LBWG memo 33

Automatic LoTSS quality metrics

Neal Jackson, Marcin Hajduk, Sethi Sagar, Alexander Kutkin,
10.02.25

Purpose of the exercise

LoTSS solutions from LINC vary in quality, mostly according to the state of the ionosphere. Good solutions typically involve phase wraps on RS baselines over timescales of > 5 minutes, whereas bad solutions can wrap on timescales of 2 minutes or less. Although the solution can normally be followed on individual channels, the phase variation often implies a large delay variation across the band at a given time, and difficulty in interpolating the solutions if the channels are relatively coarse. Fast variations in phase on RS stations also guarantees that the international stations will be unusable.

At present the user inspects the LINC target solutions, derived from calibration against a TGSS sky model, and verifies that these are reasonable before proceeding. It would be useful to be able to do this automatically and in a standard way; bad solutions, once identified, could trigger a flagging of particular antennas or a decision not to waste multiple weeks of cluster time trying to make high-resolution maps with the international stations.

The data

Solution data were extracted from the LoTSS archive during the Long Baseline Working Group's virtual busy week in February 2025. There are two types of solution file available, both as `.h5` format:

- About 75% of the database consists of solutions from data averaged to 24 2-MHz channels. Most of the testing was done on this type of solution. These solutions have been produced in the `TGSSphase` mode, with solutions in both X and Y.
- The rest of the database consists of solutions from full 488-channel data, with solutions performed using smoothness constraints along the delay direction. These are produced in the `TGSSscalarphase` mode, with single polarization solutions formed from X and Y. We have used the `TGSSscalarphase_final` solutions, produced after one cycle of selfcalibration on TGSS.

All data have 8-second integration times; some data have about 3550 time stamps (corresponding to 8 hours), others have half this (mostly lower declination fields).

In addition, some solutions are available in `.npz` file format from `killMS`; these are not considered here.

Some of the database is consistently difficult to download using `rclone`, and more than one observer noticed that $> 50\%$ of the files caused `rclone` to hang. We operated using files that could be downloaded, rather than stop to investigate this.

Testing phase variations and establishing metrics

We used 8 phase solutions, two classified by eye as "good" (L661210, L798364), two as "usable" (L2023537, L793660), one as "good in parts, otherwise borderline" (L2019040) and three as "bad" (L880032, L2031726, L2019219). Fig. 1 shows the phase vs. (frequency, time) plots for these eight solutions.

The LBCS survey uses simple metrics for evaluating data, mainly consisting of differences in phase between polarizations. This measures scatter due to limited signal, however, rather than rapidity of phase variation. In the `TGSSphase` solutions, the two polarizations are well aligned and the difference is uniformly close to zero. We therefore tested the following simple metric, for one channel and one antenna:

```
(8.0/tint) * np.median (np.abs (np.gradient (np.unwrap (data[:,antn,chan,0]))))
```

i.e. the median of the absolute gradient of the unwrapped phase on antenna `antn` in channel `chan`, for one polarization. If the integration time were different from 8 s. this would also need to be multiplied by the `8.0/tint` factor.

For the middle channel, and a 40-minute time period between minutes 160-200 (corresponding to abscissa values 1200-1500 on Fig. 1) we obtain the correspondence between eyeball evalution and this metric shown in table 1.

Observation	Eyeball	Metric
L661210	Good	0.12
L798364	Good	0.07
L2019040	Good in parts	0.12
L2023537	Usable	0.15
L793660	Usable	0.14
L8800032	Bad	0.43
L2031726	Bad	0.28
L2019219	Bad	0.28

Table 1: Eyeball evaluation vs. metric for eight solution tables.

There appears to be a good correlation between the metric and the eyeball estimation, and in particular the separation between good/usable and usable/bad appears to be consistent. Fig. 2 shows the metric for each observation, plotted against time (in one-hour chunks) and antenna (with all CS stations at the bottom, RS stations towards the top, and international stations, which have zero phase gradients owing to not having solutions, at the top).

Again a good correlation is evident: the boundary blue/green corresponds to the good/usable boundary, and green/orange to usable/bad. A likely recommendation would be to discard L880032 entirely, discard all but the first two hours of L2031726, discard the first half of L2019219, and discard selected antennas for the other observations. This can be implemented easily using the matrix generated by the script and used to make these plots.

The 488-channel mode solutions

Finally we show the solution and metric plots for the 488-channel data in Fig. 3.

The script used to do this analysis is in the `LOFAR-VLBI/lofar-lb` repository on github.

