Author: J.E. Noordam	Date of issue: Version 1.0: 1 March 2005 Kind of issue: Public	Scope: Project Documentation Doc.nr.: LOFAR-ASTRON-DOC-00000	LOFAR
	Status: Draft Revision nr.: 1.0	File: $lofar/$	

A minimum ionospherical model for low-frequency data reduction

J.E. Noordam

(jnoordam@astron.nl)

Dwingeloo

Version 1.0: 1 March 2005

Verified:

Name	Signature	Date	Rev.nr.
	o.p.v.	15th November 2009	0.9.5

Accepted:

Work Package Manager	System Engineering Manager	Program Manager
J.E. Noordam .... *date:*	C.M. de Vos .... *date:*	J. Reitsma .... *date:*

Distribution list:

Group:	For Information:
ASTRON	dOList
O.M. Smirnov (OMS)	C.M. de Vos
R. J. Nijboer (RJN)	K. van der Schaaf (KvdS)
A. G. Willis (AGW)	H. Holties
R. Assendorp (RXA)	A.G. de Bruyn
T. A. Oosterloo (TAO)	M. A. van Haarlem
G. van Diepen (GVD)	J. D. Bregman (JDB)
W. N. Brouw (WNB)	
M. Brentjens (MAB)	
M. Mevius (MXM)	

Document revision:

Revision	Date	Section	Page(s)	Modification
0.5	2004-Nov-25	-	-	Creation
1.0	Version 1.0: 1 March 2005	??	??-??	Updated class description

Abstract

An ionospheric model with the minimum number of parameters is investigated.

1 Introduction

With WSRT LFFFE observations, it is only possible to make a selfcal solution for individual telescope phases when there is a 3C source (10-100 Jy) in the field. In general, the brightest source is only 5 Jy or less (give S/N calculations). In those cases, the data must be corrected with the help of calibrator observations before and after. At 115 MHz, this would have led to a rather poor dynamic range, because even the benign ionosphere of the fall of 2004 causes appreciable phase gradients over the 3 km array. Fortunately, there was usually enough flux in the available sources to solve for this phase-gradient (one parameter).

This little anecdote contains the key to the reduction of observations with LOFAR, and other low-frequency radio telescope. The point is that we should try to find the ionospheric model with the smallest number of parameters, which still describes the observed phases with sufficient accuracy. The latter means that the accuracy should be greatest in the direction of the brightest sources, which after all have to be subtracted with the greatest accuracy.

In our original plans for LOFAR calibration, we were planning to solve for separate 2D phase screens across the FOV (main lobe and inner side-lobes) of individual stations. Such screens are entirely sufficient to account for all phenomena, and as a model they are superior to the Zernike polynomials (Cotton et al) and clumsy concepts as refraction and apparent position.

However, this approach would have required at least 3 parameters per station for a flat screen, and more to describe any curvature. This would need as many sources in the field, and our early calculations wrongly assumed that they would all have to be bright enough to yield a S/N>3 in 10 sec. Not surprisingly, this led to a rather gloomy picture. Other schemes based on calibrating 'patches' around bright sources lead to similarly unattractive numbers.

Fortunately, ionospheric phenomena are rather large-scale in space and time, and usually smooth in all these dimnensions. This means that we can get away with an ionospheric model with rather fewer parameters, and consequently with fewer bright sources to estimate their values. It is important to realise that we are not interested in the structure or the physics of the ionosphere, but just in the observables (phase, Faraday rotation) that affect our observations. Thus, we are interested in the simplest model that describes these observables with sufficient accuracy. This paper investigates what are the *minimum number of parameters* that are needed in various conditions.

A minimum ionosphere model must meet the following criteria:

1. It should be exact for the case of a *uniform blanket*, i.e. an ionosphere of uniform thickness and electron distribution, at a constant altitude above the curved Earth surface.
2. For increasingly complex situations, it should be possible to detect that the model is incomplete, e.g. by the increasing internal inconsistencies in the parameter solution.

3. Depending on the nature of such inconsistencies, there must be algorithmic recipes to automatically extend (or reduce!) the model with extra parameters.

