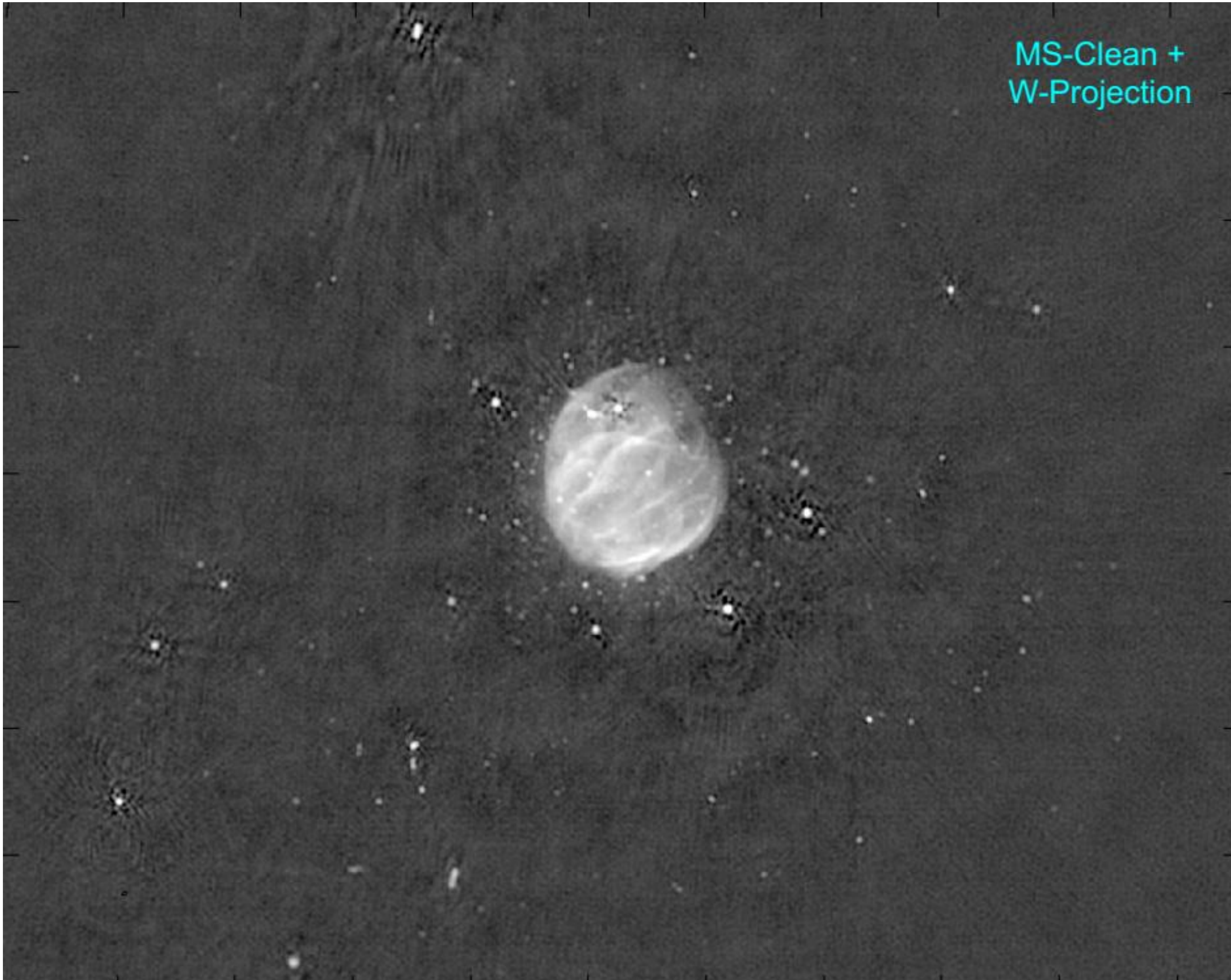
Effect is <1% in image.

