

THE VERY LONG BASELINE ARRAY GALACTIC PLANE SURVEY—VGaPS

L. PETROV¹, Y. Y. KOVALEV^{2,3}, E. B. FOMALONT⁴, AND D. GORDON⁵

¹ ADNET Systems Inc./NASA GSFC, Greenbelt, MD 20771, USA; Leonid.Petrov@lpetrov.net

² Astro Space Center of Lebedev Physical Institute, 117997 Moscow, Russia; yyk@asc.rssi.ru

³ National Radio Astronomy Observatory, Green Bank, WV 24944, USA

⁴ National Radio Astronomy Observatory, Charlottesville, VA 22903-2475, USA; efomalon@nrao.edu

⁵ NVI Inc./NASA GSFC, Code 698.2 Greenbelt, MD 20771, USA; David.Gordon-1@nasa.gov

Received 2011 January 6; accepted 2011 May 1; published 2011 June 22

ABSTRACT

This paper presents accurate absolute positions from a 24 GHz Very Long Baseline Array (VLBA) search for compact extragalactic sources in an area where the density of known calibrators with precise coordinates is low. The goals were to identify additional sources suitable for use as phase calibrators for galactic sources, determine their precise positions, and produce radio images. In order to achieve these goals, we developed a new software package, PIMA, for determining group delays from wide-band data with much lower detection limits. With the use of PIMA, we have detected 327 sources out of 487 targets observed in three 24 hr VLBA experiments. Among the 327 detected objects, 176 are within 10° of the Galactic plane. This VGaPS catalog of source positions, plots of correlated flux density versus projected baseline length, contour plots, as well as weighted CLEAN images, and calibrated visibility data are available on the Web in FITS format. Approximately one-half of objects from the 24 GHz catalog were observed at dual-band 8.6 GHz and 2.3 GHz experiments. Position differences at 24 GHz versus 8.6/2.3 GHz for all but two objects on average are strictly within reported uncertainties. We found that for two objects with complex structures, positions at different frequencies correspond to different components of a source.

Key words: astrometry – catalogs – surveys

Online-only material: color figures, machine-readable and VO tables

1. INTRODUCTION

The method of very long baseline interferometry (VLBI), first proposed by Matveenko et al. (1965), has numerous applications in the areas of high-resolution imaging, differential astrometry, absolute astrometry, space geodesy, and space navigation. Because the turbulence in the neutral atmosphere and ionospheric fluctuations set a limit of coherent averaging at typically 1–10 minutes, depending on frequency, the detection of weak sources that require longer integration is not possible.

To overcome this limitation, the majority of VLBI observations are made in phase-referencing mode: the telescopes of the array slew rapidly between a weak target and a nearby strong calibrator. The phase changes of the calibrator trace the fluctuations in the atmosphere, and when they are subtracted from the phase of the target, the residual phases are essentially free from fluctuations caused by the atmosphere, and the target integration time can be extended almost indefinitely, enabling detection and imaging of weak objects. This technique is called phase referencing.

The technique of phase referencing also allows us to determine the precise differential position of a target with respect to a calibrator with accuracy reaching 0.05 mas or better, even with moderate signal-to-noise ratios (SNRs). The advantage of differential astrometry over absolute astrometry is that the contribution of unaccounted propagation delays and errors in station positions is diluted by a factor of the target-to-source separation in radians. Either the target or the calibrator may be observed in a narrow band.

The ability to image weak sources and determine their positions accurately with respect to a nearby calibrator have made phase referencing very popular. According to Wrobel

(2009), in 2003–2008, 63% of Very Long Baseline Array (VLBA) observations used this technique. However, to make phase referencing possible, a dense catalog of phase calibrators is needed such that a suitable calibrator will be found within 2° – 3° of any target, and for precise differential astrometry, a calibrator position accuracy of a few milliarcseconds is needed. Efforts to create such a catalog of calibrators commenced in the 1980s under the NASA's Crustal Dynamic Project program (Ryan & Clark 1987) which ultimately resulted in the ICRF catalog of 608 sources (Ma et al. 1998). Later, over 6000 sources were observed in the framework of the VLBA Calibrator Survey (VCS) program (Beasley et al. 2002; Fomalont et al. 2003; Petrov et al. 2005, 2006; Kovalev et al. 2007; Petrov et al. 2007b), the VLBA RDV program (Petrov et al. 2009), and the continuing Australian Long Baseline Array Calibrator Survey program (Petrov et al. 2011). By the end of 2010, the number of known calibrators with position accuracies better than 5 mas surpassed 4600. The probability of finding a calibrator within 2° of any direction is currently 64%, and within 3° is 90%. However, the distribution of calibrators is not uniform on the sky due to several factors: the large-scale structure of the universe, the location of most observing stations in the northern hemisphere, and obscuration and confusion within 5° of the Galactic plane.

Finding calibrators in the Galactic plane region is especially difficult for several reasons. First, the region is filled with many Galactic objects, and surveys from single antennas or kilometer-sized arrays, which are needed to find calibrator candidates, avoid this region; for example, the Jodrell/VLA Astrometric Survey (Wrobel et al. 1998) and AT20G (Murphy et al. 2010). Therefore, the pool of available candidates is limited near the plane. Second, many potential candidates with flat spectra

are extended Galactic objects, such as planetary nebulae or H II regions, that cannot be detected by VLBI. Finally, the apparent angular size of extragalactic objects observed through high plasma density near the Galactic plane are broadened by Galactic scattering, and cannot be detected at low frequencies on baselines longer than several thousand kilometers.

Nevertheless, a dense grid of calibrators in the Galactic plane is needed for studies of compact galactic objects in both continuum emission (pulsars, X-ray binaries) and line emission (e.g., water masers, methanol masers, hydrogen absorption). Some extragalactic sources near the Galactic plane have been associated with *Fermi*-detected γ -ray objects (Abdo et al. 2010a, 2010b) only through VLBI calibrator surveys as suggested and successfully shown by Kovalev (2009).

In order to increase the density of calibrators in the Galactic plane and search for suitable calibrators within 2° of known water masers, we developed an observing strategy that combined the VERA and VLBA arrays. First, we systematically screened 2462 sources with declinations $> -40^\circ$ using the four-station VLBI array VERA (Honma et al. 2003) at 22 GHz (*K* band) by observing each object in two 120 s long scans. We detected 533 objects, 180 of them new at the *K* band, and these results are given in Petrov et al. (2007a). Since precise determination of parallaxes and proper motions of sources with water-maser emission is one of the main targets of the VLBI Exploration of Radio Astrometry (VERA) project, potential calibrators near known water masers were preferentially included in the observations.

This paper describes the follow-up VLBA observations at 24 GHz of 487 radio sources in order to determine their precise positions and images. We denote these observations as the VLBA Galactic Plane Survey (VGaPS), and the results of this observing campaign are described in this paper. VLBA observations, source selection, and the scheduling algorithm are given in Section 2. The data analysis procedure is presented in Section 3. For analysis of these observations we have developed a new approach for wide-band fringe search and astrometric analysis which is described in detail in Section 3.2. The validation of the results made with the new approach is given in Section 4. The images are described in Section 6, the source position catalog is listed in Section 7, and the *K*- and *S/X*-band astrometric VLBI positions are compared in Section 8. The results are summarized in Section 9.

2. SOURCE SELECTION AND OBSERVATIONS

The VLBA observations were made at 24 GHz for several reasons. First, Galactic scattering at low frequencies broadens images and degrades source positions. Even at a frequency of 2.3 GHz, many calibrators near the Galactic plane are too broadened to be useful. Second, many Galactic radio targets are associated with H₂O maser emission at 22.5 GHz. Since the correlated flux density of calibrators generally decreases with increasing frequency, a good calibrator at 2.3 and 8.6 GHz may not be sufficiently compact or bright for use at 22 GHz. Third, source positions at 22 GHz may not necessarily be the same as those at 8.6 GHz because of the effects of frequency-dependent source structure, especially for multi-component objects, and because of frequency-dependent core shifts (e.g., Lobanov 1998; Kovalev et al. 2008; Sokolovsky et al. 2011).

We considered the task of extending the pool of calibrators for Galactic astrometry more broadly: not only to increase the list of compact extragalactic radio sources within 10° of the Galactic plane, but also to re-observe known sources in the *K* band that

are either in the Galactic plane or close to known masers at higher Galactic latitudes. Water masers are the main target of the VERA project, and determination of their parallaxes and proper motions is important for studying the three-dimensional (3D) structure and dynamics of the Galaxy's disk and bulge, and for revealing the true shape of the bulge and spiral arms, its precise rotation curve, and the distribution of dark matter. Dual-beam *K*-band VERA observations require calibrators with positions known to the milliarcsecond level within 2° of the target.

2.1. Source Selection

VLBI can detect emission only from a compact component of a source, which should be bright enough to be detected above the noise background. Chances of finding a radio source sufficiently compact for VLBI detection are significantly enhanced if information about source spectra is available. For a majority of sources with flat spectra, i.e., with a spectral index of $\alpha > -0.5$ ($S \sim f^\alpha$), synchrotron emission from a compact core often dominates, so such sources are often compact. In regions more than 10° from the Galactic plane, comparison of surveys at different frequencies can be used to select the sources with flat spectra. However, in the crowded Galactic plane, source misidentification and confusion from surveys at different frequencies often result in ambiguous information about a source spectrum. Also, there are many extended sources with thermal emission, such as planetary nebulae and H II regions, that have flat spectra but are too extended for VLBI detection.

To determine if a candidate source was sufficiently strong and compact to be a VLBI calibrator, we first used the VERA array at 22 GHz (Petrov et al. 2007a). In the survey ~ 1400 objects within 6° of the Galactic plane and ~ 1000 objects within 2° of known maser sources outside the Galactic plane were observed with the four-element VERA array at baselines of 1000–2000 km in two scans each, and approximately 20% were detected. Among the 533 detected sources, 305 objects are at Galactic latitudes of $|b| < 10^\circ$, and 228 objects are at $|b| > 10^\circ$. The list of detected sources included all objects with a probability of false detection < 0.1 . These objects formed the first set of candidate calibrators.

We also added 239 objects in the Galactic plane and 44 objects outside the Galactic plane, selected on the basis of their spectra using the Astrophysical CAtalogs support System CATS (Verkhodanov et al. 2005). This database currently includes catalogs from 395 radio astronomy surveys. We selected all entries with sources within a $20''$ radius and with measurements of flux density for at least three frequencies in the range 1.4–100 GHz. We fit a straight line to the logarithm of the spectrum and then estimated the spectral index and the flux density extrapolated to 24 GHz. We selected 451 objects with extrapolated correlated flux densities in the *K*-band > 200 mJy, spectral indices flatter than -0.5 , and $|b| < 10^\circ$. Then, we scrutinized this list and removed sources within $30''$ of known planetary nebulae or H II regions and sources with anomalous spectra that indicated a possible misidentification. The remaining list was added to the pool of calibrators. This list also contained Sagittarius A* since it has never before been observed with VLBI in the absolute astrometry mode.

This set of sources formed the pool of 816 candidate objects to be followed up with the VLBA. Because of the large number of candidate objects, a priority from 1 to 4 was assigned to each source. The first priority was given to 180 new sources with correlated flux densities in the range 100–300 mJy detected

with VERA that have never been observed with VLBI before. The second priority was given to sources in the Galactic plane detected with VERA, the third priority was given to other sources detected with VERA, and the least priority was given to sources outside the Galactic plane not detected with VERA. All new sources detected with VERA were scheduled. Sources outside of the Galactic plane were included for logistical reasons. Since the Galactic plane has an inclination with respect to the equator of $62^{\circ}6$, the density of candidate sources near the Galactic plane is a non-uniform function of right ascensions. We thus included sources outside the Galactic plane in the right ascension regions that had significantly fewer candidate sources than others in order to avoid gaps in the schedules.

In addition to the target sources, we also selected 56 objects from the K/Q survey (Lanyi et al. 2010) to serve as amplitude and atmosphere calibrators. These are the brightest sources in the K band with precisely known positions and publicly available images in FITS format.

2.2. Scheduling Algorithm

A sufficiently good preliminary VLBA observing schedule for three 24 hr sessions was prepared automatically. Then it was manually adjusted in order to produce a more efficient final schedule.

The three sets of information needed in order to compile the schedule were: (1) the list of the 816 target sources with their J2000 positions and priority levels, (2) the list of the 56 calibration sources at 24 GHz, and (3) the two or three sidereal times ranges (at the array center at station PIETOWN) at which each source should be observed. These times were a function of the source declination. For example, a source with $\delta > 50^{\circ}$ could be observed three times with very flexible time ranges, whereas a source with $\delta < -35^{\circ}$, must have tight ranges in order to be observed by at least eight VLBA antennas at an elevation higher than 10° .

The scheduling goal was to observe each source for three scans of 120 s, unless it was south of declination -25° , in which case only two scans were scheduled. With an average overhead of about 45 s between sources, about 1550 scans over the three days (72 hr) could be scheduled. Every 90 minutes, four scans were reserved for atmosphere calibrators. The algorithm then began filling in observing slots, taking the highest priority sources first, and those at the lowest declinations, since these have minimal scheduling flexibility. The algorithm scheduled sources that were relatively close to one another in the sky in order to minimize slew times, which can be as long as three minutes.

The slots containing the calibrators were chosen in order to maximize the range of elevations for the VLBA antennas. In practice, this meant maximizing elevations of observed sources at SC-VLBA, MK-VLBA and PIETOWN. It was important to have one low elevation observation for all telescopes, and this could be scheduled by observing a source either in the far north and/or in the far south. All three days were scheduled at the same time, since it did not matter on which day any of the three sidereal time slots occurred for a source.

The final schedule optimization was done “by hand” and consisted of several steps. First, some sources could be placed in only one or two slots and would have to be removed if another slot could not be found. Since about 10% of the slots could not be filled using the automatic algorithm because the source list was not uniformly distributed over the sky, additional slots were usually found to complete a source’s schedule requirement. This

Table 1
VLBA Target Sources from VERA Observations

Group	Pool	Scheduled	Detected	Ratio
Galactic, VERA	305	184	140	76%
Non-Galactic, VERA	228	151	133	88%
Galactic, others	239	108	36	33%
Non-Galactic, others	44	44	15	34%
Total	816	487	327	67%

often meant bending some of the rules or moving calibration sources or blocks by five to fifteen minutes. In regions where the target source density was large, scan integrations were decreased from 120 s to 110 s. All sources with priority 1 were included.

The next stage insured that the calibration scans were optimized in order to obtain good elevation coverage for the antennas. The purpose of these observations was (1) to serve as amplitude and bandpass calibrators, (2) to improve robustness of estimates of the path delay in the neutral atmosphere, and (3) to tie the source positions of new sources to existing catalogs such as the ICRF catalog (Ma et al. 1998). The final stage of optimization tweaked the observing schedule in obvious ways in order to save slewing time since the scheduling algorithm did not minimize slew times, and because of the above adjustments to insure proper observing coverage for each source. The schedule for each session was then carefully checked using the NRAO SCHED program to insure that each scan had sufficient integration time, and no more than one of the antennas was below 10° elevation (except for sources south of -35°).

The results of the scheduling and ultimate detection for each priority group is given in Table 1. The number of target sources selected in each group is given in Column 2, those scheduled in Column 3, those detected by the VLBA in Column 4, and the detection rate in the last column. The first two groups are sources detected in the VERA Fringe Search Survey; the last two groups contain other objects. The table does not include the 56 sources used as atmosphere calibrators.

2.3. Observations

The VGaPS observations were carried out in three 24 hr observing sessions at the VLBA on 2006 June 4, 2006 June 11, and 2006 October 20. Each target source was observed in several scans: three sources in one scan, 356 sources in two scans, one source in three scans, 124 in four scans and three sources in five scans. The scan durations were 100–120 s. In total, antennas spent 57% of time on target sources. In addition to these target objects, 56 strong sources previously observed in the K band were taken from the astrometric and geodetic catalog 2004f_astro.⁶

The data digitized at four levels were recorded with a rate of 256 Mbit s⁻¹ in eight 8 MHz wide intermediate frequencies (IF) bands spread over a bandwidth of 476 MHz (Table 2). The frequencies were selected to minimize the amplitude of sidelobes of the Fourier transform of the bandpass.

