Measuring Station Beamshapes, as a function of time and frequency, and in full polarization

J.E. Noordam (*noordam@astron.nl*)

March 3, 2011

Abstract

Accurate knowledge of the station beamshapes (EJones) is crucial for the new (and existing) radio telescopes, if they are to achieve their advertised performance. This document sketches the problem in an accessible manner, and outlines a strategy to address it. The good old WSRT could play an important role in this.

1 Executive Summary

1. The next generation of radio telescopes will be (much) more sensitive and will put a (much) greater emphasis on accurate wide-field, wide-band, full-polarization imaging.

2. Therefore, 3rd generation calibration (3GC) will require a (much) more accurate knowledge of individual station beamshapes. The latter are represented by 2x2 EJones matrices. Fig 4 gives a schematic illustration of the problem.

3. It is highly unlikely that the required accuracy can be achieved by *open-loop* methods, i.e. by using a theoretical model, or an empirical beamshape that is obtained by scanning across a bright source. Fig 5 gives an estimate of the required accuracy.

4. Thus, it will be necessary to use *closed-loop* methods to measure the individual station beamshapes during the observation, as a function of time and frequency.

   (a) This is a generalization of the successful *selfcal* method, which solves for two complex gains per station, using a model of the observed source(s). These values are strictly valid only in the direction of the dominating source, but if all station beamshapes are identical they may be used for the entire field. Selfcal was introduced in 1980, and caused a spectacular improvement of several orders of magnitude (in a way, it saved the VLA).

(b) The fact that *generalized selfcal* solves for more parameters raises two fundamental issues. The first is whether there is sufficient information available, e.g. in the form of calibration beacons in the field (see fig 8). The second is whether there are sufficient equations (i.e. independent measurements) to solve for the number of parameters. The latter will not be discussed here.

(c) A potential practical problem is the considerable extra processing that will be necessary. This sets the stage for a trade-off between what is possible and what is feasible.

5. The Differential Gains (DG) method, developed by Oleg Smirnov, is such a closed-loop technique. It has been successfully demonstrated on a 21cm WSRT observation of the 3c147 field. It achieved a dynamic range of 3 million, thereby surpassing the earlier record set by NEWSTAR. For details, see fig 9.

6. ***Most importantly***, this "demonstrated" that, in a typical field, and for wavelengths longer than about 20 cm, there is enough information available in the form of calibration beacons, i.e. sources with an apparent brightness greater than a few mJy.

   (a) This fundamental calibration issue was verified for LOFAR in the very beginning. Now, the condition also seems to be met for shorter wavelengths, at least down to 20 cm. The situation is more favourable for longer wavelengths because the field is larger and the sources are brighter.

   (b) To first order, the finding is independent of station diameter D, since both the sensitivity and the field size depend on D^2, and thus cancel each other out. See fig 8.

   (c) *NB: This finding should be a great relief to the designers of SKA*, who must have been worrying about this issue. Therefore, it is a little puzzling that this was not instantly recognized.

7. The DG method should now be exploited to measure the beamshapes of all existing radio telescopes, in full polarization, and as a function of time and frequency. After half a century of poorly verified assumptions, such an exercise should be a real eye-opener, especially beyond the half-power point. It will be the subject of a specialized RadioNet workshop in Portugal later this year.

8. The DG method should also be used to measure the shape and stability of the beams of the Embrace and Apertif prototype stations, which are very important for ASTRON.

   (a) For ***Embrace***, the results would address the prime AAVP goal of demonstrating the viability of aperture arrays as a SKA station technology. Provided the Embrace prototype is modestly enlarged (to perhaps $200m^2$), and ideally made dual-pol, it is quite unnecessary to build a multi-station demonstrator in Portugal.

(b) For ***Apertif***, the results should guide the development of new calibration software while the full system is being installed on the WSRT in 2012. The currently planned tests involving one or two adjacent WSRT dishes are useful, but not sufficient. I

9. The *easiest, quickest and cheapest* way to do this is by embedding these prototypes in the WSRT as extra stations. Since the necessary online and offline software already exists, "only" a hardware interface is needed, possibly based on the old TADU backend.

   (a) The full WSRT is needed, partly for spatial resolution (to distinguish the calibration beacons), and partly because the WSRT system is highly stable and well-understood[1] .

   (b) If more than one interface can be built, multiple Embrace or Apertif stations may be included, up to a maximum of 16 stations. For instance, each newly installed Apertif frontend could be fully tested as part of a WSRT that has an increasing fraction of such frontends. This approach would be superior to, and more versatile than the planned standalone test system of 6 Apertif dishes.

10. However, ***time is of the essence***. For both prototypes, answers are really needed before the end of this year (2011). Also, the WSRT might not be available during much of 2012. Therefore, the implementation of the hardware interface with the WSRT should start as soon as possible, so that it is finished by the summer. In the meantime, specialists in the Embrace and Apertif teams can be trained in using the DG tool.

[1] Provided we can solve the recent WSRT closure problems, of course

2 Introduction

For the moment, this document consists mainly of pictures (albeit with elaborate captions, which tell much of the story), and an executive summary. The latter is meant to make the case for creating a link between Embrace (and Apertif) and the WSRT system. In a later stage, more detail will be added in the various sections.

Some nomenclature: Our little community has the confusing habit of using the word *beam* to indicate either the Point Spread Function (synthesized beam), the station (voltage or power) response beam, or the element response beam of an array element (e.g. a dipole). Throughout this document, we will use the word *PSF* to indicate the synthesized beam, and consistently prefix the word *beam* with qualifiers like *station*, *power* or *voltage.* We also use the generic word *station* not only for an aperture array (e.g. a LOFAR station) but also to indicate a parabolic dish (rather than *telescope* (WSRT) or *antenna* (VLA)*).*

3 Subtracting bright compact sources

One definition[2] of calibration is *the capability to subtract bright compact foreground sources with high accuracy.* After all, this is possible only if the parameters of the Measurement Equation, which includes both the instrument and the sky model, are known with considerable precision.

3.1 Subtraction of Cat I sources

3.2 Subtraction of Cat II sources

3.3 Frequency calibration

3.4 Intrumental polarization

4 Imaging of Cat III sources

5 Required EJones accuracy

See the caption of fig 5.

6 The Differential Gains (DG) method

See the caption of fig 9.