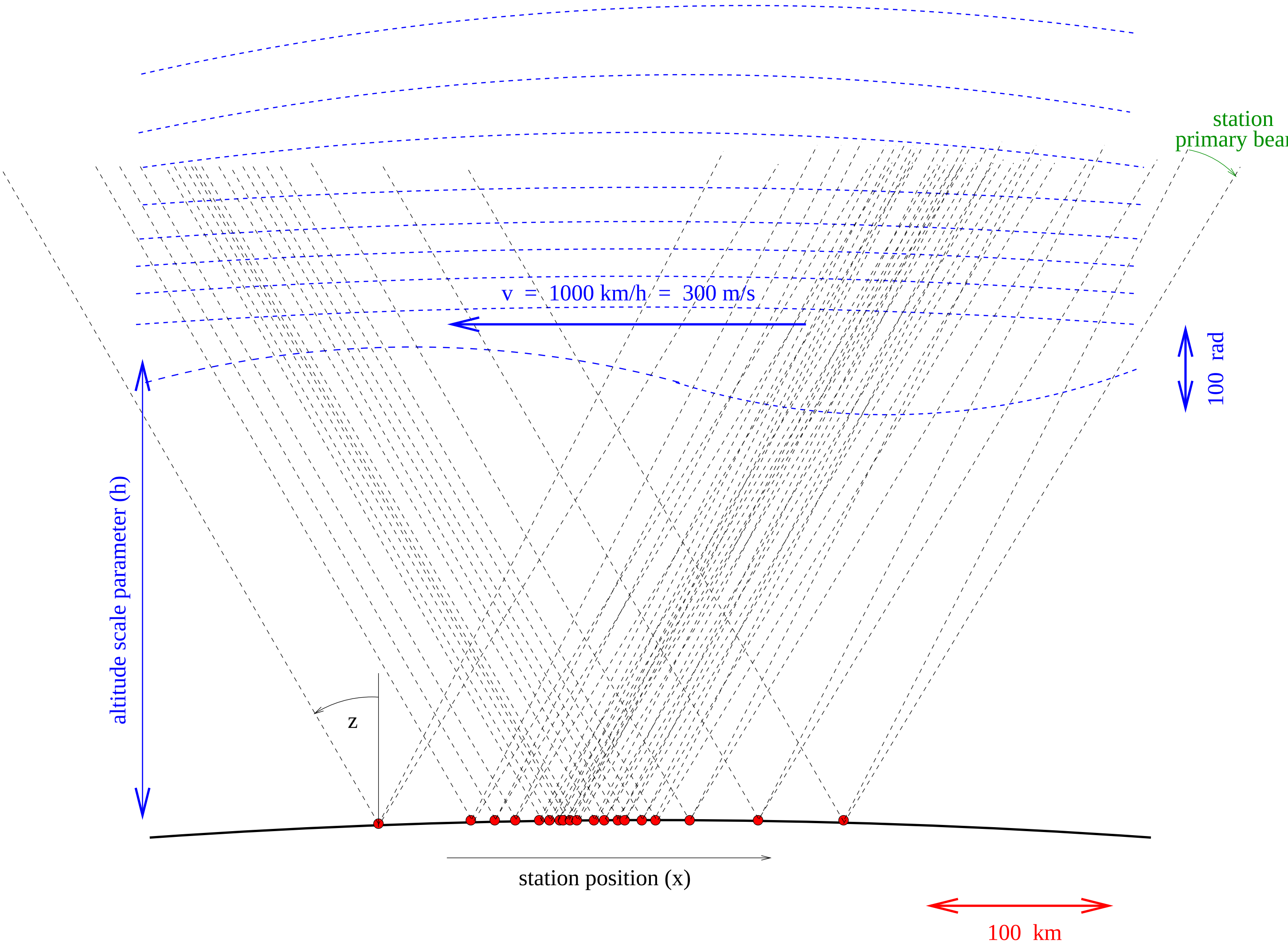

Figure 1: *Schematic diagram of the ionosphere. The main variation is caused by Travelling Ionospheric Disturbances (TID) at an altitude of 250-300 km.*

2 A large-scale (>100 km) ionospheric model

The larger the scale of the phenomena (in space and time), the simpler the model that we can get away with. Since the ionosphere often only has large-scale structure, we bfirst build a model for this case, and add complexity for dealing with more rapidly varying smal-scale structure in a later section.

