VLBA SCIENTIFIC MEMO 23

IONOSPHERIC CORRECTIONS USING GPS BASED MODELS

Craig Walker

National Radio Astronomy Observatory

Shami Chatterjee

Cornell University

December 28, 1999

ABSTRACT

Corrections for the effects of the ionosphere on VLBI phases have been tested using a geodetic project and a phase referencing test observation. Dual frequency S/X measurements on strong sources with very good positions are treated as "truth", and calibrations derived from GPS based ionospheric models are compared to them. All of the GPS models appear capable of reducing the effects of the ionosphere by a factor of roughly 2 to 5 and so provide a valuable aid for phase referencing observations. The SATLOC model does not appear to have advantages over the global models archived at CDDIS, despite a finer grid in space and time. Therefore it does not appear to be worth the effort to establish an in-house SATLOC archive.

1. Background.

The ionosphere causes by far the largest unmodeled phase offsets in VLBI data at low frequencies and can be a significant contributor even at the highest frequencies of the VLBA. Typical ionospheric delays in the test experiments reported here are on the order of 10 ns at 2.3 GHz, equivalent to about 25 turns of phase. The ionospheric delay scales with the wavelength squared, while the phase scales with wavelength. Thus the above numbers imply roughly 175 turns at 90 cm, equivalent to a geometric error of something like 160 m. At 43 GHz (7mm), the same ionosphere would still contribute about 1.3 turns of phase or about a centimeter of path. Therefore, one cannot totally ignore the ionosphere at any frequency. The ionosphere is especially problematic because it is highly variable. Day/night variations are typically an order of magnitude and shorter term variations are also large.

The geodetic/astrometric VLBI community have long used dual frequency observations at 2.3 and 8.4 GHz (S and X bands) to remove the ionosphere. They use multiple baseband channels to span on the order of 100 MHz at 2.3 GHz and 400 MHz (sometimes much more) at 8.4 GHz. These wide spanned bandwidths allow accurate "multiband" delay determinations. Then the two delays (S and X) can be be used to calculate a non-dispersive component that is the same at both bands, and a dispersive delay that scales with wavelength squared and is the result of the ionosphere. The non-dispersive delay then has the ionosphere removed and can be used for solutions for source and station positions, Earth orientation, and troposphere. Note that this scheme does not provide an absolute measurement of the ionosphere. It gives a measurement of the difference in the ionospheres over the two stations in a baseline. Plus, there are arbitrary, but hopefully constant, delay offsets between the bands at both stations (different at each station) that give an arbitrary dc offset to the measured ionospheric values. This is not a problem for the geodetic/astrometric projects because it looks like a clock offset and is treated as such. But it complicates comparison of S/X ionosphere measurements with measurements based on other types of data.

Geodetic measurements using the GPS navigation system are subject to the same ionosphere that affects VLBI since both are based on measurements of delays of radio signals propagating through the ionosphere. Their solution is also similar — the geodetic GPS observations use two frequencies within L band. The GPS community has a world wide network of receivers working all the time and so is able to gather large amounts of data about the ionosphere. Several analysis centers are now making global ionospheric models based on GPS data. Five such centers make these models available on the CDDIS (Crustal Dynamics Data Information System) data archive where they can be accessed on the internet if you have a password (which either of us can supply). The models available at CDDIS are in IONEX format and are for every 2 hours on a 5 degree grid in latitude and longitude.

In addition to the above models, NRAO has an arrangement with a private company, SATLOC, where they allow us to use ionospheric models that they provide in real time in return for us allowing them to place a receiver at our St. Croix site. Their models are calculated in real time every 10 minutes on a 2 degree grid (See VLBA Scientific Memo 22 for more details) and are intended to be used in real-time applications, mostly in agriculture. They are distributed from a satellite to special receivers at each point of use. NRAO has been provided with such a receiver which is installed at the AOC. The models are not archived unless we do it. These data are on shorter time scales and on a finer grid than the global models, but only cover the continental US. If we plan to use these models on a regular basis, we would need to establish an automated procedure to capture and archive the data, instead of doing this on a case-by-case basis as at present.

The AIPS task TECOR, written by Chris Flatters, is used to read the ionospheric models in IONEX format. It interpolates the TEC values in time and in space to the point at which the line-of-sight from each antenna to the source passes through the assumed nominal height of the ionosphere (450 km). The ionosphere is assumed to be infinitely thin. The interpolation is done in a coordinate system in which the sun direction remains fixed. Due to the large day-night differences, the ionosphere is much more constant in this frame than in one that rotates with the Earth. Note that TECOR requires an IONEX file that covers the full time range of the observations. Unfortunately the files from CDDIS start at 1:00 and go to 23:00 UT. Then another file is used for the next day. If the observations extend beyond 23:00, or start before 1:00, it is necessary to edit together the two IONEX files, which is laborious and a bit tricky. Hopefully it will be possible to get around this in the future.

In this memo, we report on comparisons of ionospheric results obtained from the GPS models and from S/X dual band observations. We treat the S/X measurements as "truth" since they directly measure the instantaneous ionosphere in the line-of-sight, subject to a constant offset that is different at each station, as mentioned above. We looked at two experiments. One was one of the regular geodetic projects from the geodetic community (RDV11). The other (TP015) was a specially designed test project in which 5 sources with reasonably high flux density and well determined astrometric positions were observed in a phase referencing style. The 5 sources were in a cluster with separations of roughly 3 to 10 degrees. The main object of the exercise was to determine if the GPS ionospheric models could provide useful improvements for phase referencing observations. Typically such observations are of target sources which are too weak to use the S/X scheme and/or are not at appropriate frequencies to measure the ionosphere directly.

In order to make the comparisons, it was necessary to derive the S/X ionospheric measurements. AIPS has some quirks that make this rather difficult. To make the corrections, we need separate multiband delays for each of S and X band. In a data set with both bands, the SN and CL tables are not structured to contain 2 multiband delays. If the bands are separated, putting the information back together is difficult. Geodetic observers deal with the ionospheric corrections outside of AIPS using SOLVE, which we could not access easily. Not wanting to struggle with the complexities of AIPS programming, we wrote a special purpose program outside of AIPS to do the derivation. That same program is able to make comparisons with GPS results and provide displays, a few examples of which will be shown below. The program is not intended for general use by the user community. In addition to the program that derives the S/X results, an additional program was written that plots the ionospheric delays, both raw, and referenced to some antenna and, ultimately to some source. This enables a direct comparison of the results from the various methods under circumstances comparable to how they would be used in phase referencing observations.

We also used the TP015 data to make phase referenced images of 4 of the target sources based on ionosphere corrected phases referenced to the fifth source. Those images were examined for quality and for astrometric accuracy. This process produced somewhat confusing results. Some are shown below, but it would seem that other effects, probably mainly the troposphere, were significantly affecting the data. For this reason, the conclusions of this memo are based mainly on the direct comparisons of measured ionospheric delays, rather than on the quality of phase referencing results.

2. The Observations.

RDV11 was observed on 1-2 October 1999 for 24 hours. This was a typical RDV run with 17 antennas delivering data. A total of 84 different sources were observed. The schedule was made with the automatic scheduling mode in SKED, which optimizes the sky coverage over each antenna. There are typically 2 or 3 subarrays in effect at any given time, although there are occasional scans with most or all antennas. The observations were fringe fitted in AIPS using separate files for S and X bands. Then the multiband delays were derived and sent to the special program described above. Only displays from that program will be shown here.

TP015 was observed on 31 March 1999 on the VLBA using the S/X system. The "target" sources were 0202+149, 0201+113, 0235+164, 0239+108 and the calibrator was 0229+131. All are taken from the USNO reference frame lists and have absolute positions with formal errors much smaller than 0.1 mas. All are easily strong enough to detect in short integrations in individual bands and to get reasonably high SNR fringe fit results. A cycle of 1.25m dwell time scans (typical total scan length including slew and setup of 1.4m) was used in which 0229+131 was observed every third scan, flanking either 0201+113 and 0202+149 or 0235+164 and 0239+108.

The geometry of the observations is shown in Figure 1. The top plot shows the relative positions and separations of the sources. Note that 0239+108 and 0235+164 are both about 7 degrees from 0229+131. This is somewhat farther than our usualy recommended maximum of 5 degrees for a phase reference source on the VLBA. The other two sources are about 3.5 degrees from the reference, which would be reasonable by our usual criteria. The bottom plot of Figure 1 shows the elevations vs. time for 0229+131 for each antenna. Most of the observations are at reasonably high elevations, although sources rise at MK during the run and get near set at SC at the end.