[2]This is the definition used in the LOFAR CDR calibration plan of 2006

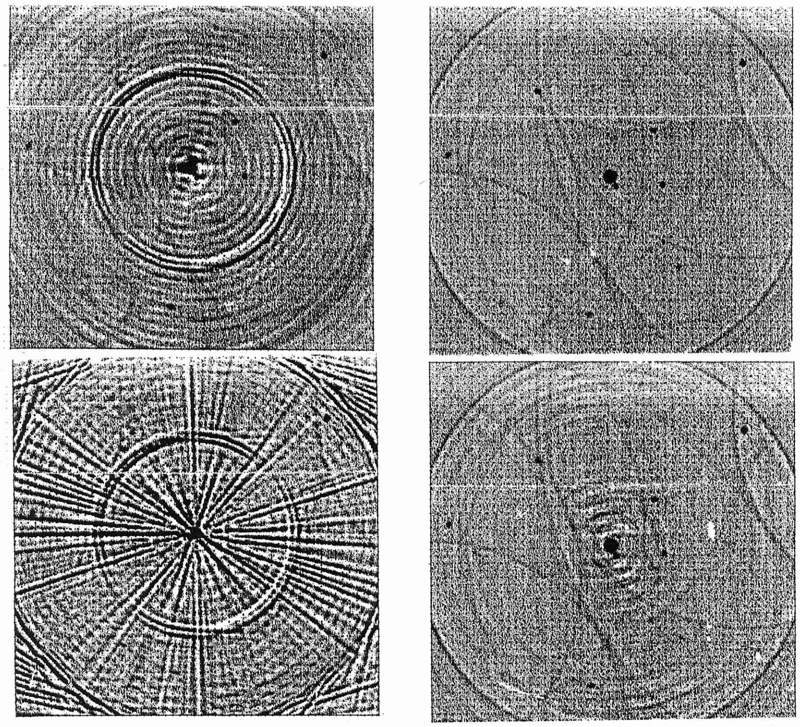

Figure 1: *The calibration quality is measured by the accuracy with which bright "foreground" sources are subtracted. The most accurate method to do this is by subtracting the predicted contribution of these sources from the uv-data, before making a residual image. The source contribution is calculated with the help of the Measurement Equation (ME), which includes the instrumental model (Jones matrices) and the Local Sky Model (LSM). Any errors in either the (intrinsic) source parameters or the instrumental parameters will cause the source to be subtracted incompletely, but in different ways. First of all, because the observed sources are the same for all stations, and constant in time, any errors in the LSM parameters will cause a source remnant in the residual image that is convolved with the* ***nominal PSF****, i.e. the PSF that is caused by deterministic imaging effects like uv-coverage, weighting and smearing. However, if the problem is caused by instrumental errors that change from station to station, and with frequency and time, the source remnant will be convolved with* ***a highly distorted PSF****, in which the sidelobes are often as high as the central peak. This is illustrated by the image above, which represents one of the first results of "redundant-spacing calibration", back in 1980. The bottom-left panel shows the remainder after subtracting the bright source 3c48 without calibration. Its distorted PSF is dominated by sidelobes caused by instumental errors (mostly troposphere). The top-right panel shows the well-behaved nominal WSRT PSF around sources that were either not subtracted at all, or subtracted with the wrong flux. Summarizing, the two classes of ME parameter errors may be distinguished from each other by means of the PSF, and tackled with different methods. NB: There are situations where the two classes of ME parameters are degenerate, for instance when the station beams do not rotate on the sky, or when they are circularly symmetric around the pointing centre. In that case, the average beamshape may be absorbed into the LSM. But this will not be the case for any deviations from this average beamshape, per station or in frequency or time.*

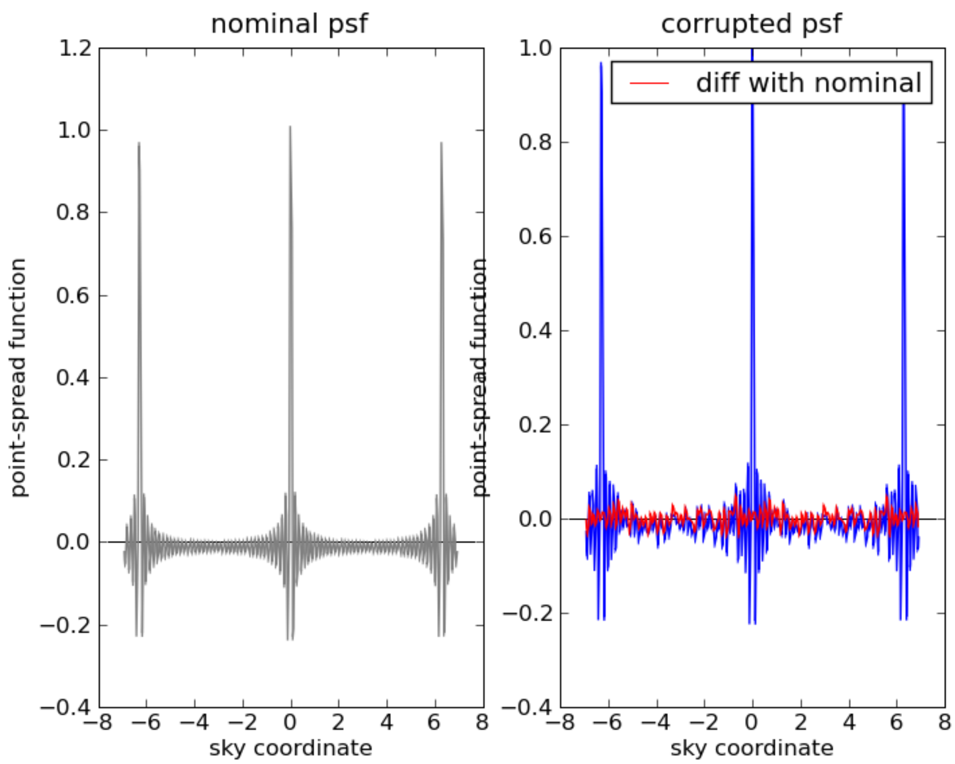

Figure 2: *Source subtraction*

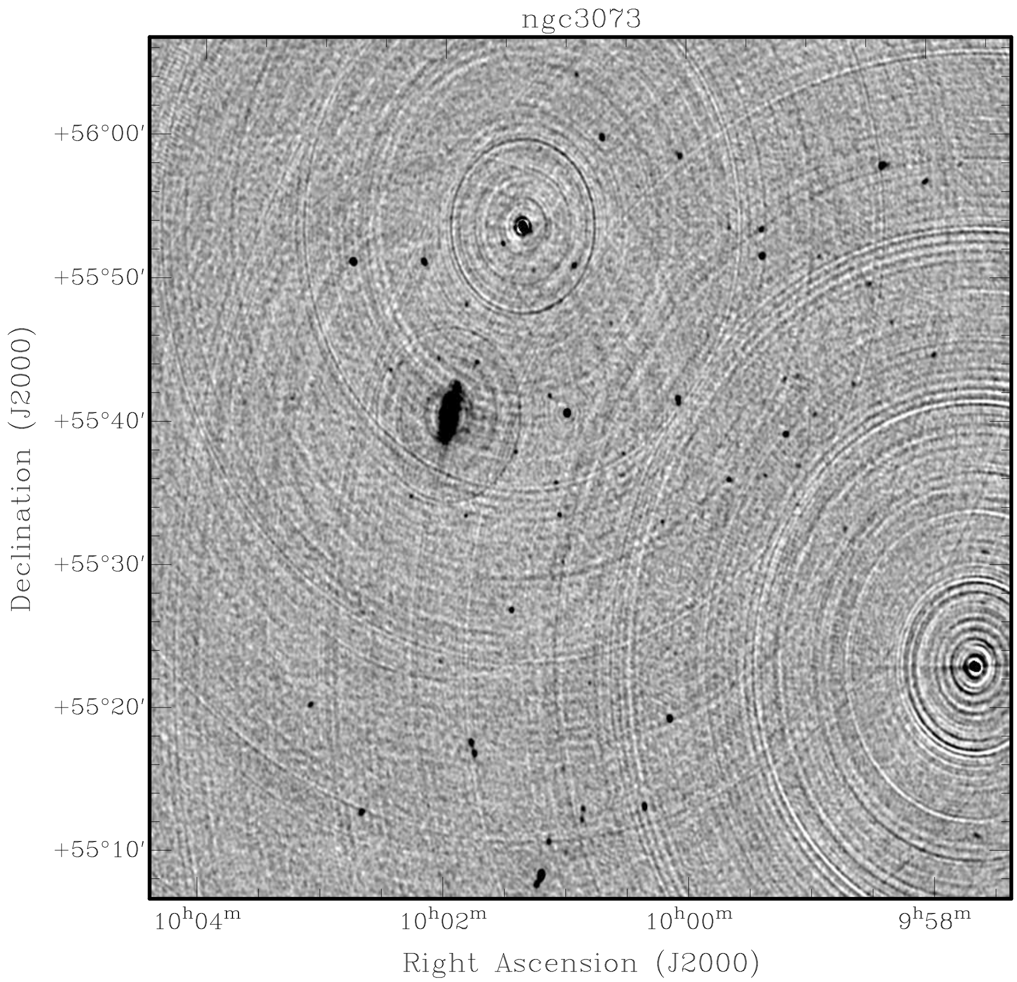

Figure 3: *A typical example of a WSRT field with off-axis dynamic range problems (Courtesy of Tom Oosterloo). Although the dominating source has been cleanly subtracted, this has not been possible to a few bright off-axis sources. This is caused partly by errors in the source model, and partly by unmodelled beamshape differences.*

