Anderson, J., Xu, M. (2018): Source Structure and Measurement Noise Are as Important as All Other Residual Sources in Geodetic VLBI Combined. - *Journal of Geophysical Research*, *123*, 11, pp. 10,162—10,190.

DOI: http://doi.org/10.1029/2018JB015550

Journal of Geophysical Research: Solid Earth

RESEARCH ARTICLE
10.1029/2018JB015550

Source Structure and Measurement Noise Are as Important as All Other Residual Sources in Geodetic VLBI Combined

James M. Anderson[1] **and Ming H. Xu**[1,2]

[1]Helmholtz Centre Potsdam, GFZ German Research Centre for Geosciences, Potsdam, Germany, [2]Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai, China

Key Points:
- Source-structure errors in geodetic VLBI were determined from actual IVS observations, not simulations
- Source-structure and measurement noise were measured using closure relationships
- Source-structure errors in CONT14 are approximately as large as all station-based errors combined

Supporting Information:
- Supporting Information S1

Correspondence to:
J. M. Anderson and M. H. Xu,
anderson@tu-berlin.de;
mhxu@shao.ac.cn

Citation:
Anderson, J. M., & Xu, M. H. (2018). Source structure and measurement noise are as important as all other residual sources in geodetic VLBI combined. *Journal of Geophysical Research: Solid Earth*, *123*, 10,162–10,190. https://doi.org/10.1029/2018JB015550

Received 24 JAN 2018
Accepted 17 OCT 2018
Accepted article online 23 OCT 2018
Published online 9 NOV 2018

Abstract The celestial sources observed by geodetic very-long-baseline interferometry (VLBI), most of which are quasars, are resolved—they have a finite angular size as measured by the interferometer—yet modern geodetic VLBI analysis techniques still treat these sources as if they are point-like. The errors introduced by source structure, the angular distribution of emission, have traditionally been believed to be a small fraction of the total measurement error budget. We analyzed the VLBI data of the CONT14 continuous observing campaign using standard geodetic and imaging techniques, along with an analysis of closures, combinations of interferometric measurements where the error contributions from station-based effects such as tropospheric and ionospheric errors, clock and cable length errors, station position errors, and so on, cancel exactly, leaving only nonclosing errors, principally source structure and measurement noise. Our statistical analysis of the closure information, based on actual VLBI observations revealed that *nearly half* (40%) of the total geodetic VLBI error budget (in terms of variance) in CONT14 came from source-structure effects, nearly the same variance as all station-based error sources (43%). The remaining error due to other nonsource-structure, nonclosing errors is also significant (17%). Previous studies, which were primarily simulation-based, did not account for the full range of effects of source structure on real measurements. Source structure *is* a major contributor to errors in geodetic VLBI, and source structure must be taken into account in the entire VLBI operational chain, from scheduling to analysis, in order to mitigate its effects on astrometric and geodetic measurements.

1. Introduction

Geodetic very-long-baseline interferometry (VLBI) is a space-geodetic technique that has been utilized by geodecists since the 1970s to measure many fundamental parameters of the Earth system, including precise coordinate tracking of sites on the Earth, the study of Earth rotation, and various other terrestrial parameters (Schuh & Behrend, 2012), as well as measuring the precise locations of special distant radio sources (typically quasars) to determine the celestial reference frame, the current version of which is the second realization of the International Celestial Reference Frame (ICRF2; Fey et al., 2015). Geodetic VLBI observations and analyses are coordinated by the International VLBI Service for Geodesy and Astrometry (IVS) as a service of the International Association of Geodesy and the International Astronomical Union.

A technical description of the geodetic VLBI principle can be found in Sovers et al. (1998). In short, coordinated observations of distant celestial sources are made by a network of globally distributed VLBI stations, with observations organized in large blocks of observing time called sessions (also known as experiments). Currently, standard IVS "rapid" observing sessions designed to measure station positions and Earth orientation parameters are performed twice per week for 24 hr each, with additional "intensive" sessions performed nearly daily for durations of typically 1 hr with smaller station networks to principally measure the difference between UT1 and UTC. Within each session, many geodetic celestial sources are observed all over the sky in a rapid sequence. The raw station data are initially processed at an IVS VLBI correlator facility, where the subsequent visibility data are used to determine group delays and rates, for each individual observation of a celestial source, for each baseline that observed that source. These data are then made available to IVS analysis centers, where the group delays and rates are analyzed to determine parameters of geodetic and astrometric interest (station and source coordinates, Earth orientation parameters, EOPs, and so on) based on the geometric orientations of the station baselines and the celestial sources.

There are many natural and instrumental processes that corrupt the delay and rate measurements, and standard geodetic-VLBI analysis packages attempt to account and correct for many of these processes, including ionospheric and tropospheric delays, station clock errors, changes in electronics and cable delays, tidal effects, thermal deformation of the antennas and other station position offsets, EOPs, and, of course, measurement noise. However, despite the fact that virtually all celestial sources observed by geodetic VLBI have structure, that is, they are resolved on angular scales to which the interferometers used for geodesy are sensitive, VLBI analysis traditionally treats sources as if they are point sources.

Current estimates of the accuracy of VLBI-derived station positions in the International Terrestrial Reference Frame are about 5 mm in each direction (Schlüter & Behrend, 2007). As the VLBI counterpart of the Global Geodetic Observing System of the International Association of Geodesy, the VLBI Global Observing System (VGOS; formerly known as VLBI2010) goals, among other enhancements, include improving the terrestrial frame position accuracies to 1 mm and velocity accuracies to 0.1 mm/year (see, e.g., Petrachenko et al., 2009; Plag & Pearlman, 2009). This ambitious goal of improving the VLBI-derived parameters by a factor of 5 encourages us to obtain a better understanding of the errors affecting VLBI measurements.

The simulations of VLBI observations affected by typical random error sources used to guide the design aspects of VGOS included various atmospheric, clock, and measurement noise levels but did not specifically investigate source structure (MacMillan & Sharma, 2008; Pany et al., 2008; Petrachenko et al., 2009; Wresnik et al., 2008). These simulations concluded that atmospheric errors are the dominant random error source for geodetic VLBI, and more recent simulation works have only included these error sources (e.g., MacMillan, 2017). Indeed, the belief or conclusion that the atmosphere is the dominant error source in modern geodetic VLBI is quite common (Halsig et al., 2016; Heinkelmann et al., 2013; Nilsson & Haas, 2010; Pany et al., 2011; Petrachenko et al., 2009).

In contrast, the effect of source structure on geodetic VLBI measurements, while understood to be significant, has been thought to be substantially smaller than other major error terms. As demonstrated by imaging surveys, such as Fey et al. (1996) and Fey and Charlot (1997), practically all geodetic VLBI sources are resolved at some level. Using a model of the effect of source structure on group delays from Charlot (1990), Fey and Charlot (1997) used CLEAN component models from observations of geodetic sources using the Very Long Baseline Array (VLBA) to determine *structure indices* for simulated observations of sources from all possible Earth-based baselines, with larger structure indices indicating that a source would probably contribute larger structure-related errors to geodetic observations. Hence, having a small structure index has been one of the criteria for a source to be a defining source for the ICRF (Fey et al., 2015; Ma et al., 1998), and the source lists used for scheduling modern geodetic VLBI experiments have been selected with at attempt to minimize source structure. The recommendation from Fey and Charlot (1997) to not use sources with structure indices of 3 or larger effectively limits the expected median absolute delay error introduced by source structure to be less than 10 ps. By attempting to limit the sources observed to have minimal structure, it is expected that source structure should not be the dominant error source. As an example, simulations by Shabala et al. (2015) using a modern source list and various realistic structure indices showed that source structure errors contribute at the level of only 10% of atmospheric and clock errors.

In this work we used actual geodetic VLBI observations—not simulations—to determine the actual levels of source structure and measurement noise in geodetic measurements and found that source structure and measurement errors, and their interaction with each other, contribute to the total error budget as much as all other error terms in modern geodetic VLBI combined. Our analysis utilized closure measurements, a special combination of interferometric measurements where standard station-based errors including atmospheric delays, clock errors, and station position errors identically cancel. In order to focus on our analysis of actual geodetic observations and the importance of source structure, we have placed our detailed discussion of closure quantities and measurement statistics in supporting information A and B (which reference Bertarini et al., 2011; Clark, 1981; Fomalont & Perley, 1999; Massi et al., 1991, 1996, 1997; Rogers, 1983; Thompson, 1984; Thompson et al., 2017; Thompson & D'Addario, 1982; Xu et al., 2017, 2016). In section 2, we present an overview of the geodetic and imaging data we analyzed for this initial exploration of the full impact of source structure on geodetic data. Our main results of the standard geodetic analysis (baseline residuals) in the context of source structure are presented in section 3. In section 4, we introduce our closure analysis for CONT14 measurements and our source-structure models and present some examples of closure measurements and source structure models to demonstrate how well the image-based models actually perform in predicting source

structure. Our closure statistics analysis, where we examine the source structure and measurement noise levels in our data and compare them to the remaining levels of station-based errors, is presented in section 5. The detailed results for individual sources are given in supporting information C for delay observables and in supporting information D for phase observables. Then, in section 6, we explain why our measurements of the impact of source structure show that source structure is actually more important than was previously believed. And finally, in section 7, we review our results and think ahead to the implications of our results and further areas of research needed for source structure. A list of symbols to assist the reader in following our discussion is located in supporting information E.

2. Data

2.1. Experiments

We analyzed the CONT14 data set of the IVS (IVS data are described by Nothnagel et al., 2015, see http://ivscc.gsfc.nasa.gov/program/cont14/ for a description of CONT14). CONT14 was a continuous observing campaign by 17 VLBI stations lasting 15 days from 2014-05-06T00:00:00Z to 2014-06-20T23:59:59Z. Observation scheduling was performed by the NASA SKED program (Gipson, 2010, see also http://lupus.gsfc.nasa.gov/software_sked.htm), one of the standard software packages used to schedule standard IVS experiments, using the same set of sources drawn from the standard geodetic VLBI source catalog for all 15 days of the experiment. The standard 2 bit, 512 Mb/s recording mode used for many modern IVS experiments was used, yielding measurement errors that are typical for modern IVS experiments. Except for the continuous 15 days of observing (the standard IVS-R1 and IVS-R4 sessions that comprise the majority of all IVS regular session take place on Monday and Thursday of each week, respectively, and only last for 24 hr, see http://ivscc.gsfc.nasa.gov/program/descrip2011.html) and the large number of stations (17) participating in the experiment (the median number of stations participating in R1 and R4 sessions in 2016 was 9), the CONT14 experiment is typical of modern geodetic VLBI experiments.

By observing the same set of sources continuously over 15 days with an unusually large number of stations, CONT14 uniquely allowed us to accurately determine the source structure of most of the 73 sources observed in the campaign. Thus, we could use the *same* measurements that were used to derive the geodetic "observables" to make images and, in turn, compare the imaging results to the geodetic analysis. We avoided problems of source structure changing with time or different source structure being observed because of differences in the arrays or differences in (u, v) coverage. The exact same source-structure information that affected the geodetic analysis was present in the imaging analysis, because the analyses were based on the same observations.

Standard dual-band geodetic observations were performed using S and X bands (~2.3 and ~8.4 GHz, respectively). The SKED software scheduled observation scan durations to achieve signal-to-noise ratios (SNRs) for each baseline in a scan of at least 20 for X band and 15 for S band, using estimates of the station sensitivities and simple source brightnesses (including a possible 2-D Gaussian estimate of source structure, although most sources were only estimated as point sources). Observations of individual sources were scheduled to stop independently for each station in a scan once all baselines including that station had reached their required SNRs in both bands. Hence, the data for all baselines for all scans should have SNRs that are roughly uniform within factors of several in the absence of source structure.

2.2. Geodetic Data and Analysis

Geodetic VLBI data sets are publicly available from several IVS data centers around the world (see http://ivscc.gsfc.nasa.gov/products-data/data.html for a list of download servers). For an overview of the standard data handling by the IVS, see Behrend (2013). The geodetic VLBI data provided by the IVS are reduced from visibilities output by the IVS correlator groups to so-called *geodetic observables*, with scalar quantities for group delay, delay rate, phase, and amplitude provided for each baseline of each scan of a VLBI experiment. The IVS data sets for CONT14 are provided in two data formats, a so-called binary *Goddard database format* and an ASCII *NGS card format* (CONT14 data sets in the vgosDB format are not yet publicly available). Data format descriptions can be found on the Internet. For example, the NGS data format can be reviewed at http://lacerta.gsfc.nasa.gov/mk5/help/dbngs_format.txt. Data sets for normal 24-hr experiments are typically only a few megabytes in size. The format precisions provided for the delay, delay rate, and phase measurements are quite high, but for amplitude only, *correlation coefficients* with typically one or two decimal digits of precision are provided by NGS data (although the formats also provide a fringe amplitude in janskys with

reasonable precision, the numerical values found in IVS data sets are typically always identically zero and are therefore useless).

We used the Goddard database data sets from the IVS for our own geodetic analysis using the Calc/Solve software, version 20170103, developed and maintained by NASA Goddard Space Flight Center (see http://lupus.gsfc.nasa.gov/software_calc_solve.htm). A general description of the Solve least-squares estimation algorithm can be found in Ma et al. (1990). By using Calc11, the a priori models for geophysical effects and precession/nutation complied with the IERS 2010 conventions (Petit & Luzum, 2010). Additional correction models included the atmospheric pressure loading model (Petrov & Boy, 2004) and the thermal expansion model of the radio telescopes (Nothnagel, 2009). The VMF1 functions (Boehm et al., 2006) were used for tropospheric mapping.

We used the highest available version (4 or 5) of the Goddard database format data for each day of CONT14 for our geodetic analysis. In these data sets, outlier flags have already been set up by the IVS analyzer who generated the databases and therefore are available for general use. The elevation cutoff angle was set to 5°. The 40 sources observed in CONT14 that are part of the ICRF2 (Fey et al., 2015) defining sources had their estimated positions referred to ICRF2 with a no-net-rotation constraint. No-net-rotation and no-net-translation constraints with respect to ITRF2014 (Altamimi et al., 2016) were also applied to 13 of the stations in the CONT14 network to realize the terrestrial reference frame.

Two kinds of parameterization were used in the solution: global parameters (based on observables from all 15 days combined), positions of the stations and radio sources; and local parameters (based on the data of each 24-hr session individually), Earth orientation parameters, including polar motion, UT1, and their time derivatives, the adjustments of two nutation parameters, offsets and rates of clock functions (1-hr segments), coefficients of the linear splines that model zenith wet delays (20-min segments), coefficients of the piecewise functions of the North and East tropospheric gradients (6-hr segments), and baseline-dependent clock offsets.

Standard data reweighting (see section 3.2) was applied during the analysis. The residuals and weights of all the observables, along with the baseline-based clock offsets, from the combined 15-day solution, as well as the outlier flags determined by the IVS analysts, were output from Calc/Solve for our further analysis here.

2.3. Visibility Data and Imaging

Using visibility data from the CONT14 experiment provided by the Bonn correlator (see ftp://ftp.mpifr-bonn.mpg.de/geoarchive/), we imaged all of the sources observed using a combination of AIPS Astronomial Image Processing System, (Bridle & Greisen, 1994; Greisen, 1990) and Python software using ParselTongue (Kettenis & Sipior, 2012; Kettenis et al., 2006). A detailed description of our imaging analysis is in preparation; here we only briefly describe our basic results. The CONT14 data were correlated with 1.28-s integration times and 0.25-MHz channel widths (yielding 32 channels per 8-MHz subband). A total of 16 subbands is present in the visibility data, 6 subbands for S band (∼2.3 GHz), and 10 subbands for X band (∼8.4 GHz). Standard radio interferometry calibration and analysis procedures were applied. A combination of manual and automated flagging was applied to remove radio frequency interference (RFI), dead subbands, and so on. The visibilities were calibrated using system temperature measurements made by the stations. Absolute gain levels were calibrated using the gain calibration values in the station logfiles for those stations providing such information—the remaining stations' gains were adjusted to match the flux densities measured by the calibrated stations.