3. The Global Ionosphere Models

We found 5 different ionosphere models on the CDDIS site. We have looked quickly at all 5, but only display 2 in the attached figures to avoid clutter. It is clear that derivation of the ionosphere from GPS data is still not a precise art. The scatter in the results from the various models is on the order of 20-50%. The differences do not remain constant with time, although the ranking does typically last through the experiment — the same model will always give the highest result etc. We are not especially familiar with the modeling methods, but guess that the differences are the result of different assumed vertical profiles and perhaps different allotment of delay to instrumental effects and ionosphere. Also, the receiver networks used are not always the same. We presume that the modeling methods will improve significantly over the next few years, so the ability to use the models to correct VLBI data should improve. One important improvement that is being studied is to add a third dimension to the models.

The following brief descriptions of the models examined are based on, or extracted from, the header information in the files.

- JPL: Global Ionospheric Maps (GIM) are generated on an hourly and daily basis at JPL using data from up to 100 GPS sites of the IGS and others institutions. The vertical TEC is modeled in a solar-geomagnetic reference frame using bi-cubic splines on a spherical grid. A Kalman filter is used to solve simultaneously for instrumental biases and VTEC on the grid (as stochastic parameters). Contact Address: gpsiono@cobra.jpl.nasa.gov .

  There were 97 stations contributing to the data set we used.

- CODE: The global ionosphere maps are generated on a daily basis by the Center for Orbit Determination in Europe (CODE), University of Berne, Switzerland. The TEC is modeled with a spherical harmonic expansion up to degree 12 and order 8 referring to a solar-geomagnetic reference frame. The 12 2-hour sets of 149 ionosphere parameters per day are derived from GPS data of the global IGS (International GPS Service) network. Contact address: stefan.schaer@aiub.unibe.ch.

  There were 80 stations contributing to the data set we used.

- EMRG: The NRCan ionosphere maps are generated daily by the Geodetic Survey Division

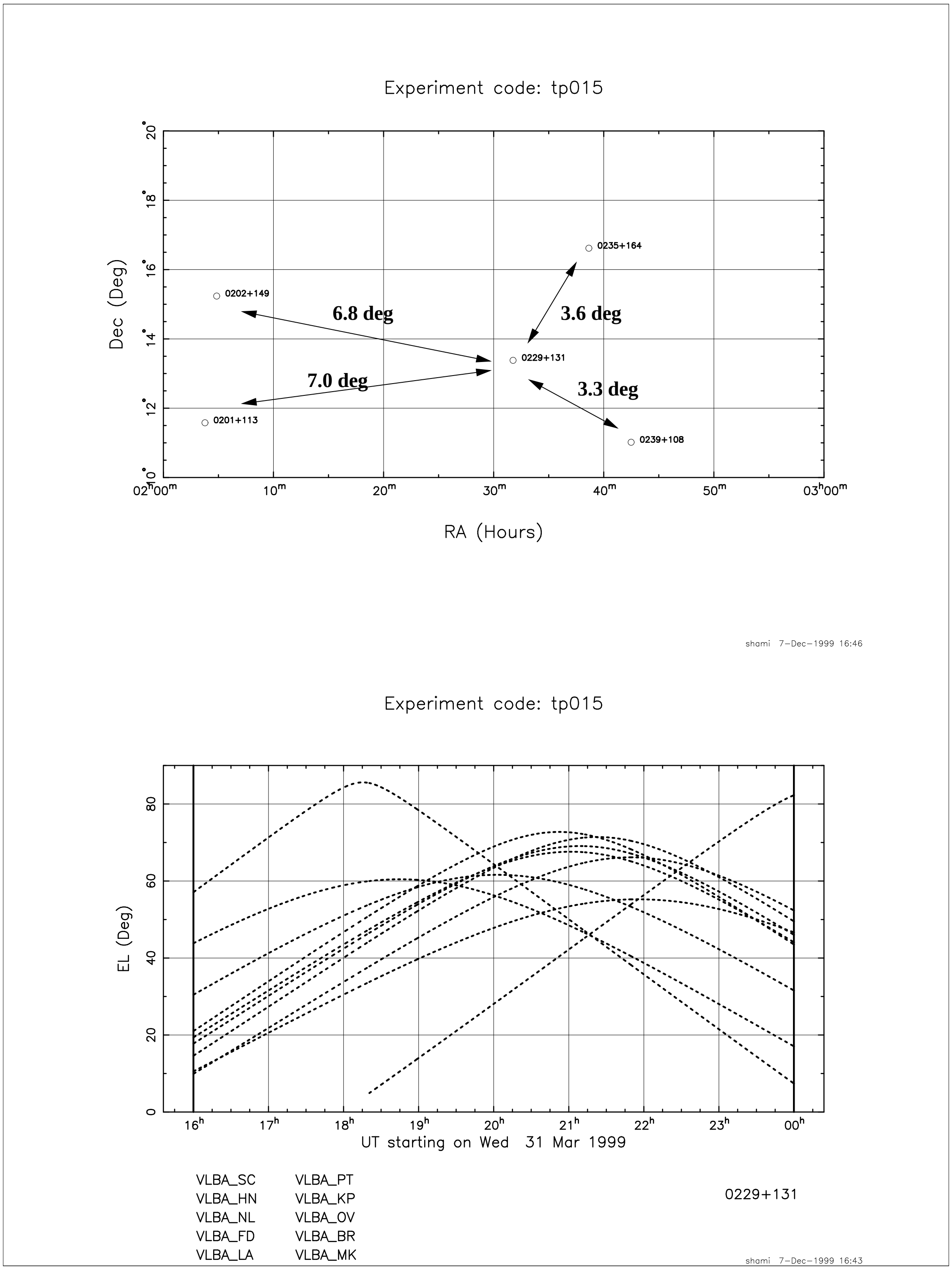

Fig. 1.— Top panel: Distribution of sources in TP015. Bottom panel: Elevation vs time for 0229+131, the central source in TP015.

(GSD) of Natural Resources Canada (NRCan). The grid point values are mean VTECs estimated in sun-fixed reference frame. Contact address: pheroux@nrcan.gc.ca.

There were 43 stations contributing to the data set we used.

- ESAG: The header information for this file is a bit confusing. But there is more information at http://nng.esoc.esa.de/gps/ionmon.html. The model is from the ESOC Ionosphere Monitoring Facility (IONMON). That facility can fit for the height of the ionosphere using an assumed profile. But the models archived at CDDIS appear to be for a fixed height. This model tends to give the lowest TEC values.

  There were 49 stations contributing to the data set we used.

- UPCG: Modeled independently for each station with a tomographic model: 72x9x2 cells in local time, latitude and height The height boundaries are: 59-739-1419 km. The estimates have been interpolated with radial basis functions (spatial smooth: 2 pixels) See for instance: Hernandez-Pajares, Juan, Sanz and Sole, JGR, Vol.103, N.A9, 20789-20796, 1998 Contact e-mail: manuel@mat.upc.es Plots at: http://maite152.upc.es/~ionex/gAGE_dip/gAGE_dip.html

  There were 83 stations contributing to the data set we used.

Note that the SATLOC models are based on approximately 15 receivers, which is much smaller than the above models, but they only attempt to cover the continental US rather than the whole globe.

4. Corrected Phases in TP015

The results shown in Figure 2 show the effects of using some of the ionosphere determinations to calibrate 2.3 GHz data from TP015. The first panel shows the uncorrected raw phases from the VLBA correlator for 5 representative baselines to LA. Considerable phase variation is seen. The correlator uses a very accurate geometric model (CALC 9.0 for TP015 - since upgraded to 9.1) and we give it source positions good to about 0.1 mas and station positions good to around a centimeter. Thus pure geometric errors should amount to a small fraction of a turn of phase. On top of this, there may be errors in Earth orientation (including UT1, pole position, nutation) at the level of around a mas and the troposphere model could be in error by a few cm, increasing at low elevation. Finally the clocks are only modeled with a linear rate so any fluctuations of higher order will show up in the phases. But all of these effects should probably add up to a turn of phase or less across the day. Instead we see many turns.

