Author: J.E.Noordam Verified by:	Date of issue: 15th November 2005 Kind of issue: Public	Scope: Research Doc.nr.: Report ASTRON-LOFAR-00227
Responsible: Approved by: A.van Ardenne	Status: Preliminary Revision nr.: 0.7	File: *G\data\tekkamer_\45HREPOR.PDF*

LOFAR Calibration and Calibratability

(jnoordam@astron.nl)

Dwingeloo

15th November 2005

Abstract

LOFAR differs from existing radio aperture synthesis telescopes in a number of important respects: The high sensitivity requires a large dynamic range, and at low frequencies the fields are very crowded and the ionosphere causes large and rapid phase variations. This document discusses the adaptation of existing calibration techniques for LOFAR. An important feature is the use of the 25% inner stations as a 'virtual core' that has enough spatial resolution to probe position-dependent effects like the ionosphere and individual station beamshapes. The tentative conclusion is that LOFAR will be 'calibratable' due to the high density of bright calibrator sources in the sky.

Contents

1 Introduction

In principle, LOFAR is just another radio aperture synthesis array, for which the existing calibration methods should be relevant. However, LOFAR is different in a number of important respects:

- Crowded fields, with up to 10^7 sources per field at the lowest frequencies. This requires long baselines to avoid *source confusion.* It also causes *sidelobe confusion* if the PSF sidelobe level is too high, so we require full uv-coverage. Crowded fields greatly increase the processing volume, mainly because a large number of sources have to be predicted and subtracted.
- Pathological ionosphere, with position-dependent phase variations of up to one radian per 10 sec at the lowest frequencies. This causes position-dependent phase errors and data decorrelation, which has to be accurately measured if we want to subtract bright sources. It also causes increasing source decorrelation towards the edges of residual images.
- More than one very bright source per field. If there is only one dominating source, selfcal automatically catches decorrelation effects, and the beamshape is not very critical. With multiple sources, this is not longer the case.
- Variable beamshapes, caused by nulling, foreshortening and electronic drifts. But also: if the beams are stationary, the sky drifts through the beam at a high rate, which causes the gain on individual sources to change during integration, and
- Dynamic range. The extreme sensitivity of the next generation radio telescopes will require a dynamic range of up to 10^7 (and sometimes more) to get to the thermal noise. This means that, apart from image-plane effects like the ionosphere and variable beamshapes, all kinds 2nd and 3rd order effects will have to be taken into account when subtracting the brightest sources.

These differences require a re-assessment and modification of existing calibration strategies (and of people's intuitive ideas!). The increased sophistication will lead to (greatly) increased data processing, and in some cases we have to worry whether there is sufficient information available to reach the required accuracy.

NB: In the present stage this document contains some tentative answers, but a larger number of questions and potential problems. It is very much meant as a catalyst to discussion.

2 The (assumed) LOFAR system

The LOFAR system is not yet fully defined at the time of writing, but ssems to converge to the following: About 100 stations, in a 'scale-free' configuration like a log-spiral, with a maximum baseline of 400 km. Each station has a diameter of about 100m, and about 100 double-polarisation receptors. The latter are dipoles for LOFAR low (10-90 MHz) and 4x4 compound antennas for LOFAR high (110-200 MHz). The inner 25%

of stations are in a 'virtual core' with a diameter of about 2 km. The spatial resolution of this core is essential for calibration of 'image-plane effects' like the ionosphere and beamshapes. It can also be used for transient detection.

The station beams are stationary. Bandwidth can be traded against number of beams as: $n_{beams} \times bw_{MHz} = 32$. The instantaneous bandwidth may be spread arbitrarily over a band of 32 MHz, in units of 1 kHz. Different beams can have different bandwidths, and perhaps even different frequencies. The number of receptors per beam may be varied in order to control the beam width somewhat.

The signals from all stations are correlated with each other. This is done independently for each station beam, with a channel bandwidth of 1 kHz. The minimum readout time is 1 msec (or even smaller with special tricks like 'coherent ...', used for pulsar work). The 1 msec readout is necessary for fringe-stopping (which is done after correlation) and smooth tracking of core beams (see below). In general, the visibility samples can be integrated for about a second, limited by integration-time smearing at the longer (400 km) baselines of sources at the edge of the field.

In principle, the stations in the virtual core are identical to the 'remote' stations, so their number can be easily changed. Their signals are used specially in two different ways: Firstly, a core beam-former makes all 50.000 beams, whch are stationary, and can be read out very quickly. Secondly the correlation products of each individual station with all core stations are combined in software to produce a set of 'station-core' visibilities for each station beam. Each member of a set is pointed at (and tracking!) a different calibrator source in the station beam. The tracking is necessary because the sky moves by a substantial fraction of a core beamwidth during 10 sec. It is done smoothly by phase-shifting the msec correlator outputs before integration.

Fig 2 gives an indication of the ultimate sensitivity that can be achieved with LOFAR

3 Calibration strategy

The corner-stone of radio aperture synthesis 'calibration' is the removal (subtraction) of bright sources from the uv-data. This is done with the well-established and highly succesful self-calibration technique (selfcal) which is a closed-loop method because it uses the data residuals themselves. Only after the bright sources have been removed it is possible to study the faint and often extended sources that are usually of greatest interest. Successful removal implies that the instrumental parameters are known with high accuracy, which is why it is equivalent to calibration in the classical sense.

All LOFAR observing modes require the removal of bright sources at some level (see fig 1 for a block diagram):

1. It is assumed that any RFI stronger than the system noise has been removed from the uv-data before correlation, e.g. at station level.

2. The **Global Sky Model** (GSM) will eventually contain more than 10^8sources. Not necessarily only from/for LOFAR: it could be a separate product. Initially compiled from existing surveys, and augmented over the life-time of LOFAR. Individual sources have variable numbers of parameters, some of which may be time-dependent. See also section 7 and A.4.

3. **Initial values** (open-loop calibration) are obtained by various means for the parameters of the Measurement Equation, especially for the ionosphere (e.g. PIM/GPS) and the station beam-shapes. The required accuracy is about 10%.

4. **Station-Core selfcal** (open-loop). Resolution and sensitivity. Using simple Sky Model (usually only one dominating source) per core beam. Get rough values for the parameters of the ionosphere model, and the individual station beam-shapes. Good enough to remove 2π phase ambiguities (ionosphere) and to predict decorrelation. Need enough bright calibrators, see fig 5. Solution time constant is 1-300 sec, but the uv-data integration time is 1 sec.

5. [Optional by-product] Use a form of tomograhy to combine all available ionosphere data into a 3D movie of the ionosphere above LOFAR.]

6. **Station-Station selfcal** (closed-loop). Using the ≈ 500 brightest sources in the Sky Model. Solve for refined values for M.E. parameters, and for Sky Model parameters of the brighter sources (see section 7). Until the residual uv-data are statistically zero and noise-like.

7. Data editing. Removal of any remaining RFI by flagging residual uv-data.

8. Make **residual uv-data** by subtracting all Sky Model sources with $s > 5\sigma$ from the measured uv-data. For deep imaging this can be up to 10^6 sources per field. The prediction operation is a major processing load, even if as many corners as possible are cut. For transient detection the number of sources is about 100 but it has to be done each second.

9. Residual imaging, i.e. making images from residual uv-data. Correct the (1 sec) residual uv-data for a particular position in the sky. Make this point the field centre by rotating the uv-data. Optionally convolve the gridded uv-data (full uv-coverage!) to reduce the field-of-view, and transform to make a 'patch' image.

10. [Optional] deconvolution of residual images. This is possible because the ionosphere-affected PSF degrades 'benignly' towards the edge of the patches.

11. Identify new sources for the Global Sky Model from long-exposure images.

12. The residual images are stored in the Residual Image Database (RID) for later combination with each other and the GSM into specific deliverables.

4 Specific observing modes

For the moment, this is just a list of possibilities that should be kept in mind and addressed in detail eventually.

- Survey imaging.
- Large-scale imaging: define subsets of station dipoles as overlapping stations.
- Pointed observations.
- Snapshot observations. Instantaneous uv-coverage.
- Line observations (differential?).
- Polarization observations.
- Simultaneous high/low observations?
- Deep imaging (e.g. reionization). Extra simultaneous bandwidth from the inner stations?
- Transient detection (astronomical, atmospheric). Down to microsec.
- Pulsar observations.