Following geodetic VLBI tradition, only right-hand-circular polarization data were recorded for the CONT14 experiment, so no corrections for instrumental polarization leakage were performed. Bandpass calibration was applied to correct complex (amplitude and phase) gain variations across the individual subbands.

For the analysis presented here, we combined the calibrated visibilities from all 15 days of the CONT14 experiment for each of the 73 sources. The data were imaged and self-calibrated from CLEAN components starting from an initial point-source model using a combination of Python code and AIPS tasks in a completely automated fashion. Although better image fidelity could be achieved by operating the self-calibration loop manually, we preferred to use this automated approach to remove human biases and ensure that all sources were treated the same way. Images were separately produced using uniform, robust, and natural weighting for the entire self-calibration process. We used the results from uniformly weighted imaging in our analysis here, as uniform weighting provides the best angular resolution in order to study the impact of source structure on the predominantly long baselines used in geodetic VLBI.

Table 1
Group1 Sources

Source name	N_{obs}	N_{scan}
(1)	(2)	(3)
0016+731	10,162	389
0059+581	14,740	523
0727−115	7,284	596
1044+719	1,495	176
1424−418	5,016	635
1921−293	1,707	353

For the 6 out of 73 sources with no or few scans with at least three stations, traditional self-calibration was not possible, and we did not attempt to determine source-structure information from these sources. For the 67 sources that we could image, the median dynamic range (defined as the ratio of the peak flux density in the image to the root-mean-square [RMS] noise level far away from the center of the image) is above 1,000, more than sufficient to measure source structure that can affect the geodetic measurements. CLEAN components from the final, best-calibrated, images of each source were recorded for further use in our analysis.

3. Geodetic Residual Analysis

3.1. Source Subsets

For our analysis of CONT14 data, we present results for individual sources, for all sources combined (labeled "All"), and for a subset of six sources that show minimal source structure effects (labeled "Group1"). The Group1 subset of sources was analyzed to determine, as well as possible given the real sources observed, the performance of the geodetic measurements in the absence of source structure, in order to see how much source structure impacted the measurements. Our Group1 sources were selected from the analysis of the X band closure delays of the CONT14 experiment performed by Xu et al. (2017). From their Table 1, we took sources with $\tau^{rms}_{closure} \leq 10$ ps (showing little structure) and $N_{tri} \geq 1,000$ (enough observations to provide good closure statistics), yielding the six sources shown in our Table 1. Three of the sources are in the Northern Hemisphere and three of the sources are in the Southern Hemisphere, so that the Group1 combination includes observations by all stations participating in CONT14.

The azimuth and elevation angle distributions for All and Group1 sources are shown in Figure 1. Although the declination distribution of our six Group1 sources resulted in no observations toward the northwest and northeast directions, the two source sets *do* have approximately the same elevation angle distributions, including many observations below 30° elevation, and the azimuthal directions have approximately the same north to south observation ratio. Since tropospheric and ionospheric fluctuations do not depend on the structure of the source being observed, the resulting delay errors should be reasonably similar for the All and Group1 source groups.

The statistical analysis of the combined six sources in our Group1 source subset provided information about how the instrumentation and geodetic analysis responded to sources that contain minimal structure. By

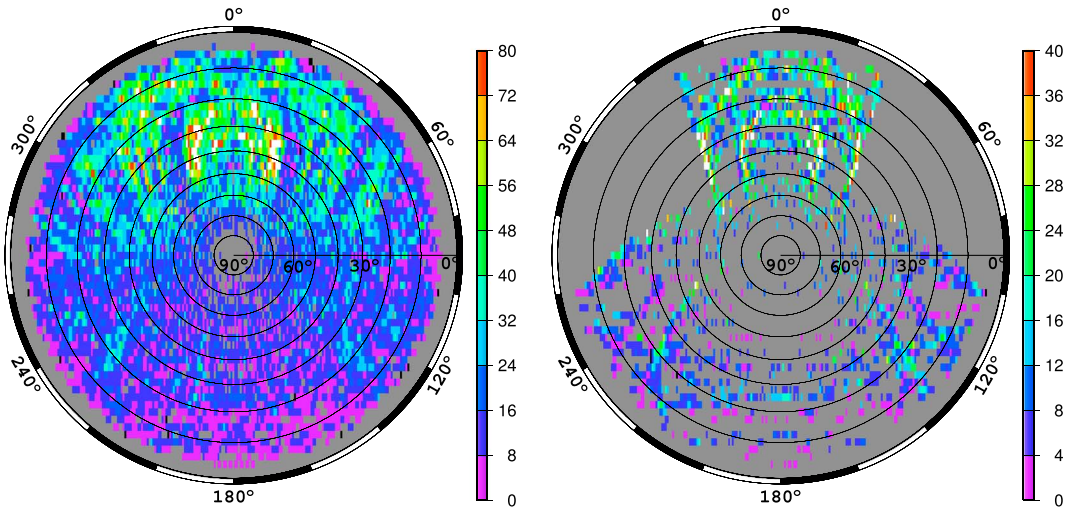

Figure 1. Sky plots showing the number and directions of observations versus azimuth and elevation for both stations in all baseline observations for All sources (left) and Group1 sources (right). The number of observations within each azimuth/elevation angle grid cell is indicated by the color bar on the right-hand side of each plot. North is at the top, east is at the right, and zenith is at the center of each plot. Although the Group1 sources have few observations to the northwest and the northeast, the elevation angle distribution of the Group1 sources is similar to the distribution for All sources.

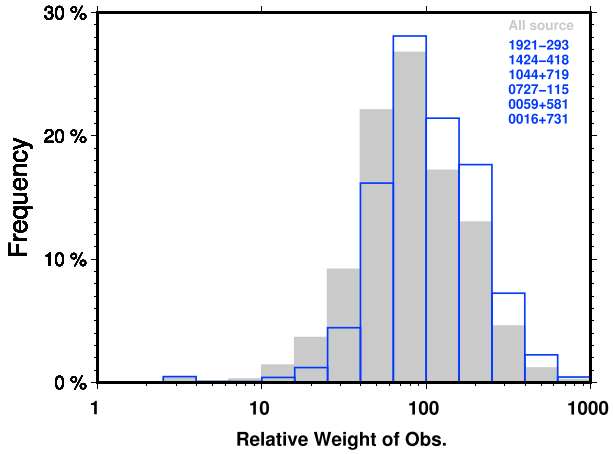

Figure 2. Histograms of the relative Calc/Solve weights for scan-baseline observations used in the analysis. The gray-filled bars show the histogram for the weights for all non-flagged data. The outline blue bars show the histogram for our Group1 sources. The histograms have been normalized such that the sum of all bars for a given source group is 1.0.

combining sources in both the Northern and Southern Hemispheres and thereby including observations from all stations participating in CONT14, Group1 enabled us to investigate the statistical properties of the typical atmospheric, instrumental, EOP, and station position offset errors that were common, in a statistical sense, to all observations.

3.2. Regular Delay Residuals

We reanalyzed the geodetic VLBI data of the 15 sessions of the CONT14 experiment using Calc/Solve. Mean source positions for the CONT14 sources were fit to the data as part of our analysis in order to remove errors originating from the effective source centroid positions during the CONT14 observations being different from their catalog positions.

As part of its standard analysis, Calc/Solve determined additional constant weighting terms for each baseline (pair of stations) for each of the 15 sessions in CONT14 in a manner that attempted to make the χ^2 term over all observations be close to 1.0, using the formal measurement errors for each observation (see Gipson, 1997). In other words, individual baselines were reweighted for each session to make the weighted residual errors match the formal uncertainties. These extra weights complicated our analysis of the closure terms, so we performed two separate analyses of the residuals (after Calc/Solve) and closures, one analysis using a weight of 1 for all unflagged data points and a weight of 0 for all flagged data points (which we called *0/1 weighting*), and a second analysis using the Calc/Solve weights (*Calc/Solve weighting*).

Figure 2 shows a comparison between the Calc/Solve weight distributions for the All and Group1 source groups. The Group1 observations tend to have somewhat higher weights than the All observations, but the distributions overlap greatly, and there are six times as many observations for All sources as there are for Group1 sources. Thus, the Group1 observations did not dominate the least-squares adjustment in Calc/Solve. Therefore, the residual levels determined for Group1 sources with Calc/Solve weighting should have contributions from station-based residual errors, such as atmospheric delay errors, clock offsets, station position errors, and EOP errors, that are representative of the errors from these same terms in the All residuals. In our 0/1 weighting analysis, the Group1 residual station-based errors should be even more representative of the All residual station-based errors.

After our Calc/Solve analysis, we output the residual ionosphere-free delays (observed minus model) from Calc/Solve and independently calculated weighted mean variances (and their corresponding standard deviation values) from equation (B.60), mean absolute residuals from equation (B.61), and median absolute residuals from equation (B.62). The ionosphere-free residual delay statistics for our analysis are presented in Table C.5 in supporting information C.3. Columns 5 and 6 show the standard deviations and mean absolute residuals for 0/1 weighting, while columns 7 and 8 show the same quantities using the Calc/Solve weights. The median absolute residuals, determined independently of the observation weights, are presented in column 9. (The main information for our All and Group1 sources is copied in Table 2.)

The standard deviations of the delay residuals over all sources were 35.3 ps for 0/1 weighting and 25.2 ps for Calc/Solve weighting, whereas the residual levels for our Group1 sources, explicitly chosen to exhibit minimal source structure, were only 24.3 ps for 0/1 weighting and 19.2 ps for Calc/Solve weighting. These values are the overall residuals from the standard geodetic analysis of the CONT14 experiment, combining residual errors from all error terms. These residual error levels show how well the existing geodetic VLBI measurement hardware, observing techniques, geophysical models, and analysis software can perform given a large number of observations performed continuously over multiple days.

These values also already give us an indication of how much source structure must be affecting the overall residual level. Assuming that the Group1 residual levels indicate the performance when no source structure is present, subtracting the Group1 residual levels from the All residual levels (in quadrature) left 25.6 ps for 0/1 weighting and 16.3 ps for Calc/Solve weighting. Since these values are of the same order as the residuals for the Group1 sources, where we assumed that no source structure was present, it is clear that source structure must have had a major impact on the CONT14 measurements. However, since some structure effects

Table 2
CONT14 Delay Statistics Summary (in Units of Picoseconds)

Data type	Band	0/1 weighting				Calc/Solve weighting			
		All sources		Group1		All sources		Group1	
		Median	Mean	Median	Mean	Median	Mean	Median	Mean
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Residual	**Iono-free**	**35.3**	**35.3**	**24.3**	**24.3**	**25.2**	**25.2**	**19.2**	**19.2**
SSMN	**Iono-free**	**22.7**	**24.3**	**11.5**	**12.4**	**17.2**	**19.0**	**9.7**	**10.6**
SS-model	Iono-free	12.5	13.7	1.5	1.6	8.6	10.1	1.3	1.5
CSSMN	Iono-free	18.3	19.6	11.0	11.9	14.0	15.4	9.3	10.2
SS-full	**Iono-free**	**19.9**	**21.2**	**3.4**	**3.5**	**14.5**	**16.0**	**2.8**	**2.9**
Station	**Iono-free**	**27.0**	**25.6**	**21.4**	**20.9**	**18.4**	**16.6**	**16.6**	**16.0**
SSMN	S	189.8	200.4	158.4	168.4	154.4	165.6	137.1	146.4
SS-model	S	31.6	35.0	4.6	4.9	26.8	31.9	3.9	4.2
CSSMN	S	184.2	194.4	157.0	166.9	149.4	160.0	135.9	145.1
SS-full	S	106.7	110.9	21.0	22.4	73.3	79.8	18.1	19.5
SSMN	X	17.5	18.8	5.7	6.1	12.7	14.3	5.0	5.5
SS-model	X	11.3	12.3	1.3	1.4	7.7	9.0	1.2	1.3
CSSMN	X	12.8	13.8	5.2	5.6	9.4	10.6	4.5	5.0
SS-full	X	16.7	17.9	2.3	2.4	11.9	13.4	2.2	2.3
SSMN	S + X[a]	24.2	25.9	14.1	15.0	18.5	20.3	12.2	13.1
SS-model	S + X[a]	12.5	13.6	1.5	1.6	8.6	10.0	1.3	1.4
CSSMN	S + X[a]	20.2	21.5	13.8	14.7	15.7	17.2	11.9	12.8
SS-full	S + X[a]	20.0	21.3	3.0	3.2	14.1	15.8	2.8	2.9

Note. The data type codes have the following meanings: Residual refers to the baseline-based residuals from our Calc/Solve analysis, SSMN refers to the statistical analysis results from the full closure values, SS-model refers to the source-structure model closure statistics, CSSMN refers to the corrected (model source structure subtracted) closure statistics, SS-full refers to the error amount from source structure plus excess measurement noise above the Group1 measurement noise, and Station refers to the station-based error level that is canceled in the closure quantities. As an aid to the reader, the rows of greatest significance have been set in a bold font.
The S + X band rows were formed by a simple addition of the variances of the S band rows and the X band rows (Var $(S+X) = 1.08^2$Var $(X)+0.08^2$Var (S)) corresponding to the calculation of the ionosphere-free delay assuming that the source structure and measurement noises are independent in the two bands.

are probably still present in the Group1 residuals and because errors from other terms may be different from observation to observation, this comparison should only be considered a rough guide to the amount of source structure present in the observations.

3.3. Structure Effects on Geodetic Residuals

An analysis of the residuals from our CONT14 geodetic analysis reveals significant dependencies on source structure.

Figure 3 shows our geodetic residuals of the CONT14 session plotted as a function of the minimum elevation angle of the two stations forming each baseline. For the sources with minimal source structure (the top row), the RMS scatter was nearly constant with elevation angle. Only in the lowest minimum elevation angle bin was a significant increase visible. For sources with larger amounts of source structure, the scatter generally increased toward lower elevation angles for all minimum elevation angles sampled. This result is important because the nearly constant RMS levels for sources with minimal source structure show that delay errors from the atmosphere did not dominate the geodetic residuals, except perhaps at the very lowest elevation angles, and other effects must dominate the residual errors.