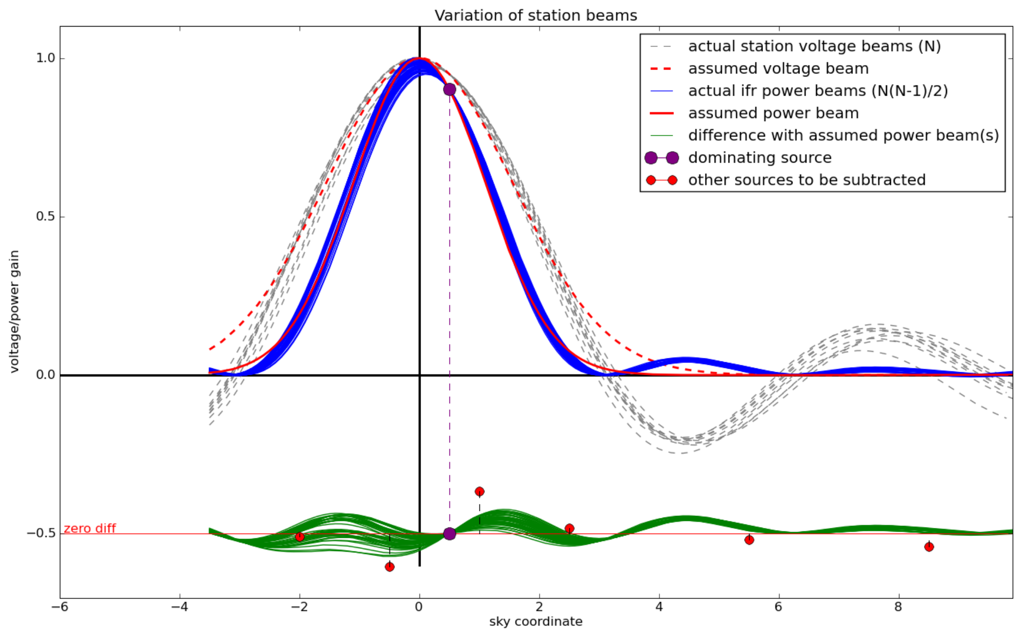

Figure 4: *In 2nd Generation Calibration (2GC), it is usually assumed that all station beamshapes are identical, for all frequencies and all times. It is applied to the data by dividing the final image by some idealized beamshape, which is obtained by means of an **open-loop** method like mathematical modelling, or scanning across a bright source. In practice, the beamshapes will vary from station to station, in frequency and time. This is schematically indicated here. Note that, through self-calibration, the station response will usually be normalised in the direction (l, m) of the dominating source in the field. Therefore, the latter may be subtracted with high accuracy. But other sources in the field will not be subtracted completely, through a combination of beamshape errors and inaccurate fluxes in their source models. This performance will no longer be sufficient for the new generation of radio telescopes, with their strong emphasis on wide-field high-accuracy observations. Third Generation Calibration (3GC) will require accurate knowledge of individual station beams, as a function of frequency and time. This implies the use of **closed-loop** methods for measuring the beamshapes continuously during the observation. Such a method is described in fig 9. Fortunately, there seems to be sufficient information available, in the form of bright calibration beacons in the field. Unfortunately, the extra cost in processing will be substantial, setting the scene for an agonizing trade-off. But at least there will be something to trade off against*

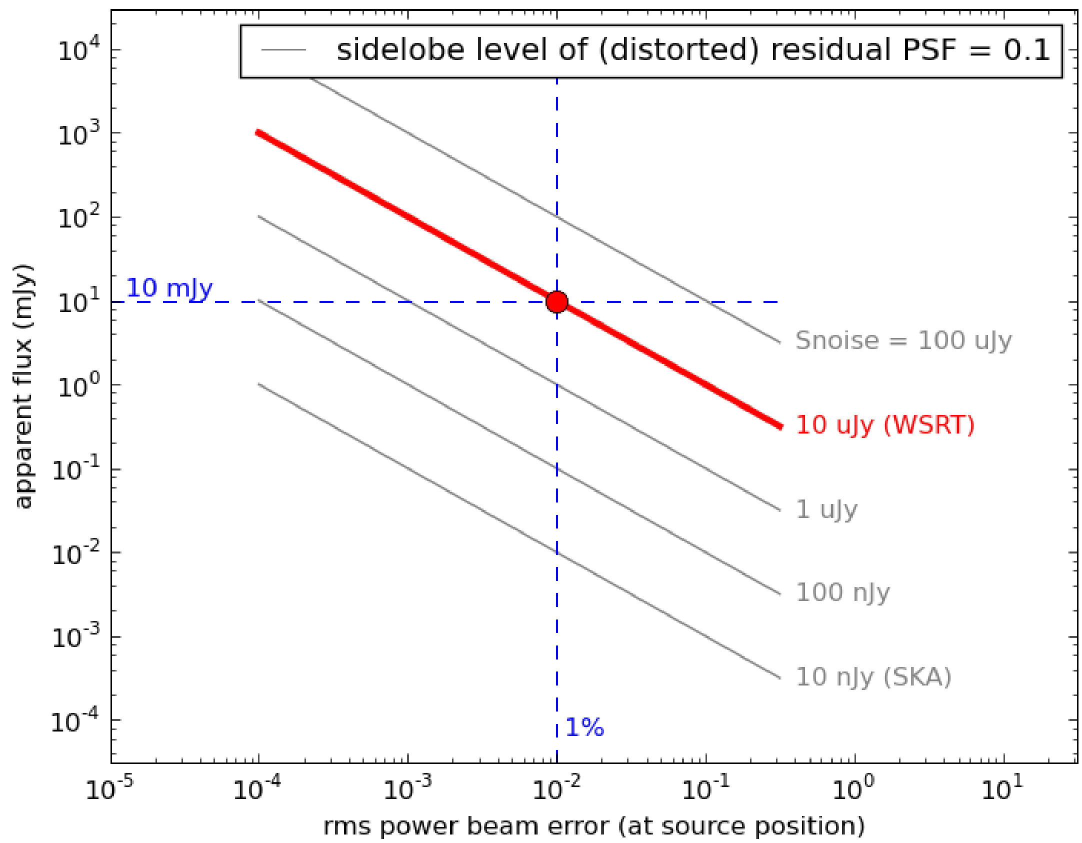

Figure 5: *An incompletely subtracted source will cause dynamic range (DR) problems if the sidelobe level of its distorted PSF (assumed to be 0.1 here, but it is probably closer to 0.5 or even higher) exceeds the thermal noise S_{noise} of the residual image. The plot indicates the apparent flux above which this is the case, given S_{noise} and the rms error in the station power beams, at the source position (l, m). The latter is the rms error over time and frequency, and over all interferometers. This plot may be used to calculate a* ***design criterion*** *for the required accuracy of the station beamshapes. For the case of the WSRT ($S_{noise} = 10\mu Jy$), sources with an apparent brightness greater than about $10mJy$ will cause cause DR problems. The plot then indicates that the rms accuracy of the passive ifr power beams should be about 1% at the positions of those sources. The passive errors in the LOFAR or SKA station beams will very likely be larger, while the DR requirement will be more stringent because their S_{noise} is smaller. This means that* ***closed-loop*** *beamshape measurements like the DG method (see fig 9) will be necessary, using the bright calibration beacons in the field. The good news is that, since the Cat I calibration beacons themselves will be substracted with the help of their own instrumental gains, the estimated beamshapes only have to be accurate enough to subtract the fainter Cat II sources (see fig 8 and 5). This represents a considerable relaxation of the requirements.*