3. DATA ANALYSIS

The data were correlated in Socorro on the VLBA hardware correlator. The correlator output contains the complex spectra

⁶ <http://astrogeo.org/rfc>

Table 2
Lower Edge IF Frequencies (in GHz)

IF Index	Frequency
IF1	24.20957
IF2	24.22257
IF3	24.26157
IF4	24.33957
IF5	24.48257
IF6	24.58657
IF7	24.65157
IF8	24.67757

of the auto-correlation function and the spectrum of the cross-correlation function for each accumulation period. The accumulation period was chosen to be 0.131072 s, and the spectral resolution was set to 0.5 MHz, i.e., 16 spectral channels per IF.

After correlation, the data were stored in a file compliant with the FITS-IDI specifications (Eric Greisen, NRAO memo N114⁷). Small a priori amplitude corrections were applied; bad data were flagged (sometimes bad data are not found until later processing); and the definitions of the reference frequencies were modified.

Further data analysis involves computation of group delays for each scan and each baseline, computation of theoretical path delay, and then fitting parameters of the linear model into the differences between the observed and theoretical delays using least squares (LSQ).

3.1. Traditional Narrow-band Fringe Fitting Algorithm

The data set then contains 128 data streams for each of the 45 baselines if all 10 VLBA antennas are operating: eight IFs, each with 16 equally spaced frequency channels. After the complex bandpass calibration, the phase difference among all of the frequency streams remains unchanged, and the relevant residual phase parameters associated with any scan are (1) the residual phase at the scan midpoint, (2) the average group delay (phase gradient with frequency), and (3) the average delay rate (delay gradient with time), for each antenna. These parameters are called the residual phase terms. These parameters are estimated with a fringe fitting procedure. Generally, one antenna is chosen as the reference—a well-behaved one near the array center—so that the set of residual phase terms is associated for each scan and all other antennas. These parameters are functions of many astrometric quantities (source position, site positions, antenna-based tropospheric path delays, Earth orientation parameters, etc.), which can be determined from analysis systems like Calc/Solve from data obtained from carefully prepared observing schedules. Source structure and dispersive phase effects (e.g., ionosphere) produce a non-linear phase/frequency relationship.

The algorithm that determines the residual phase terms is implemented in the AIPS task FRING, and in the past, all VLBA observations under the absolute astrometry and geodesy programs, such as VCS, RDV, and K/Q, were processed with the use of this software (Greisen 1988). When IFs are spread over a wide band with gaps, the AIPS algorithm determines the residual phase in two steps: first, for each of the eight IFs the residual phase, single-band group delay, and rates are independently obtained. Then, the residual phases of each IF are fit to a linear

phase versus frequency term to produce the group delay, using the AIPS program MBDLY. This procedure is described in detail in Petrov et al. (2009).

3.2. Wide-band Baseline-based Fringe Fitting Algorithm

The two-step approach has a substantial shortcoming: it requires fringe detection for each individual IF independently within a narrow band. Using all N IFs simultaneously for coherent averaging, we can detect a source with an amplitude smaller than \sqrt{N} using only one IF. This degradation of the detection limit does not pose a problem for most geodesy observations, or for absolute astrometry of bright sources, since the target SNR is usually very high, but it significantly impacts absolute astrometric experiments of weak sources. The traditional AIPS algorithm did not detect a sizeable fraction of the sources observed in the VGaPS experiment.

This limitation motivated us to develop an advanced algorithm for wide-band fringe fitting across all of the IFs within the band. For logistical reasons, instead of augmenting AIPS with the new task, we decided to develop a new software package called PIMA⁸ from scratch that is supposed to replace the AIPS for processing absolute astronomy and geodesy experiments. Here we outline the method of wide-band fringe search used for processing this experiment.

3.2.1. Spectrum Re-normalization

Digitization of the input signal and its processing with a digital correlator causes an amplitude distortion with respect to an ideal analogue system. As documented by Kogan (1995), many effects distort both cross-correlation and auto-correlation spectra exactly the same way. Therefore, if we divide the cross-correlation spectrum by the auto-correlation spectrum averaged over time and over frequencies within each IF, we will remove these distortions. However, there are two effects that affect cross- and auto-spectra differently because the amplitude of the cross-spectrum is very low and the amplitude of the auto-correlation spectrum is close to 1.

The non-linear amplitude distortion of the digitized signal was studied in depth by Kogan (1998) who derived a general expression for the correlation coefficient of the digitized signal as a function of the correlation coefficient of the hypothetical analogue signal. In the absence of fringe stopping, the output correlation coefficient ρ_{out} is expressed via the correlation coefficient for an analogue case ρ as

$$\begin{aligned} \rho_{\text{out}} = & 2\kappa \int_0^\rho \frac{1}{\sqrt{1-\rho^2}} d\rho \\ & + 2\kappa(n-1) \int_0^\rho \frac{1}{\sqrt{1-\rho^2}} \left(e^{-\frac{v_1^2}{2(1-\rho^2)}} + e^{-\frac{v_2^2}{2(1-\rho^2)}} \right) d\rho \\ & + \kappa(n-1)^2 \int_0^\rho \frac{1}{\sqrt{1-\rho^2}} \\ & \times \left(e^{-\frac{v_1^2 - 2\rho v_1 v_2 + v_2^2}{2(1-\rho^2)}} + e^{-\frac{v_1^2 + 2\rho v_1 v_2 + v_2^2}{2(1-\rho^2)}} \right) d\rho, \end{aligned} \quad (1)$$

where κ is the normalization coefficient, n is the ratio of the two levels of quantization, and v_1 and v_2 are the dimensionless digitizer levels in units of variance of the input signal. Their numerical values compiled from the paper of Kogan (1993) are presented in Table 3.

⁷ Available at <ftp://ftp.aoc.nrao.edu/pub/software/aips/TEXT/PUBL/AIPSMEM114b.PS>.

⁸ Available at <http://astrogeo.org/pima>.

Table 3
Numerical Coefficients in Integral (1) for Three Cases
of the Number of Bits per Sample: (1,1), (1,2), (2,2)

Comb. of Bits	n	v_1	v_2	κ
(1,1)	1.0	0.0	0.0	0.3803
(1,2)	3.3359	0.0	0.9816	0.05415
(2,2)	3.3359	0.9816	0.9816	0.07394

When $\rho \ll 1$, the dependence $\rho_{\text{out}}(\rho)$ becomes linear: $\rho_{\text{out}} = 2/\pi \cdot \rho \approx 0.6366\rho$ for the case of one-bit sampling, and $\rho_{\text{out}} \approx 0.8825\rho$ for the case of two-bit sampling. Therefore, distortion of the cross-spectrum is proportional to ρ_{out}/ρ , and dividing the cross-spectrum by 0.6366 or 0.8825 we eliminate the distortion of the fringe amplitude introduced by the digitization.

However, the amplitude of the auto-correlation spectrum is close to 1, and we cannot ignore non-linearity of $\rho_{\text{out}}(\rho)$. To correct the auto-correlation spectrum for digitization distortion, we follow the procedure outlined by Kogan (1995). First, the auto-correlation is inverse Fourier transformed. It should be noted that the correlator provides the auto-correlation for N spectral channels for non-negative frequencies from 0 to $N - 1$. We restore the auto-correlation at the N th channel by linear extrapolation using the $N - 2$ and $N - 1$ values of the spectrum and set the spectrum for $N - 1$ negative frequencies to zero. The dimension of the Fourier transform is $2N$. Second, the result of the transformation, the auto-correlation coefficient versus time lag, is de-tapered, i.e., divided by the self-convolution of the weighting function. The VLBA correlator normally uses uniform weighting, i.e., uses weight 1 for all points. The self-convolution of the uniform weighting is a triangle function $\wedge(i)$:

$$\wedge(i) = \begin{cases} \frac{1}{N}(N - |i|) & \text{if } |i| < N \\ 0 & \text{otherwise.} \end{cases}$$

Third, the correlation function is divided by its maximum which is found at the zero lag. Fourth, each correlation coefficient is divided by $\rho_{\text{out}}(\rho)$. Fifth, the correlation function is multiplied back by the stored value at the zero lag. Sixth, the correlation function is again tapered by multiplying it by $\wedge(i)$. Finally, we perform the Fourier transform of the corrected correlation function and get the re-normalized auto-spectrum, free from digitization distortion. The square root of the product of the auto-correlation spectra from two stations of a baseline, averaged over time and frequency within each IF, gives us an estimate of the fringe amplitude scale factor, but before dividing the cross-correlation spectrum by this scale factor, we have to take into account a specific effect of the hardware VLBA correlator.⁹ An insufficient number of bits in internal correlator registers resulted in a decrease of the amplitude when it was large, and when that happened, the auto-correlation spectra was corrupted. Kogan (1993) suggested the following model for accounting for this effect:

$$F = 1 + \frac{w}{4SAV_s}, \quad (2)$$

where w is the weight of the spectrum data equal to the ratio of processed samples to the total number of samples in an accumulation period, A is the accumulation period length, S is 2 when the correlator processed single polarization data and 1

if both polarizations were correlated, and V_s is the visibility scale factor provided by the correlator. We divide the auto-correlation spectrum by the factor F .

3.2.2. Coarse Fringe Search

The correlator output provides auto-correlation and cross-correlation spectra for each pair of baselines and each scan. The cross-correlation spectrum is computed at a uniform two-dimensional (2D) grid of accumulation periods and frequencies and is accompanied with weights that are the ratio of the number of processed samples in each accumulation period to the nominal number of samples.

The fringe fitting procedure searches for phase delay τ_p , phase delay rate $\dot{\tau}_p$, group delay τ_g , and its time derivative $\dot{\tau}_g$ which correct their a priori values used by the correlator model in such a way that the coherent sum of weighted complex cross-correlation samples c_{kj} ,

$$C(\tau_p, \tau_g, \dot{\tau}_p, \dot{\tau}_g) = \sum_k \sum_j c_{kj} w_{kj} \times e^{i(\omega_0 \tau_p + \omega_0 \dot{\tau}_p (t_k - t_0) + (\omega_j - \omega_0) \tau_g + (\omega_j - \omega_0) \dot{\tau}_g (t_k - t_0))}, \quad (3)$$

reaches the maximum amplitude. Index k runs over time, and index j runs over frequencies. ω_0 and t_0 denote the angular reference frequency within the band and the reference time within a scan and w_{kj} is weight. Function $C(\tau_p, \tau_g, \dot{\tau}_p, \dot{\tau}_g)$ is essentially non-linear, and we need a really good starting value in order to find the global maximum by traditional optimization algorithms. We can notice that term $2\pi\omega_0\tau_p$ in expression (3) does not depend on the summation indices, and $\dot{\tau}_g$ is usually small. Therefore, for the purpose of a coarse fringe search we simplify expression (3) to

$$C(\tau_p, \tau_g, \dot{\tau}_p) e^{-i2\pi\omega_0\tau_p} \approx \sum_k \sum_j c_{kj} \times e^{i(\omega_0 \dot{\tau}_p (t_k - t_0) + (\omega_j - \omega_0) \tau_g)}. \quad (4)$$

For the search of the maximum, the trial functions $C(\tau_p, \tau_g, \dot{\tau}_p)$ are computed on a dense grid of the search space $\tau_g, \dot{\tau}_p$. It follows immediately from expression (4) that $|C| = |\mathcal{F}(c_{kj})|$, where \mathcal{F} denotes the 2D Fourier transform.

The first step of the coarse fringe search is to compute the 2D fast Fourier transform (FFT) of the matrix of the cross-correlation spectrum. The first dimension of the matrix runs over time, and the second dimension runs over frequency. The sampling intervals are $\Delta t/\beta$ and $\Delta f/\gamma$, where Δt and Δf are the duration of the accumulation period and the spectral resolution, respectively, and β and γ are integer oversampling factors. The elements of the matrix that do not have measurements or have discarded measurements are padded with zeroes. The dimensions of the matrix are chosen to have a power of 2 for gaining the maximum performance of the FFT numerical algorithm.

Oversampling factors greater than 1 are used to mitigate amplitude losses. The FFT produces estimates of $|C|$ at a discrete grid of group delays and delay rates. If the maximum of the amplitude of the coherent sum of the cross-correlation function samples falls at group delays and delay rates between the nodes of the grid, its magnitude will be greater than the amplitude of $|C|$ at the nearest grid point by a factor of L .

⁹ The new VLBA software DiFX correlator does not have this problem.

The amplitude loss factor L of the coarse search matrix is the integral average over time and frequency:

$$L = \frac{1}{t_s} \int_{-t_s/2}^{t_s/2} \cos 2\pi \left(\omega_0 \tau_p - \frac{k}{\beta t_s} \right) t dt \\ \times \frac{1}{f_b} \int_{-f_b/2}^{f_b/2} \cos 2\pi \left(\omega_0 \tau_g - \frac{l}{\gamma f_b} \right) f df, \quad (5)$$

where k and l are indices of the nearest grid nodes, t_s is the scan duration, and f_b is the total bandwidth. The integral (5) is easily evaluated analytically, and the maximum losses $L = \text{sinc}(\pi/(2\beta)) \cdot \text{sinc}(\pi/(2\gamma))$ are achieved when τ_p and τ_g happen to be just between grid nodes. In a case where the oversampling factor is 1, $L = 4/\pi^2 \approx 0.405$. That loss factor effectively raises the detection limit by $1/L = 2.467$ in the worst case. We used the grid 4096×4096 which corresponds to $\beta = 4$, $\gamma = 4$ for $t_s = 120$ s, $f_b = 476$ MHz. Therefore, the maximum loss factor for our experiment is 0.959, which results in raising the detection limit by no more than 4.1%.

The group delay and delay rate that correspond to the maximum of the discrete Fourier transform of the cross-correlation matrix provide the coarse estimates of group delay and delay rate. Their accuracy depends on the grid resolution. The next step is to refine their estimates. The first stage of the fine search is an iterative procedure that computes |C| at a progressively finer 3D grid in close vicinity of the maximum using the discrete 3D Fourier transform. The third dimension is group delay rate, omitted during the coarse search. Dimensions of the transform for the group delay, phase delay rate, and group delay rate are 3, 3, and 9, respectively. At the first step of iterations, the grid runs over ± 1 element of the coarse grid for group delays and phase delay rates and in the range $\pm 2 \times 10^{-11}$ for the group delay rate. After each step of iterations, the grid centered around the maximum element shrinks its step by 2. In total, eight iterations are run. The phase of the coherent sum of the cross-correlation function determined with the iterative procedure according to expression (4) is the fringe phase with the opposite sign.

3.2.3. Probability of False Detection

The significance of the fringe amplitude depends on the level of noise. In the absence of signal, the amplitude of the cross-correlation function a has a Rayleigh distribution of:

$$p(a) = \frac{a}{\sigma^2} e^{-\frac{a^2}{2\sigma^2}}, \quad (6)$$

where σ is the standard deviation of the real and imaginary part of the cross-correlation function and n is the total number of spectrum points. In a case where all points of the spectrum are statistically independent, the cumulative distribution function of the coherent sum over n points is (Thompson et al. 2001)

$$P(a) = \left(1 - e^{-\frac{a^2}{2\sigma^2}} \right)^n. \quad (7)$$

Differentiating this expression, we find the probability density function of the ratio of the amplitude of the coherent sum of spectrum samples of the noise to its standard deviation as

$$p(a) = n \frac{a}{\sigma^2} e^{-\frac{a^2}{2\sigma^2}} \left(1 - e^{-\frac{a^2}{2\sigma^2}} \right)^{n-1}. \quad (8)$$

Table 4
Probability of False Detection as a Function of SNR

SNR	$P_f(s)$
4.96	0.3
5.19	0.1
5.61	0.01
5.99	0.001
6.68	10^{-5}

Under the assumption that all samples are statistically independent, the variance of the noise of the coherent sum of N samples is scaled as $\sigma = \sigma_s / \sqrt{t_s S_r}$, where S_r is the sampling rate of the recorded signal (6.4×10^7 in our experiment), t_s is the scan duration, and σ_s is the variance of an individual sample, 1 for a perfect system.