5 Station-Core selfcal

The station-station visibilities emerging from the correlator are combined with phase factors to form station-core visibilities between each station and the virtual core. See also fig 1. These visibilities have a 5 times lower noise than station-station values, assuming 25 stations in the core. Since they have a 20 times greater spatial resolution, determined by the primary beam of the 2 km core, they will in general be dominated by a single bright calibrator source in their fringe-stopping centre (see fig 8). Therefore these station-core visibilities probe the ionospheric phase distribution with very high spatial resolution, which is needed to characterise it in sufficient detail.

A possible two-step strategy is the following: First, a separate selfcal solution is made for each calibrator, i.e. using all station-core visibilities that see a particular calibrator. Since these will all be affected by the instrumental response of the core in the same way, the latter will drop out of the solution (at least the phase), leaving estimated values for the instrumental response in the direction of the calibrator for each station. The second step is to use these estimated values to determine the parameters of the (2D/3D?) ionospheric model. Of particular importance is the removal of 2π ambiguities, to determine the integrated ionospheric phase as seen from each station. This is necessary for the accurate prediction of visibilities over the entire frequency band, and to deal with differential refraction effects. The procedure is to flip all station phases by multiples of 2π until the solution inconsistency is minimised (VLBI fringe-fitting...?). The second step can be done in two ways: Use a large-scale model of the ionosphere and solve for its parameters and interpolate in the areas with few sources. Alternatively assume a thin layer and do tomography (AGB). The latter is open-loop, based on a shaky assumption, and there may not be enough rays through the ionosphere in the outer parts of the array. Moreover, what about the thick blanket above?

Note that we make explicit use here of the entire 4 MHz signal bandwidth, and the known dependence of the ionospheric phase on frequency. In view of the strong frequency-dependence of Faraday rotation, we have to sub-divide the band into 10-20 chunks and solve for the Rotation Measure. Since this is only a single parameter, it does not affect the S/N.

It should be noted that the same result can probably be achieved with station-station selfcal, but at much greater processing cost, and perhaps less reliability. Not only would a more complex Sky Model be needed (predict!), but there would be many more phases to flip.

Summarising, the instrumental response of station in the direction of a calibrator contains the following elements:

- Ionospheric phase. Many radians (2π ambiguities). May vary rapidly (decorrelation).
- Electronic gain. Mostly amplitude. Changes rapidly but smoothly if the station beam is stationary (the core beam must track!). Polarisation. Crosstalk? NB: The gain solution is also a rough indication whether the station is working properly, and whether it is phased up. Sidelobes?
- Sky Model errors, e.g. an incorrect flux of the calibrator, or other sources in the field. Assumed to be small.

Fig 5 shows that, given the assumed LOFAR of section 2, there are enough bright calibrator sources available to give a visibility S/N of at least 3 in 10 sec at 4 MHz bandwidth.

6 Station-Station selfcal

This is the critical part of the calibration because it deals with the observed data itself (closed-loop). The criterion for success is that the uv-data residuals are 'statistically zero', i.e. smaller than the noise and with noise-like statistics. Since it is impractical to solve for all M.E. and GSM parameters simultaneously, it is done in successive subsets. For instance:

- Ionospheric phases per station. Continuous polynomials in time, chunks of several minutes.
- Station beamshape parameters, as a function of frequency and time.
- Parameters of the brightest Sky Model sources, most of them constant in time.
- Parameters of the fainter sources. (and the faintest?)
- Etc.

Only the ≈ 500 brightest sources are used to calculate predicted visibility values. This is usually enough to get within 0.01% of the real value, because the homogeneous background of fainter sources tend to cancel each other in the complex addition. See also fig 8. The prediction error will have less influence on the parameter solution if it varies more randomly between visibilities. NB: It should be noted that useing 500 sources for prediction does NOT mean that we have explicit measurements of the instrumental response in the direction of those sources. We still solve for the same number of parameters, using uv-data residuals.

The number of parameters that can be solved for in a single solution must be smaller than the number of independent measurements (equations). With N stations, there will be $4 * N * (N - 1)/2$ measurements (4 corr/ifr). This means that we may solve for a maximum of $N - 1$ parameters for each voltage beam (2 per station). (This argument should be elaborated for the case of solving for time and frequency polynomial coeff. Also: subsequent measurements represent independent equations if the parameter-derivatives in equ 5 have changed significantly).

In some cases, it may be necessary to use only a subset of the available baselines for selfcal solution. For instance if the modelling of extended sources is so insecure that only the long-baseline predictions can be trusted.

Redundant spacings may be used to constrain selfcal solutions in a way that is independent of the Sky Model. This is only true if the redundant interferometers 'see' the same sky, i.e. have the same beam, bandpass and polarisation. The uv-points actually coincide only in the case of East-West arrays. In all other cass they only overlap when seen from a certain direction. Given all these restrictions, it is not clear how valuable the redundancy constraint really is, especially in the case of many stations....

Selfcal works better with better S/N per data sample. Fig 7 gives the station-station sensitivity for various integration times. In order to reach the required minimum value of S/N=3 (?), it will be necessary to obtain solutions for chunks of several minutes of data, even though the data integration time is 1 sec. This is done by solving for the coefficients of smooth functions (like polynomials) in time.

A potential source of concern is the possibility of noise bias. This occurs when when selfcal solutions are obtained from uv-data with a relatively low S/N that have undergone a non-linear transformation like taking the logarithm to get linear equations in the gains or the phases. The solution is to use only the visibilities themselves, which are assumed to have gaussian noise.

7 Generating and refining the Sky Model

Shapelets..

The Global Sky Model is filled initially from the existing surveys. This yields a list of parametrised source components for the brighter sources ($> 1mJy$), with accurate positions (< 1 arcsec), and fluxes extrapolated to LOFAR frequencies to perhaps 10% accuracy. During the station-station selfcal, the parameters of bright compact sources are refined by solving for them in the same way as for instrumental parameters. In addition, new sources must be identified and modelled. Unfortunately, this must be done in the image-plane, since there is no way to do it in the uv-plane. The image should be built up from many small patches, in order to minimise the edge degradation caused by ionospheric phase and variable beamshapes.We then have to deal with two major classes of objects:

- Very extended and complex objects (like Cas A) should be modelled with 'parametrised pixons', i.e. pixons that have extra parameters to describe polarisation (I,Q,U,V) and frequency and time dependence.
- Many faint sources (usually compact).

The new sources must of course be subtracted in the uv-plane, since that is the only way in which image-plane effects can be properly applied. This procedure is similar to the minor and major cycles in Clark CLEAN. If the sources are bright, their parameters will probably have to be refined by solving for them in the uv-plane. This process of finding and rough deconvolution of sources in the degraded image-plane and subtracting them in the uv-plane should converge, because....

8 Residual imaging

The only images worth making are residual images, i.e. images of the residual uv-data after subtracting the Sky Model. The reason is that strongly position-dependent instrumental effects (image-plane effects) like the ionospheric phase and the station beamshape can only be applied while *predicting* the contribution of an individual source to the uv-data. It is impossible to subtract such a source with any kind of accuracy from an image made with the original (but corrected) uv-data.

We will assume that it will be possible to predict the contributions of all Sky Model sources with such accuracy that they will be completely removed from the residual uv-data. This includes up to 10^6 sources with $S > 5\sigma$ (say), which leaves an even greater number of fainter sources. The latter cause two kinds of problems:

- Since the uv-data can only be fully corrected for one point in the sky, the image quality (source decorrelation) of the sources away from this point will be increasingly degraded by residual instrumental errors.
- There are so many faint sources per LOFAR field that their combined PSF sidelobes can cause a significant increase of the image noise (sidelobe confusion). The effect is indicated for several values of the rms PSF sidelobe level in fig 2.

The problem of source decorrelation can be dealt with by reducing the size of an individual image (by means of uv-plane convolution) so that the total variation of the residual ionospheric phase is less than one radian. The gain variations across the patch may be minimised by tracking the sky with the station beams during the observation. Multiple 'patch' images will be needed to cover the size of a station primary beam, each with its own 'correction centre'. This patch imaging approach also addresses the problem of sidelobe confusion by reducing the number of faint sources in the field. [This number is reduced further for low-resolution observations, because source confusion only adds a constant level to the image (?)]. However, it will probably still be necessary to minimise the PSF sidelobe level. The first step is to strive for full uv-coverage, which is needed in any case for effective uv-plane convolution. However, the sidelobe level may be dominated by residual ionospheric phase errors, in which case the number of independent uv-points is the determining factor. *More study is needed into PSF sidelobe levels.*

The following statements are up for grabs: The number of patches in a station beam $\propto \lambda^3$, i.e. λ^2 for the beamsize, and an extra λ because the amplitude of the ionospheric phase variations is proportional to the wavelength. The number of calibrators needed to sample an ionospheric structure (not patch!) is estimated to be about 5-10, but a minimum of 20 calibrators is needed per per station beam to estimate the electronic (gain) shape. We might make use of the fact that the latter changes more slowly and certainly more smoothly than the ionospheric phase. But it should be remembered that the sky will move w.r.t. the stationary station beam.