The fact that sources known to have little source structure had virtually no elevation-angle dependence in their residuals, whereas sources with significant source structure did have a significant elevation-angle dependence suggests that the elevation-angle dependence is just a proxy for source structure. (When the minimum

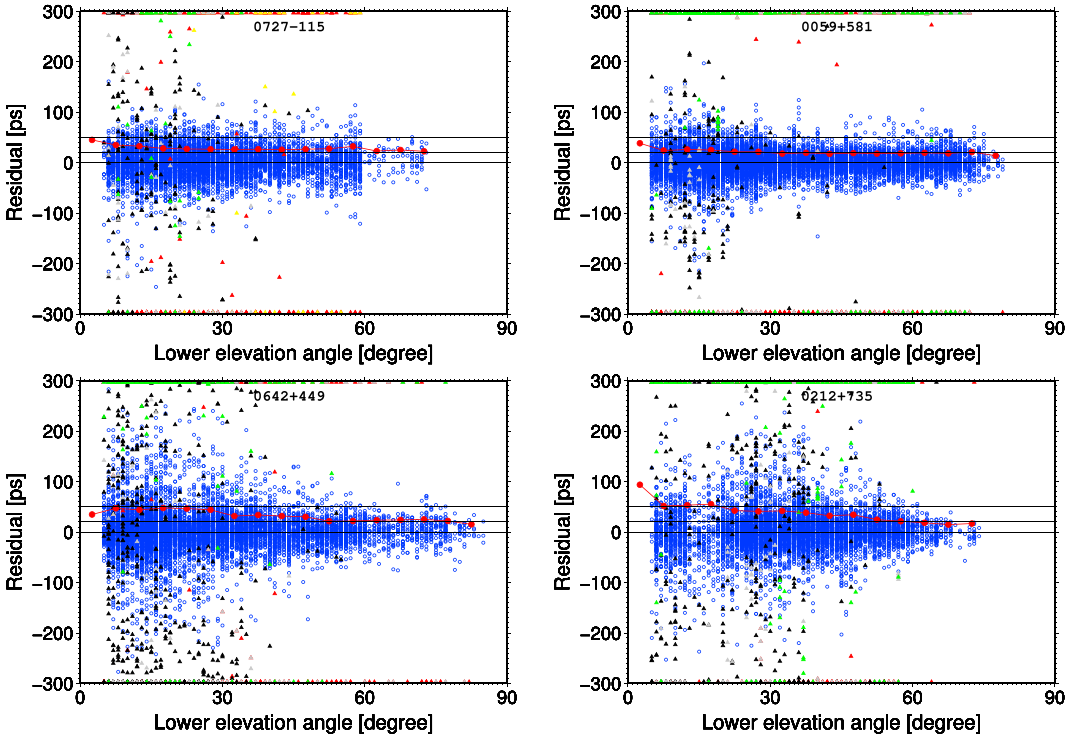

Figure 3. Plots of residual ionosphere-free group delay versus the minimum elevation angle of the two stations in each observation. The sources are 0727−115 (1.8), 0059+581 (1.9), 0642+449 (3.0), and 0212+735 (3.3), where the numbers in parentheses following the source names are the continuous structure indices of the sources for the CONT14 observations from Xu et al. (2017), with larger values indicating more significant structure. Unflagged observations are shown as blue circles. Flagged data are indicated by filled triangles, with red indicating nondetections in S band data (S band data had more noise than X band data), yellow indicating the `HOBART12−HOBART26` baseline, gray indicating flagged data involving the `MATERA` station, green indicating flagged data involving the `ZELENCHK` station, and black indicating all other flagged data. Data points with absolute residual levels greater than 300 ps are plotted along the top or bottom of each subplot. Red-filled circles connected by red lines show the root-mean-square residual scatter for data in 5° elevation angle bins. Thin horizontal black lines at 0, 20, and 50 ps are shown to guide reading the root-mean-square residual levels.

elevation angle was high, the two stations forming the baseline had to be near to each other—had a short baseline length. Geodetic VLBI sources have little structure at short (u, v) locations. For lower minimum elevation angles, the baseline lengths could be longer, where source structure is more prominent.) Source structure, not the atmosphere, must have dominated the elevation-angle dependence of the residual level for CONT14.

The geodetic residual level also showed a strong relationship to the source brightness for individual measurements. The source brightness conveys crucial information about the actual measurement noise and source-structure level in each observation. It is therefore unfortunate that the amplitudes measured by each scan-baseline observation are not used in current geodetic analysis software once delays and rates have been measured for each observation, as the amplitudes can say a great deal about what is going on in the measurements.

Significant source-structure effects caused the measured amplitudes in CONT14 to be reduced well below the total flux densities of the sources. These low amplitudes, in turn, lead to higher measurement noises. This problem became critically important when the SNRs were less than about 3, as the error distribution in phase, from which the delay and rate measurements are made, is not approximately normally distributed for low SNR. The bivariate noise distribution (real and imaginary components equivalent to phase and amplitude) results in large phase errors, with significant measurement probability for all angles, including 180° from the true phase. (By an SNR of 2, the phase uncertainty is already more than 20% larger than the expected 1/SNR radians in the limit of high SNR.) This leads to exceptionally large group-delay errors for very low SNR measurements. This problem is compounded by the significant probability of measuring random group delay values below SNRs of about 6–7 (see section 4.3.4). At the same time, large source-structure delay errors tended to

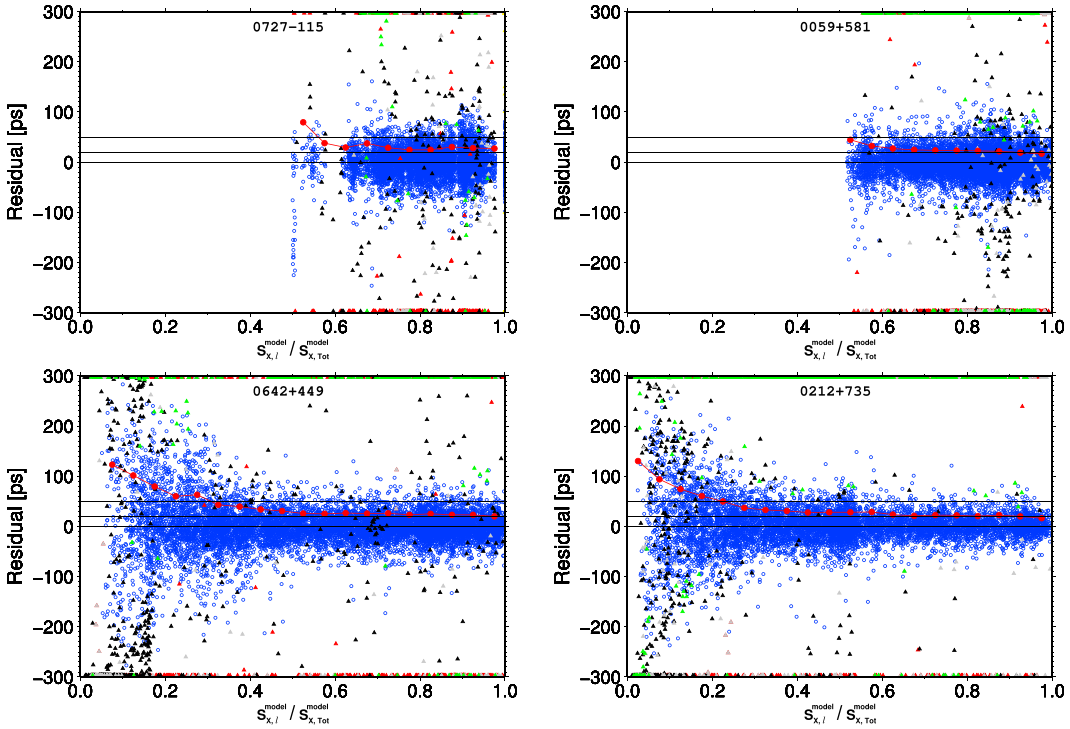

Figure 4. Plots of residual ionosphere-free group delay versus the relative X band baseline brightness for a selection of sources. The baseline brightnesses were calculated from the CONT14 imaging structure models and are shown relative to the total model flux densities of the individual sources. The sources are 0727−115 (1.8), 0059+581 (1.9), 0642+449 (3.0), and 0212+735 (3.3), where the numbers in parentheses following the source names are the continuous structure indices of the sources for the CONT14 observations from Xu et al. (2017), with larger values indicating more significant structure. See Figure 3 for the description of data point symbols. Red-filled circles connected by red lines show the root-mean-square residual scatter for unflagged data in amplitude bins spaced by 0.05 of the total flux density. Thin horizontal black lines at 0, 20, and 50 ps are shown to guide reading the root-mean-square residual levels.

occur for the same baseline (u, v) configurations that yielded low amplitudes, so that the combination of high measurement noise and high structure delays lead to high residual levels for those observations.

Figure 4 shows the group-delay residuals from our CONT14 analysis as a function of the model source brightness for each observation relative to the total flux density of each source, for a selection of sources with increasing source-structure levels. Sources with minimal structure levels (the top row) showed only modest departures in brightness from the total flux density and have nearly constant RMS residual scatter levels as a function of amplitude (the red-filled circles), excepting the measurements at the very lowest brightness levels for these sources. For the sources shown with the most source structure (only continuous structure indices of 3.0 and 3.3, so far from being sources with the most structure in the experiment!), the brightnesses for some measurements were at or below only a tenth of the total flux density for the sources. There was a steady increase in the RMS scatter levels as the baseline brightness decreased, with the scatter growing rapidly for relative brightnesses decreasing beyond 0.3. For high amplitudes, the RMS scatter is only about 20 ps. For low amplitudes, the RMS scatter was above 100 ps.

We color-coded the flagged data shown in Figures 3 and 4 to indicate data points that were made with a frequently corrupted baseline (`HOBART12−HOBART26`) or included either the `MATERA` or the `ZELENCHK` stations, which also had many outlier measurements. We plotted all other flagged data with black + marks. As Figure 4 clearly shows, the vast majority of the flagged data points not related to the above specific baseline or stations were not "outliers" in a true exceptional deviation sense. Instead, these flagged data were merely the outer reaches of a smooth scatter distribution. For the sources with the most source structure (the bottom row), the majority of these flagged points occurred at the lowest brightness levels, where the intrinsic scatter was the highest. Thus, the true RMS scatter levels for low brightness measurements were even larger than indicated by the red circles.

The observation that large numbers of VLBI baselines in CONT14 observed source flux densities that were far smaller than predicted by the scheduling software pointed out to us that there is an important omission in the modeling of source-structure delay errors in the calculation of source-structure delay as implemented in the structure index calculations of Fey and Charlot (1997) and Fey and Charlot (2000). The structure indices are calculated based purely on structure delays and do not account for the effect that source brightness has on measurement errors. The simple source-structure models effectively assume that all measurements are made with the same SNR, whereas historically, the geodetic VLBI scheduling software was only aware of a limited amount of source structure, resulting in many observations being performed with lower than intended SNR values. Thus, simple source-structure models that do not account for the extra scatter introduced by low SNR underpredict the true effects of source structure on geodetic error levels.

4. Closure Inspection

4.1. Closure Quantities

Next, we determined the residual delay error contributions from source structure and measurement noise in a different, more accurate way. We did this by studying the *closure delay* and *closure phase* statistics of the CONT14 ionosphere-free and single-band measurements. A detailed discussion of closure quantities and how we used closure quantities to estimate error contributions from source structure plus measurement noise (SSMN) is provided in supporting information A and B. In brief, using simultaneous measurements from more than two stations, closure quantities provide information about the source that was observed that is *independent* of most station-based error terms. In the case of *closure delay triangles*, the closure group delay for a triangle of three stations, *a*, *b*, and *c*, simultaneously observing the same source is given by the sum of the three baseline group delays going around the triangle in a consistent direction,

$$\tau_{\mathrm{grp},abc} \equiv \tau_{\mathrm{grp},ab} + \tau_{\mathrm{grp},bc} + \tau_{\mathrm{grp},ca}, \tag{1}$$

(equation (A.9), astronomical convention). In this summation, the effects of station-based delays—tropospheric delays, ionospheric delays, station position errors, station thermal deformation errors, clock offset errors, cable delay errors, EOP errors, errors from pointing offsets, and so on—*cancel exactly*. This happens because these delay terms appear twice for each station, once in a positive sense, and once in a negative sense, so that the sum of such terms is identically zero. The closure delay can be rewritten as

$$\tau_{\mathrm{grp},abc} = \tau_{\mathrm{grp},ab} + \tau_{\mathrm{grp},bc} - \tau_{\mathrm{grp},ac}, \tag{2}$$

(equation (A.10), astronomical convention) to account for standard station-ordering conventions. The station reference time for any station in the triangle must be the same for each baseline that station appears in, so for geodetic data, which, unlike astronomical data, use different reference times that depend on the first station in the baseline, the group delay for the *b*–*c* baseline must be adjusted to match the reference time for the *a*–*b* baseline by correcting for the delay between stations *a* and *b* by application of the Taylor series correction

$$\tau_{\mathrm{grp},bc,ab}\left(t_0\right) = \sum_{k=0}^{\infty} \frac{\tau^{(k)}_{\mathrm{grp},bc,bc}\left(t_0\right)}{k!} \left[\tau'_{\mathrm{grp},ab,ab}\left(t_0\right)\right]^k, \tag{3}$$

(equation (A.11), geodetic convention). We only apply the $k = 0$ and 1 terms for our analysis, resulting in a delay correction accurate to about 0.1 ps. See supporting information A for details about the nomenclature, additional references, and constraints related to the closure expressions above.

Through our analysis of closure delays and closure phases, we examined the geodetic measurements in a way that was independent of atmospheric errors, independent of clock errors, independent of station position errors, and independent of EOP errors, terms that are known to be significant contributors to the final residuals of geodetic VLBI measurements. The major error terms that remained in the closure quantities, the so-called *nonclosing errors*, are source structure and measurement noise. Additional nonclosing error terms such as frequency-dependent bandpass mismatch and polarization leakage were also probably significant, although not at as dominant levels as source structure and measurement noise, see supporting information A.7. In general, both frequency-dependent bandpass mismatch and polarization leakage errors depend on the source being observed, the spectral shape and polarization of the source emission for a particular baseline (u, v) coordinate (effective baseline length and orientation), and the parallactic angle of the source as seen by each antenna. In other words, these additional nonclosing errors were time and source dependent, and for the purposes of our work here, we included these effects in our broad concept of source structure.

The price we paid for eliminating the standard station-dependent error terms through closure analysis is that the closure quantities are independent of the absolute source position on the sky. Residual errors caused by the measured absolute source position being different from the reference model position used in the geodetic analysis were lost. Closure quantities are only sensitive to *relative* source structure and not at all sensitive to *absolute* source position, meaning that the source structure errors estimated from closure quantities are only lower bounds to the total source structure errors present in geodetic VLBI data. Because we fit for the absolute positions of all sources in our Calc/Solve analysis of the CONT14 data, absolute source position errors did not affect our analysis. But a standard geodetic analysis, which uses fixed absolute source positions, would contain such additional errors when sources are offset from their catalog positions, even if their relative source structures are minimal.

As described in the Supporting Information S1, we have assumed that taken over many different baselines, sources, and scans, the source structures and the measurement errors for different baselines in closures are effectively uncorrelated. For the majority of cases where this assumption might not be true, such as triangles with two long, nearly similar baselines and one short baseline, the baseline structure delays would tend to cancel in the closure, thereby resulting in *smaller* closure measurements than expected for uncorrelated values, such that our statistical analysis here would *underestimate* the true structure effect.

In summary, closure delays provide a direct and important way to estimate the error contributions from SSMN terms to the residuals of geodetic data analysis. The ionosphere-free closure delays in our analysis contribute fully to the geodetic baseline residuals—what we found in the closures must also have present in the baseline residuals.

4.2. SS-Model Closures

Then, using the CLEAN components from our uniform-weighting imaging analysis, we computed source-structure-model (SS-model) group delays and phases for each baseline for each scan for all sources in the CONT14 experiment that we could image reliably. (For an introduction to calculating source-structure group delays or phases from CLEAN components, see, e.g., Charlot, 1990, or, as we have done, derive the equations from the interferometry principles discussed in Thompson et al., 2017.) For model group delays, we computed SS-model visibilities as a function of frequency, unwrapped the phases from some arbitrary starting frequency to account for phase wraps in the source structure model, and then fitted a linear phase slope as a function of frequency to the model visibilities at the frequencies of the subbands used for geodetic analysis in the CONT14 experiment, assuming equal weights for all subbands. For model phase delays, we computed the structure phase at the reference frequency of each band. Lacking spectral information, we assumed that the spectral index for all CLEAN components was $\alpha = 0$ ($S_\nu \propto \nu^{+\alpha}$). From the MOJAVE survey of similar sources, the mean core spectral index is 0.22, while the mean jet spectral index is −1.04 (Hovatta et al., 2014). The two-component model of Charlot (1990) suggests that such spectral indices would introduce additional group delay corrections on the order of 15% of the main delay term for a strong jet component. However, the CONT14 imaging results indicate that most sources have at least tens of CLEAN components. We simulated many-component source models and found that typical spectral-index-related group delays for strong-jet sources were far lower than predicted by a two-component model, with spectral-index group delay contributions less than about 3 ps for 75% of the simulated observations. We therefore do not expect that unmodeled spectral index variations would significantly impact our results.

To again emphasize the importance of the CONT14 experiment, the CONT14 data provided a unique opportunity to combine the analysis of geodetic data made in a standard geodetic observation mode, with high-quality imaging data made *from the same observations, at the same time, using the same VLBI station network*. The source structure in our image data is the same as the source structure in the geodetic data because the measurements were simultaneous (eliminating source structure changes with time), the same frequency coverage was used (eliminating frequency-dependent source-structure changes), and the exact same station network and (u, v) coverage were used (eliminating source structure changes because of sampling different Fourier components of the source structure).