2.1 A curved blanket of uniform thickness

The simplest usable model of the ionosphere is a thin blanket of uniform thickness at an altitude of h km. The excess path L_0 in the zenith direction ($z = 0$) can be written (see [1]) as:

$$L_0 = -40.3(\frac{100}{f_{MHz}})^2 \times TEC \ \ m \tag{1}$$

where the Total Electron Content, i.e. the integral of the electron density along the propagation path, is measured in TEC units of $10^{16}\ m^{-2}$. The excess path corresponds to a phase delay, and is *negative* for the ionosphere. For an observing frequency of f = 100MHz ($\lambda = 3m$), an a typical night-time value of $TEC = 5$, the excess path length is about -200 m, or about -400 radians.

For an arbitrary zenith angle, the excess path will be longer:

$$L(z) = L_0 * S(z) \tag{2}$$

where the factor

$$S(z) = 1/\cos(\arcsin(\frac{R\sin z}{R+h})) \tag{3}$$

describes the increase in the excess pathlength through a thin layer at an altitude of h km as a function of the zenith angle. R is the Earth radius, and $S(horizon) \approx 3$ for $h = 300$km.

2.2 Large-scale structure (>100 km)

In practice, the ionosphere will always have some large-scale structure, if only the 'egg' shape following the direction of the Sun. This can be taken into account by adding some terms to L_0 that depend on ground position x[1]. So, in the zenith direction ($z = 0$), we get:

$$L_0(x) = L_0(1 + p_1 x + p_2 x^2 + p_3 x^3 + \ldots) \tag{4}$$

Here we use a low-order polynomial for simplicity, but another smooth function may do as well. Its parameters $p_k(t)$ vary as a function of time, and are to be determined (see section 3).

Using equation 2, and referreing to fig 1, we get an expression for the excess path in the direction of zenith angle z, as observed from a (station) position x:

$$L(x, z) = L_0(x - h \tan z) \times S(z) \tag{5}$$

The 'altitude parameter' h is of the order of 300 km. Its precise value is not critical in most cases, i.e. when L_0 has only a few terms. Its main function is as a 'coupling constant' that introduces z-dependency. A 'wrong' value of h will be absorbed in different values of the parameters p_k.

Note that equation 5 also predicts the variation of the ionospheric phase over the field-of-view, which will be different for different stations. This approximation will be good enough ($<$ 0.5 rad?) for most of the sources in the field, but not for the brightest sources, which have to be subtracted with the greatest accuracy. Fortunately, these sources are bright enough to have their own 'private' phase measurements, which do have the required accuracy.

3 Solving for the model parameters

We use self-calibration on bright sources in the field to obtain 'puncture-points' through the ionosphere at different positions. These measurements are then used to estimate the parameters of our ionospheric model.

3.1 The relation between z and source direction (RA, DEC)

nn

---

[1] For simplicity, the problem will be discussed in one dimension. It can easily be extended to the required 2 dimensions.

3.2 Main lobe and side-lobes

The station beam is complex. For the sources in the main lobe of the station beam, we may assume (to first order) that the phase-**differences** measured by the various interferometers are unaffected by the beam phase. Since any deviations from this assumption will vary slowly and smoothly over the field, we will assume that, in the main lobe, the beam-shape is real (amplitude only), and any phases are subsumed in the much larger ionospheric phase.

However, the story is different for sources in the side-lobes of the station beam. First of all, the beam phase will jump by π when passing from one sidelobe to another. At the very least, his will cause problems with the selfcal phase solution. Since the side-lobe patterns between stations will differ considerably, a source may be in entirely different sidelobes for different stations.

3.3 Influence of the instrumental phase (GJones)

Much less smooth than the ionosphere. Subsumes any inconsistencies in the one-source solution, so a single source **does NOT constrain the ionosphere model**.

Unknown phase-reference of the selfcal phase solution. PZD.

3.4 The brightest source in the field

If the number of stations N_s is greater than the number n_p of parameters p_k in the model a single phase solution on the brightest source will be sufficient to solve for them. **Thus, usually, only ONE bright source is needed to track the ionosphere!** This is a startling conclusion indeed.

Thus, the brightest source in the field plays an important role.

Depending on the number of stations N_s, there are three possibilities:

1. If $N_s = n_p$, the solution will be exact. This means that the curved phase-screens across the various stations are fully determined, and intersect the fat blobs associated with the brightest (100 Jy) source in the figure.
2. If $N_s > n_p$, the discrepancies between the selfcal solution and the model contain information about the incompleteness of the ionosphere model. For instance, the differences may be used to solve for the optimal value of the altitude parameter h. If that is not sufficient, more terms may be added to equation 4. Etc. (*However, this is affected by the electronic GJones phase....*)

	N_s	size	order	z_{max}	blanket	h(km)		n_p	remarks
WSRT	14	3 km	2nd	45	flat	300		4	
+WHAT	15	3 km	2nd	45	flat	300		4	
LOFAR	50	100 km	3rd	45	curved	300		5	
LOFAR	100	400 km	3rd	45	curved	300		5	

Table 1: N*umber* (n_p) *of parameters in the ionospheric phase model.*

3. If $N_s < n_p$, more than one bright source will be needed to solve for the model parameters. This will not often happen.