However, the assumption of statistical independence is an idealization. The presence of systematic phase errors, deviation of the bandpass from the rectangular shape, and other factors distort the distribution. The deviation from the statistical independence is difficult to assess analytically.

We evaluate the variance of the noise by estimating the variance over a sample of 32,768 random points in the region of the Fourier transform of the coherent sum of the cross-correlation that does not contain the signal. The indices of grid elements of the sample are produced by using the random number generator. The sample of amplitudes is ordered, and half of the points greater than the median are rejected. The variance over the remaining points is computed and an iterative procedure is launched that adds back previously rejected points in ascending order of their amplitudes and updates the mean value and variance. The iterations are run till the maximum amplitude of the next sample reaches $3.5\sigma_a$. The initial rejection and consecutive restoration of points with amplitudes $> 3.5\sigma_a$ ensures that no points with signal from the source affect the computation of σ_a and $\langle a \rangle$. The rejection of the tail of the amplitude distribution causes a bias in estimate of the mean, but the magnitude of the bias is only -2×10^{-4} , which is negligible. We define a signal-to-noise ratio as $\text{SNR} = a/\langle a \rangle$. It follows immediately from expression (6) that $\langle a \rangle = \sqrt{\frac{\pi}{2}} \sigma$.

We assume that the a posteriori distribution of the SNR $s = a/\sigma_a$ can be approximated as a function like expression (8) with effective parameters of σ_{eff} and n_{eff} :

$$p(s) = \frac{2}{\pi} \frac{n_{\text{eff}}}{\sigma_{\text{eff}}} s e^{-\frac{s^2}{\pi}} \left(1 - e^{-\frac{s^2}{\pi}} \right)^{n_{\text{eff}}-1}. \quad (9)$$

These parameters σ_{eff} and n_{eff} can be found by fitting the left tail of the empirical distribution of the SNRs. Using their estimates, we can evaluate the probability of finding an amplitude less than a if no signal is present, i.e., the probability of false detection:

$$P_f(s) = 1 - \int_0^s p(s) ds = 1 - \frac{1}{\sigma_{\text{eff}}} \left(1 - e^{-\frac{s^2}{\pi}} \right)^{n_{\text{eff}}}. \quad (10)$$

The low end of the empirical SNR distribution and the table of the probability of false detection for VGaPS experiments are shown in Figure 1 and Table 4.

3.2.4. Fine Fringe Search

Finally, the group delay, phase delay rate, group delay rate, and fringe phase at the reference frequency are estimated using

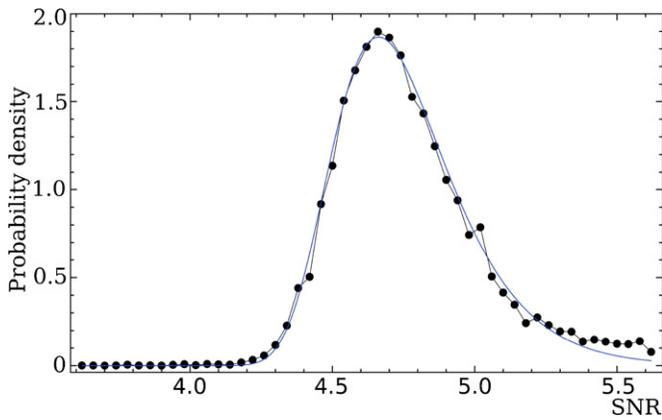


Figure 1. Low end of the empirical distribution of the achieved SNR for the fringe amplitude from results of fringe fitting VLBA data (filled circles) and the fitted curve (thin line) of the theoretical distribution for the case of no signal.

(A color version of this figure is available in the online journal.)

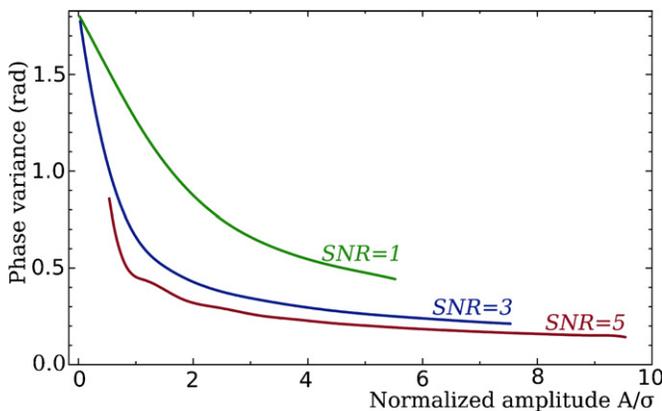


Figure 2. Variance of fringe phase as a function of the normalized amplitude A/σ_n in the presence of a signal with given SNR.

(A color version of this figure is available in the online journal.)

LSQ in the vicinity of the maximum provided its amplitude exceeds the detection limit. The goal of the LSQ refinement is to obtain realistic estimates of statistical errors of fitted parameters and to account for possible systematic errors by analyzing residuals. All cross-correlation spectrum data points of a given observation are used in a single LSQ solution with weights reciprocal to their variance.

Determining the variance of the fringe phase of an individual point is trivial only in two extreme cases: when $\text{SNR} \ll 1$, and when $\text{SNR} \gg 1$. In the first case the fringe phase distribution becomes uniform in the range of $[0, 2\pi]$ with a variance of $\pi/\sqrt{3} \approx 1.813$. In the second case expanding the expression for fringe phase in the presence of noise $\phi = \arctan \frac{S_i + n_i}{S_r + n_r}$, where S_i and S_r are the real and imaginary parts of the signal and n_i and n_r are the real and imaginary parts of the noise, into the Taylor series, neglecting terms $O(n/S^2)$, and evaluating the variance of the expansion, we get $\sigma_\phi = \sqrt{\frac{2}{\pi}} \frac{1}{S/N}$. For the general case, the problem becomes more difficult, since the σ_ϕ depends on the variance of the noise and on the amplitude of the signal nonlinearly. An analytical solution requires evaluation of complicated integrals that are not expressed via elementary functions.

We used the Monte Carlo approach to compute these variances. Let us consider a random complex set $s = A + n_r + in_i$, where A is the amplitude of the simulated signal and n_r, n_i are

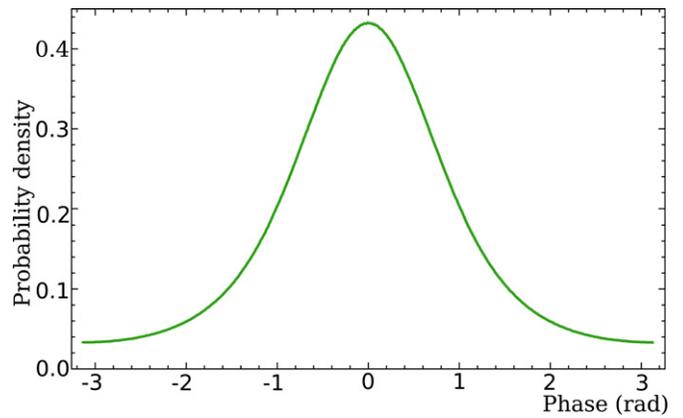


Figure 3. Probability density distribution of fringe phase with $\text{SNR} = 1$.

(A color version of this figure is available in the online journal.)

independent random variables with Gaussian distribution. Their variance σ was selected in such a way that $A = \sqrt{\frac{\pi}{2}} \sigma \text{SNR}$. Then for a given SNR, we can compute the variance of phase of s as a function of the normalized amplitude $|s|/\sigma$. This is done by generating a long series (1 billion points) of the simulated complex signal for a given SNR, computing the amplitude and phase of the time series, splitting the signal into a uniform grid of 128 bins over normalized amplitude that spans the interval $[\text{SNR} - 4.5, \text{SNR} + 4.5]$, computing the variance of the phase of the simulated signal over all points that fall into each bin, and approximating the dependence of $\sigma_\phi(|s|/\sigma)$ with a smoothing B-spline of third order over six nodes. We computed $\sigma_\phi(|s|/\sigma)$ for SNRs in the range $[0, 12.7]$ with steps of 0.1. Examples of this dependence for several SNRs are shown in Figure 2. The set of B-spline coefficients forms a 2D table with axes SNR and normalized amplitude which allows us to evaluate σ_ϕ for a given SNR and a given fringe amplitude.

It should be noted that the SNR of the coarse search is related to the amplitude coherently averaged over *all* valid cross-correlation spectrum samples, $\sum_k^{N_t} \sum_j^{N_s} w_{kj}$, where N_t is the number of accumulation periods and N_s is the total number of spectral channels. The SNR of *an individual* accumulation period (the elementary SNR) is $\sqrt{\sum_k^{N_t} \sum_j^{N_s} w_{kj}}$ times less. For our experiments, the typical reduction of the SNR is a factor of 340. This means that for almost all sources the elementary SNR will be less than 1. As we have seen previously, the distribution of fringe phases at very low SNRs is close to uniform with a variance of $\pi/\sqrt{3}$. The phase becomes uncertain due to the 2π ambiguity, and the LSQ estimation technique loses its diagnostic power. Therefore, the cross-correlation function with applied phase, phase delay rate, and group delays computed during the coarse fringe search have to be coherently averaged over time and frequency in segments large enough to have sufficiently high SNR over a segment to provide an ambiguous phase. As seen in Figure 3, the fringe phase distribution at $\text{SNR} = 1$ is already sharp enough for that. Therefore, the number of spectral channels and the number of accumulation periods within a segment is chosen in such a way that the SNR is at least 1. Marginally detected scans with $\text{SNR} = 5$ have 24 segments that average all spectral channels within an IF and over 1/3 of the scan interval.

Using all segments, we determine four fitting parameters \mathbf{p} using the LSQ:

$$\mathbf{p} = (A^T \mathcal{W} A)^{-1} A^T \mathcal{W} \phi_s, \quad (11)$$

where A is the matrix of observations, \mathcal{W} is the diagonal weight matrix and ϕ_s is the vector phases of the cross-correlation function averaged over segments.

The mathematical expectation of the square of the weighted sum of residuals R in the presence of noise ϵ is

$$E(R) = \text{Tr}(\mathcal{W}\text{Cov } \epsilon) - \text{Tr}(\text{Cov } \epsilon \mathcal{W} A (A^\top \mathcal{W} A)^{-1} A^\top \mathcal{W}), \quad (12)$$

which is reduced to $n - m$, where n is the number of equations and m is the number of estimated parameters if the weight matrix \mathcal{W} is chosen to be $(\text{Cov } \epsilon)^{-1}$.

The presence of additive errors, for example fluctuations in the atmosphere, will increase $\text{Cov } \epsilon$ and our estimate of the error variance based on the amplitude of the spectrum sample without knowledge of the scatter of the cross-spectrum phases is incomplete. A small additive noise with variance k times less than the amplitude of the signal affects the amplitude as $O(k^2)$, but it affects the phase as $O(k)$. One of the measures of the incompleteness of the error model is the ratio of the square of the weighted sum of residuals to its mathematical expectation.

We can extend our error model assuming that the weight matrix is $\mathcal{W} = \text{Cov } \epsilon^{-1} - q^2 I$, where I is the unit matrix of the same dimensions as \mathcal{W} and q is the parameter. This model is equivalent to an assumption that the used squares of weight of each segment are less than the true one by some parameter q equal to all segments. The mathematical expectation of R for this additive weighting model is

$$E(R) = n - m + q^2 [\text{Tr}(\mathcal{W}) - \text{Tr}(\mathcal{W} A \mathcal{W}^2 A)^{-1} A^\top]. \quad (13)$$

Inverting Equation (13), we find q for a given $E(R)$:

$$q = \sqrt{\frac{E(R) - (n - m)}{\text{Tr}(\mathcal{W}) - \text{Tr}(\mathcal{W} A \mathcal{W}^2 A)^{-1} A^\top}}. \quad (14)$$

Replacing the mathematical expectation of R with its value evaluated from the residuals, we can find the re-weighting parameter q for a given solution. Several iterations provide a quick convergence of $\frac{E(R)}{R}$ to 1. Applying an additive re-weighting constant results in an increase in estimates of parameter uncertainties. It may happen that q becomes imaginary. This means that \mathcal{W} was overestimated. In our analysis we set $q = 0$ when that happens.

3.2.5. Complex Bandpass Calibration

Coherent averaging of the cross-correlation spectrum over frequencies assumes that the data acquisition system does not introduce a distortion of the recorded signal, but this is usually not the case. Each intermediate frequency has its own arbitrary phase offset and group delay that may vary with time. The imperfection of baseband filters results in a non-rectangular shape of the amplitude response.

To calibrate these effects, a rail of narrow-band phase calibrator signals with a spacing of 1 MHz was injected near the receivers. Two tones per IF were extracted by the data acquisition hardware, and their phase and amplitude are available for data analysis. When the phase of the phase calibration signal is subtracted from fringe phases, the result is referred to the point of injection of the phase calibration signal and this procedure is supposed to take into account any phase changes that occurred in the signal passing through the data acquisition terminal. However, we should be aware of several complications that emerge when we try to use the benefits of the phase calibration signal.

First, the phases of two tones of the phase calibration signal separated by 6 MHz, as in our sessions, have ambiguities. Since the instrumental group delay may reach several phase turns over an 8 MHz IF, the second phase calibration tone is useless without resolving the ambiguity. Second, the phase calibration itself may have a phase offset or may become unstable if its amplitude is not carefully tuned. Therefore, we need to re-calibrate the phase calibration signal itself in order to successfully apply it to data.

We compute the complex bandpass function for each station, except the reference station, that describes the residual instrumental complex bandpass after applying the phase calibration signal. The cross-correlation spectrum needs to be divided by the complex bandpass in order to correct the instrumental frequency-dependent delay and fading of the amplitude. The procedure for evaluating the complex bandpass has several steps.

First, all data are processed applying the first tone of the phase calibration, i.e., the phase of the phase calibration signal is subtracted from each phase of the cross-correlation signal at a given IF.

Second, the bandpass reference station is chosen. Then, for each station, we find an observation at a baseline with the reference station that provided the maximum SNR during the first run. Then for each IF we average the residual spectrum over time and perform a linear fit to the residual phases to determine the instrumental group delay. Using this instrumental group delay, we extrapolate the phase of the phase calibration signal of the first phase calibration tone to the frequency of the second phase calibration tone and resolve its phase ambiguity. After that, we re-run the fringe search procedure for these scans by applying the phase calibration phase to each IF in the form of a linear function of phase versus frequency that is computed from two phase calibration tones with resolved phase ambiguity. The result of the new fringe search gives a new residual spectrum. Then we averaged the residual spectrum over time and over M segments within each IF ($M = 2$ in our experiment). The amplitude of the spectrum is normalized by dividing the average amplitude over all IFs. The phase and the amplitude of the residual averaged spectrum are approximated with a fifth degree Legendre polynomial. The result of this approximation as a complex function of frequency defines the so-called initial complex bandpass. Analysis of the residuals of rejected observations helps to diagnose malfunction of the equipment. For example, one or more video converters may fail, which may result in a loss of coherence. In that case, the part of the affected cross-correlation spectrum is masked out.