The SS-model delays and phases were not perfect representations of the observed source structure. CLEAN components, with allowable locations fixed to a regular grid spacing, have difficulty representing continuous emission distributions on the sky. Structure smaller than the interferometer beam size is also difficult to reproduce with CLEAN components but can be critically important to determining the delays observed by long

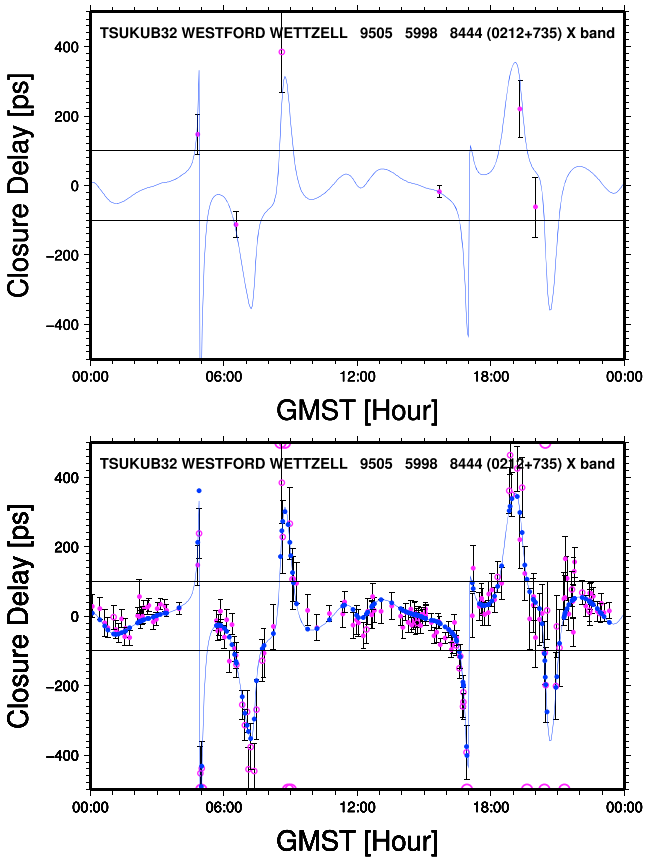

Figure 5. Plot of the closure delay at X band for source 0212+735 as a function of GMST for the triangle TSUKUB32–WESTFORD–WETTZELL. Magenta points indicate CONT14 measurements, with black bars giving the 1-σ measurement uncertainties. Filled circles indicate observations that are included in the geodetic analysis, while open circles indicate observations that were flagged by Calc/Solve on at least one of the three baselines. Measurements outside of the displayed closure-delay range are plotted as large circles at the top or bottom plot boundaries. The blue line represents the model closure structure from the imaging analysis of the CONT14 data, with blue dots indicating model values for each CONT14 measurement. The station names and the baseline lengths in kilometers are shown at the top of the plot. The thin horizontal black lines at −100 and +100 ps are intended to guide the eye. The upper plot shows six closure measurements—a typical number of closures that would be observed in a standard International VLBI Service for Geodesy and Astrometry 24-hr session. The bottom plot shows all of the closure measurements for 0212+735 by this triangle in CONT14. GMST = Greenwich mean sidereal time.

baselines. Large angular-scale features are also difficult to reproduce using CLEAN components, leading to errors for the shortest baselines. Our CLEAN components did not contain spectral information, such that over the relatively wide bandwidths employed in geodetic VLBI, the sources with large spectral index variations were not well reproduced. However, our imaging results did perform a reasonably good job of reproducing the measured source structure as we demonstrate below.

The SS-model delays and phases also did not contain all of the source-structure-related effects on the measurements. For very complicated sources (those with a great deal of source structure), our imaging-based models did not fully describe all of the structure present, for reasons mentioned directly above. The models also do not describe other nonclosing errors, including polarization leakage and bandpass mismatch. Although polarization leakage may not be a dominant error term (Bertarini et al., 2011), it *does* have a nonnegligible impact on the total error budget, and in our closure analysis with most station-based errors eliminated, polarization leakage errors were often readily apparent. (The majority of bandpass mismatch effects are caused by differences in station backends that are relatively stable in time and therefore result in constant closure offsets. Systematic baseline RFI problems also tend to result in constant closure offsets. Bandpass mismatch errors cause by source spectral index dependence on source structure vary with time and source and tend to be uncorrelated over station parallactic angle for many sources, as the source structures are uncorrelated. In contrast, polarization leakage effects show up as time variable closure offsets that are correlated over all source observations for similar parallactic angle differences between stations.) And finally, the source structure models did not take into account the *variable measurement noise* caused by the amplitude-dependence of the source structure emission observed at specific (u, v) locations. The source brightnesses at the observed (u, v) locations were only roughly estimated by the geodetic scheduling software (source brightness was typically assumed to be constant, follow a two-dimensional Gaussian form, or simply modeled as a function of baseline length), and for sources with even only modestly complicated structure, it was *common* that the measured brightnesses were only one half, one third, or even only one tenth of the brightness expected by the scheduling software. This caused the SNR of such measurements to be significantly lower than expected and the measurement noise to be far larger than desired. This variable measurement noise was source and (u, v)-location dependent. Hence, this extra measurement noise above the level expected by the scheduling software was also a source-structure-related error term that was not compensated by a postobservation source structure model.

Using our baseline SS-model delays and phases, we calculated SS-model closure delays and closure phases that described, to the accuracy of the CLEAN component models, the source structure information that the interferometer observed. By calculating SS-model closures, we eliminated the difficult problem of matching the reference point of the CLEAN components in our imaging results to the reference point of the source in the geodetic delay model. Since closures are independent of absolute source position, the image position offsets did not matter. We were left, once more, with only the relative source structure properties of the sources.

4.3. CONT14 Closure Plots

4.3.1. Introduction

The CONT14 experiment is tremendously valuable for the study of source structure in geodetic observations because it provides a large number of observations of the same sources by the same stations in a short period

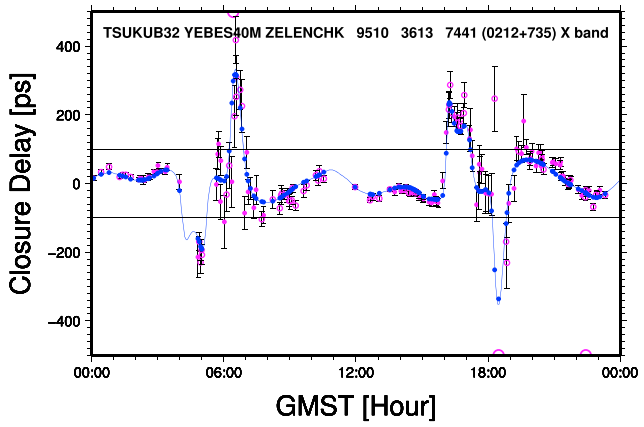

Figure 6. Plot of the closure delay at X band for source 0212+735 as a function of GMST for the triangle TSUKUB32−YEBES40M−ZELENCHK. See Figure 5 for a description of the plot symbols. The measurement noise, indicated by the 1-σ error bars, is highly variable as a function of GMST. Observations with low measurement errors typically have small model closure-delays, while large measurement errors are usually, but not always, associated with larger model closure-delays. GMST = Greenwich mean sidereal time.

of time so that the source structure does not change substantially. In contrast to earlier analyses of structure closure delays in geodetic VLBI observations such as Charlot (1990), where only up to approximately six closure delays per triangle were measured, the CONT14 data have many tens of closure measurements per triangle per source, with well-observed sources frequently having ~75 closures per triangle and most sources having more than 100 triangles. We demonstrate this in Figure 5, which shows how the large number of data points in CONT14 are important to understanding the impact of source structure in geodetic observations. Closure delay measurements and a closure model are shown as a function of the Greenwich mean sidereal time, with observations on different sidereal days wrapped by 24 hr. Assuming the source structure (on the sky) was constant over CONT14, the instantaneous closure delay of a specific triangle was determined only by the baseline orientations, and hence by Greenwich mean sidereal time, so that trends in the delay measurements that were caused by source structure should be readily apparent in this sort of plot. When only a few measurements are available, as demonstrated in the upper plot of Figure 5 where six closure delays have been selected at random from the available CONT14 data, it is difficult to judge how well the model actually fits the measurements. Looking only at the six closure delays and their uncertainties, the measurements are all within about 3-σ of zero closure and therefore could be consistent with no structure being present. Are all of the elaborate features in the model real, or has the model been adjusted to fit the measurement noise in the data? Or is significant source structure missing in the model? With only a handful of measurements, we cannot make strong conclusions.

The bottom plot of Figure 5 shows how the *many* closure measurements in the CONT14 data confirm that our models did an *excellent* job of reproducing the observations. For this triangle of stations, the source presented, 0212+735, had very strong structure extending at times beyond 400 ps in closure delay. The model did not miss any significant structure seen by the measurements nor did it predict more variable structure than seemed to be present in the data. Even for closure delay excursions of several hundred picoseconds, the model provided a faithful prediction of the measured closure values. Our results for sources other than 0212+735 were comparable, but due limited space constraints, we cannot show the full set of closure plots for all station triangles for all sources here. For the interested reader, the full set of plots can be found in Anderson and Xu (2018).

4.3.2. Variable Measurement Noise

Figure 6 shows an example of a different feature that is strikingly evident when given the large number of measurements in the CONT14 data set. The measurement errors (observation minus model) and uncertainties are highly correlated with the model source structure. Areas of significant structure and/or significant rates of change of the structure typically have far larger measurement errors and uncertainties than areas with little structure. There was a very strong anticorrelation between the single baseline measurement noise and our source model amplitudes, in agreement with our a priori expectation. These low amplitude regions are typically also associated with large magnitude closure delay/phase values or regions of closure delay/phase that are rapidly variable in time. Closure values with large error bars in our plots have at least one baseline with a low source amplitude.

4.3.3. Triangle Dependence

An example of the measurement and noise behavior of sources with significant structure for different closure triangles is shown in Figure 7. The amount of source structure seen by different station triangles can be vastly different. In general, triangles including long baselines had more structure than triangles with only short baselines for the sources observed in CONT14. (In addition to looking through all of the closure plots we have made available online, the interested reader can also consult structure correction maps provided by the Bordeaux VLBI Image Database. Our own structure correction maps based on our CONT14 imaging analysis will be published in a future manuscript. Most sources observed by the IVS have stronger structure effects on long VLBI baselines compared to short VLBI baselines.)

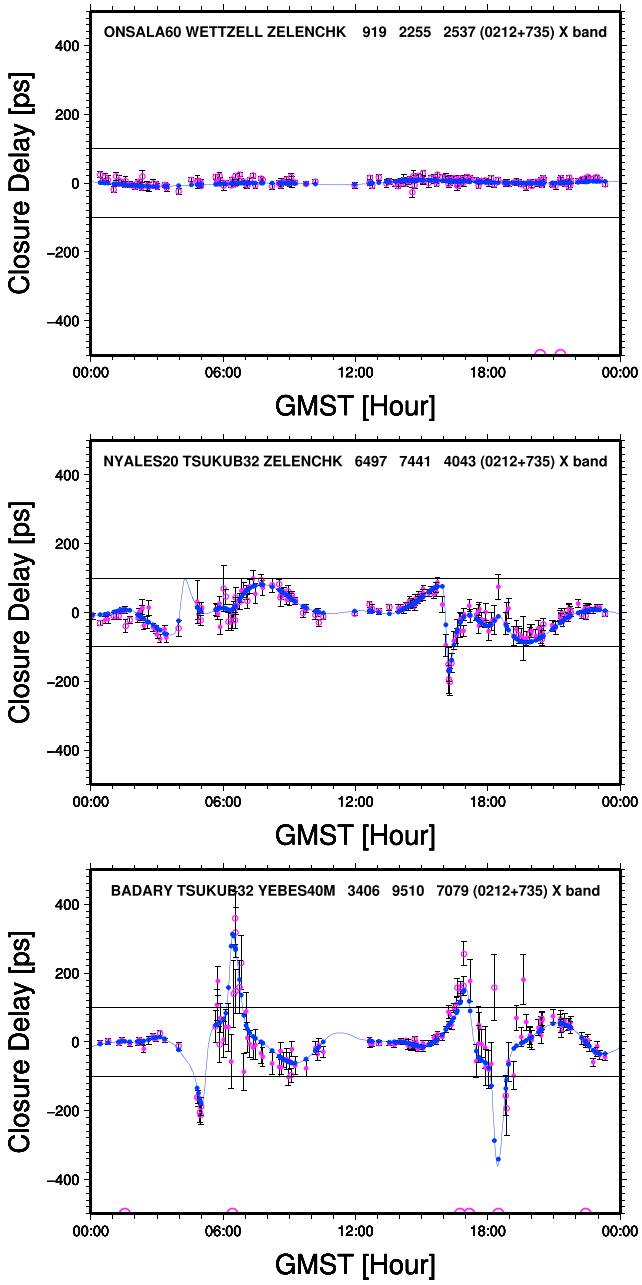

Figure 7. Plots of closure delay at X band for source 0212+735 for three different closure triangles. See Figure 5 for a description of the plot design. Increasing levels of source structure are shown in the three plots, with increasing measurement uncertainty levels. GMST = Greenwich mean sidereal time.

4.3.4. Automatic Flagging

The probability that a measurement was flagged also appeared to strongly be related to source structure, as shown in Figure 8 (see also Xu et al., 2016). The great majority of extreme outlier points (the large open circles at the top and bottom boundaries of the plot, with closure delays in some cases extending out to tens of nanoseconds) were found to occur at times when strong source structure was present—specifically when low model baseline amplitudes were predicted. These extreme outliers were therefore not caused by random instrumental errors or RFI—they must have been caused by source structure. Note that for SNR values below about 6–7, the probability of a finding a noise spike instead of the true source delay is nonnegligible, and for SNR values below ~4, the probability of measuring a random delay value is greater than 0.1 (see Thompson

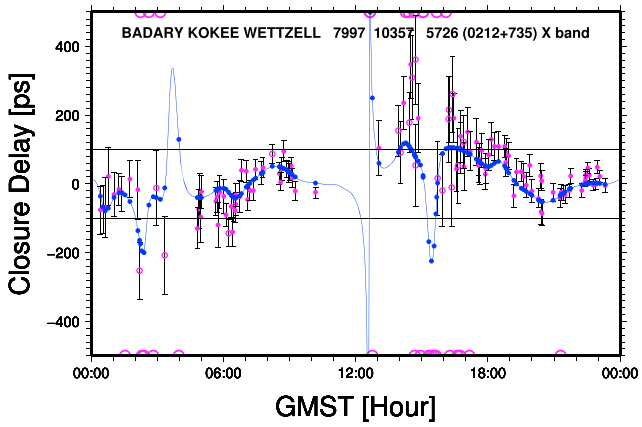

Figure 8. Plot of the closure delay at X band for source 0212+735 as a function of GMST for the triangle BADARY–KOKEE–WETTZELL. See Figure 5 for a description of the plot symbols. Closures that included at least one baseline that was automatically flagged by the Calc/Solve software, shown by the open-circle magenta points, tended to occur in regions of strong source structure. GMST = Greenwich mean sidereal time.

et al., 2017, section 9.3.4 and references therein). Lower than expected source amplitudes caused by source structure lead to huge, random delay measurement errors.