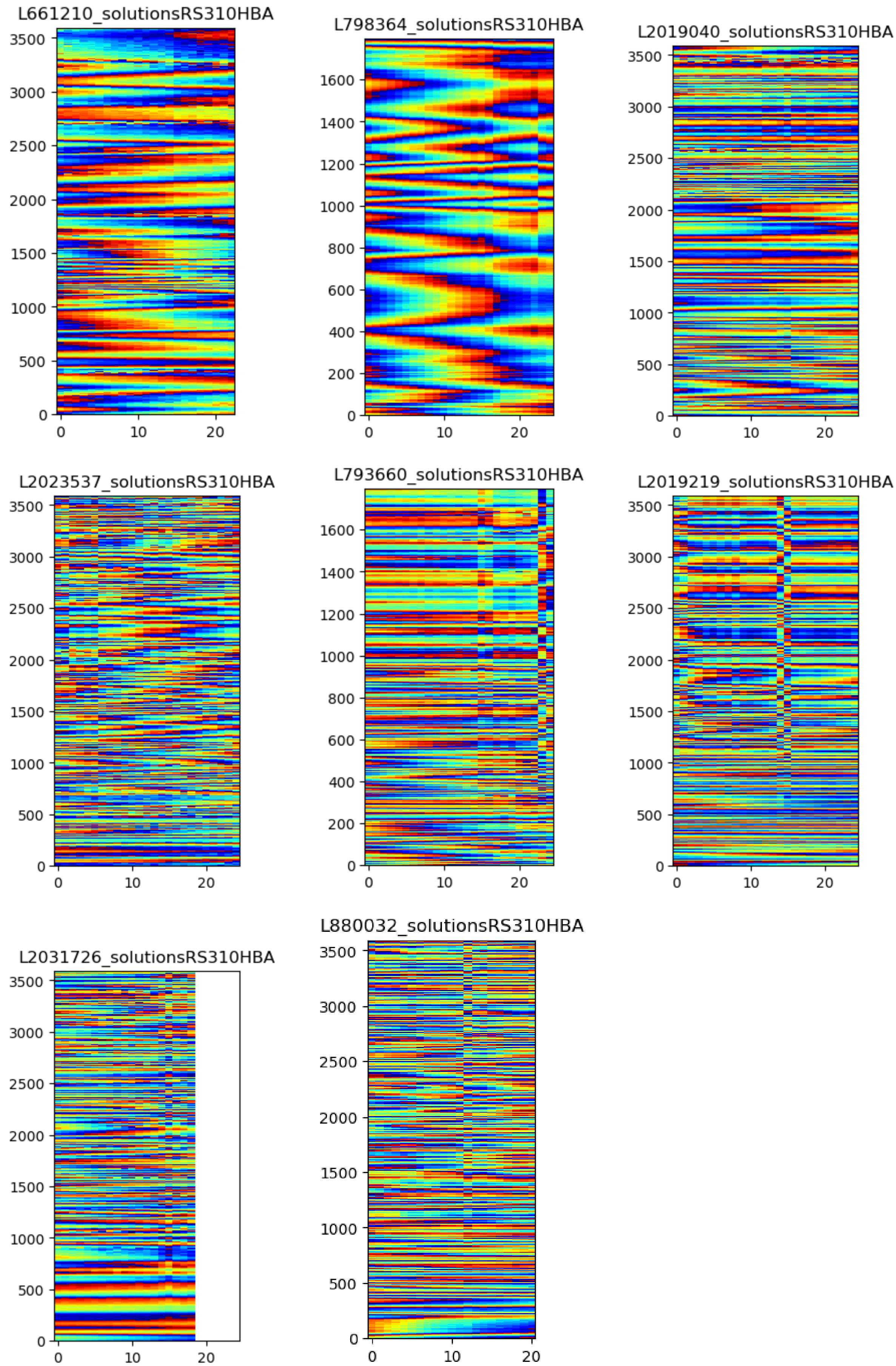

Figure 1: 24-channel solutions, in order of decreasing quality. The two on the top left were regarded as "good", at least for this antenna.