The second panel of Figure 2 shows the phases corrected using the S/X ionospheric measurements. Now the phase variations indeed meet our expectations and it is clear that a good job has been done of removing the ionosphere. Note that the phases for all sources are shown in these plots. The short term scatter is the result of differences between the sources. It is much reduced if only one source is plotted. Note also that only one ionospheric value is determined per scan so, when there is a high fringe rate, the phase slope across the scan will still be seen in the corrected data. This is apparent in several places.

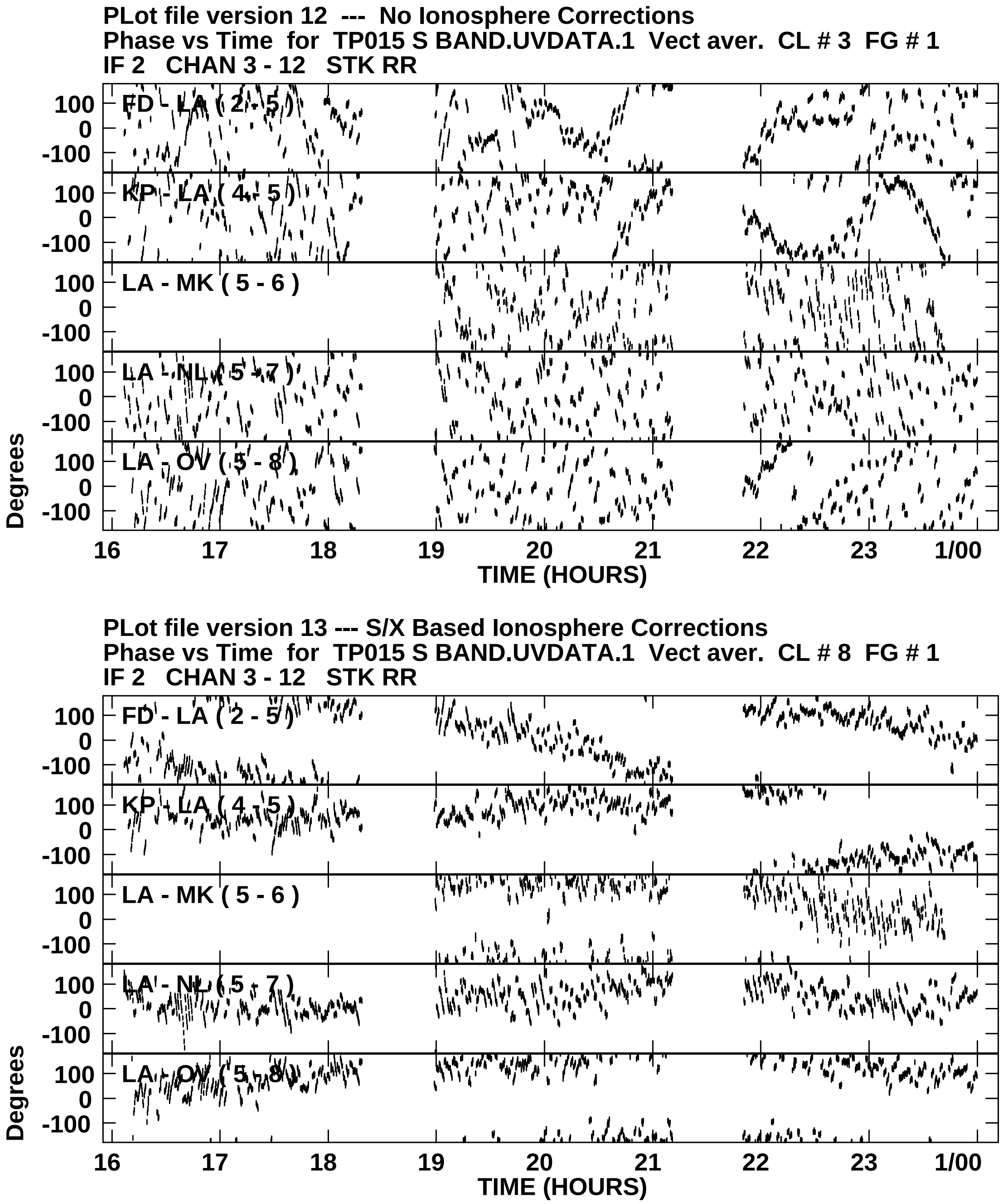

Fig. 2.— Top panel: Phase vs time for 5 baselines at 2.3 GHz from TP015. All sources shown. There have been no ionospheric corrections. Bottom panel: Phase vs time for the same data as the top panel, but using phase corrections for the ionosphere derived from the S and X band multiband delays.

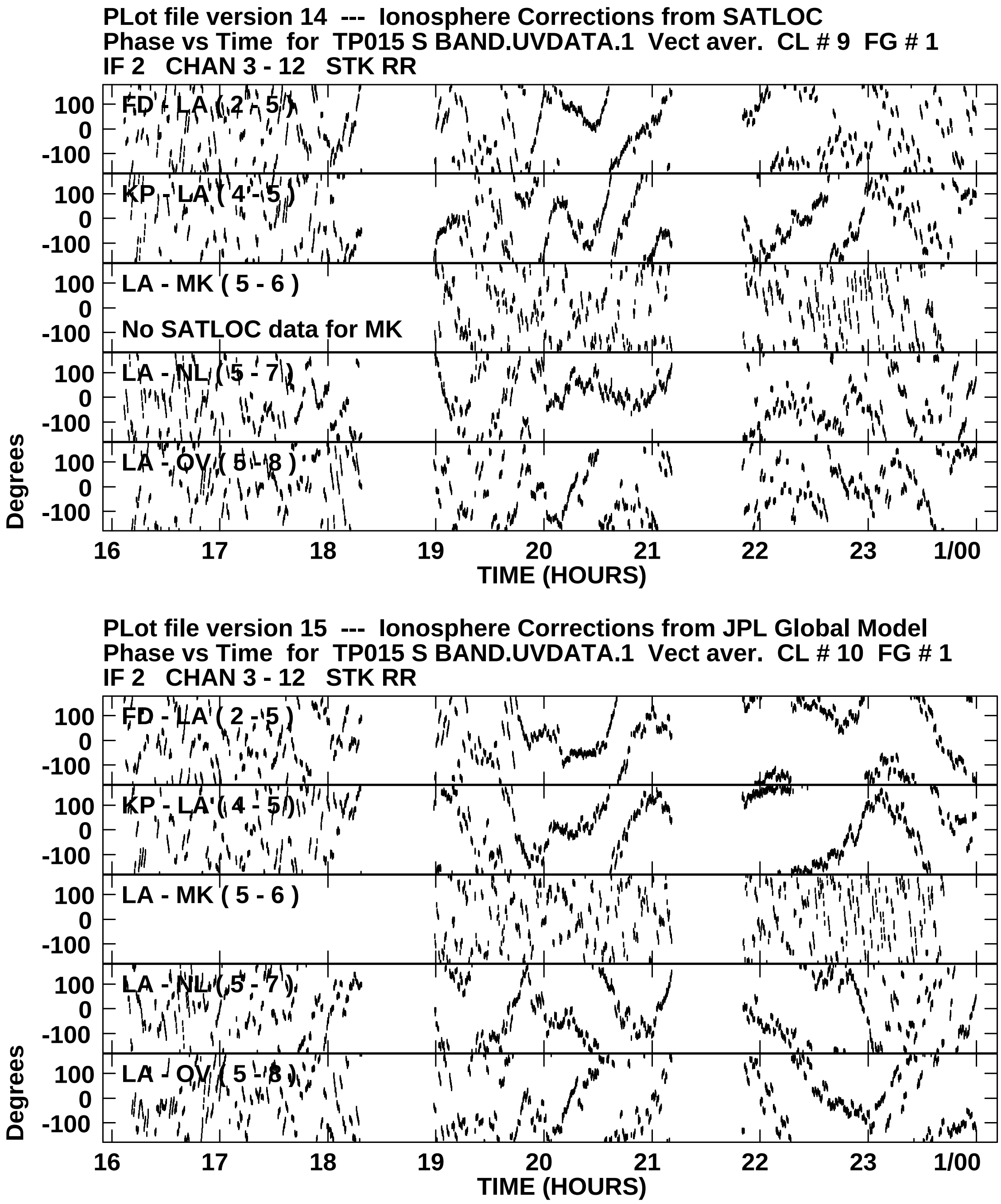

Fig. 3.— Top panel: Phase vs time for for the same data as Figure 2, but with inoospheric corrections derived from the SATLOC model. Bottom panel: Phase vs time with phase corrections derived from the JPL Global Model.