Third, we refine the complex bandpass. We select N more observations with the highest SNR from the first run at all baselines ($N = 16$ in our experiment) and re-run the fringe fitting procedure with applied phase (but not amplitude!) of the initial bandpass. We compute a residual spectrum averaged over time and M segments for each processed observation and normalize its amplitude. Then we fit a set of six coefficients of the Legendre polynomial for each station, except the reference one, for each IF, for both amplitude and phase to the phase and amplitude of the residual spectrum using a single LSQ solution. Then the residuals are computed and the observations with the maximum absolute value of residual phases and residual amplitudes are found separately. If the maximum absolute value of residual phase or residual amplitude exceeds the predefined limit, the affected observation is removed and the solution is repeated. Iterations are run until either the absolute values of

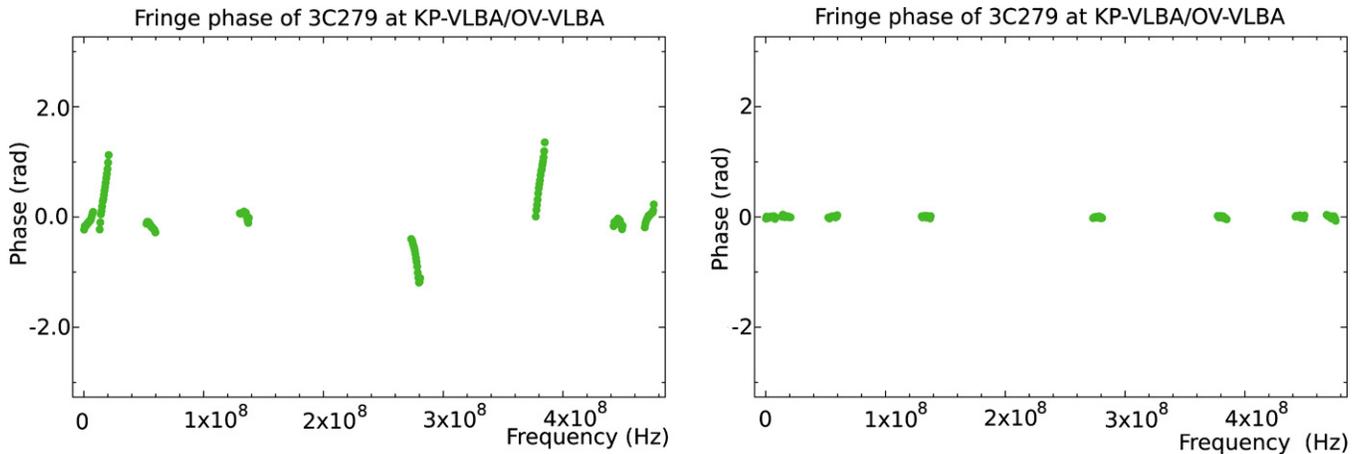


Figure 4. Residual fringe phases (radians) before (left) and after (right) applying bandpass calibration in experiment bp125b.
(A color version of this figure is available in the online journal.)

all remaining residuals are less than the predefined limit or the number of observations at a given baseline drops below $N/2$. The fitted Legendre polynomial coefficients are added to the coefficients of the initial bandpass and the result defines the so-called fine complex bandpass.

Fourth, all observations are reprocessed with the fine bandpass applied: the phase of the fine complex bandpass of the remote station of a baseline is subtracted and the phase of the bandpass of the reference station of a baseline is added before fringe fitting, and the amplitude is divided by the square root of the products of bandpass amplitudes after fringe fitting. Examples of residual phases before and after calibration are shown in Figure 4.

3.3. Computation of Total Group Delays and Phase Delay Rates

The results of the fringe search are residual phase and group delay as well as their time derivatives determined from analysis of an observation at a given baseline of a given scan with respect to the a priori delay model used by the correlator. We need to compute the total path delay related to a certain moment of time common to all observations of a scan, called a scan reference time (t_{srt}).

For logistical reasons, fringe searching at different baselines is performed independently and therefore, each observation has its own reference epoch, called fringe reference times (t_{frt}). This time epoch is computed as the weighted mean epoch of a given observation. Since stations usually start and end at slightly different times and the number of processed samples may be different, in general, t_{frt} is different at different baselines of the same scan. The scan duration may be significantly different at different stations either by a schedule design when antennas with different sensitivities participate in observations—this is often made in geodetic observations, or due to losses of some of the observing time at stations for various reasons. If we set t_{srt} as an average of t_{frt} over all baselines of a scan, it may happen that for some baselines the difference $t_{\text{srt}} - t_{\text{frt}}$ may reach several hundred seconds. Setting a common t_{srt} for as many baselines as possible is desirable since it allows computation of delay triangle misclosures and some other important statistics. On the other hand, the uncertainty of the group delay estimate is

minimal at t_{frt} . The uncertainty in group delay at t_{srt} becomes

$$\sigma_{\tau}^2(t_{\text{srt}}) = \sigma_{\tau}^2(t_{\text{frt}}) + 2\text{Cov}(\tau, \dot{\tau})(t_{\text{srt}} - t_{\text{frt}}) + \sigma_{\dot{\tau}}^2(t_{\text{srt}} - t_{\text{frt}})^2, \quad (15)$$

which is undesirable. In our work, we set the tolerance limit for the growth of the uncertainty due to the differences between t_{srt} and t_{frt} to 0.1σ or 5 ps, whichever is less. Setting this limit, we find for each observation the maximum allowed $|t_{\text{srt}} - t_{\text{frt}}|$ by solving quadratic Equation (15). In a case where all observations of a scan have overlapping intervals of acceptable scan reference times, we set it to the value that minimizes $2 \sum \text{Cov}(\tau, \dot{\tau})(t_{\text{srt}} - t_{\text{frt}}) + \sigma_{\dot{\tau}}^2(t_{\text{srt}} - t_{\text{frt}})^2$ over all observations. In the case where there are observations that have intervals of acceptable t_{srt} that are not overlapping, the set of observations of a scan is split into several subsets with overlapping acceptable t_{srt} and the optimization procedure is performed under each subset. Finally, t_{srt} is rounded to the nearest integer second.

The VLBA correlator shifts the time tag of data streams from each station to the moment of time when the wavefront reaches the center of the coordinate system. This operation facilitates correlation and allows station-based processing. The a priori path delay is computed for this modified time tag. This shift of the time tag depends on the a priori parameters of the geometric models, and therefore the total path delay produced from such a modified quantity would depend on errors of the a priori model which would considerably complicate data analysis. Therefore, we have to undo this shift of the time tag for the reference station of a baseline for further processing.

The correlator delay model for the VLBA hardware and software correlators is computed as a sum of a fifth-degree polynomial fit to the geometric delays, the linear clock offset, and the coarse atmospheric model delay over intervals of 120 s. We find the a priori delay of the baseline reference station τ_a^{rf} related to the time tag at TAI of the wavefront arrival to its phase center from the implicit equation

$$\tau_a^{\text{rf}} = \sum_{k=0}^{k=5} a_k^{\text{rf}} (t_{\text{srt}} - t_o - (\tau_a^{\text{rf}} - \tau_{\text{cl}}^{\text{rf}} - \dot{\tau}_{\text{cl}}^{\text{rf}} \tau_a^{\text{rf}} - \tau_{\text{atm}}^{\text{rf}}))^k, \quad (16)$$

which is solved by iterations. Here, t_o is the TAI time tag of the polynomial start time and $\tau_{\text{cl}}^{\text{rf}}$ and $\tau_{\text{atm}}^{\text{rf}}$ are the clock model and the atmosphere contribution of the a priori path model. We set τ_a^{rf} on the right-hand side of expression (16) to zero for the first

iteration. Three iterations are sufficient to reach the accuracy of 0.1 ps. Using the value τ_a^{rf} found at the last iteration, we compute the a priori path delay for the remote station of the baseline:

$$\tau_a^{\text{rm}} = \sum_{k=0}^{k=5} a_i^{\text{rm}} (t_{\text{srt}} - t_o - (\tau_a^{\text{rf}} - \tau_{\text{cl}}^{\text{rm}} - \dot{\tau}_{\text{cl}}^{\text{rm}} \tau_a^{\text{rf}} - \tau_{\text{atm}}^{\text{rm}}))^k. \quad (17)$$

The a priori delay rate is computed using an expression similar to Equation (17). Finally, we compute the total path delay by extrapolating the residual delay to the scan reference time:

$$\tau_{\text{tot}} = \tau_{\text{apr}}^{\text{rm}} - \tau_{\text{apr}}^{\text{rf}} + \tau_{\text{res}} + \dot{\tau}_{\text{res}} (t_{\text{srt}} - t_{\text{ftr}}). \quad (18)$$

The delay produced this way is the difference between the interval of proper time measured by the clock of the remote station between events of arrival of the wave front to the reference point of the remote antenna and clock synchronization in TAI and the interval of proper time measured by the clock of the reference station between events of arrival of the wave front to the reference point of the reference antenna and clock synchronization in TAI.

3.4. Astrometric Analysis: Delay Modeling

Our computation of theoretical time delays in general follows the approach presented in detail by Sovers et al. (1998) with some refinements. The most significant ones are the following. The advanced expression for time delay derived by Kopeikin & Schäfer (1999) in the framework of general relativity was used. The displacements caused by the Earth's tides were computed using the numerical values of the generalized Love numbers presented by Mathews (2001) following a rigorous algorithm described at Petrov & Ma (2003) with truncation at a level of 0.05 mm. The displacements caused by ocean loading were computed by convolving the Green's functions with ocean tide models using the NLOADF algorithm of Agnew (1997). The GOT00 model (Ray 1999) of diurnal and semi-diurnal ocean tides, the NAO99 model of ocean zonal tides (Matsumoto et al. 2000), the equilibrium model (Petrov & Ma 2003) of the pole tide, and the tide with a period of 18.6 years were used. Atmospheric pressure loading was computed by convolving the Green's functions with the output of the atmosphere NCEP Reanalysis numerical model (Kalnay et al. 1996). The algorithm of computations is described in full detail in Petrov & Boy (2004). The empirical model of harmonic variations in the Earth orientation parameters $\text{heo}_{20091201}^{10}$ derived from VLBI observations according to the method proposed by Petrov (2007) was used. The time series of UT1 and polar motion from the Goddard operational VLBI solutions were used as a priori.

The a priori path delays in the neutral atmosphere in directions toward observed sources were computed by numerical integration of differential equations of wave propagation through the heterogeneous media. The four-dimensional field of the refractivity index distribution was computed using the atmospheric pressure, air temperature, and specific humidity taken from the output of the Modern Era Retrospective-Analysis for Research and Applications (MERRA; Schubert et al. 2008). That model presents the atmospheric parameters at a grid $1/2^\circ \times 2/3^\circ \times 6^h$ at 72 pressure levels.

To consider the contribution of the ionosphere to the phase of the cross-correlation spectrum, note that the electromagnetic

wave propagates in a plasma with phase velocity

$$v_p = \frac{c}{\sqrt{1 - \frac{N_v e^2}{m_e \epsilon_0 \omega^2}}}, \quad (19)$$

where N_v is the electron density, e is the charge of an electron, m_e is the mass of an electron, ϵ_0 is the permittivity of free space, ω is the angular frequency of the wave, and c is the velocity of light in a vacuum. Phase velocity in the ionosphere is faster than the velocity of light in a vacuum.

After integration along the ray path and expanding expression (19), withholding only the term of the first order, we get the following expression for additional phase rotation caused by the ionosphere:

$$\Delta\phi = -\frac{a}{\omega}, \quad (20)$$

where ω is the angular frequency and a is

$$a = \frac{e^2}{2cm_e\epsilon_0} \left(\int N_v ds_1 - \int N_v ds_2 \right). \quad (21)$$

Here, s_1 and s_2 are the paths of wave propagation from the source to the first and second stations of the baseline. If $\int N_v ds$ is expressed in units of 1×10^{16} electrons m^{-2} (so-called total electron contents (TEC) units), then after having substituted values of constants, we get $a = 5.308018 \times 10^{10} \text{ s}^{-1}$ times the difference in the TEC values at the two stations.

Since the ionosphere contribution is frequency-dependent, it distorts the fringe-fitting result. Taking into account that the bandwidth of the recorded signal is small with respect to the observed frequency, we can linearize Equation (20) near the reference frequency ω_o : $\phi = -a/\omega_o + a(\omega_i - \omega_o)/\omega_o^2$. Comparing it with expression (4), we see that the first frequency-independent term contributes to the phase delay and the second term, linear with frequency, contributes to the group delay. The fine fringe search is equivalent to solving the LSQ for τ_p and τ_g using the following system of equations:

$$\tau_p \omega_o + \tau_g (\omega_k - \omega_o) = \phi_i + \frac{a}{\omega_i}, \quad (22)$$

where index i runs over frequencies and index k runs over accumulation periods.

A solution of the 2×2 system of normal equations that originates from Equation (22) can be easily obtained analytically. Gathering terms proportional to a , we express the contribution of the ionosphere to phase and group delay as $\tau_p^{\text{iono}} = -a/\omega_p^2$ and $\tau_g^{\text{iono}} = a/\omega_g^2$, where ω_p and ω_g are effective ionospheric frequencies:

$$\omega_p = \sqrt{\frac{\omega_o \left(\sum_i^n w_i \cdot \sum_i^n w_i (\omega_i - \omega_o)^2 - \left(\sum_i^n w_i (\omega_i - \omega_o) \right)^2 \right)}{\sum_i^n w_i (\omega_i - \omega_o) \sum_i^n w_i \frac{(\omega_i - \omega_o)}{\omega_i} - \sum_i^n w_i (\omega_i - \omega_o)^2 \cdot \sum_i^n \frac{w_i}{\omega_i}}}, \quad (23)$$

$$\omega_g = \sqrt{\frac{\sum_i^n w_i \cdot \sum_i^n w_i (\omega_i - \omega_o)^2 - \left(\sum_i^n w_i (\omega_i - \omega_o) \right)^2}{\sum_i^n w_i (\omega_i - \omega_o)^2 \sum_i^n \frac{w_i}{\omega_i} - \sum_i^n w_i \cdot \sum_i^n w_i \frac{(\omega_i - \omega_o)}{\omega_i}}},$$

here w_i is the weight assigned to the fringe phase at the i th frequency channel.

¹⁰ Available at <http://astrogeo.org/erm>.

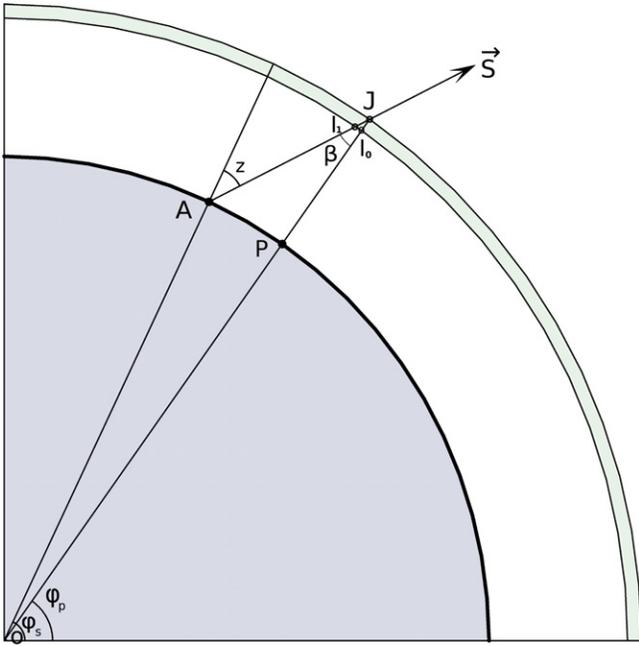


Figure 5. Ray traveling from source \vec{s} to antenna A pierces the top of the ionosphere in the point J and the bottom of the ionosphere in the point I . The ionosphere is considered a thin layer. P is the point on the Earth's surface beneath the ionosphere piercing point J .

(A color version of this figure is available in the online journal.)

They have a clear physical meaning: if the wide-band signal was replaced by a quasi-monochromatic signal with a group or phase effective ionosphere frequency, then the contribution to group or phase delay of the wide-band signal would be the same as the contribution of the quasi-monochromatic signal at those effective frequencies.

To compute the contribution of the ionosphere, we used the TEC maps from analysis of linear combinations of GPS observables made at two frequencies, 1.2276 and 1.57542 GHz. Using GPS-derived TEC maps for data reduction of astronomic observations, first suggested by Ros et al. (2000), has become a traditional approach for data processing. Analysis of continuous GPS observations from a global network comprising 100–300 stations makes it feasible to derive an empirical model of the TEC over the span of observations using the data assimilation technique. Such a model has been routinely delivered by GPS data analysis centers since 1998. For our analysis we used the TEC model provided by the GPS analysis center CODE (Schaer 1998). The model gives values of TEC in the zenith direction on a regular 3D grid with resolutions $5^\circ \times 5^\circ \times 2^h$.