Similarly, the regular points flagged by Calc/Solve (the small open circles within the plot) also appeared most often in regions of strong source structure. The majority of the flagging of data points by Calc/Solve appeared to be directly related to source structure. The Calc/Solve flagging was apparently performed in terms of absolute *baseline* delay (see also section 6.2), but since baseline delays can destructively as well as constructively add in closure delays, strong baseline structure is not always indicated by large closure magnitudes. But regions of strong closure time gradients or that have large uncertainties (resulting from low structure amplitudes) *are* associated with strong structure. Our manual examination of these data showed that the SNR was less than 9—that is, lower than scheduled but still above the SNR cutoff for true detections. This is reinforced by Figure 9, which shows similar results for yet another source with strong structure, 0738+313. Nearly all of the observations in regions of strong closure structure were flagged as outlier points by the Calc/Solve software. We can see, however, that the regular measurements (nonextreme outliers) were in very good agreement with the source-structure model. The measurement values, to within the measurement uncertainties, were correct and appropriate given the source structure present and were not outliers in the sense of unusual deviations from a reasonable error distribution. On the other hand, because no corrections for the actual source structure were made in the standard geodetic analysis and had these observations not been flagged, the Calc/Solve residual level would have been far larger.

In contrast, Figure 10 shows example closure plots for two of the sources that make up our Group1 subset. For all closure triangles for these sources, the closure plots show a nearly (but *not completely*) flat structure with only a small scatter about the model structure curve. Very few points were flagged by Calc/Solve—there were few extreme outlier points, and very few of the regular measurements were flagged.

4.4. S Band

S band data had more source structure and also greatly higher measurement noise. Figure 11 shows four example closure triangles showing S band results for source 0212+735. The vertical scale is much larger than used for the X band plots, as the S band data were far noisier than X band data. When the measurement noise was not excessively large, as shown in the upper plots in this figure, the closure measurements tracked the model structure quite well. However, some VLBI stations, and in particular, some VLBI baselines, had far larger measurement uncertainties. With only six subbands and being located at a relatively low frequency, geodetic VLBI S band measurements are susceptible to RFI, and when multiple S band subbands are flagged because of RFI, especially the subbands at the far ends of the band, the measurement noise even for high brightness observations cases can be huge. The behavior of measurement noise can easily be shown and investigated through our closure analysis.

Even for our Group1 sources, with little intrinsic source structure, the measurement noise at S band is quite large, as shown in Figure 12. Even when the S band measurement noises were scaled by a factor of 0.08 for the ionosphere-free combination, they were larger than the corresponding X band noise levels. Here the same two example Group1 sources using the same closure triangles as shown for X band in Figure 10. From the large number of measurements in the CONT14 session, we could see that the closure-delay measurements were, in fact, consistent with little S band structure being present, but the measurement noise was extremely high.

Our closure measurements showed that amount of source structure and the amount of measurement noise were both significantly larger for S band than their corresponding values for X band for CONT14. Southern Hemisphere stations tended to have larger S band measurements errors than Northern Hemisphere stations; hence, southern sources had larger typical S band measurement errors than northern sources. When the frequency-scaling terms for the ionosphere-free combination were applied, the S band simple source structure contributed little to the total residual level, but the S band measurement noise continued to make a major contribution to the total error budget. (Our source-structure models were based on imaging data that,

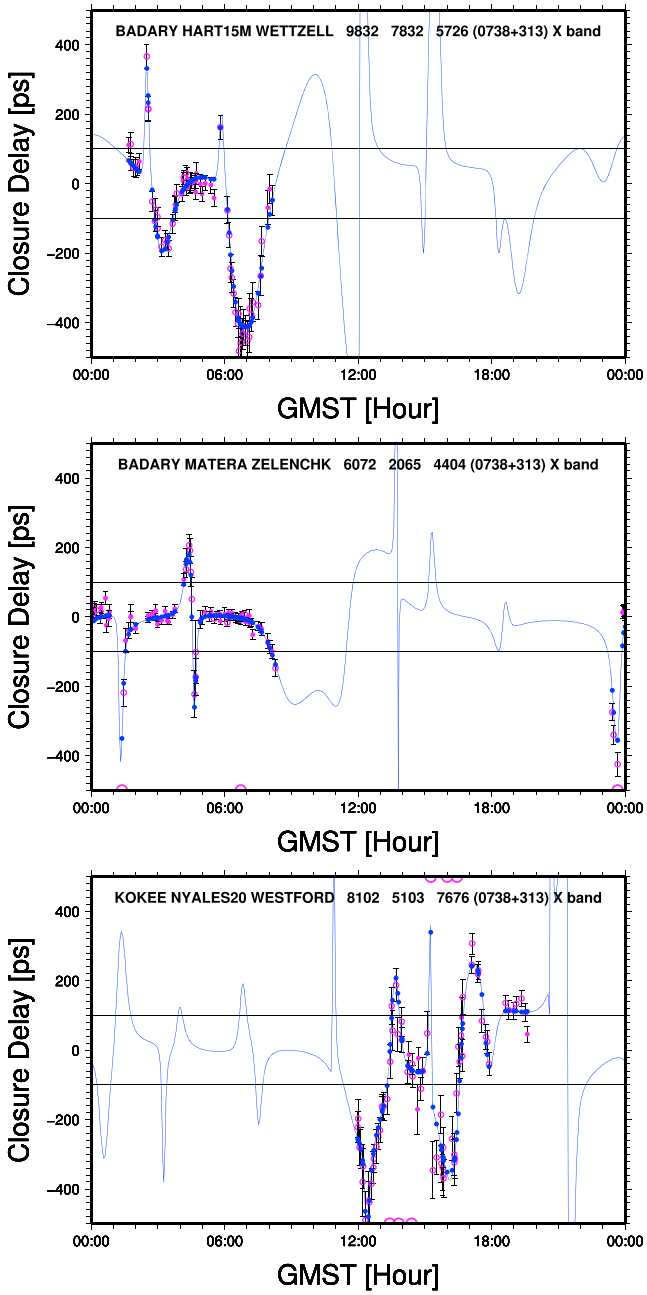

Figure 9. Plots of closure delay at X band for source 0738+313 for three different closure triangles. See Figure 5 for a description of the plot design. Increasing levels of uncertainty for increased source structure are seen in the three plots. GMST = Greenwich mean sidereal time.

in turn, were based on visibility [phase] data. The imaging results were far less sensitive to missing subbands, such that our S band structure models are still quite good.) And since measurement noise was correlated with source structure, S band source structure therefore continued to contribute to the ionosphere-free residuals.

5. Closure Delay Statistics

5.1. SSMN Closure Delays

We then performed a statistical analysis, described in detail in supporting information B, of the closure values, caused principally by (relative) SSMN, in order to estimate the contribution of source structure and measurement noise in the geodetic, baseline-based, observation residuals. In the case of 0/1 weighting, where all contributing baselines have the same weight, this statistical estimation was exact—the baseline delay vari-

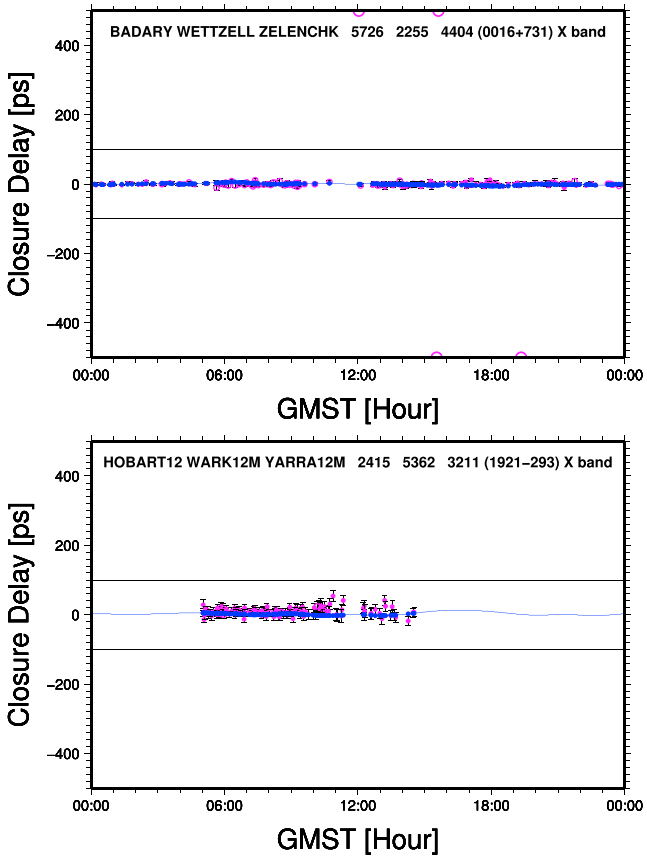

Figure 10. Plots of closure delay at X band for a selection of triangles for two of the Group1 sources, 0016+731 and 1921−293, as a function of GMST. See Figure 5 for a description of the plot design. Sources and triangles with little structure had few observations flagged. GMST = Greenwich mean sidereal time.

ance has an exact mathematical relationship to the closure statistics. In the case of other weighting schemes such as the Calc/Solve weighting, where the weighting of different baselines within a triangle can be different, the estimated baseline variance determined from the closure statistics was the statistically most likely value under the assumption that the weights are independent of the closing errors. Closure delays from each individual scan were analyzed to determine the variances from SSMN terms in the observables (baseline delays) of each scan. These scan-based SSMN variances were then combined over all observables for all scans to give the overall SSMN variance for the entire experiment. The large number of stations that participated in CONT14, the scheduling of most scans to include at least three stations (less than 3% of the scheduled observables were single-baseline scans), and the long duration of the experiment giving many scans for each source made the CONT14 experiment the best geodetic experiment on which we could perform this type of analysis.

The overall SSMN variance estimate based on the mean scan variances—giving the exact SSMN variance in the 0/1 weighting case—was calculated from equation (B.15). We refer to this as our *mean* estimate. Equations (B.17), (B.19), and (B.21) give overall SSMN variances using the minimum (*min*), median (*median*), and maximum (*max*) closure delay values, respectively, calculated over all stations and scans. The min estimate gives the minimum SSMN variance in the overall delay residuals—the SSMN residuals *cannot* be any smaller than this value. Similarly, the max estimate gives the maximum possible SSMN variance in the overall delay residuals. And finally, the median estimate gives a robust estimate of the probable overall SSMN variance in the case of significant outlier points in scans with many closure triangles. For the Calc/Solve weighting case and observation outlier points that were not flagged to zero weight during the analysis, this median estimate could be a better estimate of the true overall SSMN variance using the Calc/Solve weights than the mean estimate.

Table C.6 in supporting information C.3 provides our analysis results for the overall ionosphere-free delay variances (written as standard deviations) for sources with closure triangles in CONT14. We provide the min,

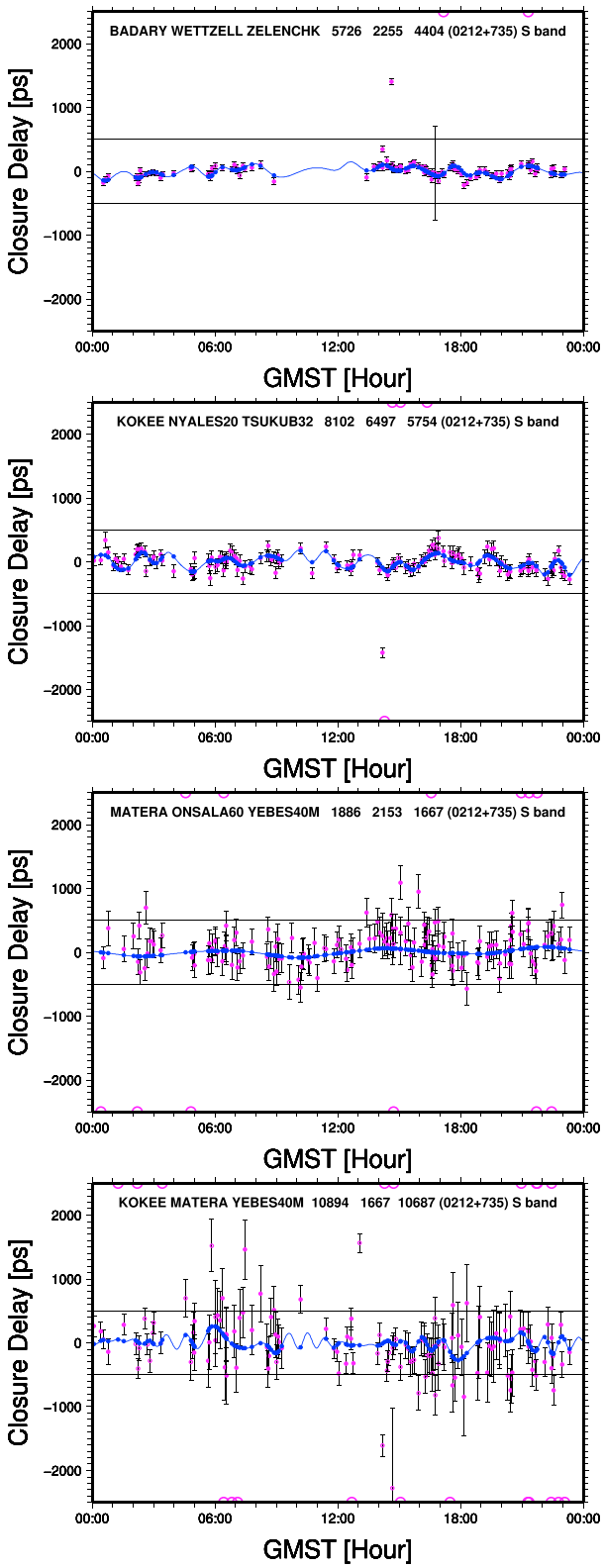

Figure 11. Plots of closure delay at S band for source 0212+735 as a function of GMST. See Figure 5 for a description of the plot design. Closure triangles with various levels of source structure and measurement uncertainty are shown. The agreement between the model and the observations was quite good, when the scatter was small enough to detect the structure. GMST = Greenwich mean sidereal time.

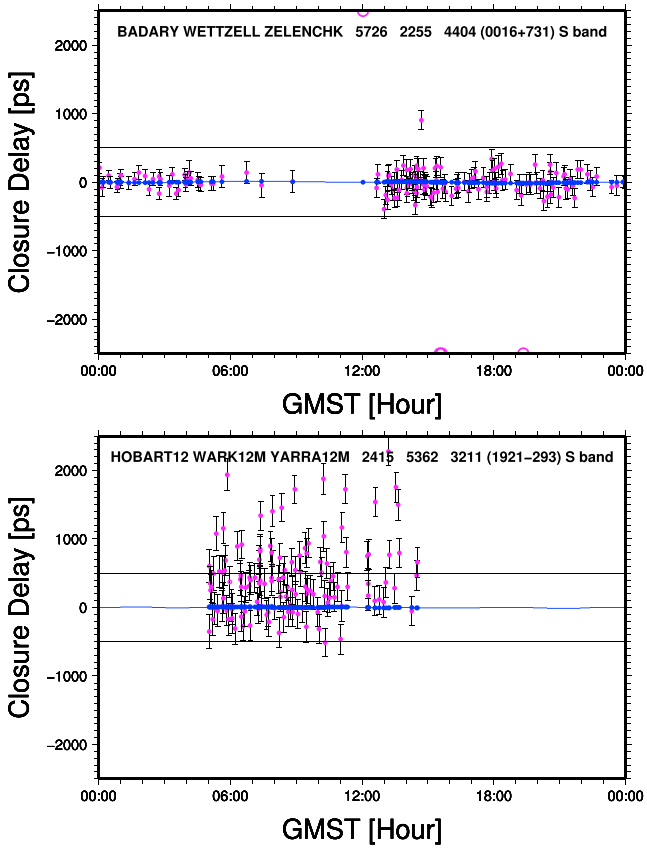

Figure 12. Plots of closure delay at S band for a selection of triangles for two of the Group1 sources, 0016+731 and 1921−293, as a function of GMST. See Figure 5 for a description of the plot design. The sources and closure triangles are the same as in Figure 10. Little source structure was present, but the measurement noise was quite large.
GMST = Greenwich mean sidereal time.

max, median, and mean statistical estimates for each source individually, as well as our All and Group1 sets of sources. The mean (median) SSMN standard deviations were 24.3 ps (22.7 ps) for 0/1 weighting and 19.0 ps (17.2 ps) for Calc/Solve weighting for All sources, as summarized in Table C.7. The SSMN errors clearly constituted major portions of the total geodetic error budget. As discussed in supporting information B.1, our statistical analysis assumed that source structure was uncorrelated among baselines within closure triangles, and (anti)correlated source-structure delays for small or modest structure sources can be *underestimated* by our closure analysis. Therefore, our SSMN values should be considered as lower limits to the actual amount of relative source structure and measurement noise in the observations.