The results of calibrating the phases with the SATLOC regional model and with the JPL global model are shown in Figure 3 in the top and bottom panels respectively. Neither has flattened the phase slopes anywhere near as effectively as the S/X results. But the phase slopes have been reduced and, most importantly, the differences between sources have been significantly reduced. Note that the SATLOC model does not cover MK so that station is not corrected in the upper panel. MK corrections are derived based on the JPL model, but there are significant offsets between the S/X results and the model results for this station.

The phase plots demonstrate that the ionosphere does indeed cause large phase effects and that the S/X data can be used to remove them rather effectively. They also demonstrate that the GPS models improve the phases, including the relative phases between sources, but the corrections are not as good as those derived from the S/X data.

5. Direct Comparison of Delays in TP015

Examination of the phase plots above, or images discussed below did not clearly demonstrate the effects of the use of the GPS models. So a program was written to directly compare the different ionospheric delays in 3 different ways. Figure 4 shows one of the stations for which the models give the best corrections — NL. The upper left panel shows the total dispersive delay predicted by the JPL, CODE, and SATLOC models, adjusted to apply to 2.3 GHz. The CODE model truncates at 23 hours — this was to avoid the problem with editing the IONEX files mentioned earlier. The effort was made to edit the JPL model files and the SATLOC model was derived in a different way so this was not a problem. The other panels are one per source.

The lower plot in each of the other panels is the data for this station with the interpolated data for LA, for the same source, subtracted. For this plot, the variations between the S/X data and the models should track if the models are good. But there is still an arbitrary constant offset because the possible offsets in the VLBI data, described earlier, are station dependent and are therefore not removed by the referencing. The S/X data have been shifted so that the average matches the average in the models, but still any constant offset should not be a concern. As can be seen, the slopes are generally correct, but none of the models track the S/X data exactly.

The upper plot in each panel is the data from the lower plot, but now with the values for the reference source removed. This is what would actually be used in a phase referencing observation. Now the constant terms, which are the same for all sources, have been removed by the referencing and the data from the S/X and the models should match if the corrections are good.

Figure 5 and Figure 6 show similar plots for OV and MK. At OV, the models deviate from the S/X results in some fairly clear ways. At MK, the SATLOC model provides no coverage and the other two have even more significant deviations from the S/X data than at OV. But in all cases, it is clear that correcting the data using the models would be better than making no corrections at all. A rough impression from examining the plots, plus examining the RDV11 data, is that the models will provide a factor of about 2 to 5 improvement over not using any ionospheric correction.

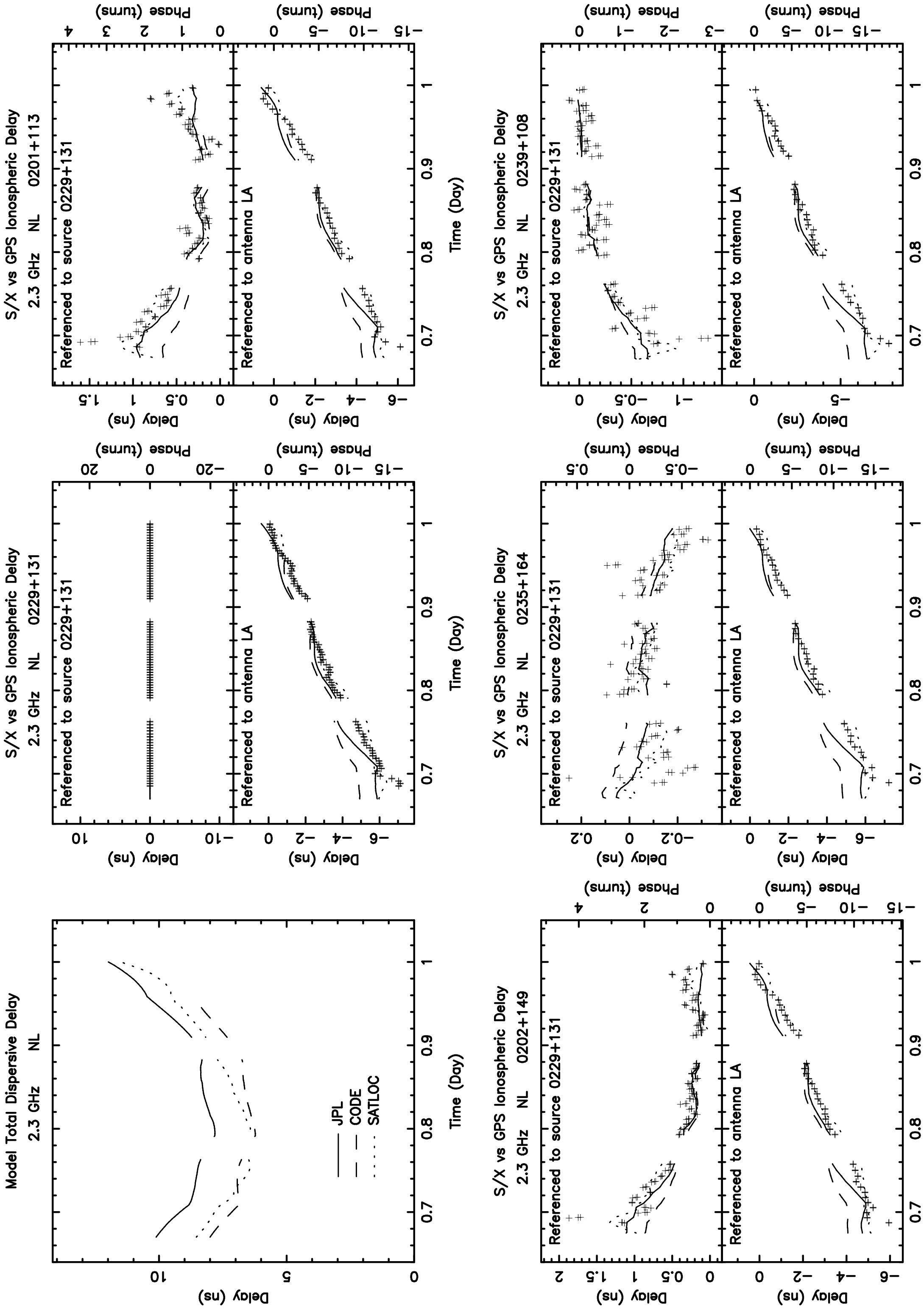

Fig. 4.— Comparison of S/X delays with ionospheric models for data from NL. The + signs indicate the S/X measurements. See the text for a detailed description.