For the purpose of modeling, the ionosphere is considered a thin spherical layer with constant height H above the Earth's surface. The typical value of H is 450 km. In order to compute the TEC from GPS maps, we need to know the coordinates of the point at which the ray pierces the ionosphere—point J in Figure 5. First, we find the distance from the station to the ionosphere piercing point $D = |AJ|$ by solving triangle OAJ . Noticing that $|OA| = R_\oplus$ and $|OJ| = R_\oplus + H$, we immediately get

$$\beta = \arcsin \frac{\cos E}{1 + \frac{H}{R_\oplus}},$$

$$D = R_\oplus \sqrt{2 \frac{H}{R_\oplus} (1 - \sin(E + \beta)) + \left(\frac{H}{R_\oplus}\right)^2}, \quad (24)$$

where E is the elevation of the source above the horizon.

Then the Cartesian coordinates of point J are $\vec{r} + D\vec{s}$. Transforming them into polar coordinates, geocentric latitude and longitude, we get arguments for interpolation in the 3D grid. We used the 3D B-spline interpolation by expanding the TEC field into the tensor products of basic splines of the third degree. Interpolating the TEC model output, we get the TEC through the vertical path $|JI_o|$. The slanted path $|JI_1|$ is $|JI_o|/\cos\beta$. Therefore, we need to multiply the vertical TEC by $1/\cos\beta(E)$, which maps the vertical path delay through the ionosphere into the slanted path delay. Here we neglect the ray path bending in the ionosphere. We also neglect Earth's ellipticity, since the Earth was considered spherical in the data assimilation procedure of the TEC model.

Combining equations, we get the final expression for the contribution of the ionosphere to path delay:

$$\tau_{\text{iono}} = \pm \frac{a}{4\pi^2 f_{\text{eff}}^2} \text{TEC} \frac{1}{\cos\beta(E)}, \quad (25)$$

where f_{eff} is the effective *cyclic* frequency, the plus sign is for the group delay, and the minus sign is for the phase delay.

Computation of the theoretical path delay and its partial derivatives over model parameters is made using VTD software.¹¹

3.5. Astrometric Analysis: Parameter Estimation

Astrometric analysis is made in several steps. First, each individual 24 hr session is processed independently. The parameter estimation model includes estimation of (1) clock functions presented as a sum of a second-degree polynomial and a linear spline over 60 minutes, (2) residual zenith atmosphere path delay for each station presented as a linear spline, (3) coordinates of all stations, except a reference station, and (4) coordinates of the target sources. The goal of the coarse solution is to identify and suppress outliers. The main reasons for outliers are (1) errors in determining the global maximum of the fringe amplitude during fringe search and (2) false detections. Both errors decrease with increasing the SNR. Because of this, we initially run our solution by restricting to $\text{SNR} \geq 6$, and then restore good detections with SNR in the range of [5, 6].

After identifying outliers and removing them from our solution, we apply estimates of the parameters of the a posteriori model to outliers, which allows us to predict the path delay with accuracy better than 500 ps, except for sources that had fewer than two detections. Then we re-run the fringe search for outliers and restrict the search window to ± 1000 ps with respect to predicted delay. We also lower the SNR detection limit to 4.8, since the number of independent samples in the restricted search window, and therefore, the probability of false detection at a given SNR is less. This procedure allows to restore from 50% to 80% of observations marked as outliers in the previous step. The weights of observables were computed as $w = 1/\sqrt{\sigma_o^2 + r^2(b)}$, where σ_o is the formal uncertainty of group delay estimates and $r(b)$ is the baseline-dependent re-weighting parameter that was evaluated in a trial solution to make the ratio of the weighted sum of the squares of residuals to its mathematical expectation to be close to unity using the technique similar to that used for fine fringe search.

Finally, we run a global VLBI solution that uses all available observations to date, 7.56 million, from 1980 April through 2010 August in a single LSQ run. The estimated parameters are as follows.

¹¹ Available at <http://astrogeo.org/vtd>.

1. *global* (over the entire data set): coordinates of 4924 sources, including target objects in the VGaPS campaign, positions and velocities of all stations, coefficients of B-spline expansion that model non-linear motion of 17 stations, coefficients of harmonic site position variations of 48 stations at four frequencies: annual, semi-annual, diurnal, semi-diurnal, and axis offsets for 67 stations.
2. *local* (over each session): tilts of the local symmetric axis of the atmosphere (also known as “atmospheric azimuthal gradients”) for all stations and their rates, station-dependent clock functions modeled by second-order polynomials, baseline-dependent clock offsets, and the Earth orientation parameters.
3. *segmented* (over 20–60 minutes): coefficients of linear splines that model atmospheric path delays (20 minutes segment) and clock functions (60 minutes segment) for each station. The estimates of clock functions absorb uncalibrated instrumental delays in the data acquisition system.

The rate of change for the atmospheric path delays and clock functions between adjacent segments was constrained to zero with weights reciprocal to 1.1×10^{-14} and 2×10^{-14} , respectively, in order to stabilize solutions. We apply no-net rotation constraints on the positions of 212 sources marked as “defining” in the ICRF catalog (Ma et al. 1998) that requires the positions of these sources in the new catalog to have no rotation with respect to the position in the ICRF catalog to preserve continuity with previous solutions.

The global solution sets the orientation of the array with respect to an ensemble of ~ 5000 extragalactic remote radio sources. The orientation is defined by the series of Earth orientation parameters and parameters of the empirical model of site position variations over 30 years evaluated together with source coordinates. Common sources observed in VGaPS as atmosphere and amplitude calibrators provide a connection between the new catalog and the old catalog of compact sources.

3.6. Astrometric Analysis: Assessment of Weights of Observations

As follows from the Gauss–Markov theorem, the estimate of parameters has minimum dispersion when observation weights are chosen reciprocal to the variance of errors. The group delays used in the analysis have errors due to the thermal noise in fringe phases and due to mismodeling theoretical path delay in the atmosphere

$$\sigma^2 = \sigma_{\text{th}}^2 + \sigma_{\text{io}}^2 + \sigma_{\text{na}}^2, \quad (26)$$

where σ_{th}^2 is the thermal noise, and σ_{io}^2 and σ_{na}^2 are the contribution of the ionosphere and the neutral atmosphere to the error budget, respectively.

3.6.1. A Priori Errors of the GPS Ionosphere Model

The first term, σ_{th}^2 , was estimated during the fringe fitting. The second term can only be evaluated indirectly. Sekido et al. (2003) used six dual-band intercontinental VLBI sessions at 10 stations in 2000 July and compared TEC values from GPS with TEC estimated from VLBI observables. They drew a conclusion that the errors in path delay derived from the GPS TEC model were at the range of 70 ps in the zenith direction at 8.6 GHz. However, the ionosphere path delay is a non-stationary process. Therefore, great caution should be taken in an attempt to generalize conclusions made from analysis of a small network over a short time period. The non-stationarity of ionospheric

fluctuations implies that an exact expression for the variance of the ionosphere fluctuations during any given period does not exist, and any expression for the variance is an approximation.

Since 1998 June when the GPS TEC maps became available through 2010 August, more than 2000 dual-band S/X VLBI sessions under geodesy and absolute astrometry programs were carried out, including 92 sessions at the VLBA. It can be easily shown that the contribution of the ionosphere in the X band, τ_{xi} , can be found from the linear combination of group delay observables with coefficients that are expressed through effective ionosphere frequencies at these bands, ω_x and ω_s , defined in expression (23):

$$\tau_{xi} = (\tau_x - \tau_s) \frac{\omega_s^2}{\omega_x^2 - \omega_s^2}. \quad (27)$$

We used this data set to evaluate the errors of the contribution of the ionosphere to group delays derived from GPS TEC maps, considering the ionosphere contribution from dual-band VLBI observations as true for the purpose of this comparison. We computed the ionosphere contribution from the GPS model and from VLBI observations for each session. The root mean squares (rms) of the differences of the contribution VLBI–GPS was computed for all sessions and all baselines. We sought regressors that can predict rms(VLBI–GPS). We expect the short-term variability of the ionosphere at scales less than several hours to dominate the errors of the GPS model. The sparseness of the GPS network and limited sky coverage result in missing high frequency spatial and temporal variations of the ionosphere. The turbulent nature of the ionosphere path delay variations suggests that the rms of the ionosphere model errors due to missed high frequency variations will be related to the rms of the low frequency variations either as a linear function or as a power law. After many tries we found that the following parameter can serve as a regressor: $\text{RG} = \sqrt{\text{rms}_{g1}^2 + \text{rms}_{g2}^2}$, where rms_{gi} is the rms of the GPS path delay at the i th station of a baseline during a session. The rms_{g1} and rms_{g2} are highly correlated at short baselines and $\sqrt{\text{rms}_{g1}^2 + \text{rms}_{g2}^2} > \text{rms}(g2 - g1)$. At long baselines the ionosphere contribution de-correlates. Therefore, the dependence of rms(VLBI–GPS) versus RG will depend on the baseline length and possibly on other parameters. Figure 6 shows examples of this dependence for a short baseline and for a long baseline.

It is remarkable that rms(VLBI–GPS) versus RG fits reasonably well with a linear function. We computed rms(VLBI–GPS) for each baseline and fitted it to the linear model $F + S \cdot \text{RG}$. The floor of the linear fit has a mean value around 20 ps (Figure 7). As expected, the slope of the fit increases with baseline length as shown in Figure 8. The growth is linear up to baseline lengths of 2000 km, which apparently correspond to the decorrelation of the paths through the ionosphere. The growth of the slope beyond baseline lengths of 2000 km is slower and it shows more scatter.

We use this dependence to predict the rms of the GPS ionosphere model errors. For each station that participated in the experiment we computed the rms of the ionosphere variations in zenith direction. Then we express the predicted variance of the GPS ionosphere model errors as

$$\sigma_i^2 = \left(F_b \frac{f_i^2}{f_{\text{eff}}^2} \right)^2 + \left(\tau_{\text{iono},1} - \frac{\bar{\tau}_1^z}{\cos \beta(E_1)} \right)^2 S_b^2 + \left(\tau_{\text{iono},2} - \frac{\bar{\tau}_2^z}{\cos \beta(E_2)} \right)^2 S_b^2, \quad (28)$$

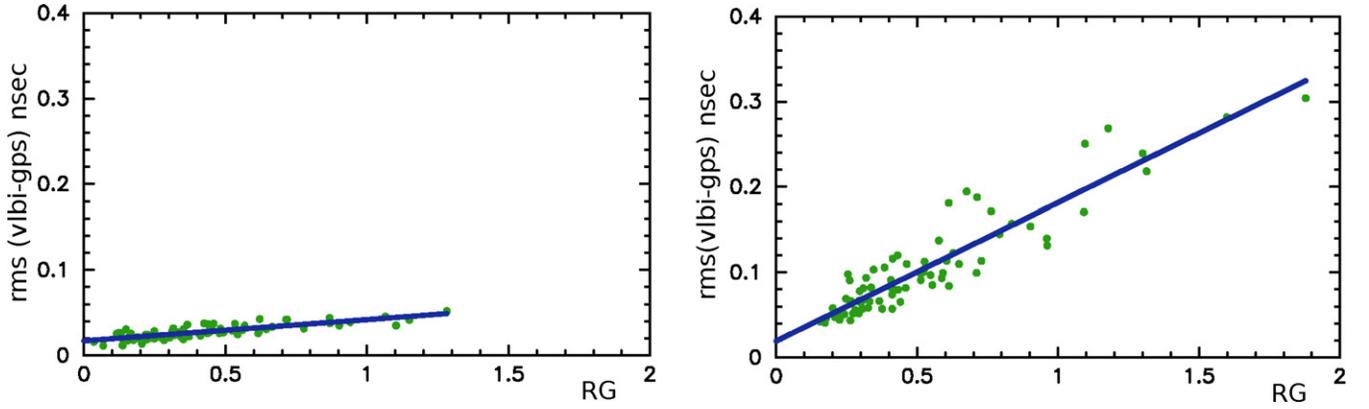


Figure 6. Dependence of the rms of the differences VLBI minus GPS as a function of $RG = \sqrt{\text{rms}_{g1}^2 + \text{rms}_{g2}^2}$ for baseline FD-VLBA/PIETOWN, 565 km long (left) and for baseline LA-VLBA/MK-VLBA, 4970 km long (right). The quantity rms_{gi} is the rms of GPS path delay at the i station of a baseline during a session. Each green dot corresponds to one VLBI session. The thick blue straight line is a linear fit through the data.

(A color version of this figure is available in the online journal.)

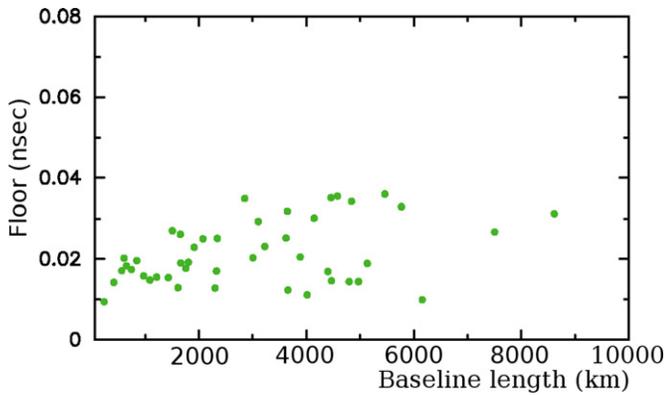


Figure 7. Floor of the regression model of the dependence of $\text{rms}(\text{VLBI-GPS})$ of $RG = \sqrt{\text{rms}_{g1}^2 + \text{rms}_{g2}^2}$ for all VLBA baselines as a function of the baseline length.

(A color version of this figure is available in the online journal.)

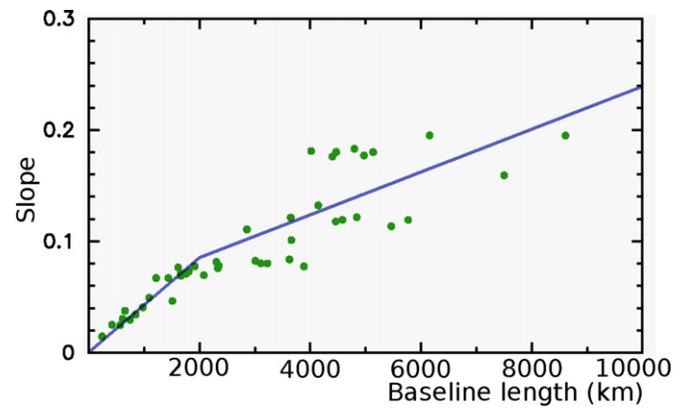


Figure 8. Slope of the regression model of dependence of $\text{rms}(\text{VLBI-GPS})$ of $RG = \sqrt{\text{rms}_{g1}^2 + \text{rms}_{g2}^2}$ for all VLBA baselines as a function of the baseline length. The straight lines show a linear approximation of the slope for two ranges of the baseline length: below and over 2000 km.

(A color version of this figure is available in the online journal.)

where $\tau_{\text{iono},i}$ is the ionosphere path delay at the i th station computed using the GPS TEC maps, $\bar{\tau}_1^z$ is the zenith ionosphere path delay from GPS TEC maps averaged over a 24^h period with respect to the central date of the session, F_b and S_b are the parameters of the linear model $\text{rms}(\text{VLBI-GPS})$ versus RG for a given baseline, f_i is the frequency for which the model was computed (8.6 GHz), and f_{eff} is the frequency of the experiment for which the model is applied (24.5 GHz). The term $\tau_{\text{iono},i} - \frac{\bar{\tau}_1^z}{\cos \beta(E_i)}$ is the difference between the ionosphere path delay from GPS at a given direction and the average ionosphere path delay scaled to take into account the elevation dependence. This term is an approximation of rms_{g1} used for computation of RG .