5.2. SS-model Closure Delays

Using these SS-model closure delays, we computed overall SS-model variances for the entire experiment in the same way that we calculated the SSMN variances (see supporting information B.3, especially the text surrounding equation (B.38)). This gave us an estimate of the minimum source structure effects present in the closure-delay-derived variance estimates. As shown in Table C.7, the model delay scatter for Group1 sources was just 1.6 ps for 0/1 weighting (1.5 ps for Calc/Solve weighting), whereas the All sources had a scatter of 13.7 ps for 0/1 weighting (10.1 ps for Calc/Solve weighting), in our mean statistics. Although the levels for All sources are significant, they are only a modest fraction of the total SSMN error budget. As with the SSMN values, the SS-model values should be considered as lower limits to the actual amount of relative source structure observations.

5.3. CSSMN Closure Delays

Taking the SS-model observation delays, we corrected the geodetic residual delays for the source structure we knew was present from our imaging analysis, providing corrected SSMN (CSSMN) delay residuals. As described in supporting information B.3, following an analogous procedure to our computations of the overall SSMN and

SS-model variances, we calculated CSSMN variances. These estimates showed the remaining closure errors that were not attributable to the simple SS-model delays from our imaging analysis.

Our SSMN, SS-model, and CSSMN variances (written as standard deviations) for the ionosphere-free delays are presented in Table C.7 in supporting information C.3. To conserve space, only the median and mean statistical estimates are shown. The minimum and maximum statistical estimates are only useful in demonstrating the range of values allowed by the data and otherwise say little about the probable values given the data.

The mean CSSMN scatter for Group1 sources was 10.2 ps for Calc/Solve weighting, while the All sources had a mean CSSMN scatter of 15.4 ps. The measurement noise clearly dominated the simple model structure for the Group1 sources. Similarly, the measurement noise for All sources was significantly higher than for the Group1 sources. Sources with significant source structure had significantly higher measurement noise in the CONT14 experiment even though all observations of all sources were scheduled to meet the same minimum SNR requirements. Using 0/1 weighting, the contrast between Group1 and All sources was even stronger. In other words, some of the measurement noise is caused by the presence of source structure.

5.4. Analysis of Delay Statistics

In Table 2, we summarize our statistical analysis of the CONT14 observation delay residuals and closure delay quantities, presenting the key results for our All and Group1 source subsets. The geodetic-analysis "Residual" row data were taken from the variance (written as standard deviation) column (5) of the ionosphere-free analysis results shown in Table C.5. (We duplicated the residual variances in both the mean and median closure analysis columns in Table 2, since the residual variance is independent of how the closure analysis was performed.) The ionosphere-free SSMN data are taken from Table C.6, and the ionosphere-free SS-model and CSSMN data are taken from Table C.7.

Our Group1 sources had lower levels of source structure compared to typical sources in CONT14, but they should have had similar error levels for station-based (closing) errors such as atmospheric delay errors, clock offsets, and station position errors. Based on our examinations of plots of the measured geodetic closure delays for CONT14 compared to the SS-model delays, the SS-model predictions were a good match to the measurements for these sources (see section 4). Since the Group1 sources had little structure, the measured amplitudes were never more than a factor of 2 below the amplitude levels expected by the scheduling software, and the SNR values achieved for the Group1 observations were all reasonably high. This means that the CSSMN variances for Group1 represent the typical closure variances one would expect in modern geodetic VLBI experiments in the absence of source structure. The Group1 CSSMN variances still include bandpass mismatch and polarization leakage, which made a nonnegligible contribution to the remaining error budget, so some source-structure-related effects were still left in the Group1 CSSMN variances. But, lacking suitable information to correct for polarization or spectral shapes at this point, we ignored this problem and treated the Group1 CSSMN variances as *upper limits* to the expected, source-structure-free closing variances for all sources.

Subtracting the Group1 CSSMN variances from the SSMN closure variances, for both the All and Group1 source groups, we obtained the full source-structure-related measurement variance (SS-full) with the source-structure-free measurement noise and nonclosing errors of the Group1 sources subtracted. The SS-full values represent our best estimate of the complete level of source-structure-related errors in the CONT14 geodetic delay analysis, combining the effects of both direct source structure delays and the excess measurement noise above the levels measured in compact sources.

Next, we subtracted the SSMN closure-based variances from the scan-baseline residual variances to get the station-based residual error level (the "Station" row). This row gives the combined level of all other usual station-based errors (atmosphere, clocks, station positions, cable delays, and so on), with the effects of relative source structure and measurement noise subtracted. Note that our station variances still include the effects of the absolute reference positions of the sources being different from the geodetic source model positions, but since we determined the mean source positions over the entire CONT14 experiment as part of our geodetic analysis, these absolute position errors should be small. Although we expect the station-based variances to be *similar* for the All and Group1 source subsets, we do not expect them to be exactly equal, especially for the 0/1 weighting analysis where the mean source positions were determined using the Calc/Solve weighting. Additional differences between the All and Group1 station-based variances can be present because of the automatic flagging of outlier points by the geodetic analysis software. The Group1 sources, with their low

source-structure delay variations, may enable the analysis software to more easily detect outliers, for example. On the other hand, sources with few observations or sources with many outliers or strong source structure may be unable to properly identify true outliers. For observations with large source-structure and measurement errors, such as occur in the All subset, station-based parameters such as tropospheric parameters and clock offsets will be less constrained, allowing larger station-based errors in the All subset. For both the 0/1 weighting and the Calc/Solve weighting, the combined station-based residual errors for the All and Group1 source subsets agree with each other reasonably well.

Then, looking at the All source results for our ionosphere-free analysis, the SSMN variances were approximately the same size as the station-based variances for the mean closure delay variance estimates, the best statistical estimates for the SSMN errors, for both the 0/1 (24.3 ps vs. 25.6 ps) and the Calc/Solve (19.0 ps vs. 16.6 ps) weighting schemes. For the more robust median SSMN variances that are less sensitive to the distribution of errors in the measurements, the SSMN variances were slightly lower than the station-based variances, with 22.7 ps versus 27.0 ps for the 0/1 weighted station-based variances and 17.2 ps versus 18.4 ps for the Calc/Solve SSMN variances. Both the mean and the median statistics show that nonclosing errors, SSMN, were nearly as large, or possibly larger than, all station-based errors combined.

In addition to our ionosphere-free analysis, we also analyzed source structure separately for the S band and X band measurements and present these results in Table 2 (the full analysis results are presented in Tables C.1 – C.4). Then, assuming that the S and X band errors are independent, we calculated the theoretical S + X ionosphere-free combination statistics assuming the standard ionosphere-free calculation factors of 1.08 for X band and 0.08 for S band (see supporting information A.5). Several features of our results are noteworthy. The simple source structure (SS-model) errors were a factor of about 3 larger in S band than X band in terms of picoseconds. However, in the ionospheric-free calculation, the X band simple source structure dominated because of the ionosphere-free calculation factors. In contrast, the S band measurement noise for our Group1 sources (roughly indicated by the CSSMN values) was more than twice the X band contribution to the ionosphere-free calculation ($135.9 \cdot 0.08 \approx 10.9$ vs. $4.5 \cdot 1.08 \approx 4.9$). The source structure effects in the ionosphere-free combination were set mostly by the X band contribution, while the measurement noise was dominated by the S band contribution. Separately, the source-structure effects at S and X bands were statistically independent. The statistical values presented in the "Iono-free" rows of Table 2 were calculated by combining the S and X band structure effects for each observation per baseline, whereas the "S + X" rows combined the overall variances from the "S" and "X" rows in the table. The fact that the SS-model and SS-full results from both computations agree to a fraction of a picosecond indicates that the S and X band structure effects were statistically independent over all observations. On the other hand, the SSMN and CSSMN values for the ionosphere-free calculation are 1.3 ps to 2.8 ps smaller than the S + X variance combination, indicating the nonclosing errors such as bandpass mismatch and polarization leakage do partially cancel in the ionosphere-free combination

5.5. Analysis of Phase Statistics

We also calculated source-structure and measurement noise effects based on an analysis of closure phase. VLBI phase measurements are far more precise than group delay measurements, so there is a great interest in the geodetic VLBI community in using phase delay measurements for future geodetic analysis, once the measurements and analysis techniques are accurate enough to avoid phase-wrap ambiguities. An analysis of the source-structure effects on phase-delay measurements is therefore of significant interest, to see whether source structure is more, just as, or less problematic.

We summarize our closure phase-delay analysis in Table 3 (the full analysis results are presented in Tables D.1 – D.4), showing the individual band analysis results for closure-based structure for S and X bands in units of degrees of phase and the ionosphere-free combination of the statistical results (S + X) in units of picoseconds. The phase-delay results should be directly compared to the corresponding group-delay results presented in Table 2.

The variances (standard deviations) of the closure phases were small enough, both for S band as well as X band, that there was little danger of making phase-wrap errors in the analysis. Although formally the phase variances should take the two-dimensional nature of measurements in the complex plane into account, given the small variances for most sources, we neglected this extra complexity for our analysis here.

Table 3
CONT14 Phase Statistics Summary

Data type	Band	Unit	0/1 weighting				Calc/Solve weighting			
			All sources		Group1		All sources		Group1	
			Median	Mean	Median	Mean	Median	Mean	Median	Mean
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
SSMN	S	°	10.7	12.1	7.6	9.0	9.9	11.3	7.1	8.4
SS-model	S	°	5.4	5.7	1.4	1.7	5.2	5.7	1.3	1.5
CSSMN	S	°	8.7	10.2	7.2	8.5	8.0	9.4	6.7	8.0
SS-full	S	°	7.9	8.6	2.4	3.0	7.3	8.0	2.3	2.6
SSMN	X	°	11.0	12.0	4.3	5.1	9.0	10.1	3.9	4.8
SS-model	X	°	9.2	9.7	1.0	1.0	7.3	8.1	0.8	0.9
CSSMN	X	°	5.2	6.2	3.9	4.8	4.4	5.4	3.6	4.5
SS-full	X	°	10.3	11.0	1.8	1.7	8.2	9.0	1.5	1.7
SSMN	S + X[a]	ps	4.1	4.5	1.7	2.0	3.4	3.8	1.6	1.9
SS-model	S + X[a]	ps	3.3	3.5	0.4	0.4	2.7	2.9	0.3	0.4
CSSMN	S + X[a]	ps	2.1	2.4	1.6	1.9	1.8	2.1	1.5	1.8
SS-full	S + X[a]	ps	3.8	4.0	0.7	0.7	3.0	3.3	0.6	0.7

Note. The data type codes have the following meanings: SSMN refers to the statistical analysis results from the full closure values, SS-model refers to the source-structure model closure statistics, CSSMN refers to the corrected (model source structure subtracted) closure statistics, and SS-full refers to the error amount from source structure plus excess measurement noise above the Group1 measurement noise.
The S + X band rows were formed by a simple addition of the variances of the S band rows and the X band rows after converting from degrees of phase to units of time delay, and allowing for the ionosphere-free band multiplication factors (so Var $\left(\tau_{\text{grp}}^{\text{iono-free}}\right) = 1.08^2\text{Var}\left(\phi_X\right) / \left[360° \cdot 8.4\text{GHz}\right]^2 + 0.08^2\text{Var}\left(\phi_S\right) / \left[360° \cdot 2.2\text{GHz}\right]^2$).

In units of degrees, the X band source structure was larger than the S band source structure (see the SS-model and SS-full rows). Taking the conversion factors to get the ionosphere-free phase delay statistics in units of time into account, the X band standard deviations are weighted about 3.5 times more than their S band counterparts. Thus, the X band source-structure phases dominate the ionosphere-free results. On the other hand, the measurement noise (see the CSSMN rows) was larger by about a factor of two for S band than X band, in units of degrees. But in contrast to the group-delay results, the ionosphere-free measurement noise in units of time still had a larger contribution from the X band measurements than from the S band measurements.

Comparing the group- and phase-delay results, the ionosphere-free group-delay measurement noise was about a factor of 8 times larger than the phase-delay measurement noise. Although this ratio is lower than one would expect based purely from the observing frequencies and band bandwidths, it still represents a tremendous improvement for phase-delay measurements over group delay. Interestingly, the simple source-structure effects (SS-model) are more than a factor of 3 smaller in units of picoseconds for the ionosphere-free phase-delay than they are for the ionosphere-free group-delay. For the full source-structure effects (SS-full, direct source structure phases/delays plus excess measurement noise), the phase-delay results are a factor of about 4.7 better than for group delay. Source structure is statistically significantly weaker in phase-delay measurements than it is in group-delay measurements. A move to the analysis of phase-delay measurements for standard S/X geodetic VLBI would not only greatly reduce the measurement noise, it would also greatly reduce the impact of source structure.

An analysis of the CSSMN values showed that using phase measurements, the measurement errors increased fractionally less for the All sources compared to the Group1 sources. The impact of source structure on phase noise was less severe than its impact on group-delay measurements. Furthermore, the group-delay simple source-structure effects (SS-model) were about two thirds the size of the measurement errors (CSSMN), but the simple source-structure for phase-delay was about 50% larger than the measurement noise in the ionosphere-free combination. Source structure can be better determined using phase measurements than though group delay.

6. Limitations of Previous Source Structure Studies

6.1. Structure Index Background

Based on the source-structure work of Charlot (1990), Fey and Charlot (1997), Ma et al. (1998), Fey and Charlot (2000), and Fey et al. (2015), other groups found that, for example, source structure contributes only 8 ps of delay error for typical IVS observations (Sovers et al., 2002) and source structure is negligible for European geodetic VLBI (Tornatore & Charlot, 2007). Why are our results so different from the conclusions of other studies?

The basic explanation is that through our analysis of closure quantities, we were able to measure the *full* effect of (relative) source structure, whereas previous studies only considered the effects of *simple* source structure—the simple delay models based on source structure. Previous studies did not account for the increased measurement noise in the actual VLBI observations that was caused by source structure, primarily the low measured brightnesses. Additional source structure effects, including the limitations in the source-structure delay model such as spectral curvature over the finite observation bandwidth and errors resulting from imperfect determinations of the source structure from imaging measurements, also contributed to the true amount of source-structure effect on the measurement residuals being higher than estimated by previous studies, but the main difference revealed by our closure analysis was the measurement-noise dependence on source structure.

6.2. CONT14 Residual Error Distributions

The outlier flagging analysis performed by Calc/Solve works on baseline measurements, not closures, and therefore, the outlier flagging must contend with all of the closing (station-based) errors in addition to all of the nonclosing (SSMN) errors.

As shown by Xu et al. (2016), the VLBI delay error distribution is *not* normally distributed—it exhibits a strong tail at high error values. Figure 13 illustrates typical residual error distributions of the CONT14 sources for sources showing small, large, and intermediate residual scatter levels in Table C.5. We show histograms of the solution residuals, rather than histograms of the source structure model delay errors, to emphasize the full errors present in actual geodetic data sets.

All sources showed a significant tail of residual absolute delay error, extending up to 25 ns (note the units of *nanoseconds*!). In section 4.3.4, we have shown that many of these points were caused by source structure (although the measurement errors are far larger than the model structure delays); the remaining points were assumed to be corrupted measurements. Through various means, the geodetic analysis software tried to identify such points to be flagged. (More information about these points will be provided below.) However, in the range of absolute residual error of 0 ps to 500 ps, each source presented a unique residual distribution. With a factor of 15 times as many observations per source in the CONT14 experiment as compared to regular 24-hr IVS sessions, the residual distributions for CONT14 were of far better quality for interpretation. Sources with low residual scatters had steep declines with absolute residual level in the range from 0 ps to 200 ps, while sources with large residual scatter had shallower declines, as expected.