This one too: It may be possible deconvolve a residual image, because the PSF degrades towards the edge of the patch in a benign anf known way (at least the degradadtion caused by ionospheric phase errors). Useful? Maximum of 100 sources (AGB argument).

9 The need for full uv-coverage

Full uv-coverage is expensive to obtain, even if we use Multi-Frequency Synthesis (MFS). The reasons to require full uv-coverage are: Firstly to reduce the rms PSF sidelobe level, and thus to avoid sidelobe confusion caused by the large number of faint ($S < 5\sigma$) sources in the field that cannot be subtracted from the uv-data. Secondly to allow uv-plane convolution to reduce the field of view of 'patch' images. However, a low PSF sidelobe level is less important in a patch image, since the number of faint sources is smaller. In addition, the PSF sidelobe level will probably not be dominated by missing uv-points, but by residual ionospheric phase errors and/or residual primary beam gain errors (especially if they do not track the sky).

Since uv-data can only be fully corrected for a single point in the sky (assuming that we know all instrumental errors), the degradation by residual errors increases towards the edge of a patch image. Thus it is better to have many small patch images, rather than a few larger ones. However, small patch images require a large uv-plane convolution function. It is conceivable (?) that such a large function covers so many uv-points that full uv-coverage is less essential.

In summary, the issue of full uv-coverage is intimately connected to the patch size. More study is urgently needed into the various contributions to PSF sidelobe levels, and into the necessary conditions for uv-plane convolution and patch image formation. This should lead to an estimation of the optimal patch size as a function of observing frequency. (E.g: since the ionospheric phase errors vary less at higher frequencies, this allows larger patches).

It is also necessary to determine the processing and storage cost of more or less patches. This should include the possibility of using uv-data that have been integrated longer or have wider bandwidth.

10 FFT beamformers and tracking beams

The two main beamformers are at station level and at core level. The sky moves so quickly, and core beams are so narrow, that they must track their calibrator sources in station-core selfcal. This is easily achieved in software when forming station-core visibilities, by combining station-station visibilities with suitable phase factors. It is not clear whether an additional core beamformer will be required, e.g. for all-sky transient detection.

A difficult question is whether the stations should have FFT beamformers, which form all beams on the sky but do not track the sky, or DFT beamformers which form a single beam at a time which may track. (NB: If it is possible to somehow rotate the beam-pattern of a FFT beamformer around the celestial North Pole, this would represent tracking without rotation w.r.t. the sky, so that sould change this discussion. For the moment we will assume that that is impractical). The problem with a non-tracking station beam is that sources will be observed with strongly variable gain (and polarisation!) during an extended observation. This can be taken into account when subtracting Sky Model sources, but will cause problems in residual images (see also section 9). A DFT beamformer will also cause off-axis gain variations, because the beam rotates w.r.t. the sky and changes shape with elevation and RFI nulling, but they will be much smaller than with a non-tracking beam.

Tracking should be smooth...! Just like nulling should not cause abrupt beam-changes. Requirement?

Should the beamshape be more or less invariant w.r.t. frequency? What are the sensitivity implications?

One way to deal with non-tracking primary beams is to observe a field simultaneously with adjacent beams. By combining the residual uv-data for these beams when making a residual image, the residual gain variations will be smaller because the adjacent beams represent a more uniform gain over the field.

The suggestion has been made that a small number (<10) of DFT beam-formers is cheaper than an FFT beamformer, in components and nr of operations, but that the cost of the latter may come down in time. Thus, one might consider to start with 2 DFT beams per station, and replace them with FFT beamformers in a few years time. It is also possible to initially equip only the core stations with FFT beamformers, e.g. for all-sky transient detection. Altervatively, one may buffer the signals from individual dipoles in the core stations, to be used for specialised beamforming whenever a transient has been detected (somehow).

11 Polarization

Influence of galactic Faraday screen (no Stokes I, but signoficant Q,U, especially on the shorter baselines). Effect of the ionosphere.

Non-orthogonal dipoles (in projection at least). Mutual coupling.

12 Transient detection

Refer to the note by Robert Braun.

Cosmic ray showers.

Piggy-back operation?

13 Processing considerations

13.1 Required features

- Global Sky Model. Parametrisation of components. Pixons? Images?
- Flexible M.E. parametrisation. Including polynomial coefficients.
- Solver that can solve for arbitrary subsets of M.E. and GSM parameters, using abitrary subsets of uv-data.
- Solver that can judge the quality of the solution and take appropriate action.
- A good ionosphere model (f,t)
- A good station beamshape model (f,t). Sidelobes!
- Efficient implementation of the predict operation. Different classes of objects, e.g. bright/faint. The 500 brightest properly with full sophistication, the rest with uv-plane convolution (?).
- Patch imager. uv-plane convolution. Combination into deliverables.

13.2 Potential bottlenecks

- Modelling of very complex sources (e.g. foreground polarization)

- Predict/subtract.
- uv-plane convolution.

14 Conclusions

- The cornerstone of LOFAR calibration is the removal of bright sources, using selfcal.
- Station-core selfcal is an effective device to solve for 2π ambiguities in the ionospheric phase.
- There are sufficient bright calibrators available for station-core selfcal, given the assumed LOFAR.
- The only image worth making is a residual image. All the sources in the Sky Model should be subtracted first.
- Patch imaging is necessary to avoid source decorrelation towards the edge of residual images, and to reduce sidelobe confusion by limiting the number of un-subtracted faint sources in the field.
- More study is needed into PSF sidelobe levels. The influence of missing uv-data, residual ionospheric errors, tracking and stationary station primary beams, etc.
- Galactic foreground polarisation is very difficult to model. It might cause considerable problems in estimating instrumental paremeters, and in subtracting it from the residual uv-data.

15 Questions

- Do we need equal-sized beams for different frequencies?
- It is not clear whether we should require full uv-coverage.
- It is not clear whether the primary beams should track the sky.
- What is the optimum patch size? And the processing tradeoffs?
- What really determines PSF sidelobe levels? And how important is it?
- Should RFI nulling be constrained to cause only smooth beamshape varaiations?
- Does a primary beamformer track the sky smoothly? Is that necessary?
- Do we want pigg-back transient detection? What does that require?
- Co-location: pros and cons.

References

[1] J.D.Bregman, *Concept design for a low frequency array,* April 2000, SPIE Proceedings Astronomical Telescopes and Instrumentation 2000, Radio Telescopes, Vol 4015

[2] K. van der Schaaf, *Updated LOFAR System Definition*, April 2001

[3] J.E.Noordam, *Guidelines for the LOFAR Station Configuration*, March 2001

[4] P. Fridman, *Genetic algorithms for tied array phase calibration*, 2000

[5] G. Fehmers, T*omography of the Ionosphere*, Sept 1996, PhD Thesis, Technical University Eindhoven

[6] J.P. Hamaker, J.D.Bregman, R.D.Sault, *Understanding Radio Polarimetry I*, 1995

[7] J.E.Noordam, *The Measurement Equation of a Generic Radio Telescope*, Febr 1996, AIPS++ implementation note nr 185, see http://aips2.nrao.edu/docs/aips++.html

[8] J.P. Hamaker, *Understanding Radio Polarimetry IV, 2000*

[9] R.Braun, *Strategies for Transient Detection with LOFAR,* March 2001, Internal LOFAR Note

[10] ...

Acknowledgements

The initial LOFAR calibration scheme was proposed by J.D. Bregman. Many people besides him and the author have contributed to polishing it during often fierce but always enjoyable discussions over coffee and in front of whiteboards. In no particular order: A.G. de Bruyn, R.Braun, C.Lonsdale, J.P.Hamaker, P.Fridman, C.M. de Vos, M.P. van Haarlem, K. van der Schaaf, A. van Ardenne, ...

A Appendix: Self-calibration (selfcal)

The success of LOFAR calibration relies on the closed-loop technique of *self-calibration* (selfcal) to remove the effects of bright sources. Given a model of the observed brightness distribution (Sky Model) and a model of the observing instrument, one may predict the values of the measured visibility data. Selfcal uses the difference between measured and predicted visibilities to solve iteratively for instrumental and Sky Model parameters. The process is finished when the differences (residuals) are statistically zero and noise-like. An image made from the residuals can be used to find new (fainter) sources for inclusion in the Sky Model. The final residual image will only show thermal noise, and sources that were too faint or too extended to be included in the Sky Model. Often these will be the sources that we are most interested in.