3.5 The other bright sources in the field

S/N>3 in 10 sec(?).

If the model is correct, their (separate) selfcal phase solution should produce phase that are consistent with it. In fact, the model may be used to choose the unknown phase bias of such a solution (by adding suitable constraint equations to the solution).

Any remaining discrepancies are an indication that the ionospheric model is not (yet) correct. This information can be used in various ways...

3.6 Combining the results from several beams

An overall ionospheric model may play a useful role here...

4 Small-scale structure (<100 km)

In principle, one might add more terms to $L_0(x)$ in equation 4. However it will probably be necessary to add cross-terms:

4.1 Ultra-small scale structure (<10 km)

Stop observing...?

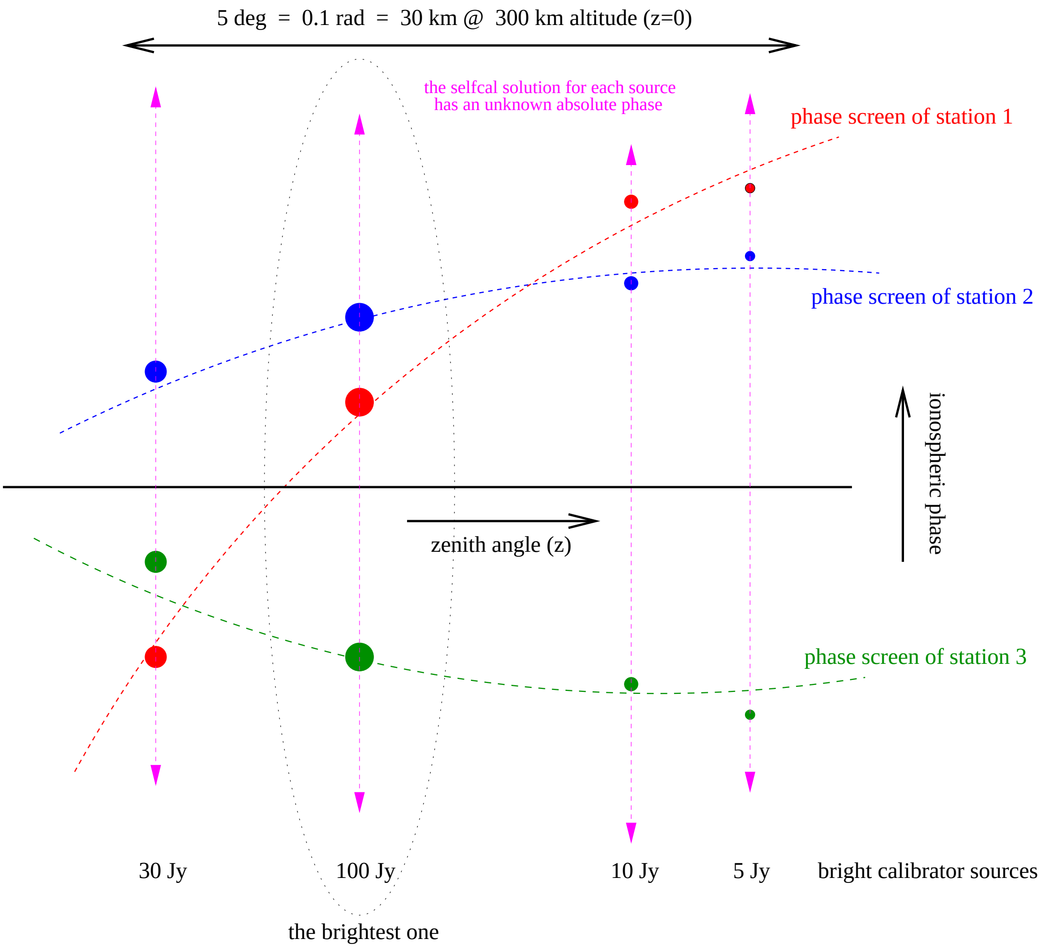

Figure 2: *Bright calibrator sources are used to solve for the* n_p *parameters of the ionospheric phase model. It is important to realise that, in most cases, only one bright source is needed to constrain the ionospheric phase model. On the contrary, the model is used to constrain the selfcal phase solutions for the Tfainter sources. The remaining discrepancies with the latter are merely used to determine the completeness of the model. Nevertheless, even an incomplete model may be entirely sufficient for the subtraction of Cat II sources, as long as the brighter Cat I sources are subtracted with their 'own' phases.*