The vertical axes are log-scaled, so that normally distributed residuals should show a single linear slope as a function of residual level that is related to the standard deviation of the distribution. But Figure 13 shows that all of the sources presented had variable slopes as a function of residual error. All of the sources had a relatively shallow distribution over the 0 ps to 20 ps range, that steepened beyond 20 ps, before flattening somewhat again around 70 ps to 100 ps. This shallower distribution extended out to several hundred picoseconds or more before becoming completely flat and extending out to tens of nanoseconds. The slopes from ~20 ps to several hundred picoseconds, and the relative number counts of observations per bin to the peak of the distribution, were source dependent. The shallow tail distributions at high residual errors caused the overall standard deviations to be significantly larger than would be predicted by the median absolute residual delay.

The geodetic VLBI analysis software spent a considerable effort in identifying outlier points and flagging them away (equivalent to setting their weights to 0). This removed the huge nanosecond residuals that were often related to true measurement errors/blunders, but it also removed a significant portion of the hundreds of picoseconds residuals that appear to be related to source structure in these figures, as shown in section 4.3.4. Effectively, all residuals above an absolute value of about 250 ps were flagged, regardless of the cause of the residual level, for all sources. Very strong structure in any source was automatically flagged by Calc/Solve.

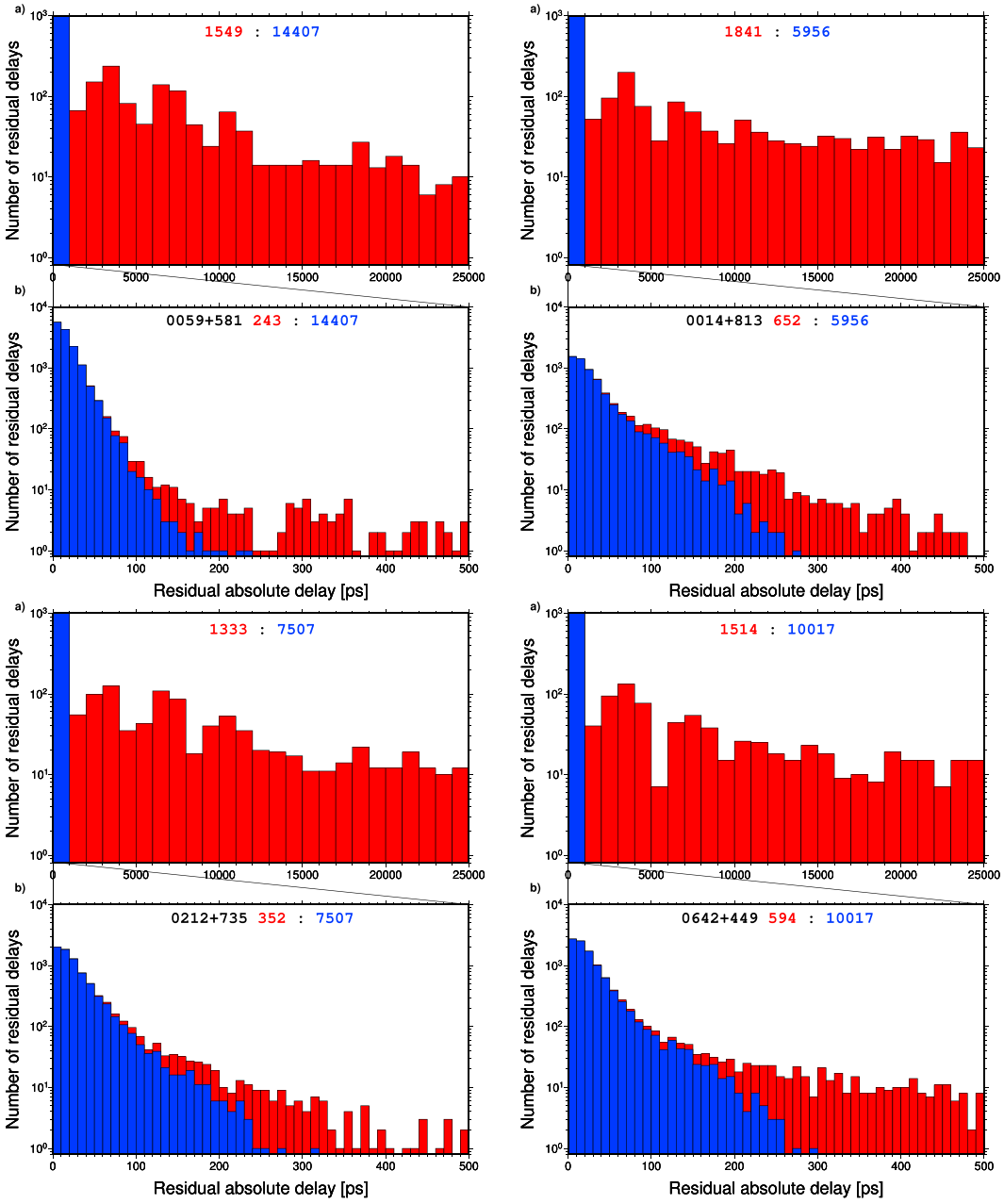

Figure 13. Histograms of absolute ionosphere-free residual-delay error in the Calc/Solve CONT14 solutions. Coarse histograms for each source show delay errors out to 25 ns, with histograms showing the fine details of the 0 ps to 500 ps absolute residual delay region set immediately below the coarse histograms. Sources were selected to show low (0059+581, top left), high (0014+813, top right), and intermediate (0212+735, bottom left, and 0642+449, bottom right) residual root-mean-square levels. Red values indicate measurements (observables) flagged by Calc/Solve, while blue values indicate points retained in the Calc/Solve analysis. All sources had significant numbers of flagged observations above an absolute residual level of 200 ps. The blue histogram area for 0059+581 shows a steep drop-off with increasing delay error, whereas the slope for 0014+813 is much flatter. Many 0014+813 observations have been flagged in the 100 ps to 250 ps region, but it is not clear from this plot whether the points are truly outliers or are part of the real source structure distribution. The high residual error tail for 0642+449 below 1 ns is much more significant than for 0212+735, even though both sources had approximately equal residual root-mean-square levels after flagging. All sources show significantly nonnormal error distributions for the measurements. Red and blue numbers in each subplot give the number of flagged and unflagged observations, respectively, in the plots.

VLBI observations that tended to show large source-structure delays also tended to have low source brightnesses—and thus high measurement noises. The flagging of measurements above some absolute residual level therefore preferentially eliminated the largest source-structure errors in data sets and correspondingly preferentially eliminated measurements from sources with the most complicated structure. Hence, sources with the largest structure errors had relatively more observations flagged out than minimal-structure sources. (The fact that many of the sources with the most structure in CONT14 also had very large total brightnesses did not impact this process, since the scheduled scan durations were adjusted to provide a fixed SNR assuming a simple—typically just a single 2-D Gaussian—source-structure model.)

The statistical analysis performed in Calc/Solve to attempt to identify outliers also appeared to depend on the number of observations available. Utilizing 17 stations, many more stations than standard IVS sessions, CONT14 produced 3–4 times more observations per session than usual. The large number of observations has, we believe, allowed the Calc/Solve software and analyst to perform a "better" job of flagging out observations containing large structure delays than was usually possible.

Considerable effort was also spent by Calc/Solve in determining a set of weights to minimize the impact of high-residual observations that are not completely flagged away. This process used residual delay offsets determined at intermediate processing steps to reweight the data, effectively reinforcing the initial intermediate parameterization. Details of the Calc/Solve reweighting process can be found in Gipson (1997). The flagging and downweighting of large residual points in Calc/Solve brought the overall RMS level of the final residuals down, but although the final parameterization was more precise, it was not necessarily more accurate. In any case, the Calc/Solve flagging resulted in a simple source-structure model scatter of just 13.7 ps for 0/1 weighting, and with the Calc/Solve weights applied, this scatter was only 10.1 ps. These scatter levels were significantly *lower* than predictions of source structure effects using median absolute deviation (MAD) continuous structure indices from the Bordeaux VLBI Image Database. We estimated simple source-structure residual levels of 14.2 ps using the Bordeaux VLBI Image Database MAD values directly and 21.1 ps by assuming normally distributed source-structure errors with the corresponding MAD values, scaling by the number of observations per source.

Thus, even though the CONT14 experiment included more sources with "large" amounts of structure than other typical IVS experiments in recent years, most of the strong structure effects in those sources were actually flagged away, and strong structure effects probably contributed "less" per measurement to the final CONT14 analysis than such structure contributes to other IVS experiments.

6.3. Understanding Prior Analysis Results

We examined the analyses of a selection of previous results studying source structure in geodetic VLBI observations in order to understand why earlier estimates of the impact of source structure are so different from our results. Tornatore and Charlot (2007) examined source structure for European baselines. They computed structure indices from their imaging results only for projected baselines less than 4,780 km, compared to the standard maximum baseline length of two Earth radii (12,742 km) for structure indices indented for use for global geodetic VLBI sessions. Naturally, the greatly reduced baseline lengths used in the European session as compared to standard global sessions resulted in significantly reduced estimates of source-structure effects.

In contrast, the analysis of Sovers et al. (2002) was performed on observations including more globally distributed stations, with an important caveat—the 10 so-called *RDV* (Research and Development VLBI) sessions they analyzed had more than half of the stations made up of VLBA stations. The 10-station VLBA has many short (by geodetic VLBI standards) baselines in North America. Sovers et al. (2002) used imaging data derived from the visibility data of the same observations, imaging each 24-hr session independently, to derive source-structure corrections. By applying the structure corrections, they were able to remove about 8 ps of error in quadrature. First, we note that the 8 ps is only a lower limit to the simple source structure actually present in the measurements; 8 ps is the amount of source-structure error that they were able to remove. We presume that errors in the determination of the source-structure models, such as measurement noise and limited (u, v) coverage resulting from only one 24-hr session's worth of data being available, limited their ability to correct for source structure in the geodetic data. The true amount of simple source-structure must therefore be even larger than 8 ps.

Second, we return to the issue of the baseline configurations used for the observations. With observations dominated by the 10 VLBA stations in North America, the observations in the sessions analyzed by Sovers

et al. (2002) involved many shorter than usual baselines, resulting in less source-structure being observed when weighted over all observations, than a typical global IVS R1 or R4 session, utilizing typically longer baselines, would have. Combining these two effects, we would expect R1 and R4 sessions observing similar sources to have significantly more simple source-structure than 8 ps. This hypothesis needs to be tested in the near future.

Finally, we note again that these studies focused on only the simple source-structure effects and ignored the effects that source structure introduces on the measurement noise.

7. Conclusions

Source-structure-related errors in geodetic VLBI are about as large as all other error sources combined.

Analyzing the CONT14 data set using closure analysis and using Calc/Solve weighting, we found that nonclosing errors, predominately coming from source structure and measurement errors, account for 19.0 ps of the residual error (standard deviation). Given the total residual error level of 25.2 ps, this only leaves 16.6 ps of error (in quadrature) for all station-based errors combined. Using model source-structure information determined from imaging the same observations, we found that the simple model prediction of source structure, that is, only considering the linear delay impact of the source-structure model and ignoring amplitude, frequency curvature, and other nonclosing effects, predicts just 10.1 ps of source-structure error. But, using the six of the most-compact sources to correct for the typical measurement noise level, we found that the full effect of source structure, including amplitude and other effects, is at least 16.0 ps. The full effect of source structure on geodetic residuals is about as large as all station-based errors combined.

Given that we only measured relative source structure and eliminated absolute position errors from our analysis, our results are only lower limits to the full amount of source structure that was present in CONT14.

The full impact of source structure is far greater than the simplistic structure delay models predict, principally because those models have ignored the decrease in source brightness for baselines that see significant source structure, and the corresponding loss of SNR and increase in measurement noise, that accompanies lower brightnesses than are expected by the scheduling software. Given good models of source structure, the structure delays can be corrected in geodetic VLBI analysis. The loss of SNR cannot be corrected. Future geodetic VLBI observations need to include far better source-structure models for scheduling VLBI observations, in order to prevent lower than expected amplitudes from being common. The bivariate noise distribution causes phase errors, and hence group delay errors, to be larger at low SNRs than a simple proportionality to inverse SNR, so that correct brightness predictions will be even more important for VGOS observations, where the scheduled SNR for each band might only be 7, versus the current scheduled SNR of 20 for X band (Petrachenko & Bérubé, 2007). Allowing the same fractional decrease in brightness would yield proportionally larger noise levels when the scheduled SNR is lower and dramatically increase the probability of finding extreme (random) delay outliers, which rapidly become more probably below SNR values of 6–7. Alternatively, station-based, rather than baseline-based, fringe-fitting, taking full account of source structure on each baseline and using the full frequency-dependent complex visibilities, could be significantly less sensitive to low-brightness baselines when many stations are present.

The scheduling of geodetic VLBI sessions should be improved to better accommodate the effects of source structure on the final measurements, in order to both account for the structure delay scatter introduced by source structure, as well as to better account for the lower baseline flux densities resulting from source structure. Improved scheduling will require better models of the expected structure for individual sources to be maintained.

Closure analysis is insensitive to the reference directions of the sources. Not needing to match the geodetic and imaging reference directions simplified our analysis here, but *correcting* observations for source structure in order to better determine geodetic and astrometric parameters of interest requires knowing the structure with respect to stable reference directions. One challenge is to determine how to identify the existing geodetic reference directions with respect to the existing source structure. Another challenge is to identify a stable reference direction for each source that can routinely be determined in spite of the time and frequency dependence of the source structure, with jet component creation and motion as well as potential time-variable core shift and core brightness variability, to an accuracy far smaller than the imaging beamsize. Although the quoted formal position uncertainties for astronomical studies fitting 2-D Gaussian components to high SNR single-frequency data can be below 100 microarcseconds at geodetic frequencies (see, e.g., Schinzel, 2011),

comparisons to other fitting techniques suggest that actual positional errors can be much larger (Mertens & Lobanov, 2015).

Future studies will need to determine how well the image-plane structure needs to be determined in order to adequately correct the geodetic observables and how well standard geodetic observations can determine source structure. For example, CONT14 has roughly a factor of 60 times more phase and amplitude imaging constraints after calibration than typical 24-hr IVS sessions because of its large number of participating stations and observing for 15 days. Individual IVS sessions as they are currently scheduled might not be capable of providing sufficient information to adequately determine source structure, and temporally nearby sessions will probably need to be combined to provide suitable structure constraints. Significant collaboration with the astronomical VLBI community is required to make progress in these areas.

On the other hand, baseline structure amplitudes, which control the measurement noise, are not dependent upon the determination of the reference direction, and the structure models to aid in scheduling observations to avoid low SNR measurements or to aid in weighting measurements might be optimally constructed in a different manner from structure models intended to predict phase and delay corrections.

Beyond a phase gradient (group delay), higher-order spectral curvature can be caused by various effects, including ionospheric delay and ionospheric Faraday rotation as well as source structure. For baselines in a triangle that use different subbands (different subbands flagged at different stations) or otherwise weight the subbands differently (sensitivity issues), this results in closure errors. We intend to investigate such effects on our closure analysis and residual group-delay analysis in the future.

We plan to conduct our analysis on other geodetic VLBI data sets, including R1 and R4 sessions in addition to the CONT17 experiment, in order to evaluate the level of source-structure-related errors in more typical IVS data sets. Because the high source-structure measurements of sources showing strong structure in our CONT14 analysis were typically flagged away during the Calc/Solve analysis, it is not clear that more typical sessions that do not include such strong structure sources necessarily have lower structure effects.

We also intend to investigate our ability to remove the effects of source structure as well as physically possible. For the best results, we expect that VLBI data will need to be corrected for source structure at the visibility stage prior to determining group delays or phase delays. We intend to investigate the temporal stability of geodetic VLBI sources in order to know how distant in time, from the epoch when source structure was measured, such corrections continue to provide adequate calibration for geodetic scheduling and analysis. This may, in turn, require a better definition of how well source structure has been measured in an image, such as perhaps casting the quality in terms of expected geodetic delay residuals rather than in terms of component brightnesses relative to the image background noise.