Selfcal will only converge to the 'correct' result if the dominant radio sources are compact and well-separated by the resolution of the array. It also helps that we can assume that most of the flux is inside the primary beam and that the radio sky is positive. The number of independent parameters to be solved for should be as small as possible, and certainly much smaller than the number of measured data samples. They should also be orthogonal to each other, even if only in terms of having different time-constants. A very significant reduction in the number of parameters is obtained if it may be assumed that all instrumental parameters are station-based rather than interferometer-based. This is so crucial that it is sometimes called the *selfcal assumption.*

A.1 The Measurement Equation (M.E.)

Until recently, the selfcal of radio telescopes has been based on a pragmatic and approximate instrumental model, in which frequency dependence and polarization were often added only as an afterthought. It was also assumed that instrumental effects were the same for all sources in the field, or that the primary beamshape is identical for all stations and does not vary in time. The new generation radio of telescopes like LOFAR and SKA are so sensitive that they will only be able to achieve their full dynamic range if they use the more rigorous instrumental model pioneered by [6] and [8]. The resulting Measurement Equation is a generic full-polarisation matrix formalism that describes how a visibility value $V_{ij}(f,t)$ is arrived at when observing a brightness distribution consisting of k sources I_k with an interferometer ij between stations i and j. If we ignore interferometer-based effects for the moment, the basic structure is:

$$V_{ij}(f,t) = \sum_k (J_{ik} \otimes J_{jk}) * S * I_k \tag{1}$$

in which the operator $\otimes$represents the Kronecker matrix product. Each Sky Model source I_k is a 4-element vector of Stokes parameters $(I,Q,U,V)_k$. The 4x4 Stokes matrix S determines the polarisation representation (linear or circular). The effects of station i are fully described by the 2x2 Jones matrix J_i:

$$J_{ik}(f,t) = \begin{pmatrix} j_{ik}^{11}(f,t) & j_{ik}^{12}(f,t) \\ j_{ik}^{21}(f,t) & j_{ik}^{22}(f,t) \end{pmatrix} \tag{2}$$

The matrix elements are complex mathematical expressions of instrumental parameters, each of which may depend on time, frequency and source position. Optionally a Jones matrix can be decomposed into products of other 2x2 matrices, each of which describes a separate instrumental effect of a station:

$$J_{ik}(f,t) = G_i * D_i * E_{ik} * P_i * F_{ik} * K_{ik} \tag{3}$$

The matrices G_i and D_i and P_i are *uv-plane effects*, i.e. they do not depend on the source distribution. They represent electronic gain, polarisation 'leakage' and receptor orientation w.r.t. the sky respectively. The matrices E_{ik} and F_{ik} and K_{ik} are *image-plane effects* because they depend in the position $(l.m)$ of source I_k. They represent the shape of the primary beam, the ionosphere and the Fourier kernel respecively.

A.2 Solving for arbitrary subsets of parameters

The structure and parametristion of the Measurement Equation is discussed in more detail in [8]. It should be stressed that the formalism appears to be essentially complete, i.e. it is valid for all existing and planned radio telescopes, even the more exotic ones. However, the expressions for the various matrix elements may be different in each case, and may also be refined as our understanding of certain effects evolves. Therefore the LOFAR calibration system should should be set up in such a way that it can solve for arbitrary sets of parameters in arbitrary mathematical expressions. The first step towards this goal is to use a generic solver that does not have any knowledge about the instrument. Another way of writing eqution 1 is:

$$V_{ij}(f,t,l,m) = V_{ij}(p_1, p_2, p_{3,}, ..., p_n) \tag{4}$$

in which the $p_k(f,t,l,m)$ are parameters of the M.E. and the Sky Model. Selfcal solves for arbitrary subsets of the $p_k$by solving sets of simultaneous equations of the form:

$$V_{ij}^{meas} - V_{ij}^{pred} = \sum_k \frac{\delta V_{ij}^{pred}}{\delta p_k} \Delta p_k \tag{5}$$

Thus, the solver only has to be supplied with the left-hand values and the gradients $\delta/\delta p_k$ to do its job. Ideally, it should be capable of detecting when the system is ill-conditioned, e.g. because the parameters in a subset are not orthogonal, and taking suitable action.

It should be noted that, if some parameters p_k vary smoothly with time, direction or frequency, this can be expressed by means of smooth functions like low-order polynials.The new parameters to be solved for are the coefficients of these functions. This approach is not oly more robust because it is less sensitive to noise, but also because it reduces the number if independent parameters. It also allow us to allow for the effect of rapid variations within an integration time when predicting values (e.g. decorrelation). Last but not least, it may offer for substantial savings in processing, for instance in the predicting of visibility values.

The noise on the input data propagates to different parameters in different ways, depending on the magnitude of the coupling coefficients (derivatives) etc. Elaborate.....

A.3 Decorrelation

Decorrelation: If the phase of a source changes during an integration time, its (contribution to the) measured visibility will have a smaller amplitude because of decorrelation. An example is the well-known phenomenon of integration-time smearing, in which sources far from the phase-centre of the image are broadened (radially) because the phase changes more for the long baselines than for the short ones. Because the effect only depends on the source position and integration time, and the phase change is uniform, it can be predicted with high accuracy for each source in the field.

If the phase changes (uniformly) by Δa rad during the integration time, the resulting visibility value will have a phase that is the mean phase, while its amplitude is multiplied by

$$g_{decorr} = \frac{1}{\Delta a}\int_{-\Delta a/2}^{\Delta a/2} cos\phi d\phi = \frac{2}{\Delta a} sin(\frac{\Delta a}{2}) = sinc(\frac{\Delta a}{2}) \tag{6}$$

For $\Delta a = 1$ rad, this amounts to $g_{decorr} = sin(0.5)/0.5 = 0.96$. This means that sources appear to vary in flux by up to 4% at a time-scale of minutes. (Position variation?) What is worse, different sources will vary differently.

LOFAR is designed against a ‘worst case’ ionospheric phase change of 1 radian in 10 sec, in a particular direction.

A.4 The Global Sky Model (GSM)

One of the main deliverables of LOFAR will be a global model of the visible sky at LOFAR frequencies. It will eventually contain parametrised representations of all sources that can be identified in LOFAR images, i.e. down to 3σ above the thermal noise. Initially, it can be filled with existing surveys (WENSS, NVSS), which provide the positions of the brightest sources, whose flux can be extrapolated to LOFAR frequencies with an accuracy of 10% or so.

The GSM plays an important role in calibration. It provides the bright calibrators for station-core selfcal, the 500-1000 bright sources needed to predict visibility values for station-station selfcal with sufficient accuracy, and the many sources that have to be subtracted before residual imaging.

Overview of source parameters: (RA,DEC), IQUV(f,t), extent, spectral index, rotation measure. The majority of sources will have a minimum number of parameters (e.g. non-variable, unpolarised point sources with an average spectral index), but especially the bright ones will have to be modelled more accurately. Some examples of potentially troublesome sources are:

- Galactic foreground polarisation. Galactic cosmic-ray electrons spiralling around the the galactic magnetic field-lines cause a very smooth diffuse large-scale background, which is highly linearly polarised. The total intensity (I) is too smooth to be seen by an interferometer, but the linear polarisation (Q,U) is much more structured because of Faraday rotation by the inhomogeneous foreground ISM. **This is a potential show-stopper,** because it is almost impossible to model with any accuracy. Our hope must be that it is relatively weak at LOFAR frequencies, and that it is virtually invisible to the longer baselines.
- Supernova remnants. These are very extended, and have a complicated frequency structure.

A.5 The ionosphere model

Fig 3 gives a schematic model of the ionosphere. For LOFAR calibration, the most problematic part is the layer of Travelling Ionospheric Disturbances (TID), which cause phase variations that vary rapidly over time and position. The orders of magnitude have been taken from ionospheric studies that have been carried out over the last several decades.

Scattering and defocussing....

Decorrelation...

The ionospheric model

...Closed-loop system, i.e. on the sky itself (3D tomography only delivarable to the ionospheric community).

Nr of calibrators per station beam is determined by the TID sampling density (not the patch). Make sure that 2π ambiguities are solved correctly (by solving for continuous ionospheric structure). What is the accuracy (mean,rms) of the measurement, and is that enough? Integration times longer than 1 sec?

A.6 The station beamshape model

For a smooth model of the beamshape, at least 20 calibrators are needed (S/N?). Not possible to link simultaneous beam solutions?