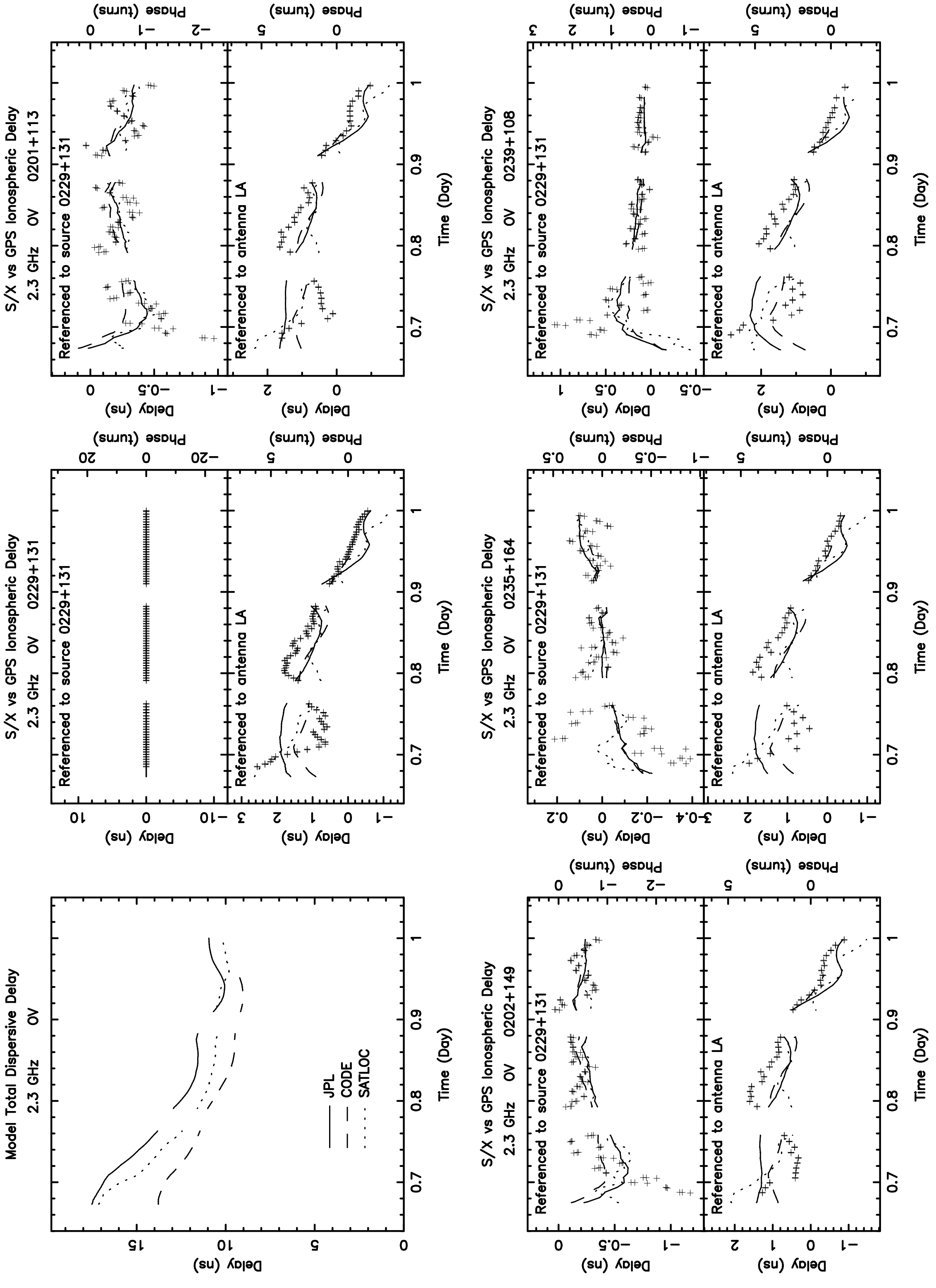

Fig. 5.— Comparison of S/X delays with ionospheric models for data from OV. The + signs indicate the S/X measurements. See the text for a detailed description.

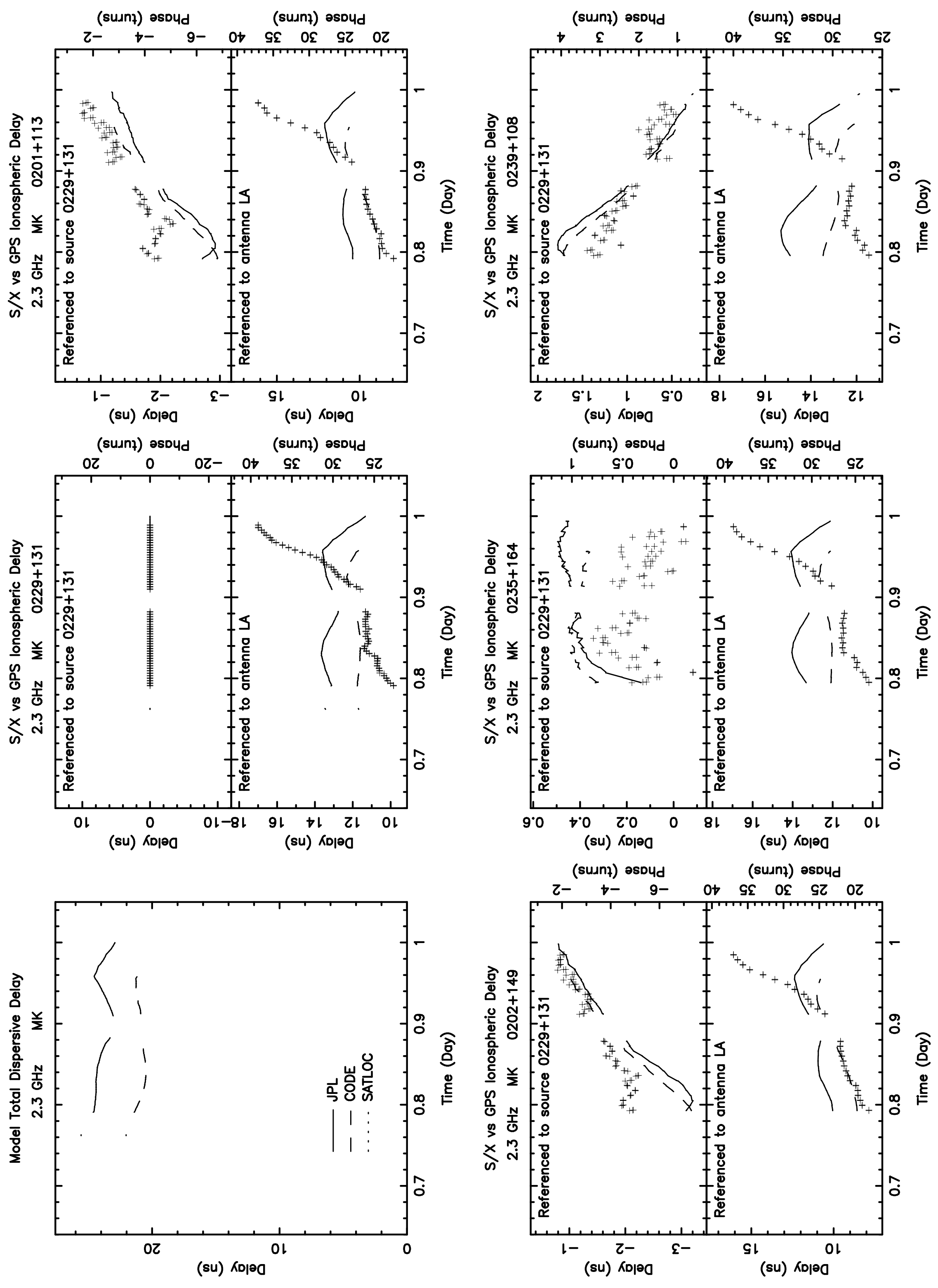

Fig. 6.— Comparison of S/X delays with ionospheric models for data from MK. The + signs indicate the S/X measurements. See the text for a detailed description. Note that the SATLOC data does not cover MK.

6. RDV11 Results

Figure 7 shows data from 4 baselines from the RDV11 global geodesy observation (the first 4 by alphabetical order). For each baseline, three panels are shown. The bottom panel shows the fitted multi-band delays for S and X bands. The middle panel shows the ionospheric and non-dispersive ("Geo") delay derived from the S and X band delays. Also the ionospheric delay derived from a global model, the JPL model in this case, is shown ("GPS Delay"). The top panel is the difference between the S/X and GPS model ionospheric delays. Information printed within the second panel gives the average delays and scatter for the S/X and GPS ionospheric delays. Similar information in the top panel gives the average value and scatter for the S/X - GPS difference. Note that there is an arbitrary DC offset because of the unknown offset between S and X band receivers in the VLBI data, as described earlier. Pay closest attention to the scatters. Note that the stations in the figure are the Brewster WA VLBA site (BR), the Fort Davis TX VLBA site (FD), the Gill Creek AK geodetic site (GC — near Fairbanks Alaska), The Hancock NH VLBA site (HN), and the Kokee Park HI geodetic site (KK — on Kauai in Hawaii). The apparent large, short term, scatter in raw data is a result of scheduling short scans on sources all over the sky. The scatter represents the differences along the line-of-sight for observations at different elevations and azimuths.

There are a few things to notice. First, the ionospheric delay is clearly much larger at S band than at X band, as expected. Delays of 10-20 ns are common. This is much larger than any other source of variable delay offsets at that band. There is a strong diurnal effect, also as expected. There is a strong latitude dependent effect, which is expected, but perhaps not widely appreciated. This can be seen by contrasting the BR-GC baseline (both high latitude) results with the BR-KK baseline (one low latitude station). It is also clearly seen in the GPS model data.

The GPS model does not give perfect correction of the data, but is much better than making no correction. Typically the scatter in the S/X-GPS difference is 2 to 5 times smaller than the scatter in the raw S band delays or in the S/X derived ionospheric delays.