3.6.2. A Priori Errors of the Path Delay in the Neutral Atmosphere

Rigorous analysis of the errors of modeling the path delay in the neutral atmosphere is beyond the scope of this paper. Assuming the dominant errors of the a priori model are due to high frequency fluctuations of water vapor at scales less than 3–5 hr, we seek a regression model in the form of dependence of the rms of errors on the total path delay in the non-hydrostatic component of the path delay. We made several trial runs using all 123 observing sessions at the VLBA under geodesy and absolute astrometry programs with reciprocal weights modified

according to

$$\sigma_{\text{used}}^2 = \sigma^2 + \left(a \cdot \frac{\tau_s}{\tau_z} \right)^2. \quad (29)$$

Here, τ_s is the contribution of the non-hydrostatic constituent of the slanted path delay and τ_z is the non-hydrostatic path delay in the zenith direction computed by direct integration of the equations of wave propagation through the atmosphere using the refractivity computed from the MERRA model, and a is the coefficient. We found that when coefficient $a = 0.02$ is used, the baseline length repeatability, defined as the rms of the deviation of baseline length with respect to the linear time evolution, reaches the minimum. We adopted the value 0.02 in our analysis of VGaPS experiments. For typical values of τ_z , the added noise is 8 ps in the zenith direction and 80 ps at 10° elevation.

3.6.3. Ad Hoc Added Variance of the Noise

For each baseline and each session we also computed an ad hoc variance of observables that, added in quadrature, makes the ratio of the weighted sum of squares of post-fit residuals to their mathematical expectation close to unity in a similar way as we updated fringe phase weights. Expression (14) was used

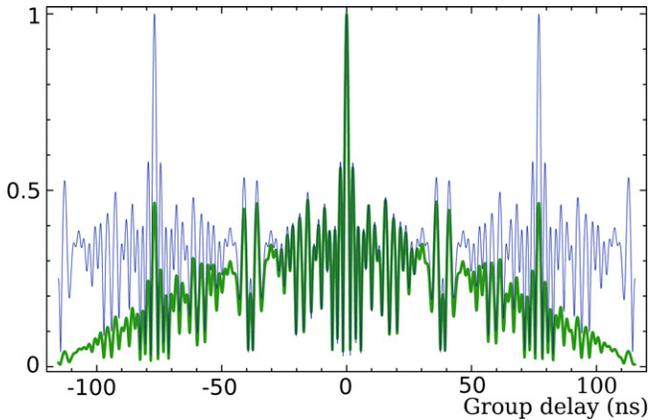


Figure 9. Coherent sum of cross-correlation amplitude, normalized to its value at the maximum as a function of group delay shift with respect to the maximum for VGaPS observing sessions. The thick green line shows result of processing with PIMA. The thin blue line shows results of processing with AIPS. The pattern of the thin blue line repeats with a period of 76.923 ns.

(A color version of this figure is available in the online journal.)

to compute this variance. This ad hoc variance was applied to further inflate the a priori observable uncertainties that have already been corrected for the inaccuracy of the a priori model of wave propagation through the ionosphere and the neutral atmosphere according to expressions (28) and (29). In contrast to expressions (28) and (29), the baseline-dependent ad hoc variance is elevation-independent.

4. VALIDATION OF THE WIDE-BAND FRINGE SEARCH ALGORITHM

Using PIMA, we detected 327 target sources versus 136 targets detected with AIPS, since the AIPS detection limit is lowered by a factor of $\sqrt{8} = 2.83$. Thus, the yield of the experiments was improved by a factor of 2.4—a very significant improvement that well justified our efforts to create a new software package for processing astrometric observations.

Another difference is that the results of processing the data with AIPS have ambiguities in group delay that are reciprocal to the minimum difference between intermediate frequencies. This means that the group delays are $\tau_g + K\tau_s$, where $\tau_s = 76.923$ ns and K is an arbitrary integer number. At the same time, the results of processing with the wide-band fringe fitting algorithm do not suffer this problem. The reason for group delay ambiguities is that the narrow-band fringe fitting algorithm implemented in AIPS first coherently averages the data within each IF, and in the second step of fringe fitting it processes a rail of narrow-band signals. The Fourier transform that describes the dependency of the amplitude of the coherent sum of the cross-spectrum on group delay has a periodicity that is reciprocal to the minimum frequency separation of IFs, 76.923 ns in our case. The wide-band fringe fitting algorithm does not average the spectrum. Therefore, the periodicity of the Fourier transform of the coherent sum of the cross-spectrum in the wide-band algorithm is equal to the spectral resolution, i.e., 2 mks for VGaPS experiments. Figure 9 illustrates this difference.

The lack of group delay ambiguities has a profound effect on determining source positions with poorly known a priori coordinates. In the presence of group delay ambiguities, we had to solve for source positions first using less precise so-called narrow-band group delays determined by arithmetic averaging group delays computed for each IF independently. The accuracy of these source position estimates is often insufficient to reliably

Table 5
Solution Statistics from 33 Global RDV Sessions
Processed with AIPS and PIMA

Statistics	AIPS	PIMA	
		SNR _{min} = 5.0	SNR _{min} = 10.0
No. of obs used	467 769	531 299	472 717
fit wrms	18.40 ps	21.22 ps	19.65 ps
No. of sources	776	800	773
wrms $\Delta\Psi \cos \epsilon_o$	0.10 mas	0.12 mas	0.14 mas
wrms $\Delta\epsilon$	0.10 mas	0.10 mas	0.12 mas
Bas. rep. at 5000 km	4.81 mm	4.75 mm	4.96 mm
Bas. rep. at 9000 km	8.54 mm	8.08 mm	7.95 mm

Notes. The statistics in the central column were computed using all observations. The statistics in the right column were computed using observations with SNR > 10.

resolve group delay ambiguities, especially in the presence of narrow-band group delay outliers. When the number of used observations is large, say more than 10, the data redundancy allows us to detect the presence of incorrectly resolved ambiguities and fix the problem, but if the number of observations is small, chances are the error in group delay ambiguity resolution will not be noticed. In the past, source position estimates made with less than eight observations were considered unreliable. The use of the wide-band fringe fitting algorithm eliminates this problem entirely. Reprocessing the old data revealed that group delay ambiguities for a considerable number of sources were indeed resolved incorrectly, which resulted in source position errors as large as 4'! In contrast, the wide-band fringe fitting algorithm provides reliable estimates of source positions using a minimum redundancy of three observations.

Since the wide-band fringe search algorithm is new, we would like to be sure that the new algorithm does not introduce new systematic errors with respect to the old one. As a validation test, we re-processed a set of 33 VLBA absolute astrometry/geodesy experiments under the RDV program (Petrov et al. 2009) and 12 VLBA absolute astrometry experiments under the K/Q program (Lanyi et al. 2010). Each test experiment had a duration of 24 hr.

The RDV experiments were observed on a global network, including all 10 VLBA stations, with dual-band receivers at 8.6 GHz (X band) and 2.3 GHz (S band), with four IFs allocated to the X band and four IFs allocated to the S band. Fringe fitting, outlier elimination, re-weighting, and in the case of AIPS, group delay ambiguity resolution, were made independently using PIMA and AIPS. Subsequent data reduction and parameter estimation were made using identical setups. Estimated parameters of the solution were the same as in processing the VGaPS sessions, except for treatment of site positions: they were treated as local parameters, i.e., estimated for each session independently.

The statistics of the solution for 33 global RDV sessions using AIPS and PIMA are shown in Table 5. The weighted root mean squares (wrms) of the post-fit residuals is larger in the PIMA solution for two reasons. First, the PIMA solution contains 14% more points, mainly with SNRs in the range [5.0, 10.0], that were undetected by the traditional AIPS algorithm. The errors of these observables are systematically larger. We reran the PIMA solution and excluded all points with SNR in either the X or S band less than 10. The difference in wrms post-fit residuals was significantly reduced. The second reason is that the group delay formal errors were inflated in PIMA processing to make the ratio of post-fit residuals of fringe fitting to its mathematical expectation close to 1. This was not done for the AIPS solutions.

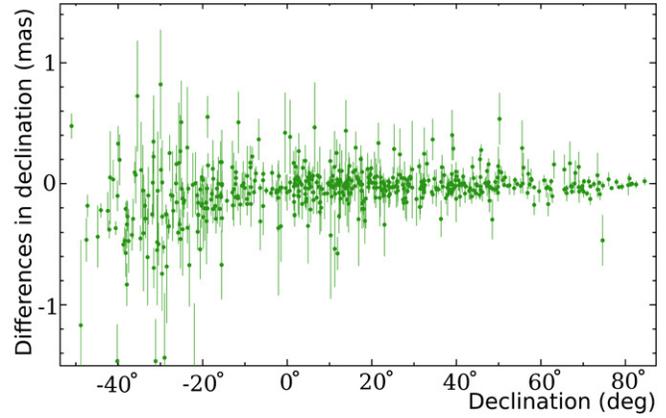
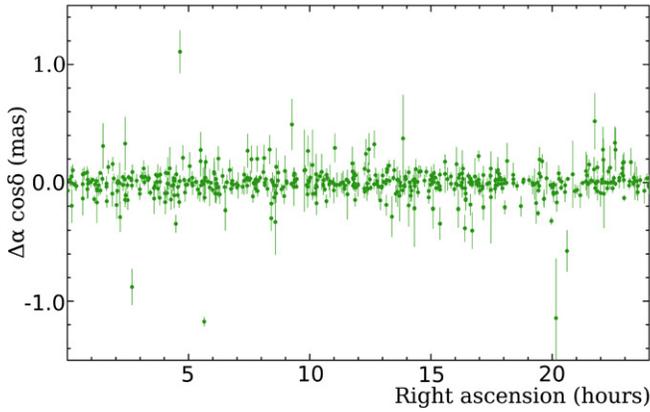


Figure 10. Systematic differences in source coordinate estimates between solutions using group delays derived by PIMA and by AIPS. Left plot shows $\Delta\alpha \cos \delta$ (α), right plots shows $\Delta\delta$ (δ).

(A color version of this figure is available in the online journal.)

Table 6
Differences Between AIPS and PIMA Positions
of 528 Sources Observed in RDV Experiments

Statistics	AIPS–PIMA SNR _{min} = 5.0	AIPS–PIMA SNR _{min} = 10.0
wrms $\Delta\alpha \cos \delta$ (α)	0.072 mas	0.020 mas
wrms $\Delta\delta$ (δ)	0.088 mas	0.032 mas

Notes. The differences in the left column were computed using all observations. The differences in the right column were computed using observations with SNR > 10.

Another test of goodness of the solution and possible presence of systematic errors is the so-called baseline length repeatability test (Petrov et al. 2009). We computed the wrms of baseline length estimates with respect to the fitted linear model of their evolution with time. The dependence of the baseline length wrms with the length of baselines L is fitted by function $R(L) = \sqrt{A^2 + B^2L}$. Values of $R(L)$ at $L = 5000$ km and $L = 9000$ km are presented in Table 5. We also computed the wrms of the deviations of estimates of daily offsets of nutation angles in longitude, $\Delta\psi$, and nutation in obliquity, $\Delta\epsilon$, with respect to the empirical nutation expansion heo_20091201. The statistics in Table 5 show the satisfactory agreement between AIPS and PIMA solutions and do not reveal any systematic differences.

Since the goal of VGaPS was to derive source positions, comparison of the positions from AIPS and PIMA processing is important to evaluate the level of systematic differences. We thus computed the differences in source coordinates $\Delta\alpha \cos \delta$ (α) and $\Delta\delta$ (δ). We restricted our analysis to 528 sources that had more than 64 observations in both AIPS and PIMA in order to avoid effects of a greater number of observations available in the PIMA solutions. Plots of these differences are shown in Figure 10 and position comparisons are shown in Table 6. The table contains the results from 528 sources observed in RDV sessions with at least 64 used observations. The second column shows statistics of differences between AIPS and PIMA with an SNR cutoff of 5. The third column shows the differences between AIPS and PIMA solutions with an SNR cutoff of 10. These differences are only about 20% of a typical position uncertainty.

In the right plot of Figure 10, we can see small systematic declination differences at low declinations. Similar differences but produced from the PIMA solution made with an SNR cutoff of 10.0, shown in Figure 11, help us to understand the origin

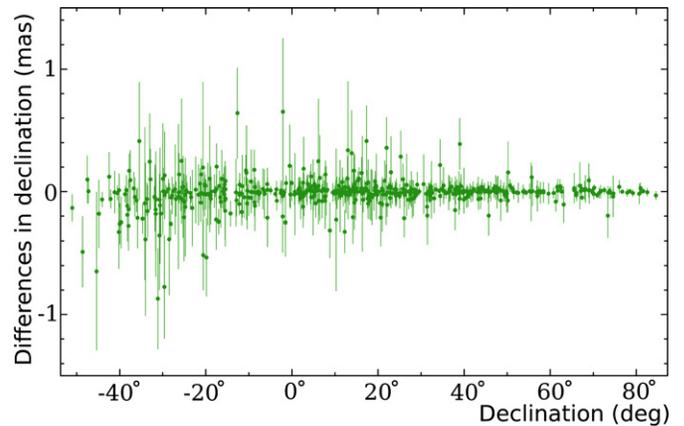


Figure 11. Systematic differences $\Delta\delta$ (δ) in source coordinate estimates between solutions using group delays derived by PIMA with SNR cutoff 10 and by AIPS. (A color version of this figure is available in the online journal.)

of this pattern. When the SNR cutoff is raised to 10, the wrms of differences are reduced by two to three times to 0.02–0.03 mas and the systematic pattern disappears. For comparison, the average formal uncertainties for declinations are 0.12 mas and 0.07 mas for right ascensions scaled by $\cos \delta$, so the systematic errors are small. Sources at low declinations observed on a VLBI array located in the Northern Hemisphere are necessarily taken at low elevations. Including additional low SNR observations at low elevations in a solution, which are unavailable in the AIPS solutions, changes the contribution of systematic errors due to mismodeling path delay in the neutral atmosphere that are higher at low elevations. At present, it is not clear whether including observations with SNRs in the range 5–10 increases systematic errors, or the opposite. However, the magnitude of the differences, less than 0.5 mas as declinations in the range $[-50^\circ, -25^\circ]$, does not raise a concern.

Finally, we have re-analyzed 12 VLBA experiments under the K/Q program. The frequency setup in the K/Q and VGaPS campaigns was identical. Lowering the detection limit by a factor of $\sqrt{8} \approx 2.83$ with the use of the wide-field fringe fitting algorithm greatly helped. Processing the data with PIMA failed to detect only 8 out of the 340 observed sources, with 61 non-detections in data processing with AIPS. We compiled Table 7 from statistics of the K/Q solution in a form similar to Table 5. Analysis of source position differences did not reveal

Table 7
Solution Statistics from 12 VLBA Experiments at 24 GHz
under the K/Q Program Processed with AIPS and PIMA

Statistics	AIPS	PIMA	
		SNR _{min} = 5.0	SNR _{min} = 14.1
No. of obs used	104 887	139 213	105 761
fit wrms	19.50 ps	23.08 ps	19.19 ps
No. of sources	279	332	276
wrms $\Delta\Psi \cos \epsilon_o$	0.13 mas	0.10 mas	0.09 mas
wrms $\Delta\epsilon$	0.18 mas	0.13 mas	0.14 mas
Bas. rep. at 5000 km	5.56 mm	4.58 mm	4.58 mm
Bas. rep. at 9000 km	9.29 mm	7.36 mm	7.32 mm

Notes. The statistics in the central column were computed using all observations. The statistics in the right column were computed using observations with SNR > 10.