In general, the available number of calibrators is higher, needed for ionosphere, see above. If enough, solve for individual receptor complex gains: this is the only (?) way to link the solutions for simultaneous station beams in different directions. 20 calibrators times 8 beams gives 100 data-points, which is sufficient to solve for 100 receptor gains (real, not complex!).

In priciple one might solve for the complex gains of the 100 individual receptors per station. There are often enough (> 100) calibrators available, especially if more than one station beam is used to solve for the same receptors. This is also a way to link the solutions for different station beams. However, in practice this is complicated by the fact that it is not easy to separate the ionospheric and beamshape effects, and we also do not know the individual receptor beamshapes with sufficient accuracy. Is is more effective to solve for some complex gain model across the station beam with the smallest possible number of parameters. (main lobe, sidelobes?).

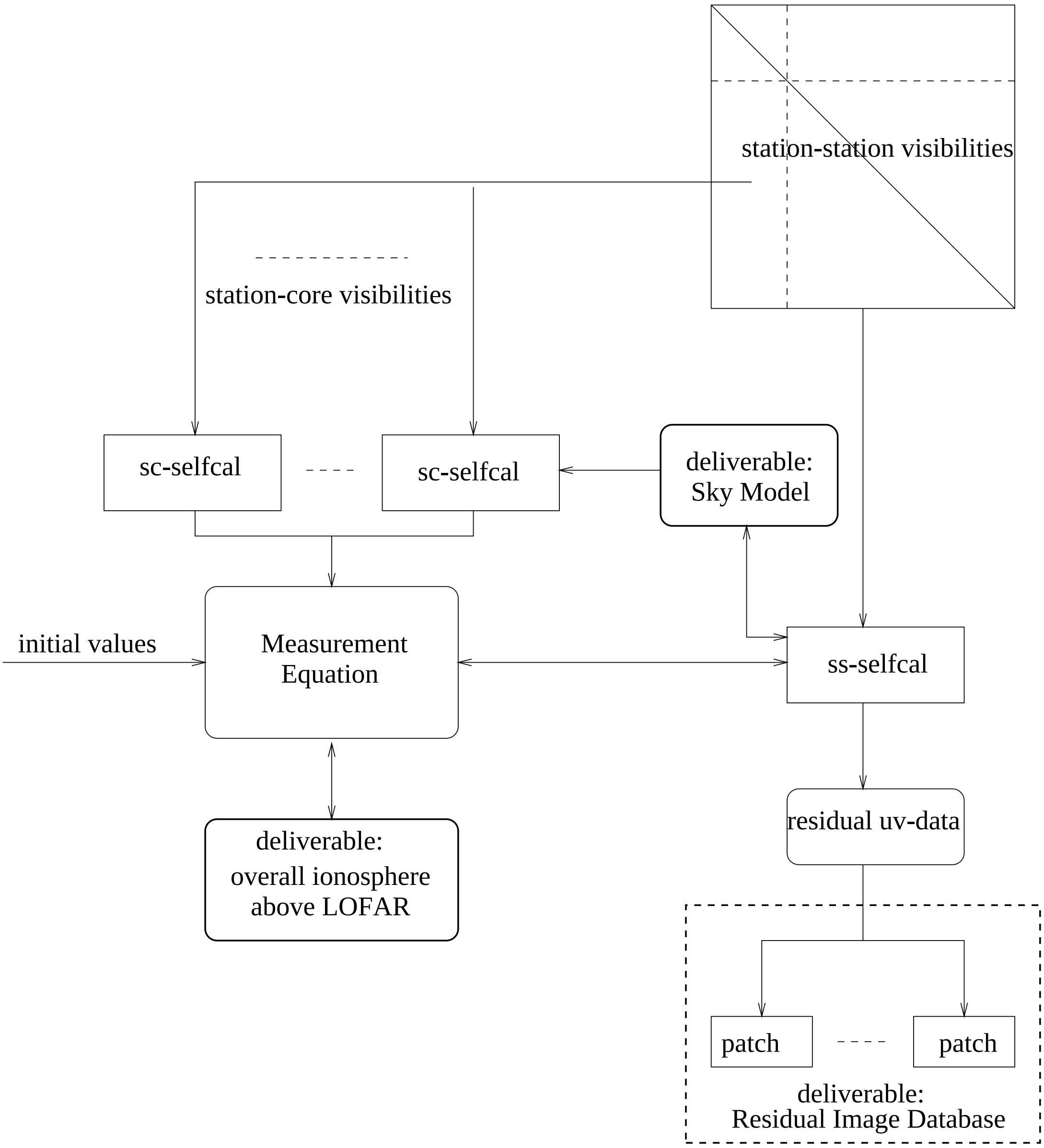

Figure 1: *Schematic overview of the LOFAR calibration strategy, as explained in section 3. Specific observing modes use variations on this theme.*

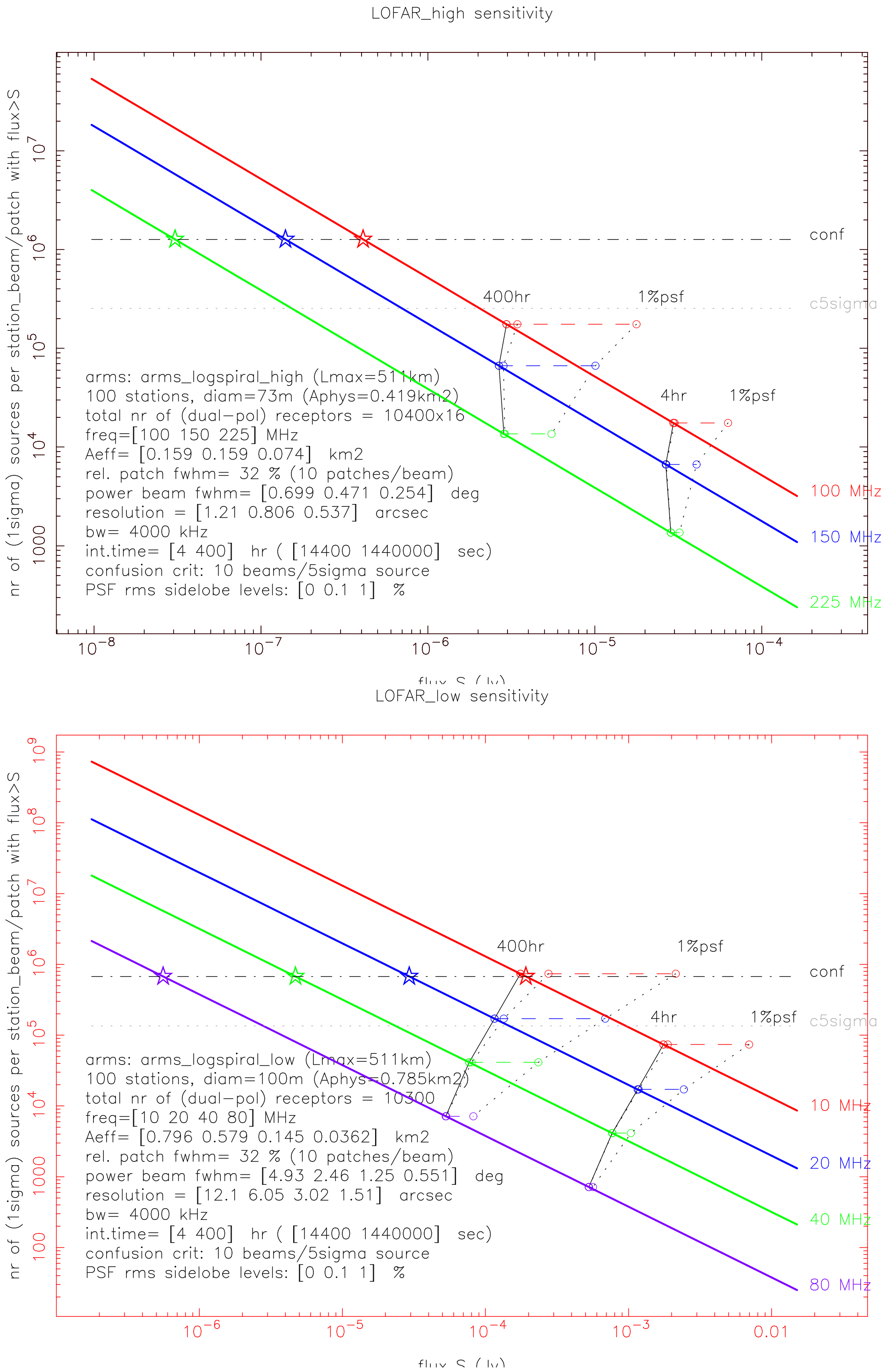

Figure 2: *Sensitivity of the assumed LOFAR (see section 2) at various observing frequencies, at various total observing times. The logN-logS plot emphasizes the fact that LOFAR fields are very crowded. Note that the number of sources relate to a 'patch' field, which is smaller than the station primary beam. Sidelobe confusion will be greater than the thermal noise if the rms PSF sidelobe-level is too large or if the field is too large.*