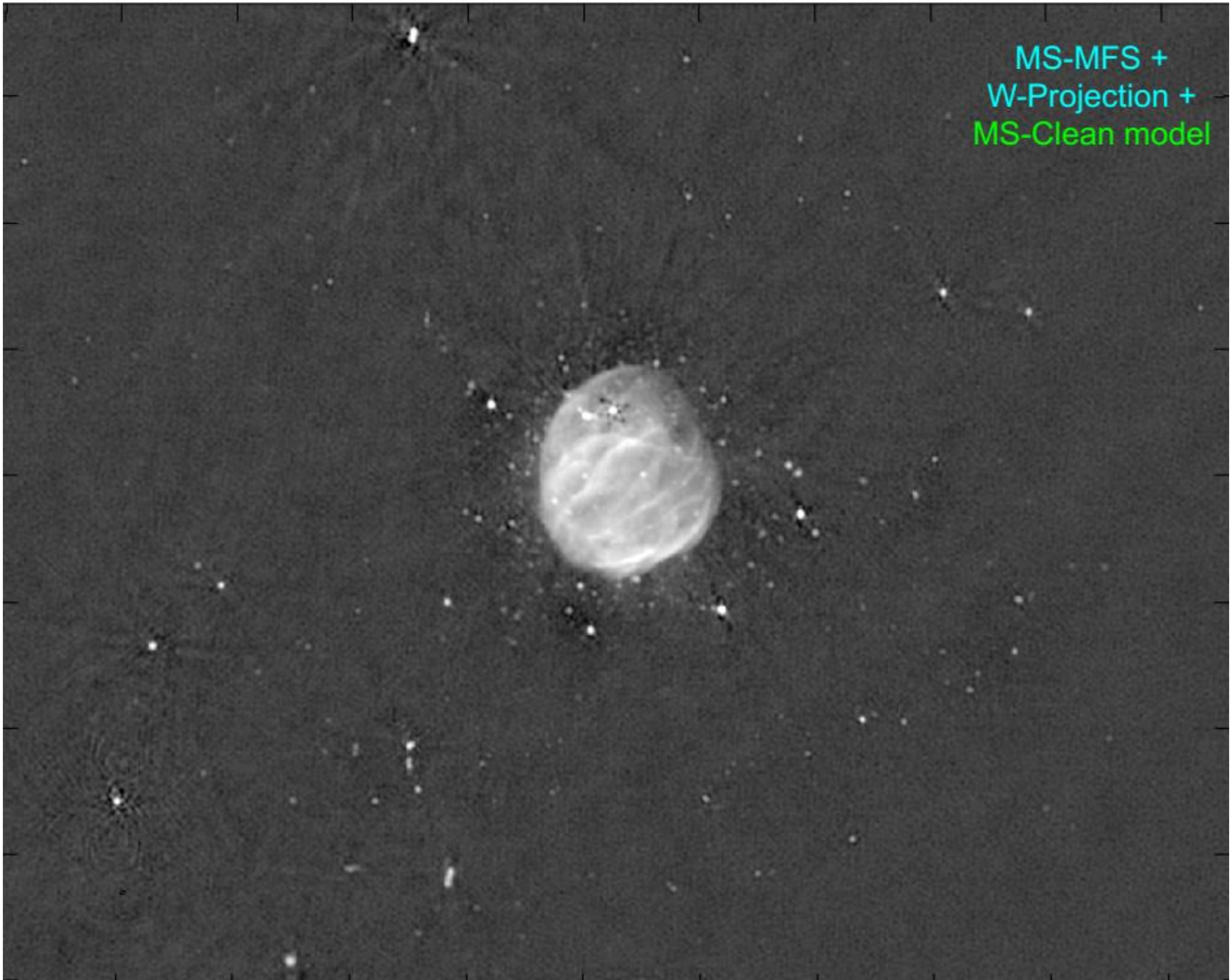
Cause??

Sometimes in applying calibration produced an incorrect weight in the data. Not present in the original data.

These problems can sneak up on you. Beware.


SNR G55.7+3.4
1256, 1384, 1648, 1776 MHz
Algorithm Choices
Only MS-Clean


MS-Clean +
W-Projection


MS-MFS +
W-Projection +
MS-Clean model

Very high DR LOFAR image of 3C196 field   Pandey

4x8h nights

12°x12°
50" PSF

80 Jy peak
80 μJy noise

→DR 10^6 : 1

Note here how bad the off-axis sources appear due to varying station beams