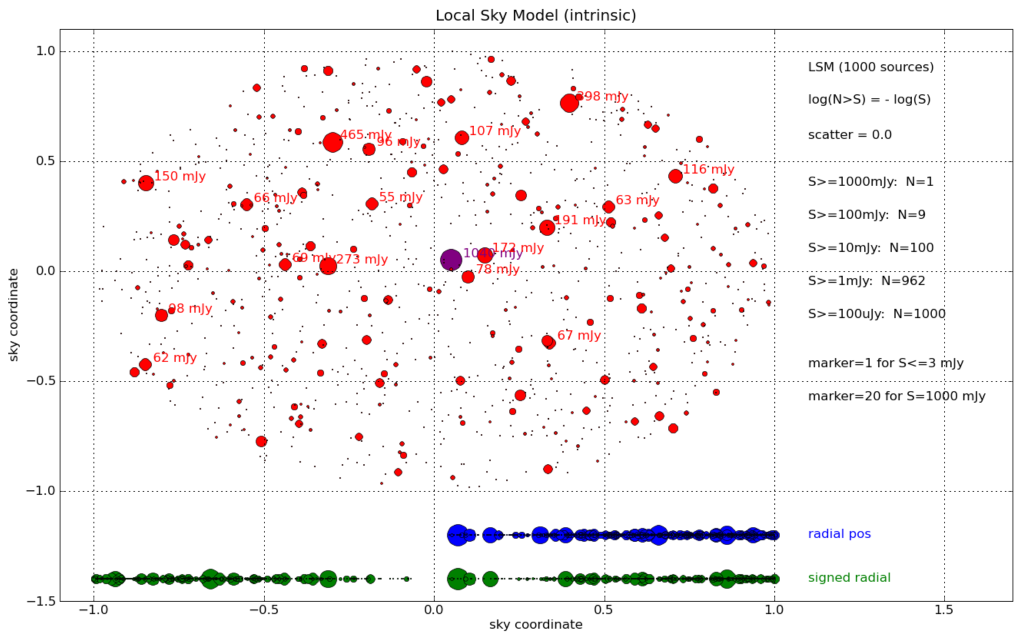

Figure 6: *A realistic distribution of (intrinsic) source brightnesses in a typical field follows the well-known expression log*$(N \geq S) = -log(S)$*, where* $(N \geq S)$ *is the number* N *of sources with flux greater or equal than* S*. This means that there are* α *times as many sources with fluxes* $\leq S/\alpha$ *than there are sources with flux* $\leq S$*. The fluxes are represented by logarithmic marker sizes, where sources with* $S \leq 0.1mJy$ *have a marker of size=1 (not all of them are shown). Obviously, the contribution of a source to a visiblity sample will be determined by the apparent flux, which will be attenuated by a power beam. The latter is the product of the voltage beams of two stations. Along the bottom of the plot, the same source fluxes are plotted against their radial position, i.e. their distance to the pointing centre. The green dots are "signed" radial positions, in the sense that they are negative for sources with a negative horizontal sky coordinate. The latter representation is used in figures 8 and 12 to indicate the total number of calibration beacons that are available in all directions to sample the 2D station beamshapes.*