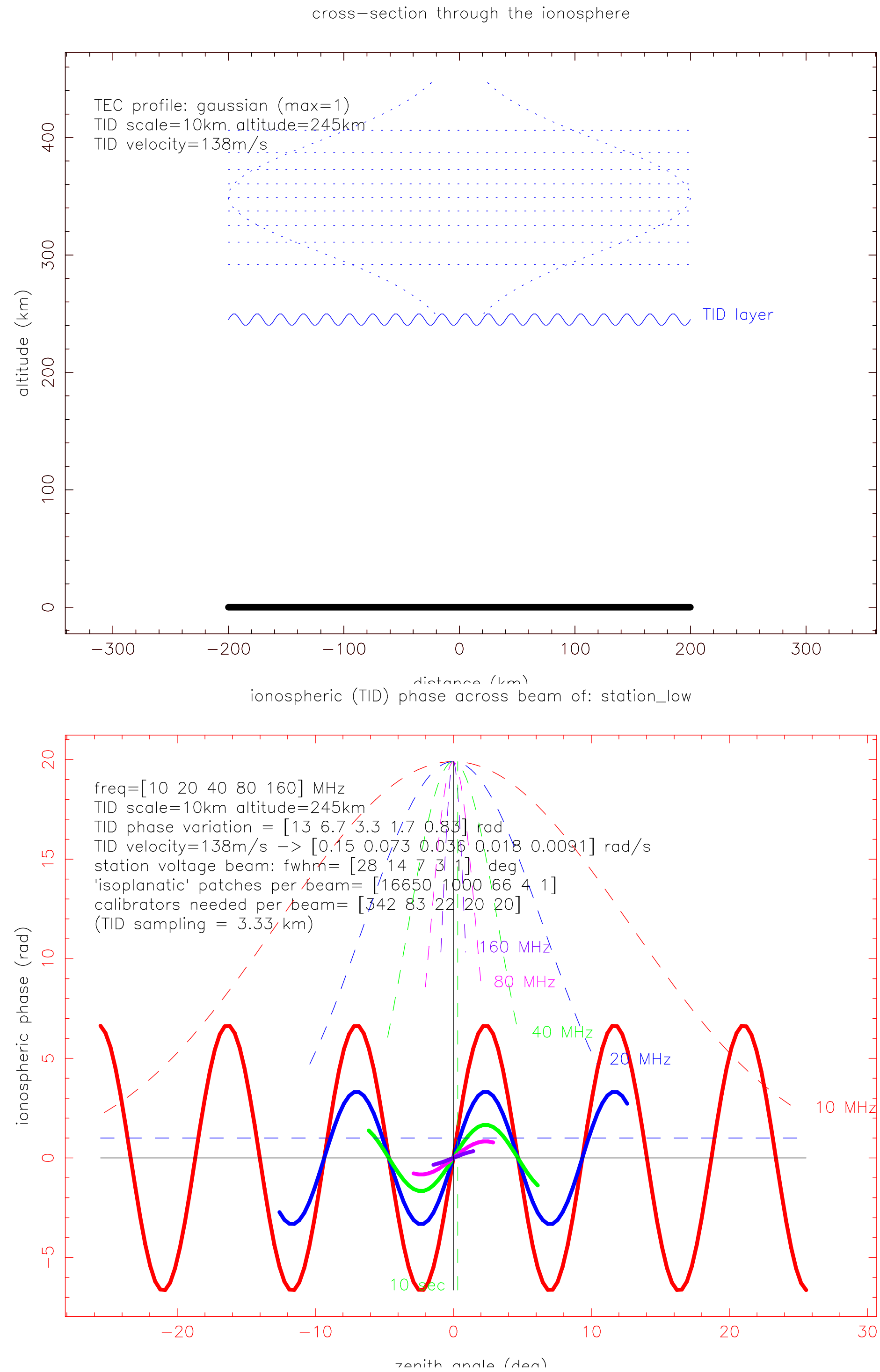

Figure 3: *Schematic model of worst-case ionospheric behaviour, as measured by various authors over the last 2-3 decades. The moving TID layer causes the strongest phase variations in space and time, as indicated by the (highly idealised) sine-wave. The bottom panel shows the phase variation over a station beam at various observing frequencies. It is 'demonstrated' that the ionospheric phase in a particular direction will not change b[illegible]dian in 10 sec (vertical dashed line), which is required for LOFAR calibratability. Even then, [illegible]ting decorrelation will have to be taken into account in the selfcal predict.*

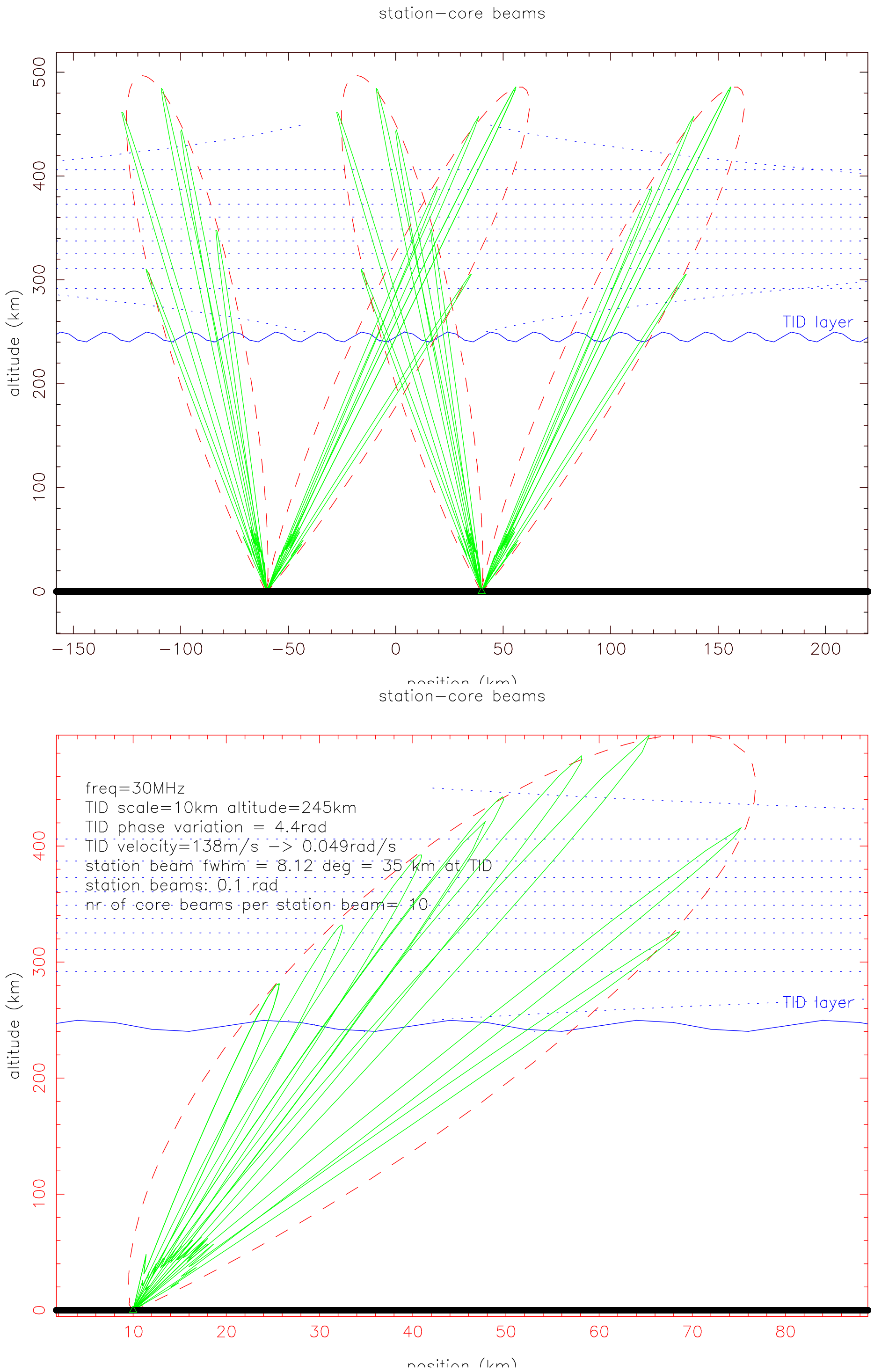

Figure 4: *Station-core selfcal is used to obtain initial values for station beamshapes andionospheric phase distribution, especially resolving any 2π ambiguities. Each narrow core beam is dominated by a single bright calibrator source, with known flux. About 20 core beams are needed per station beam to estimate their beamshapes as a function of time. More core beams are needed at the lower frequencies to estimate the ionosphere. See*