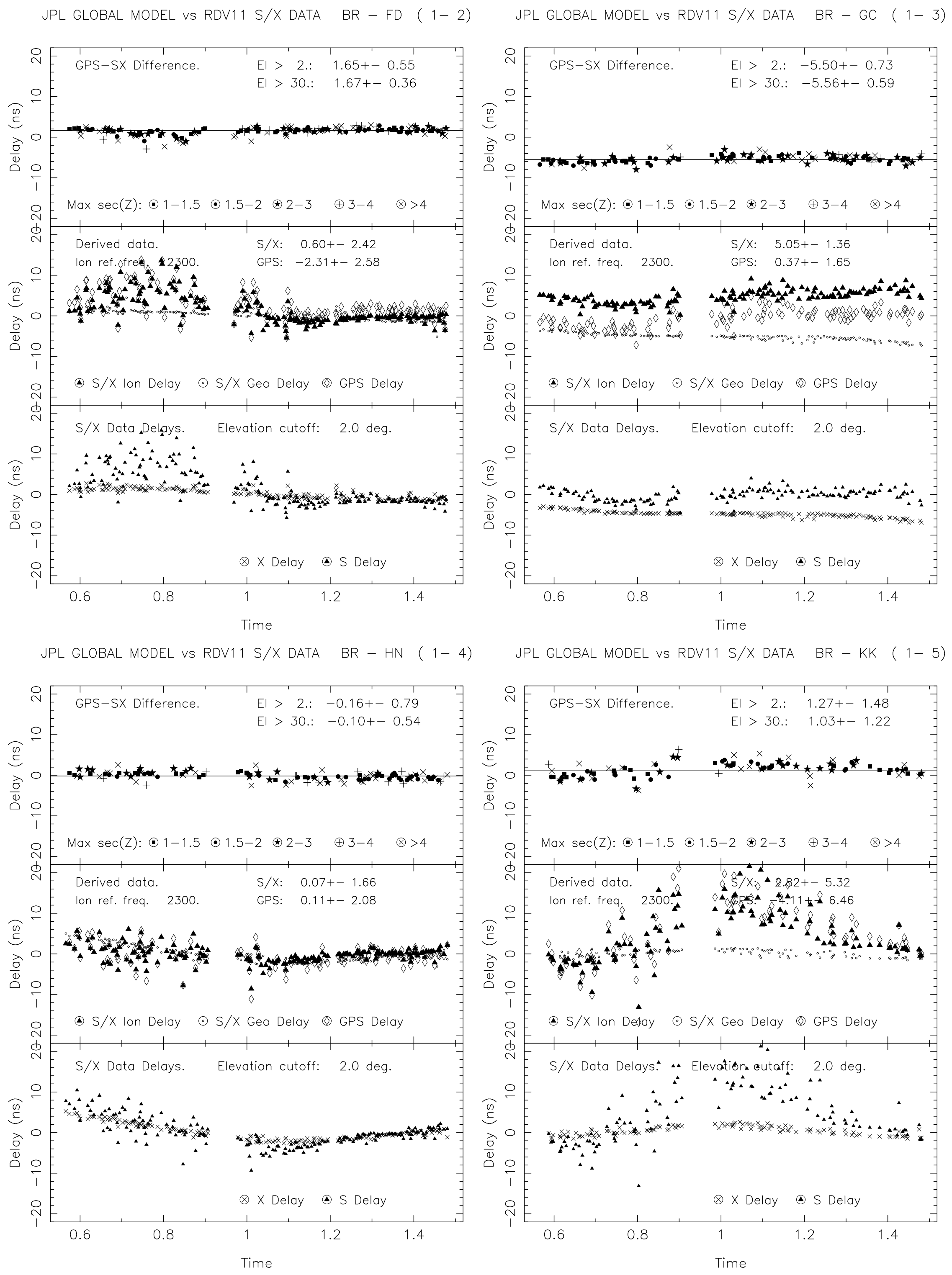

Fig. 7.— Delays at S and X bands, derived ionospheric and geometric delays along with the JPL model delays, and the S/X-JPL difference in ionospheric delay for 4 baselines in RDV11. The large scatter is the result of a schedule that rapidly moves between sources all over the sky.

7. TP015 Phase Referenced Imaging Results

We have used the ionosphere corrected data to make phase referenced images of the "target" sources in TP015. Figures 8 and 9 show such images for two of the sources at S band. Four images are shown of each source, differing only in the type of ionospheric calibration applied, as labeled above the plots. All images were made with IMAGR using a robustness of 4, which is close to natural weighting. It is clear that all of the ionospheric calibrations are better than no calibration. It is also clear that the S/X based calibration is best.

Often the goal of phase referenced observations is to measure positions. Figures 10 and 11 show positions derived for the source in each of the phase referenced images at S and X band respectively (the X band images were made with robustness of 1). The cross is the phase center, at the location of the position from the geodetic catalog. Recall that these positions have formal errors of less than 0.1 mas based on massive fits to many geodesy observations. The derived positions from the images are based on least squares fits of a gaussian to the peak in the image (JMFIT in AIPS). The relative flux density of each fitted gaussian is indicated by the size of the circle marking the source location. In the S band plots, the points marked with "NON" are based on data with no ionospheric calibration. In one case (0201+113), the image with no ionospheric calibration did not have a single peak that could be identified as the source. The "SX" points are based on the S/X calibration excluding the 2 stations (SC and MK) for which SATLOC data was unavailable or was unreliable, and the "SXA" data includes all stations. The "JP" and "JPL" points are the 8 and 10 station results with the JPL global model used for calibration. The "GPS" points are based on SATLOC calibration and 8 station data. In the X band plot, "NO" is the point with no calibration, "SX" and "JP" are the 10 station SX and JPL calibrated points and the "SA8" is the 8 station result using SATLOC calibration.

In most cases, the uncalibrated images give worse positions than the calibrated images. The recovered flux density is also lower. In one case, the phase referencing did not work at all without ionospheric calibration. In no case was the phase referencing perfect. The imaging results are subject to other model errors besides the ionosphere. The dominant such errors are likely to be from the troposphere which we have made no effort to remove. One reason that the geodetic positions are likely to be much better than our phase referenced positions is that serious efforts to remove tropospheric delay offsets and all other significant effects (EOP, clocks ...) are part of the usual geodetic data processing.

One somewhat disturbing feature of the position measurement results is that the spread of positions is much larger at S band than at X band. This is likely related to the larger beam size at S band and such an effect is reasonable when there are significant uncorrected errors. But the S/X results should have the ionosphere removed quite effectively and the troposphere should affect both S and X band the same in terms of delay, which is what matters for a position measurement. A frequency dependent spread might be expected if the longest baselines are sufficiently poorly calibrated that the source could be anywhere over roughly a beam area. But we don't believe we are in that regime.