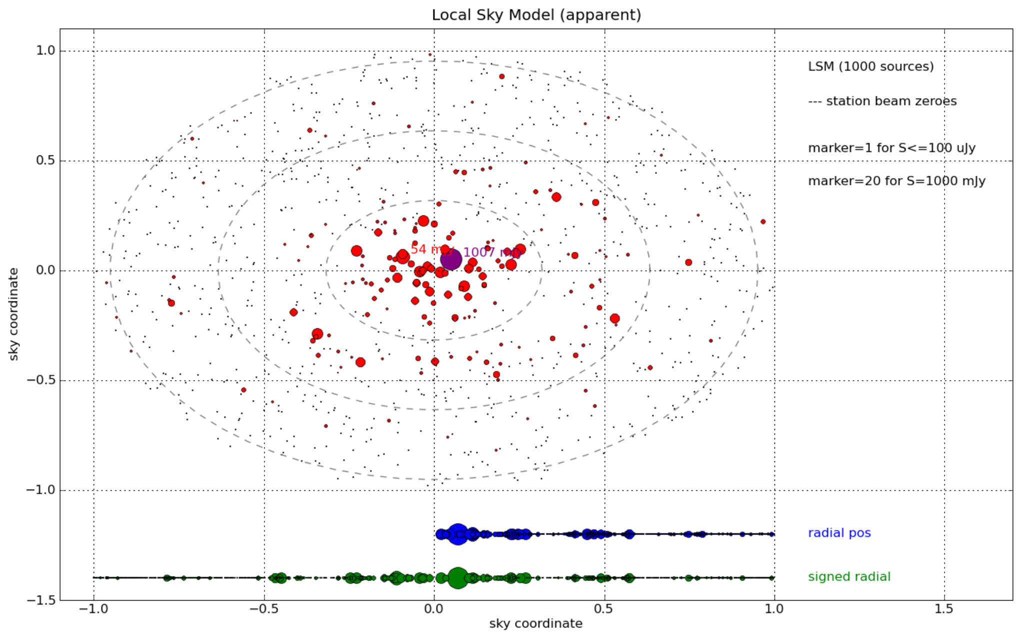

Figure 7: xxxx

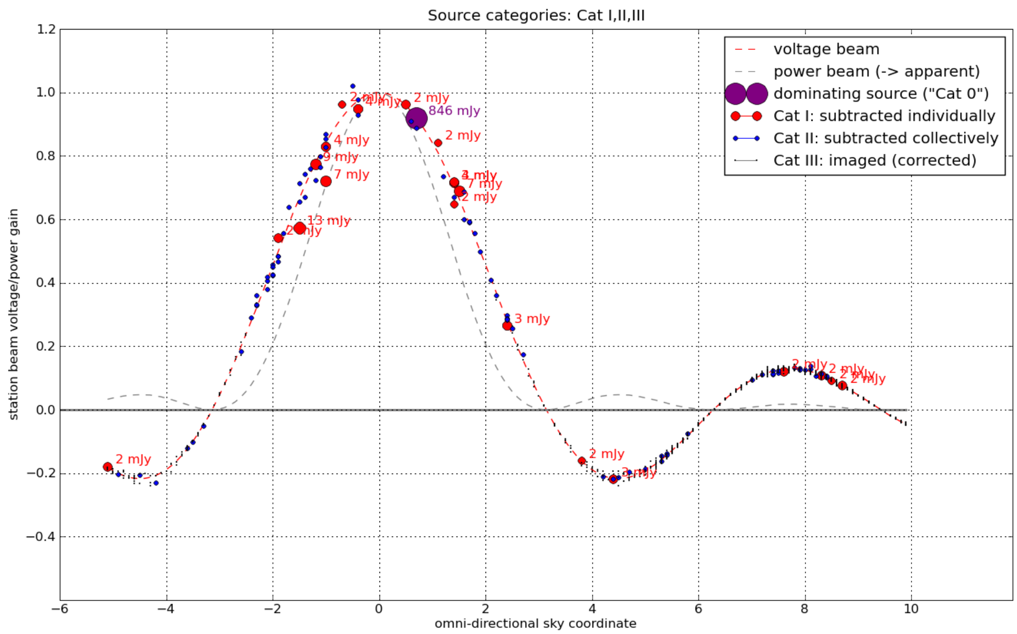

Figure 8: *When discussing beamshape calibration, it is useful to distinguish 3 categories of sources. The brightest sources in the LSM are* ***Cat I*** *sources (red circles). They are bright enough to be used as calibration beacons. Methods like the DG method (see fig 9) determine the specific ME parameter values in their direction* (l, m)*. These are used to subtract them from the uv-data with maximum accuracy, and also to estimate beamshape (EJones) parameters. We suspect that "every source that is bright enough to cause trouble is bright enough to be tackled individually" (i.e. as a Cat I source). All the other (fainter) sources in the LSM are* ***Cat II*** *sources (blue). They are subtracted from the uv-data with the help of the interpolated station beamshapes. Finally, there are many* ***Cat III*** *sources (black dots). They are too faint (*$< 5\sigma$ *in the final image) to be detected for inclusion in the LSM, and thus they cannot be subtracted from the uv-data. In order to image them properly, the beamshapes must be applied to the residual uv-data during the gridding process (A-projection). In this picture, to illustrate the relation of the various source categories to the beamshapes, the apparent fluxes of the sources are plotted along a station voltage beam. The scatter represents uncertainties in the LSM source fluxes. Obviously, such LSM flux errors in the bright Cat I sources (red) affect the estimation of station voltage beamshapes, while LSM flux errors in the fainter Cat II sources (blue) limit the accuracy with which they are subtracted from the uv-data.*

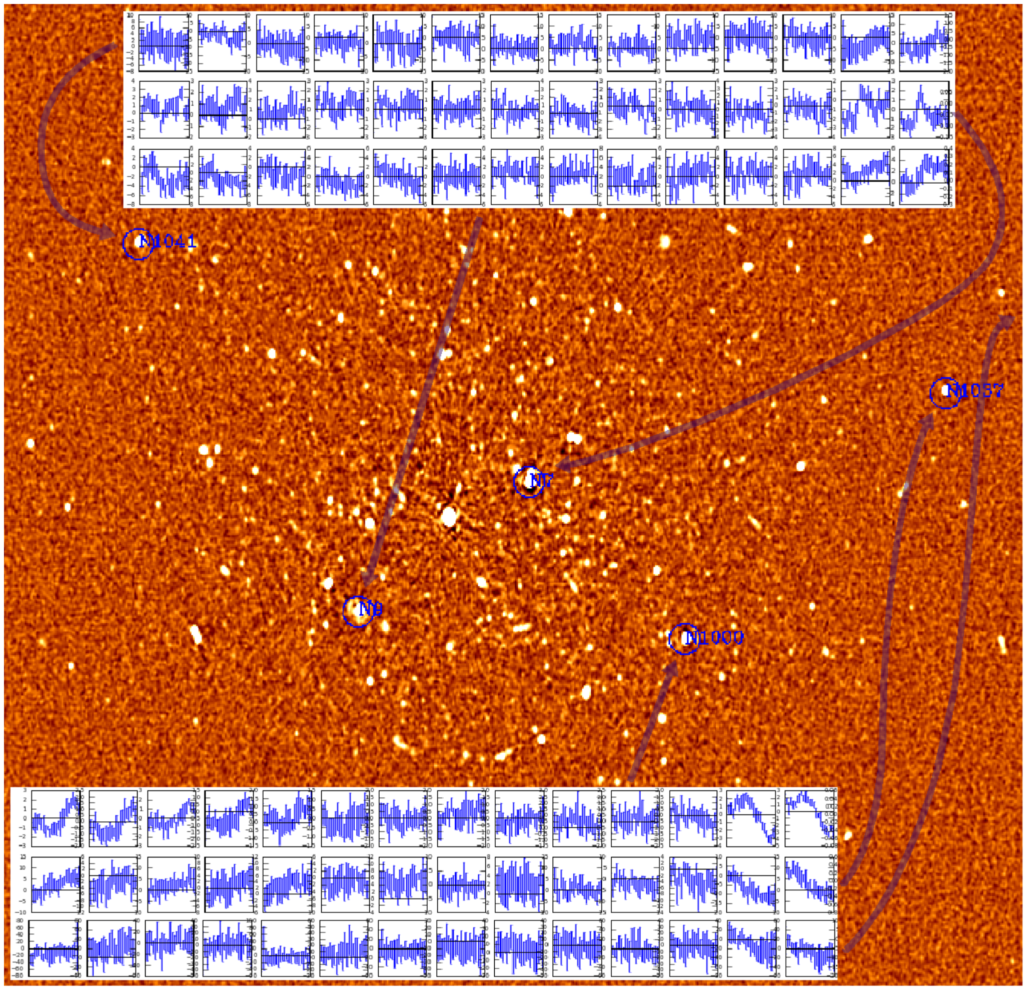

Figure 9: *The Differential Gains (DG) method, developed by Oleg Smirnov, is the first practical method to measure actual station beamshapes (EJones matrices) in a closed-loop fashion, i.e. during the observation. First, it takes out the rapid (tropospheric) gain variations in the direction on the dominant source (in this case 3C147, 21 Jy). Then it solves for the differential gains (DG) in the direction of a number of fainter "calibration beacons" in the field. Because the DGs vary only slowly, it is possible to integrate over 30 minutes or more. Especially if one does not require the DGs to be constant during that interval, but solves for the coefficients of low-order polynomials in time and frequency (as MeqTrees does). The insets show the estimated DGs for the 6 calibration beacons, for each of the 14 WSRT stations. Note that the S/N is sufficient to measure real effects. In fact, the vertical bars are not noise, but the variation in the solution for different frequency channels. The estimated DGs are used first of all to subtract the calibration beacons themselves with maximum accuracy. But they may also be used to solve for the parameters of individual station beams as a function of time and frequency, as is illustrated in fig 11. These are then interpolated to subtract fainter sources from the uv-data. They are also used to correct the residual uv-data as part of the imaging progress (A-projection).*