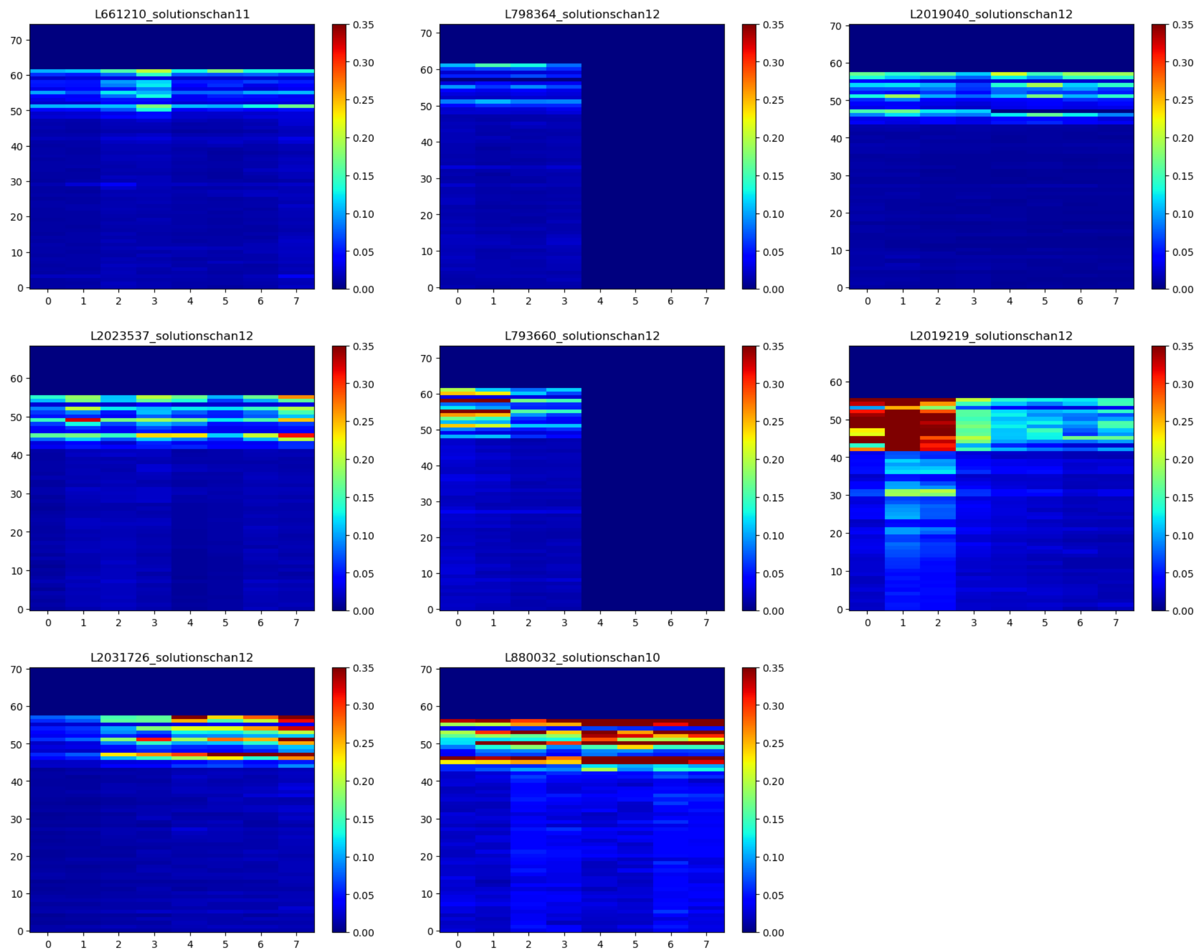

Figure 2: Plots of the suggested metric against time in hours (x) and antenna (y). All CS stations are at the bottom, RS stations towards the top, and international stations, which have zero phase gradients owing to not having solutions, at the top

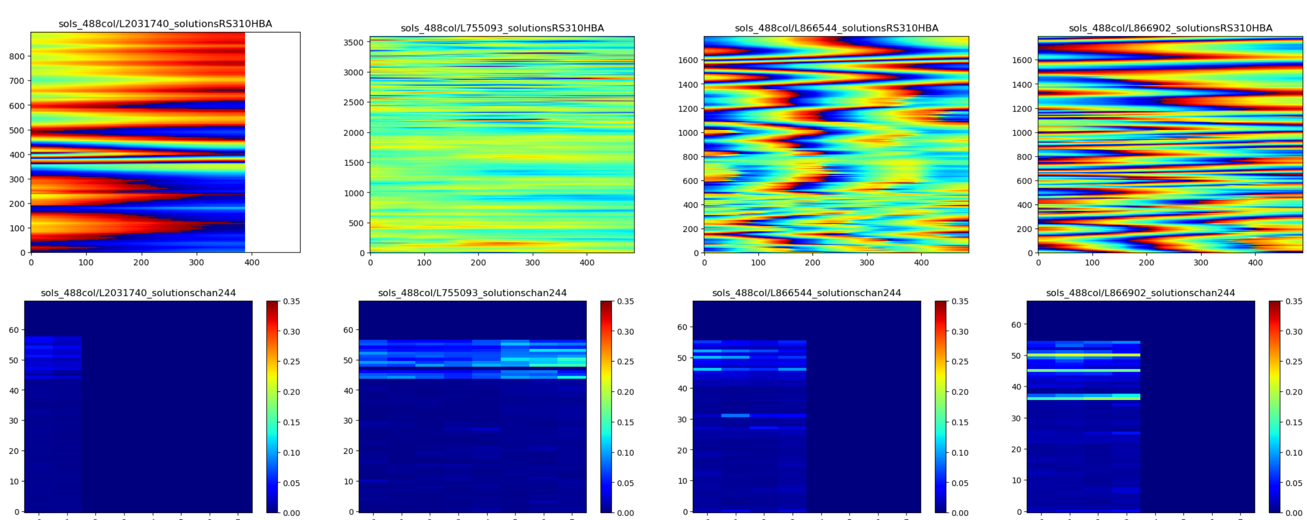

Figure 3: Plots of the solutions and suggested metric, as in Figs. 1 and 2, for the central channel of the 488-channel solutions.