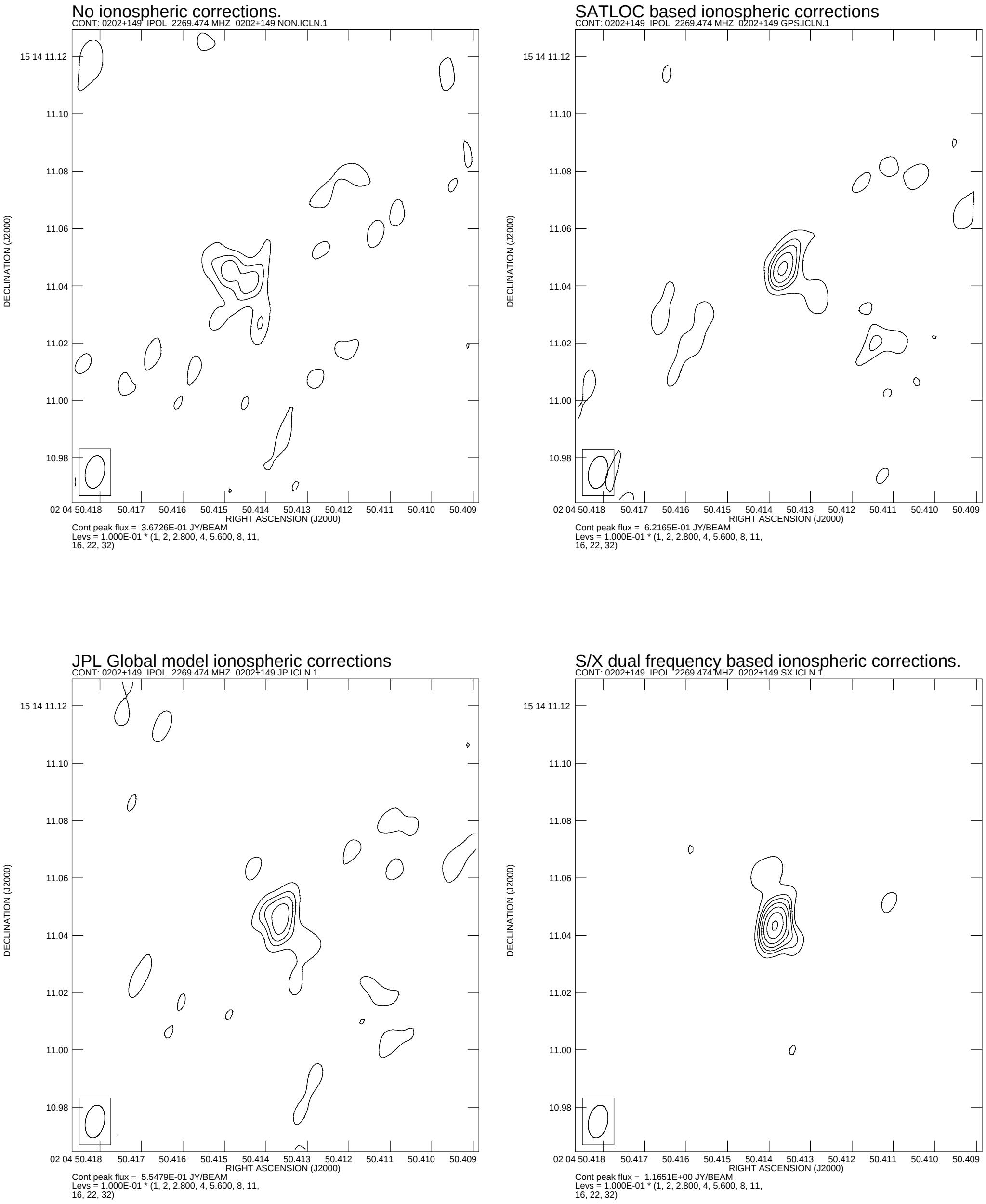

Fig. 8.— Images of 0202+149 made at S band with different ionospheric phase corrections, as labeled.

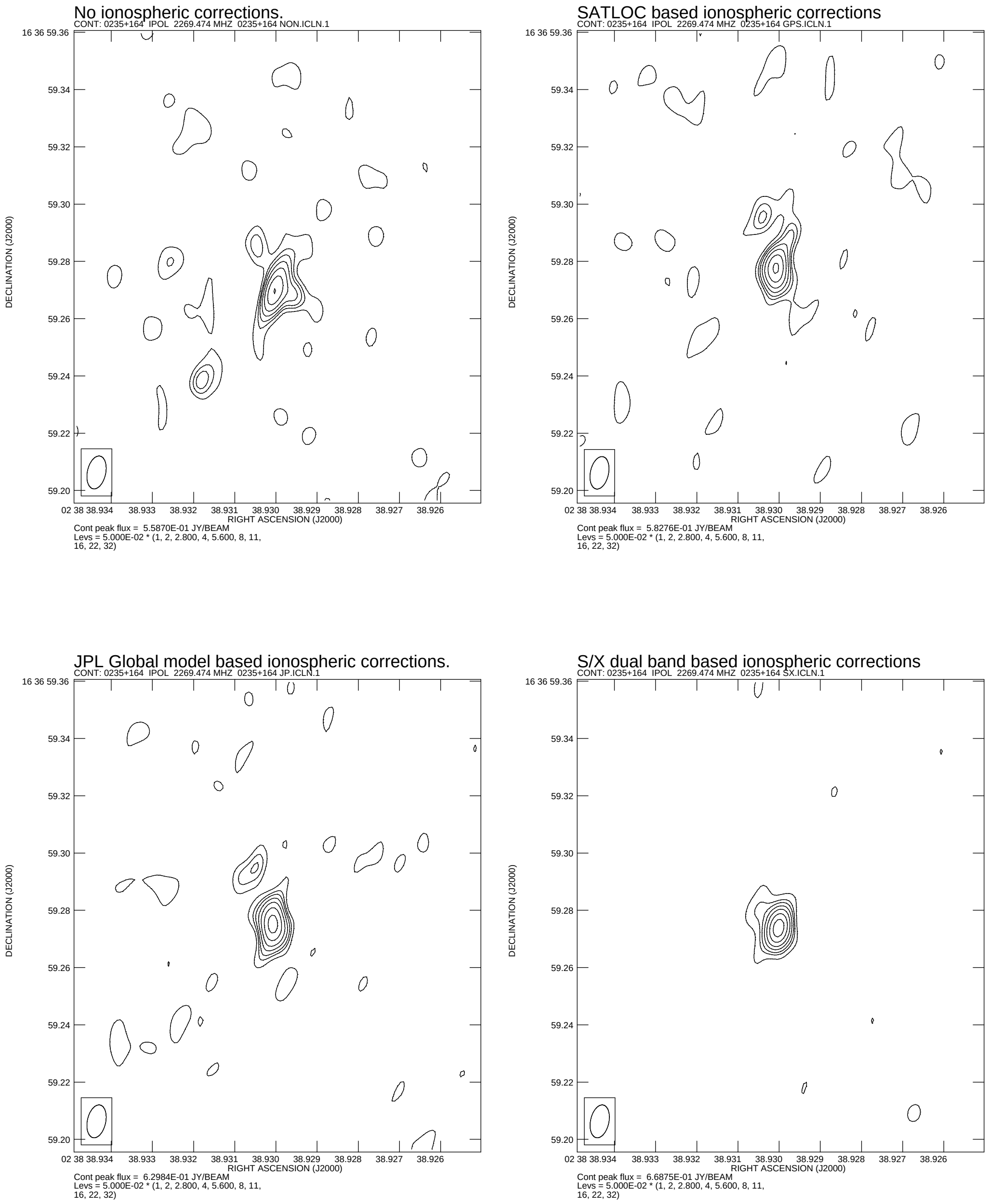

Fig. 9.— Images of 0235+164 made at S band with different ionospheric phase corrections, as labeled.