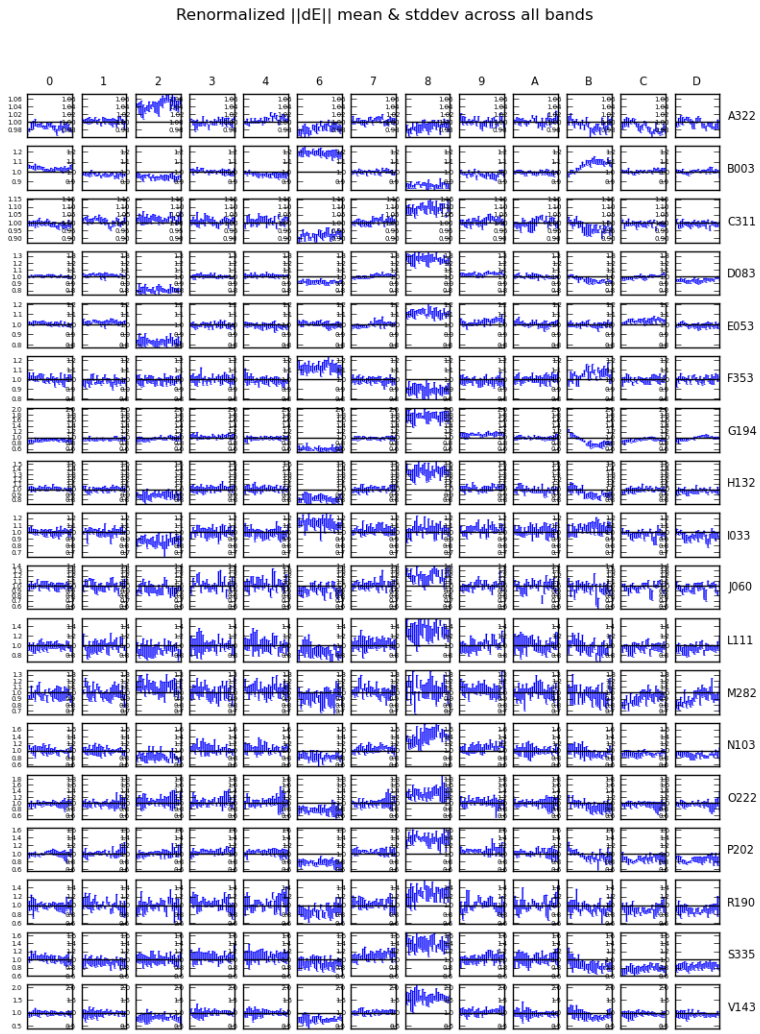

Figure 10: *In order to verify the DG method, a controlled experiment was done, using the "QMC2" field which has 20 calibration beacons and a brightest source of 200 mJy (100 times fainter than 3c147 in fig 9). Deliberate (but secret) pointing errors were introduced for a number of WSRT stations. Here are the Differential Gains (amplitudes only) for the 14 WSRT stations, for the 20 calibration beacons. Note that the source names have a 3-digit number. The first two indicate the azimuthal position, while the third indicates the distance from the field centre. The plots have many fascinating details, which will not be discussed here. They have been included here for comparison with fig 11. Suffice it to say that, with the help of these DGs, all 20 beacons were cleanly subtracted.*

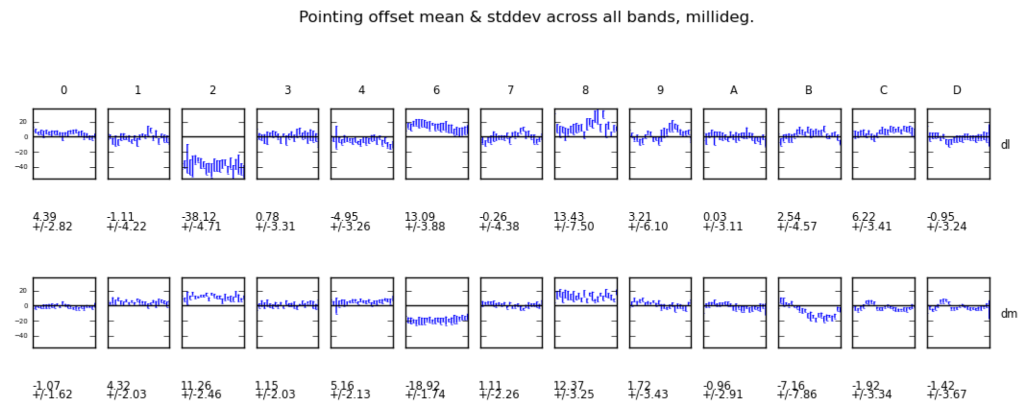

Figure 11: *Rather than solving for Differential Gains in the direction of calibration beacons in the field (see figs 9 and 10), it is also possible to solve directly for the parameters of some mathematical model of the WSRT station beams. This is where it gets interesting. In this case, the model was a circularly symmetric* $cos^3(r)$ *voltage beam, with parameters* $\Delta l_i(t)$ *and* $\Delta m_i(t)$ *representing station pointing errors (in milli-degrees). As before, the integration time was 30 min, and the error bars correspond to the stddev over the 8 spectral windows (i.e. not the noise!). The* ***deliberate*** *constant pointing errors of stations RT2, RT6, RT8 are clearly detected, as is the* $cos(HA(t))$ *pointing error of RTB. Note that the solutions are less "noisy" than the DG solutions in fig 10, which is expected because we are solving for fewer parameters (i.e. 2 per station, rather than 20). But the remaining structure in the solutions suggests that the beam behaviour cannot be explained by pointing errors alone, so more sophisticated beam models are needed. This is confirmed by the resulting image (not shown here), which has more residual structure around the calibration beacons than the one in fig 10. Scrutinizing plots like the above may offer clues as to other effects that should be taken into account. For instance, the solutions for the mispointed stations have somewhat larger error bars, suggesting that the frequency dependence could be modelled better. And the close similarity of the RTC and RTD solutions might suggest that one or more calibration beacons are resolved for the longer baselines, and asymmetrical.*

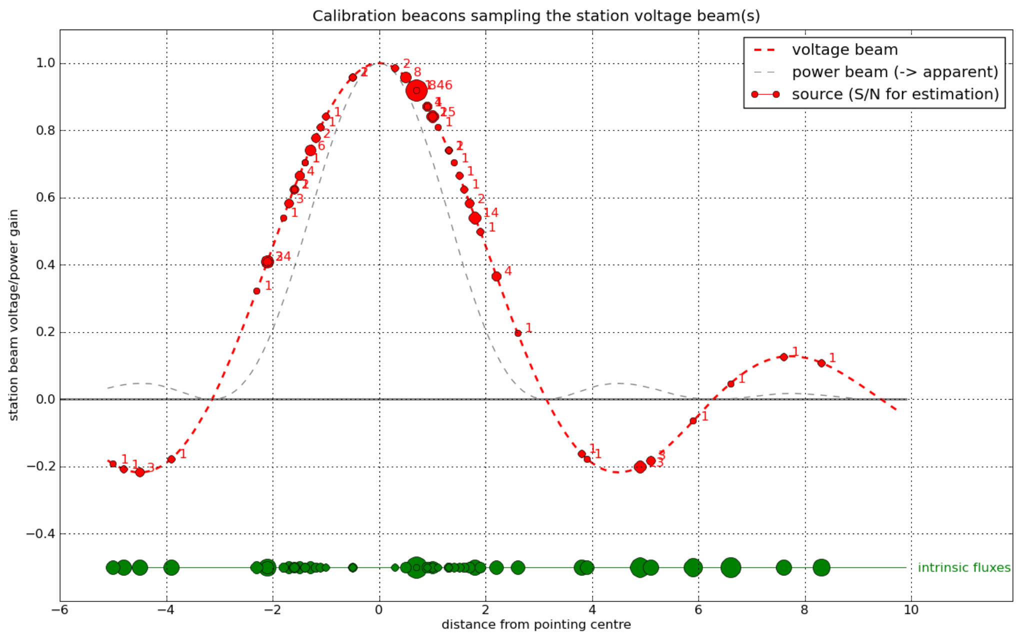

Figure 12: *Closed-loop estimation of station voltage beams depends crucially on the availability of sufficient* ***calibration beacons*** *in the field, to sample the 2D beamshapes over some required area. The beacons are radio sources with sufficient (apparent) brightness ($\gtrapprox$ a few mJy) to allow the estimation of instrumental gains in their specific direction (l,m), with reasonable S/N. In this plot, the intrinsic fluxes of the sources in the LSM (see fig 6) are attenuated with the station power beam, to produce apparent fluxes. These are plotted along the station voltage beam, to indicate how well its 2D area is sampled by the calibration beacons. Note that the horizontal axis is the radial distance to the pointing centre.*