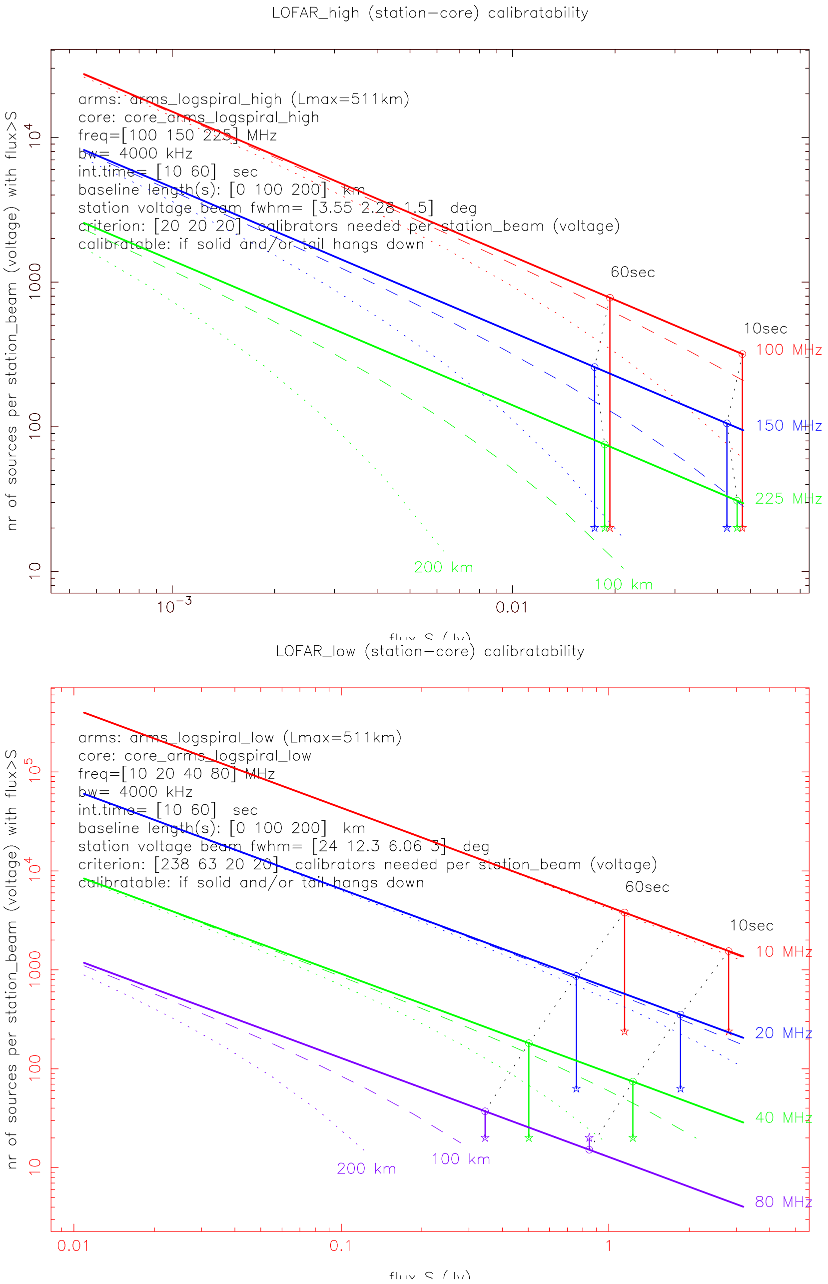

Figure 5: *The dots on the logN-logS lines indicate the number of available sources per station beam that are bright enough to give S/N>3 in 10 sec with a bandwidth of 4 MHz. Station-core calibratability (i.e. using combinations of stations with the virtual core) requires that there are at least such 20 sources per station voltage beam, as indicated by the stars. At the lowest frequencies, the required number of calibrator sources is determined by the ionospheric 'seeing cell' size. Unfortunately, many of the brighter sources are extended, which causes their visibility to decrease for long baselines. The dashed and dotted lines indicate this effect for baselines of 100 and 200 km. The conclusion is that it will be difficult to calibrate the stations far from the core at 80 and 225 MHz.*

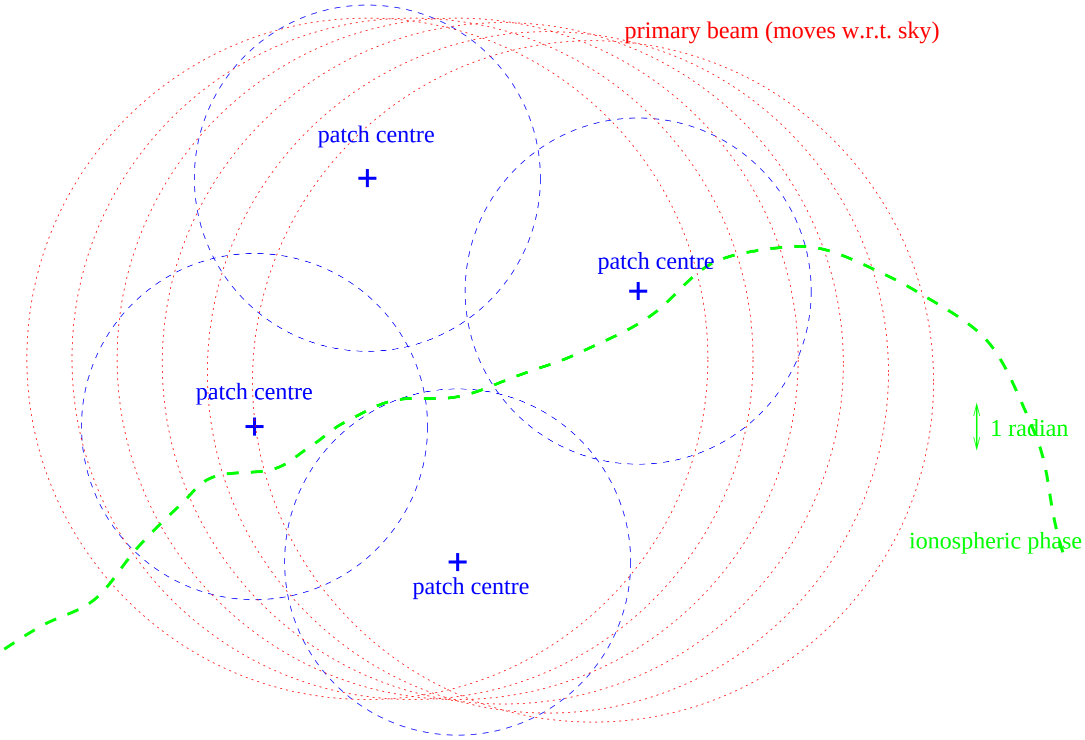

Figure 6: *Patch imaging. After subtracting the bright sources, images are made of the residual uv-data. Since uv-data can only be fully corrected for a single point in the sky, the image quality degrades towards the edge because of increasing residual ionospheric phase errors. Therefore, the image size must be reduced (by convolution in the uv-plane) to 'patches' over which the residual phase errors are less than one radian. The size of a station primary beam is indicated for comparison. Note that, contrary to a patch, the station beam may move w.r.t. the sky.*