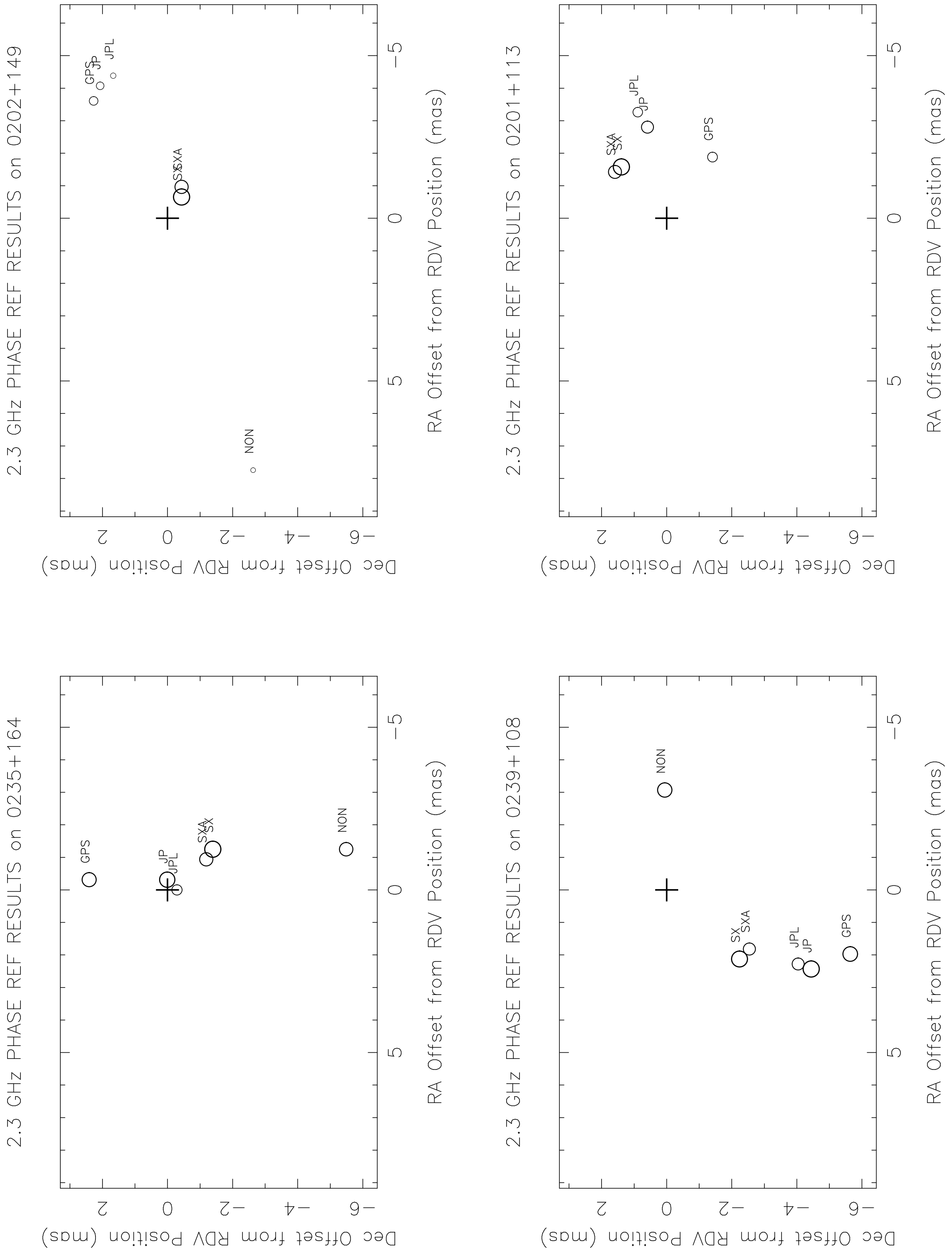

Fig. 10.— Positions of sources measured with least squares fits to image peaks in S band phase referenced images made with various ionospheric corrections. See the text for details.

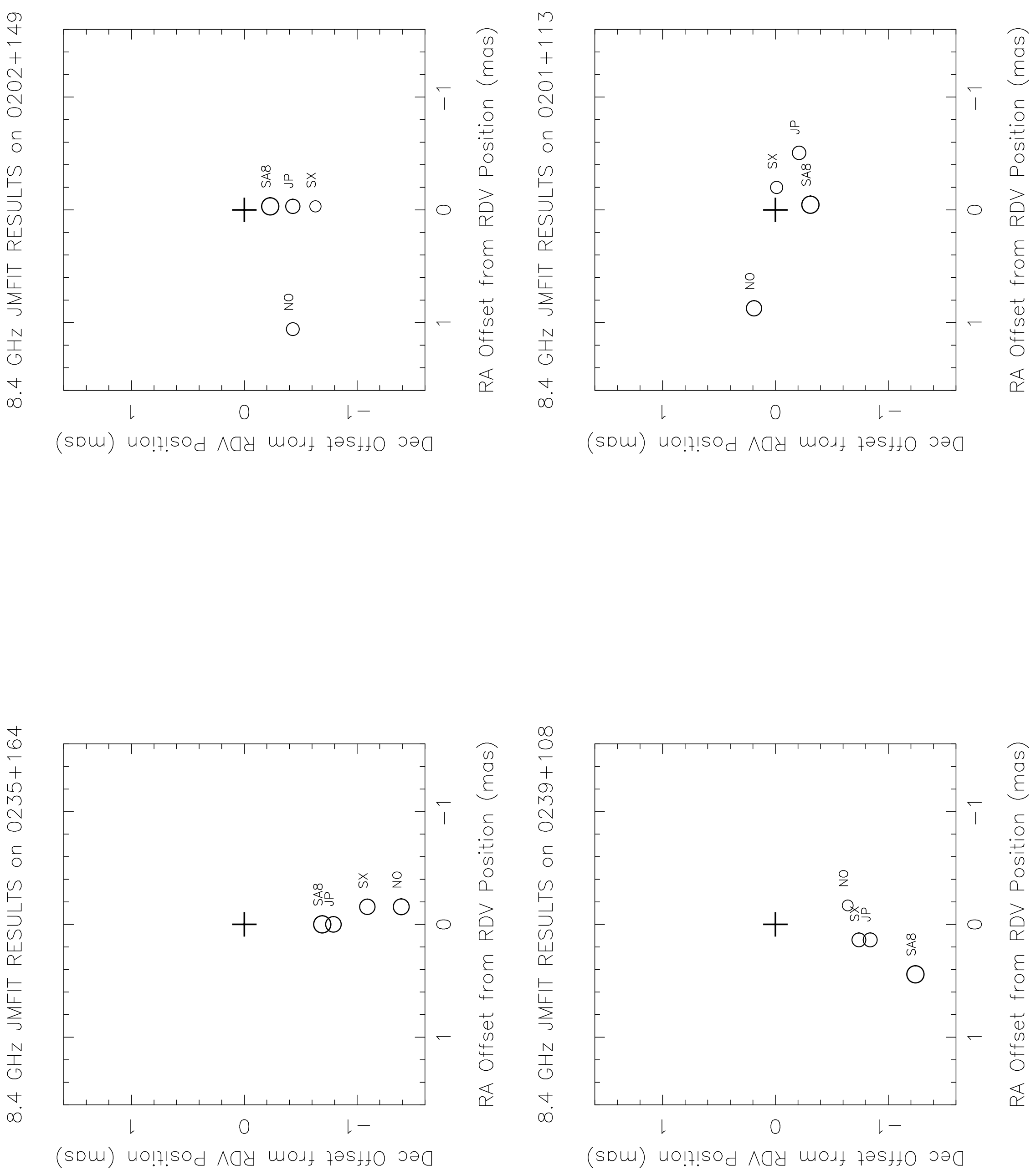

Fig. 11.— Positions of sources measured with least squares fits to image peaks in X band phase referenced images made with various ionospheric corrections. See the text for details.

8. Conclusions

We have presented results from two observations in which we were able to compare S/X derived ionospheric corrections with corrections based on GPS data. We come to the following conclusions:

- If you want to remove the ionosphere accurately, you should observe strong sources and design the experiment with enough spread in frequency to measure the ionosphere. Dual band S/X observations are the traditional way to do this, but other pairs (requiring new hardware), and even just multiple IF's at L band and below, could be used. Such a correction scheme cannot be used on weak sources, which eliminates its use for a majority of phase referencing observations.
- The GPS models provide useful ionospheric corrections. Phase offsets due to the ionosphere, which dominate at low frequencies, can be reduced by a factor of roughly 2 to 5 using such models, but cannot be corrected perfectly.
- Ionospheric modeling is in a sufficiently primitive state that differences between models from different analysis centers are commonly 20-50%. Model errors at this level are consistent with the quality of phase improvements seen in the VLBI data.
- The SATLOC models do not appear to have any advantage in accuracy over the global models despite the finer grid and higher time density. The global models are easier to obtain and cover all stations and so are better for our purposes.

Based on the above conclusions, we recommend:

- We should NOT make the gathering and archiving of SATLOC models an observatory priority. It may be worth keeping the receiver functional for any users who might wish to gather the data. But this can be on an as-needed basis.
- Ionospheric corrections based on the global models should become a common part of VLBI calibration, especially at low frequencies and perhaps at all frequencies. Since these models are only provided on time scales long compared with typical coherence times, they should not degrade data, and they are likely to help.

Gathering the models from CDDIS and applying them is rather easy — an approximately 10 minute job if it is not necessary to edit together multiple files. Therefore, providing the data as a service to the users is not a high priority, but would be a convenience. We need to discuss whether to apply the models in the correlator, pass them in calibration transfer, or do nothing at the correlator. Given that the user might wish to use a different model from whatever is passed, we recommend the calibration transfer option. The model would be applied as part of post-processing. This option may require a new AIPS table to hold the model because currently the information is only included in CL tables which are often destroyed and rebuilt for VLBI data (even version 1). In any case, one small but high priority task is eliminate the need to edit IONEX files for observations near the day boundary.

It is likely that the GPS ionospheric models will improve with time. One significant improvement for our application is likely to come from 3D modeling. We need to watch for this to become available and be prepared to have our software accept such models. Related to this, TECOR currently assumes an infinitely thin ionosphere. A possible future development would be to use a more realistic profile. This might allow better results to be derived even from the current 2D models, let alone from future 3D models.

We have not studied the use of data from the individual receivers at each site. Such data would have the advantage that it is subject to the same height effects as the VLBI data and so could potentially provide better corrections. But there are problems with the data, or at least our use of it. Our first look had results clearly in error by an order of magnitude. The problems may lie in the area of corrections for the offsets between the channels on the satellites and in the receivers, but that is not certain. It could be something as simple as a format conversion problem. Considerable effort will be required to learn how to use such data effectively. That is beyond the scope of this memo but is an approprate project for future work.