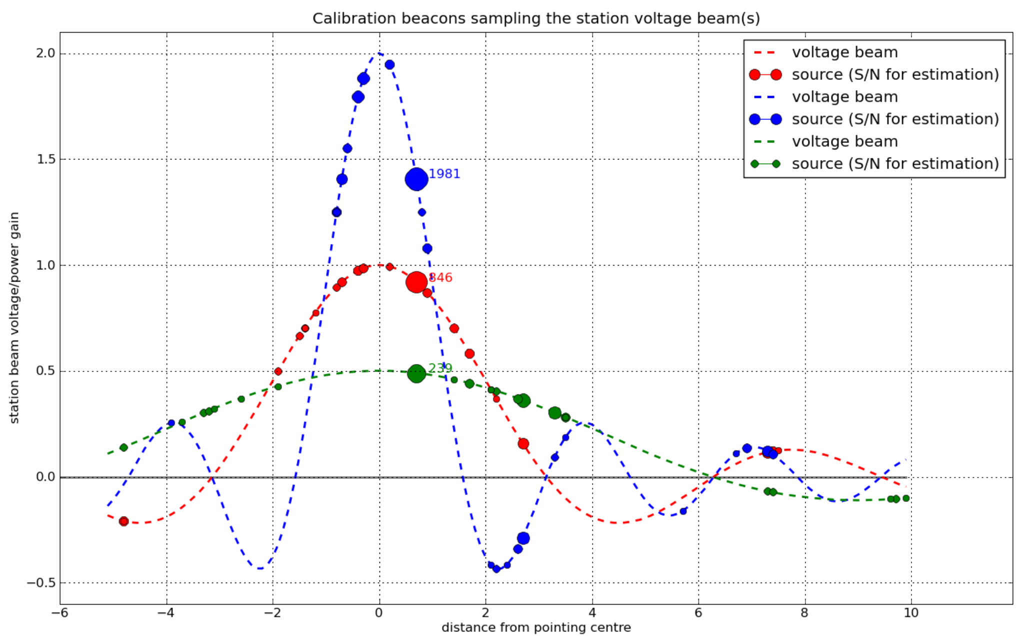

Figure 13: *To first order, the availability of calibration beacons is independent of station diameter D, since both the sensitivity and the field size depend on D^2, and thus cancel each other out. This is illustrated here for three different beam widths. Thus, if there are enough beacons for the WSRT @ 21cm (see fig 9), there should be enough for any telescope with comparable T_{sys}, for all wavelengths longer than 21cm. This finding is of crucial importance for the calibratability of SKA.*

7 Using the WSRT as a testbed

7.1 Measuring WSRT station beams

7.2 Measuring the Embrace station beam

7.3 Measuring the Digestif station beam

7.4 Measuring the Apertif station beams

8 Conclusions

Appendix: EJones modelling

Appendix: Estimating M.E. parameters

Estimating instrumental (Jones matrix) parameters

Estimating LSM source model parameters

Appendix: Designing station beamshapes

References

[1]

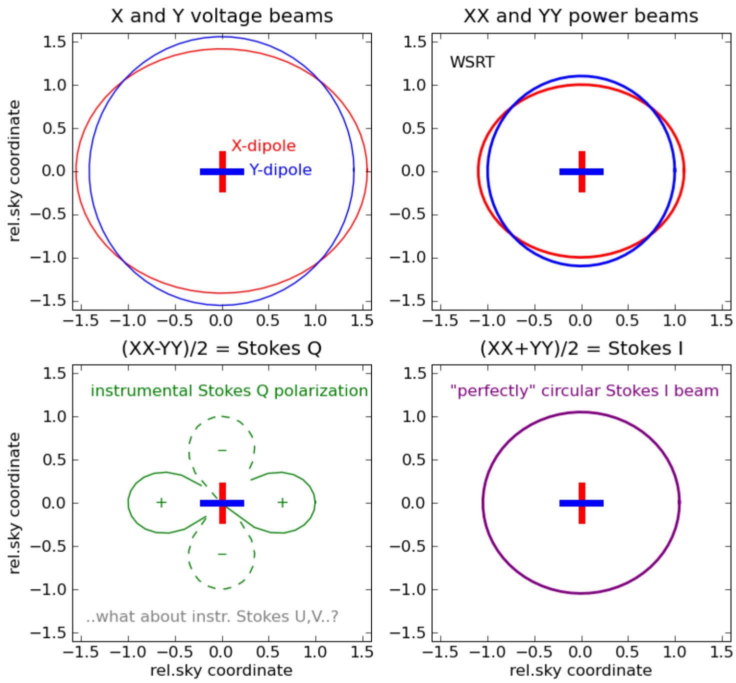

Figure 14: *A station beam actually consists of two receptor beams. In the case of the WSRT, these are the response beams of the two perpendicular dipoles (X and Y), in combination with the parabolic dish. Instrumental polarization is caused by differences between receptor beams, and interactions between them. A dipole placed in the focus of a perfectly round dish has a slightly elliptical voltage beam, due to the elongated shape of the dipole. This is illustrated schematically (and somewhat exaggerated for clarity) in the top left panel. XX and YY visibilities "see" power beams (top right) that are the product of two receptor voltage beams of different stations. To first order, the sum of the elliptical XX and YY beams will be circular, leading to a circularly symmetric Stokes I beam (bottom right). But the differences between the XX and YY beams cause a clover-leaf of position-dependent instrumental Stokes Q polarization (bottom left). It should be emphasized that this is only the simplest, most readily understood case. Instrumental Stokes U and V polarization is much less understood, even for the the WSRT. And even the WSRT has (small) "leakages" between X and Y beams. Other telescopes present more complex cases. For instance, due to the off-axis placement of the receivers, the R and L voltage beams of a VLA station are less circular, and not co-located (beam squint). Moreover, due to the Alt/Az mount, the entire asymmetric pattern rotates on the sky during the observations. LOFAR will obviously be much more complicated still. However, the good news is that all this behaviour can be modelled perfectly by means of a 2x2 complex EJones matrix, as part of the Measurement Equation. And thanks to the DG method, instrumental polarization can now be measured during the observations, using polarized and un-polarirized calibration beacons.*

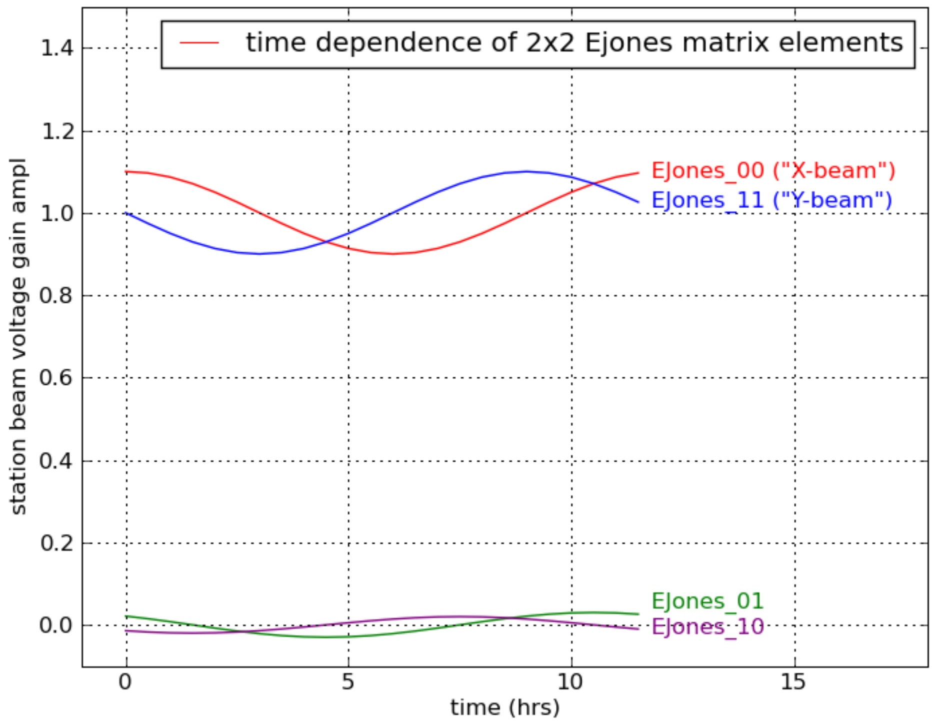

Figure 15: *The picture shows the variation of the four EJones matrix elements as a function of time, for the case where the beam rotates on the sky (i.e. for an alt-az mount or an aperture array). Note that the beamshapes are complex functions, but interferometer phases cancel in the case of identical station beams. Obviously the magnitude of the variation depends on the source position (l,m) in the beam, and on the asymmetry of the beam w.r.t. the pointing centre. For instance, it will be greater for the VLA, with its "beam squint" caused by its off-axis receivers. And worst of all for aperture arrays like LOFAR stations, whose rotating beams are elongated by projection effects at low elevations. The same projection effects also cause much more complicated instrumental polarization. The time variation of the beamshape patterns is minimized if the beam is stationary on the sky, as is the case with equatorial mounts (e.g. WSRT) or by means of mechanical counter-rotation (ASKAP). This greatly simplifies calibration, and should lead to better results with (much) less processing.*