Table 8
Differences between AIPS and PIMA Positions of 244
Sources Observed in the K/Q VLBA Experiments

Statistics	AIPS–PIMA	
	SNR _{min} = 5.0	SNR _{min} = 10.0
wrms $\Delta\alpha \cos \delta(\alpha)$	0.062 mas	0.060 mas
wrms $\Delta\delta(\delta)$	0.108 mas	0.096 mas

Notes. The differences in the left column were computed using all observations. The differences in the right column were computed using observations with SNR > 10.

any pattern of systematic errors. The statistics of these differences presented in Table 8 do not exceed formal uncertainties of positions which are 0.08 mas and 0.14 mas in right ascensions scaled by $\cos \delta$ and declination, respectively. The first column is four sources with an SNR cutoff of 5, and the second column with an SNR cutoff of $5 \cdot \sqrt{8} = 14.1$. The baseline length repeatability and the wrms of nutation offset time series from PIMA solutions are 20%–30% smaller. We do not have an explanation for why the PIMA solution produces noticeably better results for 24 GHz observations, but no significant improvement was found from analysis of S/X RDV experiments.

The above results of analysis of validation runs processing 0.6 million observations at K, X, and S bands, collected during 1080 hr of recording at the VLBA and the global VLBI network with both AIPS and PIMA demonstrate that the new wide-band algorithm for group delays does not introduce any significant systematic errors while detecting more sources because it evaluates group delays using the coherent sum of the data across the wide-band. We conclude that PIMA has passed major validation tests.

5. INVESTIGATION OF SYSTEMATIC ERRORS IN SOURCE POSITIONS

Using single-band data, ionosphere path delay mismodeling may produce systematic position errors. Lanyi et al. (2010) showed that in their analysis of K-band observations systematic errors reached several mas and had a tendency to be larger at low declinations. In Section 3.6.1 we evaluated the rms of random errors caused by ionosphere path delay mismodeling. However, inflating weights to account for the variance of errors in general does not guarantee that source positions will have no systematic errors.

To evaluate the magnitude of possible ionosphere driven systematic errors we made the following Monte Carlo simulation.

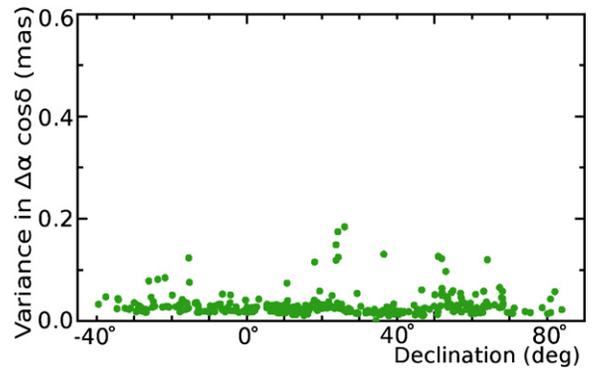


Figure 12. Modeled systematic errors $\Delta\alpha_i(\delta) \cos \delta$ driven by the mismodeling ionosphere path delay contribution evaluated from the Monte Carlo simulation. (A color version of this figure is available in the online journal.)

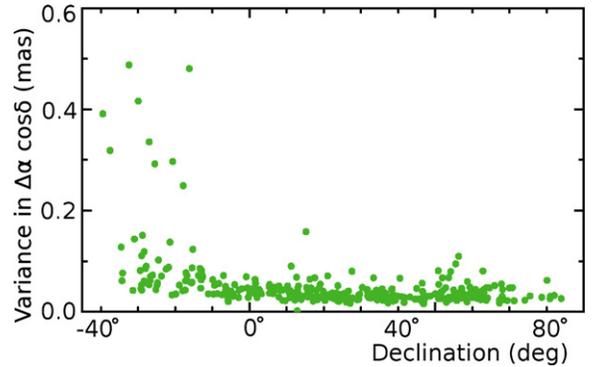


Figure 13. Modeled systematic errors $\Delta\delta_i(\delta)$ driven by the mismodeling ionosphere path delay contribution evaluated from the Monte Carlo simulation. (A color version of this figure is available in the online journal.)

We added to the theoretical path delay the zero-mean Gaussian noise $M \cdot N(0, \sigma_i)$, with a variance σ_i computed according to expression (28). The noise was magnified M times in order to make the contribution of ionosphere path delay errors dominant over other sources of errors. We made 64 analysis runs of VGaPS data using different seeds of the random noise generator. The magnification factor 100 was used. Thus, we produced 64 estimates of position of each source with added noise. We computed the rms of position estimates of each target source and divided it by M . To the extent of the validity of expression (28), these rms's represent expected errors due to the inadequacy of the ionosphere path delay models based on using the TEC models derived from GPS analysis. Plots of $\Delta\alpha_i(\delta)$ and $\Delta\delta_i(\delta)$ errors are shown in Figures 12 and 13. For 90% of the sources, errors are at the level of 0.02–0.04 mas. $\Delta\delta(\delta)$ increases to 0.1 mas at declinations less than -20° , and for some sources may reach 0.4 mas. The disparity in systematic errors for sources at comparable declinations reflects the disparity in the number of observables used in the solution.

For assessment of remaining systematic errors we exploited that fact that 56 known sources were observed as amplitude and atmospheric calibrators. Positions of these sources are known from previous dual-band the S/X observations with accuracies better than 0.1 mas. We split the set of 56 calibrators into two subsets of 28 objects and ran two additional solutions. In the first solution we suppressed 28 calibrators in all sessions except VGaPS and determined their positions solely from VGaPS. In the second solution we did the same with the second subset. Considering that the positions of calibrators from numerous

S/X observations can be regarded as true for the purposes of this comparison, we treated the differences as VGaPS errors.

We computed the χ^2 per degree of freedom statistics for the differences in right ascensions and declinations $\Delta\alpha$ and $\Delta\delta$ and sought additional variances v_α and v_δ which, being added in quadrature to the source position uncertainties, will make them close to unity:

$$\frac{\chi_\alpha^2}{\text{ndf}} = \frac{\sum_{k=1}^{k=n} \Delta\alpha_k^2 \cos^2 \delta_k}{n \sum_{k=1}^{k=n} \sqrt{\sigma_{\alpha,k}^2 \cos^2 \delta_k + \alpha_{i,k}^2 + v_\alpha^2 \cos^2 \delta_k}}$$

$$\frac{\chi_\delta^2}{\text{ndf}} = \frac{\sum_{k=1}^{k=n} \Delta\delta_k^2}{n \sum_{k=1}^{k=n} \sqrt{\sigma_{\delta,k}^2 + \delta_{i,k}^2 + v_\delta^2}}. \quad (30)$$

The denominator in Equation (30) is a mathematical expectation of the sum of squares of differences, provided the estimates of source positions are statistically independent.

We found the following additive corrections of the uncertainties in right ascensions scaled by $\cos \delta$ and for declinations, respectively: $v_\alpha = 0.08$ mas and $v_\delta = 0.120$ mas.

The final inflated errors of source positions, $\sigma_\alpha^2(f)$ and $\sigma_\delta^2(f)$, are

$$\sigma_\alpha^2(f) = \sigma_\alpha^2 + v_\alpha^2 + \alpha_i^2 / \cos^2 \delta,$$

$$\sigma_\delta^2(f) = \sigma_\delta^2 + v_\delta^2 + \delta_i^2. \quad (31)$$

The positions of some sources may indeed be different due to the core-shift effect (Kovalev et al. 2008; Porcas 2009). Treatment of this shift as VGaPS errors makes our estimates of re-weighting parameters and therefore reported final inflated errors somewhat too conservative.

6. IMAGING

For imaging purposes we performed the a priori calibration of the data following a traditional method, using AIPS (Greisen 1988). In the future, we plan to introduce all extra steps required for an accurate amplitude calibration into PIMA as well.

We followed the usual AIPS initial VLBA calibration procedure involving a priori amplitude calibration with measured antenna gain curves and system temperatures as well as sampling-based calibration adjustments. Atmospheric absorption is significant at 24 GHz. We have estimated its effect using system temperature data covering the whole range of elevations and weather information in order to adjust visibility amplitudes for the opacity. Typical values of the opacity were found to range between 0.03 and 0.1 for different VLBA telescopes and observing epochs. We performed phase calibration using the phase calibration signal injected during observations and fringe fitting. A separate solution for station-based group delay and phase delay rate was made for each frequency channel (IF). As the final step of calibration, bandpass corrections were determined and applied.

Our observations were scheduled around 24 GHz since the K-band continuum performance is better at this frequency, away from the water line. Unfortunately, at the time of these observations most of the VLBA telescopes had no gain curve measurements close to 24 GHz. That has changed since 2007 when regular 24 GHz gain curve measurements started to be performed at all VLBA stations. For all VLBA antennas except MK-VLBA and HN-VLBA we have used gain curves measured at 22.2 GHz while for the former the curves at 23.8 GHz were applied. Antenna efficiency and the noise diode spectrum change

with frequency (see, e.g., Petrov et al. 2007b, as well as results of VLBA gain curve measurements at 22 and 24 GHz after 2007).¹² This is one of the main sources of the total amplitude calibration uncertainty. Additionally, IFs in our experiment are widely spread (Table 2) which might introduce extra amplitude shifts. We used strong flat-spectrum sources in the sample in order to estimate global relative amplitude correction factors for different IFs but did not find any to be larger than 10% with high confidence. No extra frequency channel specific amplitude corrections were applied to the data.

After a priori calibration, data were imported to the Caltech DIFMAP package (Shepherd 1997), visibility data were flagged, and images were produced using an automated hybrid imaging procedure originally suggested by Greg Taylor (Pearson et al. 1994) which we optimized for our experiment. The procedure performs iterations of phase and amplitude self-calibration followed by CLEAN image reconstruction. We were able to reach the VLBA image thermal noise level for most of our final CLEAN images. Examples of three images for compact and resolved objects are shown in Figure 14.

Total errors of our measurements of correlated flux density values for sources stronger than ~ 200 mJy were dominated by the accuracy of the amplitude calibration described above. This considers the error of amplitude calibration as not exceeding 15% and this estimate is confirmed by our comparison of the flux densities integrated over the VLBA images with the single-dish flux densities which we measured with RATAN-600 in 2006 for slowly varying sources without extended structure. Details of the RATAN-600 single-dish observing program including the method of observations and data processing can be found in Kovalev et al. (1999, 2002).

7. THE CATALOG OF SOURCE POSITIONS

Of 487 sources observed, three or more detections were found for 327 objects. After careful identification and removal of outliers due to the incorrect selection of the global maximum for weak sources with $\text{SNR} < 6$, we selected 33,452 observations out of 59,690 from three VGaPS experiments for analyzing in the single LSQ solution together with 7.56 million other VLBI observations. The semi-major error ellipses of inflated position errors for all sources except 0903+154 vary in the range 0.21–20 mas with the median value of 0.85 mas. The histogram of position errors is shown in Figure 15.

The VGaPS catalog is listed in Table 9. Although positions of all 5047 astrometric sources were adjusted in the LSQ solution that included the VGaPS sources, only coordinates of 327 target sources observed in the VGaPS campaign are presented in the table. The first column gives calibrator class: ‘‘C’’ if the source is recommended as a calibrator or ‘‘U’’ if it has an unreliable position, since there were less than five detections and there is a risk that the secondary maximum of the coherent sum of weighted complex cross-correlation samples has been picked and has not been flagged out. The second and third columns give the IVS source name (B1950 notation) and IAU name (J2000 notation). The fourth and fifth columns give source coordinates at the J2000.0 epoch. Columns 6 and 7 give source position uncertainties in right ascension and declination in mas after applying inflated errors according to Equation (31) (without $\cos \delta$ factor), and Column 8 gives the correlation coefficient between the errors in right ascension and declination. The number of group delays used for position

¹² http://www.vlba.nrao.edu/astro/VOBS/astronomy/vlba_gains.key

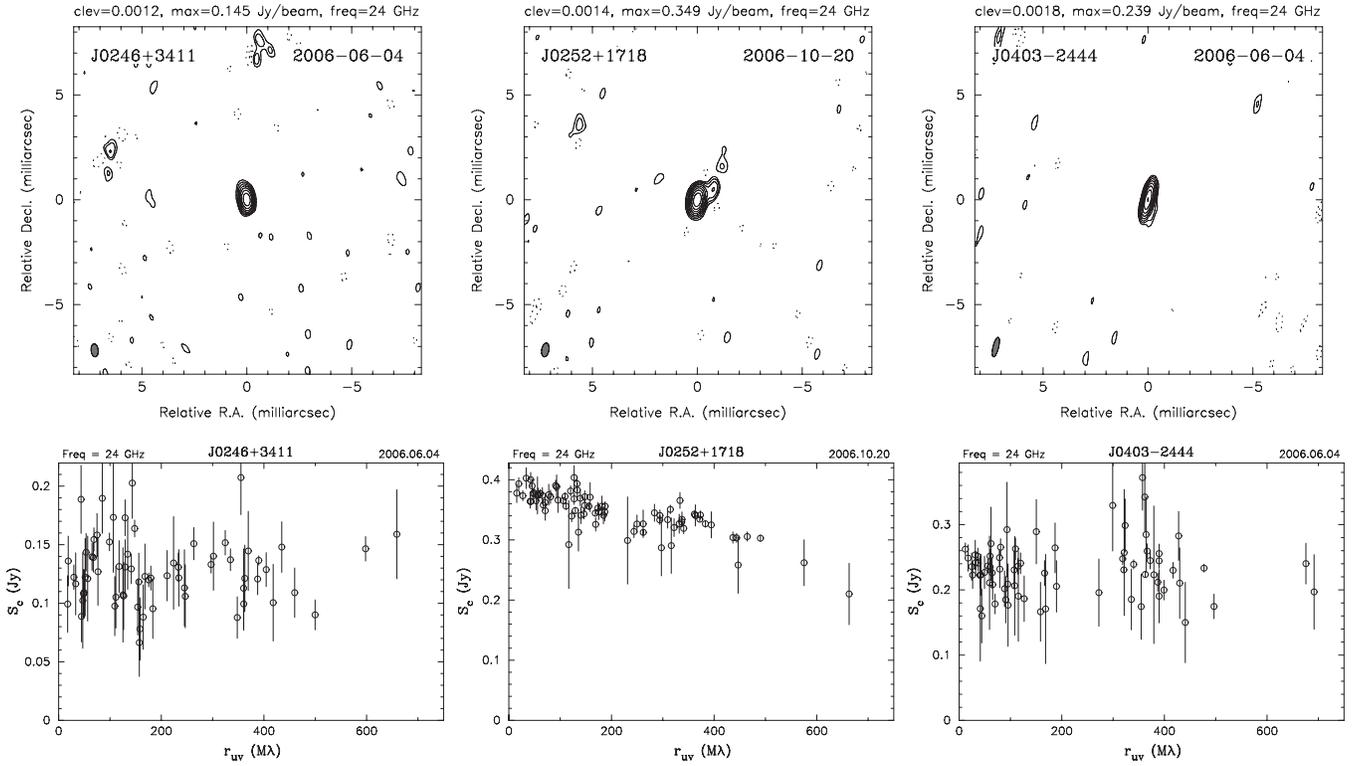


Figure 14. From top to bottom. *Row 1:* naturally weighted CLEAN images at 24 GHz. The lowest contour is plotted at the level given by “clev” in each panel title (Jy beam^{-1}), the peak brightness is given by “max” (Jy beam^{-1}). The contour levels increase by factors of two. The dashed contours indicate negative flux. The beam is shown in the bottom left corner of the images. *Row 2:* dependence of the correlated flux density on projected spacing. Each point represents a coherent average over one 2 minute observation on an individual interferometer baseline. The error bars represent only the statistical errors.