5 Use of the 'frozen' pattern

The ionospheric phase often manifests itself as a 'frozen' pattern, which changes only slowly while travelling with a velocity of 500-1000 km/h (165-330 m/s). Obviously, we should try to take advantage of this phenomenon, just like we should try to use all *a priori* knowledge about the system.

NB: There appear some doubts about the existence of a usefully frozen pattern. Thus it is fortunate that it does not play a large roles in our calibration scheme.

Solving for the travel velocity (v_x, v_y). This is a byproduct of the time-behaviour of the station phase.

6 Faraday rotation

The Faraday rotation is related to the overall ionospheric phase via the Total Electron Content (TEC). However, it also depends on the angle between the line-of-sight and the local Earth magnetic field. We have the choice of treating the Faraday rotation separately (but similarly) to the phase, or to connect them somehow....

Total electron content:

$$TEC = \int N_e dl \; m^{-2} \tag{6}$$

Faraday rotation angle:

$$\theta = RM \times \lambda^2 \; rad \tag{7}$$

in which the Rotation Measure (RM):

$$RM = \int N_e \vec{B}.\vec{dl} \; rad/m^2 \tag{8}$$

The Earth ionosphere can cause a RM of up to $3-4 \; rad/m^2$.

7 Frequency dependence

We assume that the ionospheric phase is $\propto \lambda^2$. This *a priori* knowledge can be used first of all to improve the S/N of the solution for the parameters of the ionospheric delay function $L_0(x)$. It can also be used to distinguish between the (large) ionospheric phase and any (small) phase-factor in the shape of the station primary beams (to first order we assumed that the beamshape is real, i.e. a pure amplitude effect). Any beamshape phase will probably be $\propto \lambda$. Finally, we should consider the electronic phase (GJones), which will be different for different stations, but the same for all sources in the field.

8 The full Measurement Equation

Assume that all phases are the ionosphere, and all amplitudes the beam...? Problem, since station electronbic phases may be arbitrary, this would affect the all-sky smoothness of the ionospheric phase. So, should we absorb these individual variations in a GJones matrix, i.e. an uv-plane efffect that is valid for all sources?

How do we relate the ionospheric phase (relative/absolute?) with the Faraday rotation?

The important thing is, as always, to have a minimum-parameter model that is used to subtract the Cat II sources, while we subtract the Cat I sources with their own parameters, for maximum accuracy. From the differences between these two sets we can estimate the error in the residual image....

9 The WNB point: sidelobe gain effects...?

The claim is that, for some reason, the phase and gain solutions may not be separated. If that might turn out to be the case, it is not clear whether this would be a secondary effect (comparable to the difference made by the socalled 'complex solution' in NEWSTAR.

10 The JPH point: matrix vs scalar selfcal

Everything that is being discussed here is entirely consistent with the full-polarisation Measurement Equation. In particular, the instrumental and ionospheric effects are described with direction-dependent Jones matrices.

11 The role of external ionosphere data

PIM and other physical models.

Reflection measurements (ionosonde, chirp sounder).

Transmission measurements (two-frequency GPS).

All these provide an accuracy of about a radian at best, in the absence of TID's and other short-term phenomena. They will be useful to get a better starting position for the LOFAR phase-locking procedure.

The latter is a continuous selfcal phase solution on one or more bright sources. It starts with 3 stations in the core. The other stations are included one by one, moving outwards from the centre. The phase of this new station is varied by steps of 2π, until the solution is consistent (phase-lock). Its phase will then be continuously tracked while the remaining stations are included, and during the subsequent observations.

12 MeqTree implementation

MeqTrees are remarkably suitable for the implementation of the kind of ionospheric models described here, and for solving for their parameters.

13 Conclusions

References

[1] Thompson, Moran and Swenson, *Interferometry and Synthesis in Radio Astronomy* (1994)

Contents