These studies will address some basic concerns for the VLBI community and give some practical ideas on how to move forward in the way of correcting structure effects in actual measurements and thereby improve the geodetic performance of VLBI.

Acknowledgments

We would like to thank David Gordon for advice on optimizing our analysis using Calc/Solve. The results published in this paper were produced using the data owned by the International VLBI Service (IVS) and its international self-funded member organizations. The raw data can be obtained from Internet servers described in the relevant sections of this article. Plots showing the closure measurements and source-structure models for all sources and triangles can be found in Anderson and Xu (2018). This research has made use of material from the Bordeaux VLBI Image Database (BVID; http://www.astrophy.u-bordeaux.fr/BVID/content.html). Figures were produced using the Generic Mapping Tools package (Wessel et al., 2013). We would like to thank the anonymous reviewers whose comments helped to improve this manuscript.

J. M. Anderson was funded by the project Ties between kinematic and dynamic reference frames (D-VLBI) (SCHU 1103/4-1) as part of the DFG Research Unit Space-Time Reference Systems for Monitoring Global Change and for Precise Navigation in Space funded by the German Research Foundation (FOR 1503). Ming H. Xu was funded by the Young Fund from the National Natural Science Foundation of China (Y646041001).

References

Altamimi, Z., Rebischung, P., Métivier, L., & Collilieux, X. (2016). ITRF2014: A new release of the International Terrestrial Reference Frame modeling nonlinear station motions. *Journal of Geophysical Research: Solid Earth*, *121*, 6109–6131. https://doi.org/10.1002/2016JB013098

Anderson, J. M., & Xu, M. H. (2018). Closure plots for "Source structure and measurement noise are as important as all other residual sources in geodetic VLBI combined", *GFZ data services*. Potsdam, Germany: GFZ German Research Centre for Geosciences. https://doi.org/10.5880/GFZ.1.1.2018.001

Behrend, D. (2013). Data handling within the international VLBI service. *Data Science Journal*, *12*, WDS81–WDS84. https://doi.org/10.2481/dsj.WDS-011

Bertarini, A., Roy, A. L., Corey, B., Walker, R. C., Alef, W., & Nothnagel, A. (2011). Effects on geodetic VLBI measurements due to polarization leakage in S/X receivers. *Journal of Geodesy*, *85*, 715–721. https://doi.org/10.1007/s00190-011-0478-7

Boehm, J., Werl, B., & Schuh, H. (2006). Troposphere mapping functions for GPS and very long baseline interferometry from European Centre for Medium-Range Weather Forecasts operational analysis data. *Journal of Geophysical Research*, *111*, B02406. https://doi.org/10.1029/2005JB003629

Bridle, A. H., & Greisen, E. W. (1994). *The NRAO AIPS project—A summary* Technical Report AIPS Memo 87, The National Radio Astronomy Observatory. Retrieved from https://doi.org/ftp://ftp.aoc.nrao.edu/pub/software/aips/TEXT/PUBL/AIPSMEMO87.PS

Charlot, P. (1990). Radio-source structure in astrometric and geodetic very long baseline interferometry. *Astronomical Journal*, *99*, 1309–1326. https://doi.org/10.1086/115419

Clark, B. G. (1981). *Orders of magnitude of some instrumental effects* (Technical Report VLA Scientic Memorandum No. 137), National Radio Astronomy Observatory. Retrieved from http://library.nrao.edu/public/memos/vla/sci/VLAS137.pdf

Fey, A. L., & Charlot, P. (1997). VLBA observations of radio reference frame sources. II. Astrometric suitability based on observed structure. *Astrophysical Journal Supplement Series*, *111*, 95–142. https://doi.org/10.1086/313017

Fey, A. L., & Charlot, P. (2000). VLBA observations of radio reference frame sources. III. Astrometric suitability of an additional 225 sources. *Astrophysical Journal Supplement Series*, *128*, 17–83. https://doi.org/10.1086/313382

Fey, A. L., Clegg, A. W., & Fomalont, E. B. (1996). VLBA observations of radio reference frame sources. I. *Astrophysical Journal Supplement Series*, *105*, 299–330. https://doi.org/10.1086/192315

Fey, A. L., Gordon, D., Jacobs, C. S., Ma, C., Gaume, R. A., Arias, E. F., et al. (2015). The second realization of the International Celestial Reference Frame by very long baseline interferometry. *Astronomical Journal*, *150*, 58. https://doi.org/10.1088/0004-6256/150/2/58

Fomalont, E. B., & Perley, R. A. (1999). Calibration and editing. In G. B. Taylor, C. L. Carilli, & R. A. Perley (Eds.), *Synthesis imaging in radio astronomy II*: Astronomical Society of the Pacific Conference Series (pp. 79–110).

Gipson, J. M. (1997). *Constraints, covariances and* χ^2 *in Solve* (*Technical Report*), National Aeronautics and Space Administration. https://doi.org/http://lacerta.gsfc.nasa.gov/mk5/help/reway_03_help.ps.gz

Gipson, J. (2010). An introduction to SKED. In R. Navarro, et al. (Eds.), *Sixth international VLBI service for geodesy and astronomy. Proceedings from the 2010 General Meeting, "VLBI2010: From Vision to Reality"*, (pp. 77–84). Hobart, Tasmania, Australia: National Aeronautics and Space Administration.

Greisen, E. W. (1990). The astronomical image processing system. In G. Longo & G. Sedmak (Eds.), *Acquisition, processing and archiving of astronomical images* (pp. 125–142).

Halsig, S., Artz, T., Iddink, A., & Nothnagel, A. (2016). Using an atmospheric turbulence model for the stochastic model of geodetic VLBI data analysis. *Earth, Planets, and Space*, *68*, 106. https://doi.org/10.1186/s40623-016-0482-5

Heinkelmann, R., Söhne, W., & Schuh, H. (2013). Comparison of GNSS (EUREF) and VLBI (EVGA) tropospheric delays, *EGU general assembly conference abstracts* (Vol. 15, pp. EGU2013–7942). Vienna, Austria. Retrieved from http://adsabs.harvard.edu/abs/2013EGUGA..15.7942H

Hovatta, T., Aller, M. F., Aller, H. D., Clausen-Brown, E., Homan, D. C., Kovalev, Y. Y., et al. (2014). MOJAVE: Monitoring of jets in active galactic nuclei with VLBA experiments. XI. Spectral distributions. *The Astronomical Journal*, *147*, 143. https://doi.org/10.1088/0004-6256/147/6/143

Kettenis, M., & Sipior, M. (2012). Parseltongue: AIPS Python interface. Astrophysics Source Code Library. http://ascl.net/1208.020

Kettenis, M., van Langevelde, H. J., Reynolds, C., & Cotton, B. (2006). Parseltongue: AIPS talking Python. In C. Gabriel, C. Arviset, D. Ponz, & S. Enrique (Eds.), *Astronomical data analysis software and systems XV* (Vol. 351, pp. 497–500). San Lorenzo de El Escorial, Spain: Astronomical Society of the Pacific Conference Series (San Francisco).

Ma, C., Arias, E. F., Eubanks, T. M., Fey, A. L., Gontier, A.-M., Jacobs, C. S., et al. (1998). The International Celestial Reference Frame as realized by very long baseline interferometry. *Astronomical Journal*, *116*, 516–546. https://doi.org/10.1086/300408

Ma, C., Sauber, J. M., Clark, T. A., Ryan, J. W., Bell, L. J., Gordon, D., & Himwich, W. E. (1990). Measurement of horizontal motions in Alaska using very long baseline interferometry. *Journal of Geophysical Research*, *95*, 21,991–22,011. https://doi.org/10.1029/JB095iB13p21991

MacMillan, D. S. (2017). EOP and scale from continuous VLBI observing: CONT campaigns to future VGOS networks. *Journal of Geodesy*, *91*, 819–829. https://doi.org/10.1007/s00190-017-1003-4

MacMillan, D., & Sharma, R. (2008). *Sensitivity of VLBI2010 simulations to parameterization of input simulated turbulence, clock, and white noise* (Technical Report 2008-011v01), International VLBI Service for Geodesy and Astrometry. Retrieved from https://ivscc.gsfc.nasa.gov/publications/memos/ivs-2008-011v01.pdf

Massi, M., Comoretto, G., Rioja, M., & Tofani, G. (1996). Baseline errors on European VLBI Network measurements. II. Instrumental polarization. *Astronomy and Astrophysics Supplement Series*, *116*, 167–176.

Massi, M., Rioja, M., Gabuzda, D., Leppaenen, K., Sanghera, H., Ruf, K., & Moscadelli, L. (1997). Baseline errors in European VLBI Network measurements. III. The dominant effect of instrumental polarization. *Astronomy and Astrophysics*, *318*, L32—L34.

Massi, M., Tofani, G., & Comoretto, G. (1991). Baseline errors on VLBI measurements. I—Analytic description. *Astronomy and Astrophysics*, *251*, 732–736.

Mertens, F., & Lobanov, A. (2015). Wavelet-based decomposition and analysis of structural patterns in astronomical images. *Astronomy and Astrophysics*, *574*, A67. https://doi.org/10.1051/0004-6361/201424566

Nilsson, T., & Haas, R. (2010). Impact of atmospheric turbulence on geodetic very long baseline interferometry. *Journal of Geophysical Research*, *115*, B03407. https://doi.org/10.1029/2009JB006579

Nothnagel, A. (2009). Conventions on thermal expansion modelling of radio telescopes for geodetic and astrometric VLBI. *Journal of Geodesy*, *83*, 787–792. https://doi.org/10.1007/s00190-008-0284-z

Nothnagel, A., International VLBI Service for Geodesy and Astrometry (IVS), Alef, W., Amagai, J., Andersen, P. H., Anderson, J., Andreeva, T., et al. (2015). The IVS data input to ITRF2014, *GFZ data services*. Potsdam, Germany: Helmoltz Centre. https://doi.org/10.5880/GFZ.1.1.2015.002

Pany, A., Böhm, J., MacMillan, D., Schuh, H., Nilsson, T., & Wresnik, J. (2011). Monte Carlo simulations of the impact of troposphere, clock and measurement errors on the repeatability of VLBI positions. *Journal of Geodesy*, *85*, 39–50. https://doi.org/10.1007/s00190-010-0415-1

Pany, A., Wresnik, J., & Böhm, J. (2008). *Vienna VLBI2010 PPP simulator* (Technical Report 2008-012v01): International VLBI Service for Geodesy and Astrometry. Retrieved from https://ivscc.gsfc.nasa.gov/publications/memos/ivs-2008-018v01.pdf

Petit, G., & Luzum, B. (2010). IERS conventions (2010). IERS Technical Note, 36, https://www.iers.org/IERS/EN/Publications/TechnicalNotes/tn36.html.

Petrachenko, B., & Bérubé, M. (2007). *VLBI2010 source map alignment simulation* (Technical Report 2007-008v01), International VLBI Service for Geodesy and Astrometry. Retrieved from https://ivscc.gsfc.nasa.gov/publications/memos/ivs-2007-008v01.pdf

Petrachenko, B., Niell, A., Behrend, D., Corey, B., Boehm, J., Charlot, P., et al. (2009). *Design aspects of the VLBI2010 system. Progress report of the IVS VLBI2010 Committee* (Technical Report): International VLBI Service for Geodesy and Astrometry, NASA/TM-2009-214180. Retrieved from https://space-geodesy.nasa.gov/docs/2009/TM-2009-214180.pdf

Petrov, L., & Boy, J.-P. (2004). Study of the atmospheric pressure loading signal in very long baseline interferometry observations. *Journal of Geophysical Research*, *109*, B03405. https://doi.org/10.1029/2003JB002500

Plag, H.-P., & Pearlman, M. (2009). *Global geodetic observing system*. Berlin: Springer-Verlag. https://doi.org/10.1007/978-3-642-02687-4

Rogers, A. E. E. (1983). Subject of closure errors. The MK III experience and the VLBA specifications (Technical Report VLB Array Memo No. 253): National Radio Astronomy Observatory. https://library.nrao.edu/public/memos/vlba/main/VLBA_253.pdf

Schinzel, F. K. (2011). Physics and kinematics of the Parsec Scale Jet of the Quasar 3C345 (PhD thesis), Max-Planck-Institut für Radioastronomie. frank@schinzels.de, http://kups.ub.uni-koeln.de/4400

Schlüter, W., & Behrend, D. (2007). The International VLBI Service for Geodesy and Astrometry (IVS): Current capabilities and future prospects. *Journal of Geodesy*, *81*, 379–387. https://doi.org/10.1007/s00190-006-0131-z

Schuh, H., & Behrend, D. (2012). VLBI: A fascinating technique for geodesy and astrometry. *Journal of Geodynamics*, *61*, 68–80. https://doi.org/10.1016/j.jog.2012.07.007

Shabala, S. S., McCallum, J. N., Plank, L., & Böhm, J. (2015). Simulating the effects of quasar structure on parameters from geodetic VLBI. *Journal of Geodesy*, *89*, 873–886. https://doi.org/10.1007/s00190-015-0820-6

Sovers, O. J., Charlot, P., Fey, A. L., & Gordon, D. (2002). Structure corrections in modeling VLBI delays for RDV data. In N. R. Vandenberg & K. D. Baver (Eds.), *International VLBI Service for Geodesy and Astrometry: General meeting proceedings* (243 pp.) Tsukuba, Japan.

Sovers, O. J., Fanselow, J. L., & Jacobs, C. S. (1998). Astrometry and geodesy with radio interferometry: Experiments, models, results. *Reviews of Modern Physics*, *70*, 1393–1454. https://doi.org/10.1103/RevModPhys.70.1393

Thompson, A. R. (1984). *Tolerances on polarization mismatch* (Technical Report VLB Array Memo No. 346), National Radio Astronomy Observatory. Retrieved from https://library.nrao.edu/public/memos/vlba/main/VLBA_346.pdf

Thompson, A. R., & D'Addario, L. R. (1982). Frequency response of a synthesis array—Performance limitations and design tolerances. *Radio Science*, *17*, 357–369. https://doi.org/10.1029/RS017i002p00357

Thompson, A. R., Moran, J. M., & Swenson, G. W. Jr. (2017). *Interferometry and synthesis in radio astronomy* (3rd ed.). Cham, Switzerland: Springer International Publishing AG. https://doi.org/10.1007/978-3-319-44431-4

Tornatore, V., & Charlot, P. (2007). The impact of radio source structure on European geodetic VLBI measurements. *Journal of Geodesy*, *81*, 469–478. https://doi.org/10.1007/s00190-007-0146-0

Wessel, P., Smith, W. H. F., Scharroo, R., Luis, J., & Wobbe, F. (2013). Generic mapping tools: Improved version released. *Eos, Transactions American Geophysical Union*, *94*(45), 409–410. https://doi.org/10.1002/2013EO450001

Wresnik, J., Pany, A., & Böhm, J. (2008). *Impact of different turbulence, clock and white noise parameters on VLBI2010 simulation results with OCCAM Kalman filter* (Technical Report 2008-007v02), International VLBI Service for Geodesy and Astrometry. Retrieved from https://ivscc.gsfc.nasa.gov/publications/memos/ivs-2008-007v02.pdf

Xu, M. H., Heinkelmann, R., Anderson, J. M., Mora-Diaz, J., Karbon, M., Schuh, H., & Wang, G. L. (2017). The impacts of source structure on geodetic parameters demonstrated by the radio source 3C371. *Journal of Geodesy*, *91*, 767–781. https://doi.org/10.1007/s00190-016-0990-x.

Xu, M. H., Heinkelmann, R., Anderson, J. M., Mora-Diaz, J., Schuh, H., & Wang, G. L. (2016). The source structure of 0642+449 detected from the CONT14 observations. *Astronomical Journal*, *152*, 151. https://doi.org/10.3847/0004-6256/152/5/151