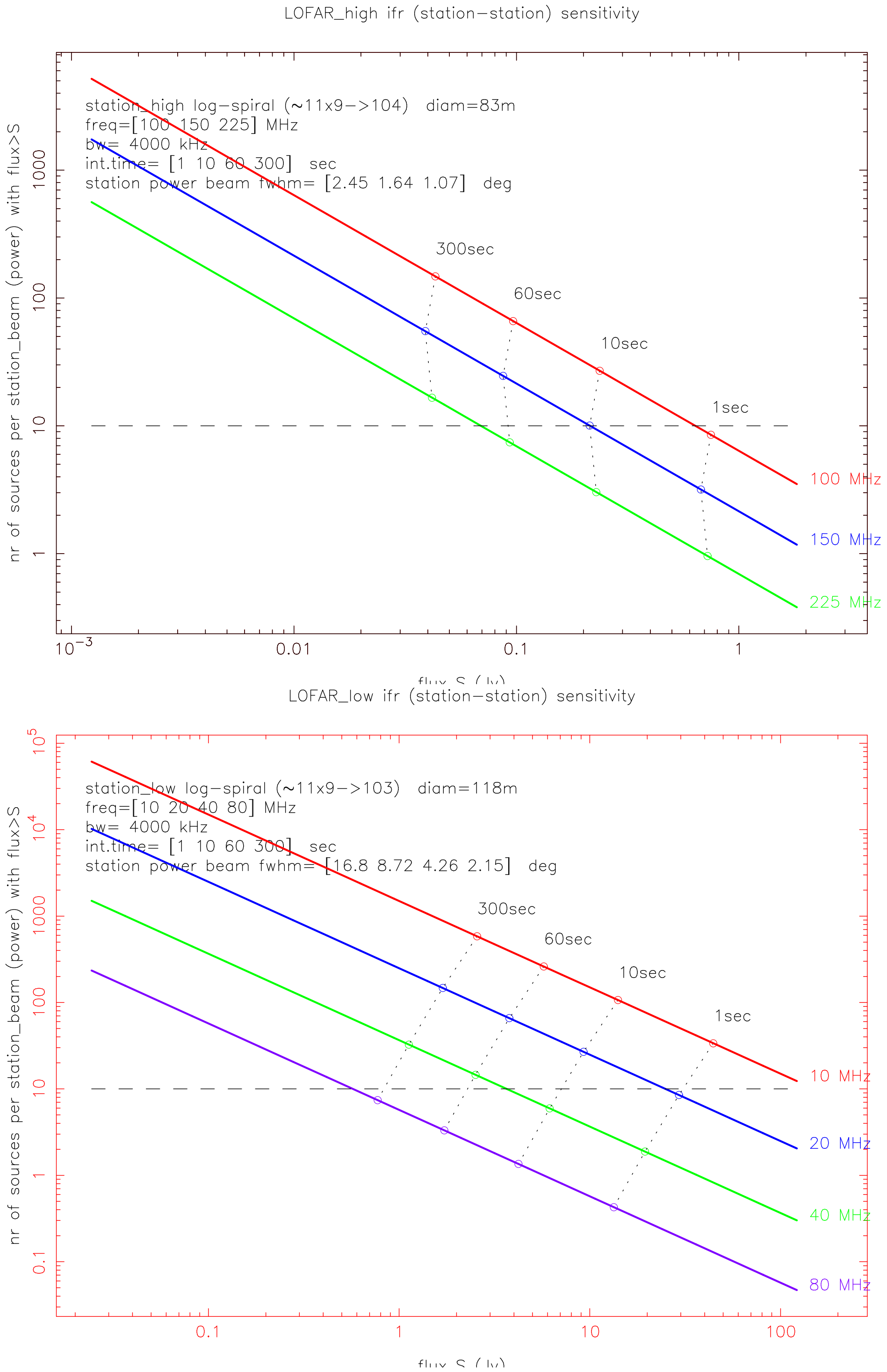

Figure 7: *Interferometer (station-station) sensitivity. The dots indicate the number of sources above the noise for various frequencies and integration times. The dots should lie above the dashed line (10 sources) which should give the S/N>3 needed for station-station selfcal.*

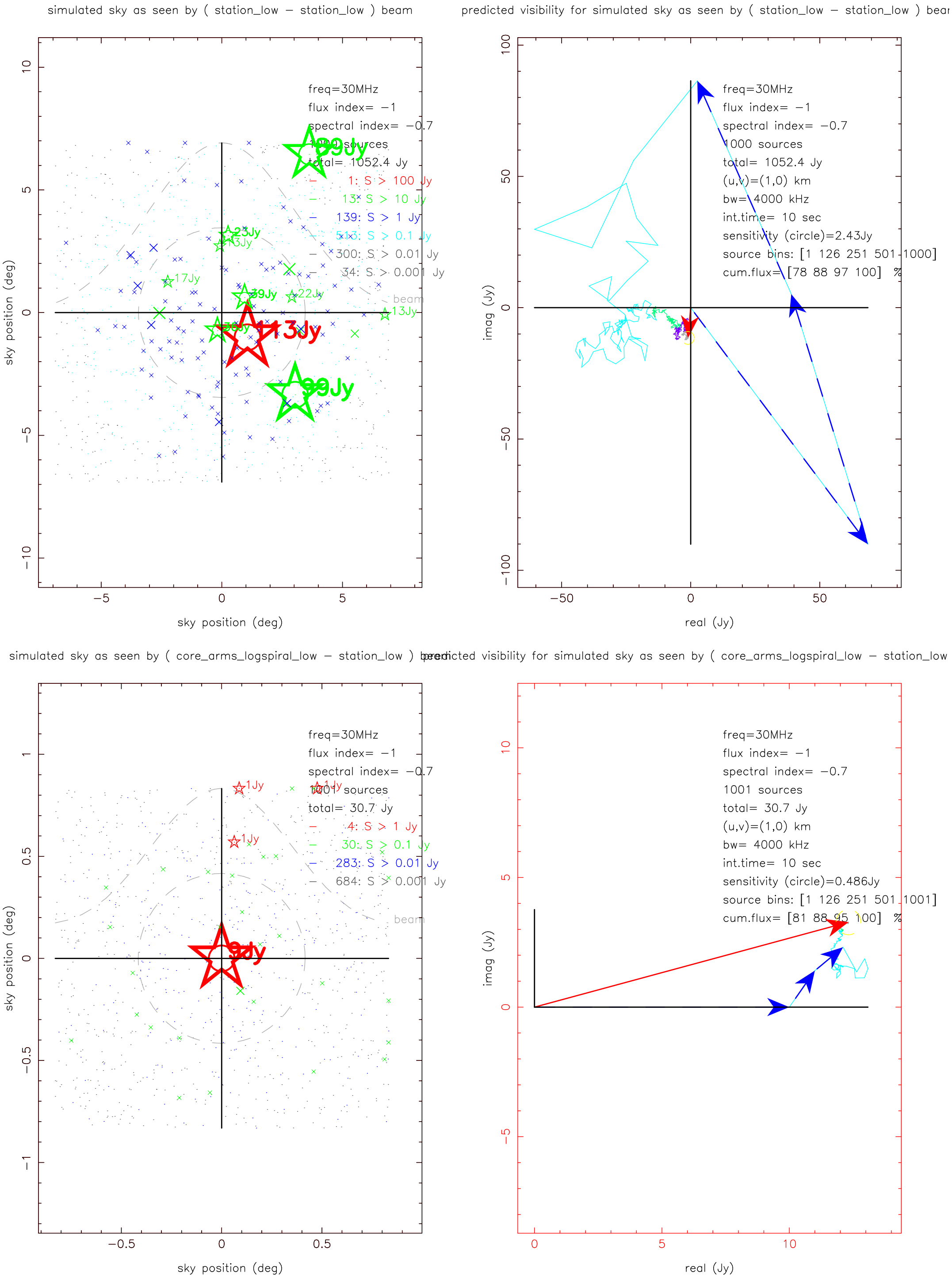

Figure 8: *Some simulated source distributions and their predicted visibilities. At the top is a field as seen by a regular station-station interferometer. At the bottom is a station-core interferometer, pointed at a dominating calibrator source. Note that in both cases the source distribution has a significant effect on the actual visibility values.*

measeq.eps not found!

Figure 9: *The Measurement Equation describes the relationship between a Sky Model and the measured visibility data. The socalled 'selfcal assumption' is that most instrumental effects are 'antenna-based', i.e. they can be traced back to individual antennas (stations). The 'interferometer-based' effects are assumed to be negligeable. Optionally, the 2x2 Jones matrix of an individual antenna may be written as a product of 2x2 matrices that each describe a particular instrumental effect. In any case, the various matrix elements are complex expressions of parameters, which must be solved by selfcal. Image-plane effects like beamshape and ionosphere cannot be applied to the uv-data, but only while predicting source contributions.*

`polyparm.eps not found!`

Figure 10: *Wherever possible, the parameters of the Measurement Equation should be expressed as smooth functions (e.g. polynomials) in time, frequency or spatial coordinates. This results in a considerable reduction in the number of parameters that has to be solved for. It also reduces the processing, since the selfcal predict does not have to be done for individual frequency channels, but only for as many frequency-bins as there are polynomial coefficients (usually less than 10).*