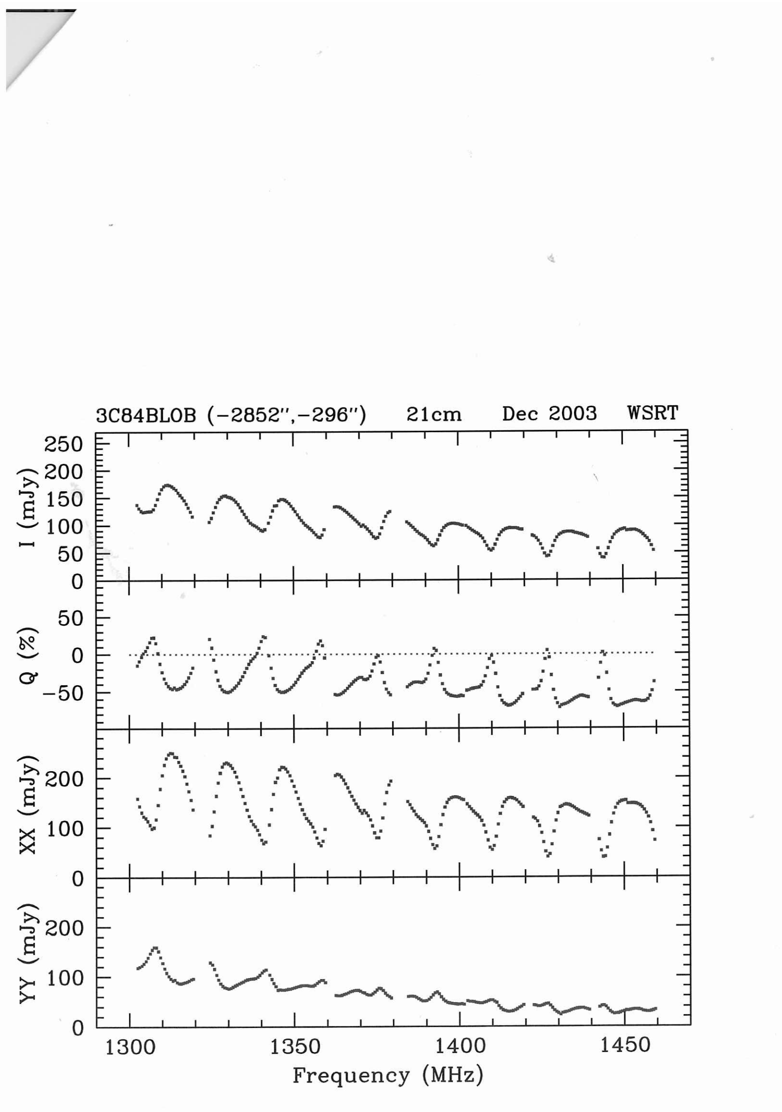

Figure 16: *Station beamshapes have a strong frequency dependence. The largest effect is geometric, i.e. the beam width is determined by the station diameter divided by the wavelength. But often there are also other, less predictable effects. As an example, the plot shows the infamous* ***17 MHz ripple*** *across the band of the WSRT, ostensibly caused by standing waves between the focus box and the dish apex (the effect seems to be much smaller for the Apertif frontends, since these hardly reflect the signal). For LOFAR, the electronic beamforming, with its mutual coupling and ground-plane effects, causes frequency dependencies that are not yet fully understood (and may not be easily predictable, or even repeatable). A worrying aspect is that such effects are not very "smooth", so their modelling will require more parameters. Of the four dimensional variables (l,m,f,t), frequency dependence will probably be the least smooth, and will thus be the most difficult (and expensive) to calibrate.*

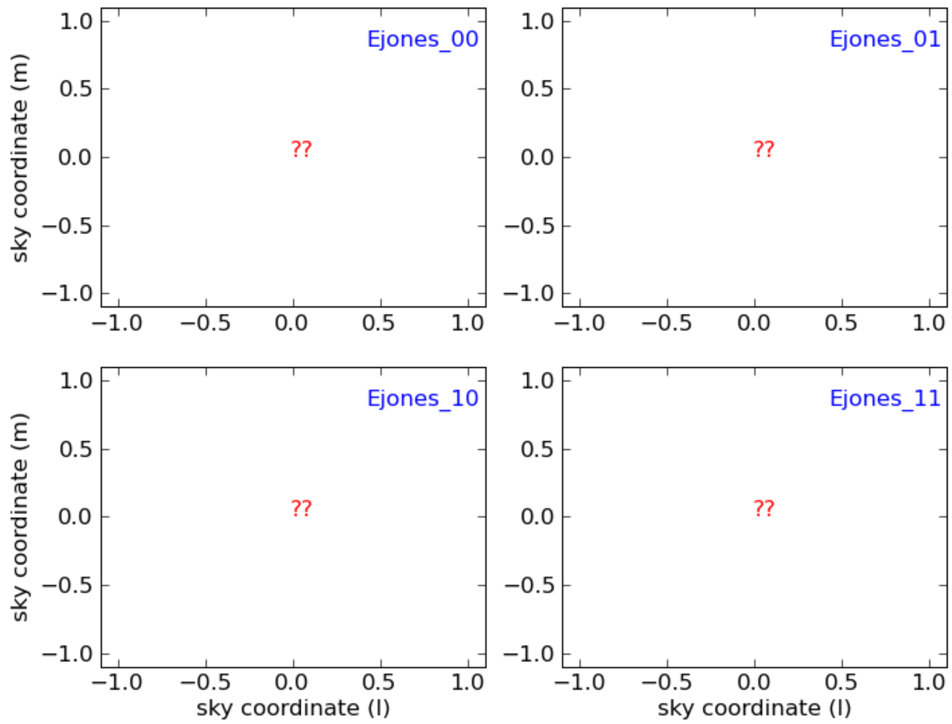

Figure 17: *In order to measure station beamshapes with the help of calibration beacons in the field, parametrized beamshape models are required. The parameter values are then estimated by comparing the predicted values with the measured ones at the position of the beacon sources. Depending on the type of station, and the required accuracy, beamshape models may range from simple gausssians to highly irregular shapes. They must be developed gradually. However, it is possible to lay down some ground rules.* ***Firstly***, *four separate mathematical expressions are needed, for the* 2×2 *elements of each station EJones matrix, which represents the voltage beam. Obviously, these four expressions may share some parameters.* ***Secondly***, *beamshapes should be defined in a sky coordinate system, e.g. (RA,DEC) or (l,m). This makes it much easier to relate it to the positions of calibrator beacons. Moreover, if (l,m) is used, this avoids complications associated with the zenith, or the celestial pole.* ***Thirdly,*** *all solvable parameters should be functions of frequency and time, e.g. low-order polynomials. This takes care of unpredictable beamshape variations during the observations.* ***Fourthly,*** *the mathematical expressions should have the general form of a series of suitable base-functions, whose coefficients serve as parameters. This makes it possible to trade off processing cost against accuracy by choosing the number of terms. Obviously, the base-functions should be chosen in such a way that the number of solvable parameters is minimized. An attractive choice is shapelets, which are very versatile and easily implemented, and are eminently suited to the kind of shapes that station beams tend to have.*