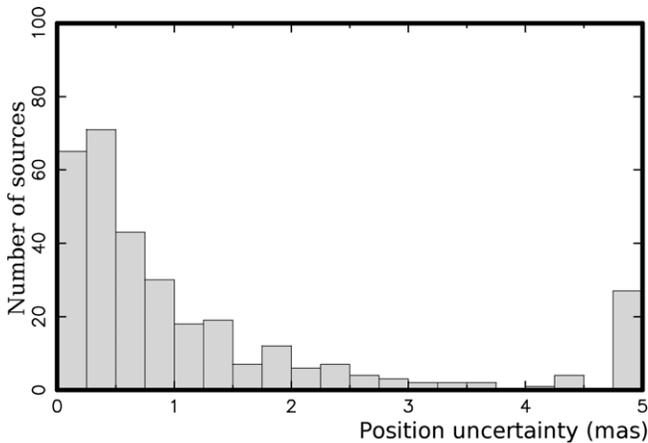


Figure 15. Histogram of the semi-major axes of inflated position error ellipses among 327 target sources in the VGaPS catalog. The last bin shows errors exceeding 4.75 mas.

determination is listed in Column 9. Column 10 gives the estimate of the flux density integrated over the entire map. This estimate is computed as a sum of all CLEAN components if a CLEAN image was produced. If we did not have enough detections of the visibility function to produce a reliable image, the integrated flux density is estimated as the median of the correlated flux density measured at projected spacings less than $70 \text{ M}\lambda$. The integrated flux density is the source total flux density with spatial frequencies less than $12 \text{ M}\lambda$ filtered out, or in other words, the flux density from all details of a source with size less than 20 mas . Column 11 gives the flux density of unresolved components estimated as the median of correlated flux density values measured at projected spacings greater than $400 \text{ M}\lambda$. For

some sources no estimates of the unresolved flux density are presented, because either no data were collected at the baselines used in calculations, or these data were unreliable.

An online version of this catalog is available.¹³ For each source it has four references: to a FITS file with CLEAN components of naturally weighted source images, to a FITS file with calibrated visibility data, to a postscript map of a source, and to a plot of correlated flux density as a function of the length of the baseline projection to the source plane.

In Table 10, we present a priori coordinates, total flux densities extrapolated to 24 GHz and spectral index estimates for 160 target objects that have not been detected in the VGaPS experiment. Some of these sources were detected in other VLBI astrometry experiments in the S/X bands.

8. COMPARISON OF K - AND S/X -BAND ASTROMETRIC VLBI POSITIONS

We searched the VLBI archive and found that among our target sources, 206 were observed with S/X at the VLBA under VCS and RDV programs before 2010 November. We investigated the differences in K -band observations against independent S/X observations. We restricted our analysis to 192 objects that had uncertainties from X/S and K -band solutions less than 5 mas. Figures 16 (top and bottom) show the differences in right ascensions and declinations.

8.1. Special Cases: 3C 119 and 3C 410

Positions of two sources, J0432+4138 (also known as 3C119) and J2020+2942 (3C410) are found to be significantly off,

¹³ <http://astroge.org/vgaps>

Table 9
VGaPS Catalog

Class (1)	Source Name		J2000.0 Coordinates		Errors (mas)			No. of Obs. (9)	Correlated Flux Density (in Jy)	
	IVS (2)	IAU (3)	Right Ascension (4)	Declination (5)	$\Delta\alpha$ (6)	$\Delta\delta$ (7)	Corr (8)		Total (10)	Unres (11)
C	2358+605	J0001+6051	00 01 07.099852	+60 51 22.79800	0.94	0.51	-0.115	123	0.11	0.11
C	2359+606	J0002+6058	00 02 06.696680	+60 58 29.83950	4.42	1.92	-0.168	26	0.09	<0.08
C	0002+541	J0005+5428	00 05 04.363368	+54 28 24.92414	0.60	0.47	0.167	79	0.34	0.11
C	0003+505	J0006+5050	00 06 08.249784	+50 50 04.41150	0.68	0.81	-0.041	64	0.16	0.14
C	0005+568	J0007+5706	00 07 48.468649	+57 06 10.43705	1.81	2.18	0.631	46	0.08	0.09
C	0012+610	J0014+6117	00 14 48.792125	+61 17 43.54198	0.57	0.29	-0.034	153	0.25	0.16
C	0024+597	J0027+5958	00 27 03.286191	+59 58 52.95899	0.56	0.34	-0.214	136	0.23	0.16
C	0032+612	J0035+6130	00 35 25.310605	+61 30 30.76122	0.94	0.52	-0.038	117	0.13	0.10
C	0034-220	J0037-2145	00 37 14.825799	-21 45 24.71171	1.17	2.78	-0.834	59	0.09	0.09
C	0039+568	J0042+5708	00 42 19.451680	+57 08 36.58569	0.39	0.25	0.046	162	0.48	0.32
C	0041+677	J0044+6803	00 44 50.759596	+68 03 02.68540	0.67	0.29	-0.163	154	0.23	0.19
C	0044+566	J0047+5657	00 47 00.428864	+56 57 42.39373	0.53	0.39	0.006	154	0.18	0.13

Notes. Units of right ascension are hours, minutes, and seconds. Units of declination are degrees, minutes, and seconds.

(This table is available in its entirety in machine-readable and Virtual Observatory (VO) forms in the online journal. A portion is shown here for guidance regarding its form and content.)

Table 10
List of 160 Sources That Have Not Been Detected in VGaPS Observations

Source Names (1)	Right Ascension (2)	Declination (3)	Gal. Lat. (4)	Flux (5)	Sp. Ind. (6)	No. (7)	
(1)	(2)	(3) (h m s)	(4) ($^{\circ}$ $'$ $''$)	(5) (deg)	(6) (mJy)	(7)	
2359+548	J0002+5510	00 02 00.470	+55 10 38.00	-6.8	121.	-0.03	2
0003+669	J0006+6714	00 06 10.000	+67 14 38.30	5.0	0
0009+655	J0012+6551	00 12 37.671	+65 51 10.82	3.5	195.	-0.59	8
0010+722	J0013+7231	00 12 58.750	+72 31 12.76	10.1	501.	0.02	8
0017+590	J0020+5917	00 20 24.550	+59 17 30.50	-3.1	324.	-0.02	6
0018-194	J0021-1910	00 21 09.370	-19 10 21.30	-79.6	0
0028+592	J0031+5929	00 31 03.120	+59 29 45.30	-3.0	3855.	1.52	2
0041+660	J0044+6618	00 44 41.300	+66 18 42.00	3.7	119.	-0.91	7
0107+562	J0110+5632	01 10 57.553	+56 32 16.93	-5.9	267.	-0.66	10
0113+241	J0116+2422	01 16 38.067	+24 22 53.72	-37.8	168.	-0.03	7
0128+554	J0131+5545	01 31 13.860	+55 45 13.20	-6.4	150.	-0.06	5

Notes. Columns 1 and 2 show IAU B1950 and J2000 source names, Columns 3–5 show a priori sources positions at the J2000.0 epoch, Column 6 shows extrapolated a priori total flux density at 24 GHz, Column 7 shows coarse estimate of the spectral index, and Column 8 shows the number of measurements of flux density found in the CATS database that were used for evaluation of the flux spectral index and extrapolation the flux density.

(This table is available in its entirety in machine-readable and Virtual Observatory (VO) forms in the online journal. A portion is shown here for guidance regarding its form and content.)

namely 38.3 and 38.0 mas, respectively. The significance of this offset, 55 and 72 times the inflated uncertainties, is too high to be explained by known error sources. RATAN-600 observations have shown that both of them continuously show steep radio spectra and are slowly variable. On VLBI scales, they were found to have significantly extended structures (e.g., Figures 17 and 18).

One of these compact steep spectrum radio sources, the object J0432+4138, is well studied at parsec scales. Recent high-dynamic range images at 5 and 8.4 GHz and a full bibliography on historic observations can be found in Mantovani et al. (2010). This source shows several bright components at 24 GHz. Two of them, namely components “A” and “C” (Figure 17), are located 40.6 mas apart. The feature C is stronger, with a total flux density of 0.54 Jy, but more extended. The FWHM of a circular Gaussian component fitted to the feature is found to be 1.4 mas. The feature A is dimmer, 0.13 Jy, but more compact—0.3 mas.

We note that here and below in this section the total flux density reported for different source features is calculated as a sum of all CLEAN components representing the corresponding structure. The Gaussian components fit is being performed in the visibility plane.

It is evident that the VGaPS 24 GHz observations referred to the position of the *weaker* A component, while the *S/X* observations referred to the position of the C component. This is counter-intuitive. Since component C is resolved, its contribution to the fringe amplitude is small at long baseline projections. At short baseline projections component C dominates; at long baseline projections component A dominates, but since the partial derivative of group delay with respect to source position is proportional to the baseline length projection on the source tangential plane, the contribution of long baseline dominates in estimates of source position. According to Mantovani et al. (2010), the total flux density of components A and C at 8.4 GHz

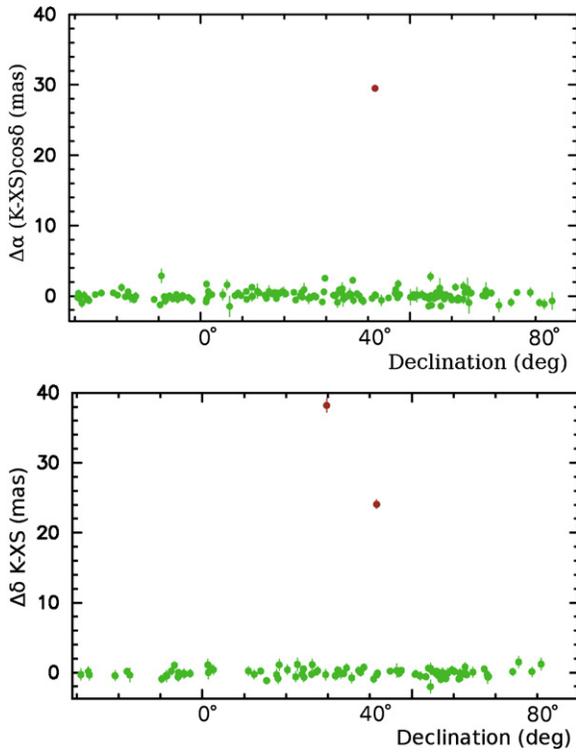


Figure 16. Differences in right ascensions scaled by $\cos \delta$ (top) and declinations (bottom) from *K*-band VGaPS vs. *S/X* historical VLBA observations among 164 sources with uncertainties less than 2 mas. The positions of the two sources differ significantly.

(A color version of this figure is available in the online journal.)

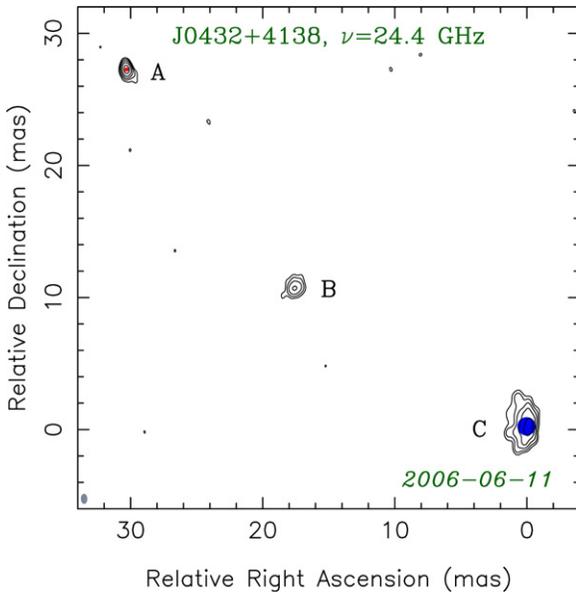


Figure 17. Naturally weighted *K*-band VGaPS CLEAN image of 3C119. The lowest contour of $4.1 \text{ mJy beam}^{-1}$ is chosen at three times the rms noise, and the peak brightness is 97 mJy beam^{-1} . The contour levels increase by factors of two. The dashed contours indicate negative brightness. The beam's FWHM is shown in the bottom left corner of the images in gray. Red and blue spots indicate the positions and sizes (FWHM) of circular Gaussian model components for the features "A" and "C," respectively.

(A color version of this figure is available in the online journal.)

in 2001 was 70 mJy and 1121 mJy, respectively. This large difference between components' flux densities is also supported by the previous VCS1 observations of this object in the *S/X* bands.

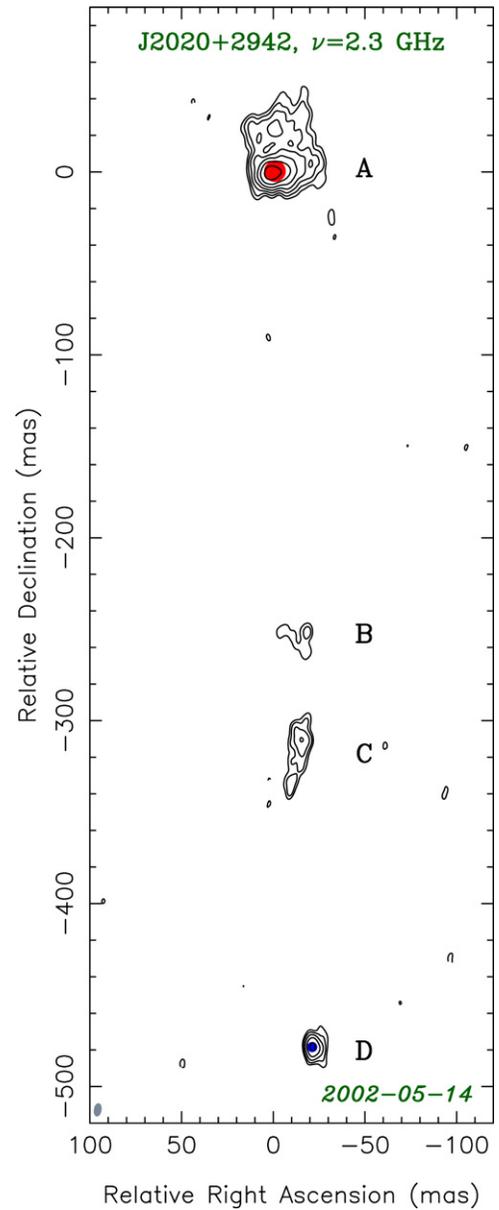


Figure 18. Naturally weighted *S*-band CLEAN image of 3C410 redone by us using VCS2 VLBA observations from 2002 May 14. The lowest contour of $2.5 \text{ mJy beam}^{-1}$ is chosen at three times the rms noise, the peak brightness is $242 \text{ mJy beam}^{-1}$. The contour levels increase by factors of two. The dashed contours indicate negative brightness. The beam's FWHM is shown in the bottom left corner of the images in gray. Red and blue spots indicate the positions and sizes (FWHM) of circular Gaussian model components for the features "A" and "D," respectively. It should be noted that feature A is very extended and the single Gaussian component does not represent it well.

(A color version of this figure is available in the online journal.)

Component C is less resolved at 8.4 GHz and it dominates over A even at long baselines. Therefore, coordinates of a source in the *X* band are closer to the C component. We checked this directly by suppressing observations at baselines longer than 800 km in a *K*-band trial solution. The position estimate became close to the *S/X* position. It is worth noting that the difference in position in the *K* and *X* bands are $29.2 \pm 0.4 \text{ mas}$ in right ascension and $24.1 \pm 0.4 \text{ mas}$ in declination, while the offset of component A with respect to component C in the *K* band is slightly larger: $30.25 \pm 0.10 \text{ mas}$ and $27.05 \pm 0.15 \text{ mas}$. The *K*-band and *X*-band positions are not exactly the position of

Table 11
Position Difference of J2020+2942 at Different
Frequencies with Respect to Its Position in the X Band

Band	R.A. Shift	Decl. Shift
X	0.0 ± 0.58 mas	0.0 ± 0.71 mas
S	25.77 ± 4.11 mas	478.79 ± 4.39 mas
X/S	-2.29 ± 0.54 mas	-36.78 ± 0.66 mas
K	0.25 ± 0.41 mas	1.45 ± 0.70 mas

Note. The offset in right ascension is scaled by $\cos \delta$.

components A and C, since in both solutions the contribution of the second component is small but not entirely negligible.

The nature of the difference in positions of J2020+2942 is similar. It was observed in the VCS2 experiment on 2002 May 15 (Fomalont et al. 2003). It had 63 detections in the S band and 72 detections in the X band. The position of this source in the S band from analysis of only S-band data applying the ionosphere contribution from the GPS TEC model shows a very large offset of $0''.5$ with respect to the X-band position (refer to Table 11). The errors of the ionosphere contribution in the S-band during the solar maximum affected position estimates considerably, however, not to that extent. A comparison of the positions of 130 sources with $\delta > 0$ from the solution that used only S-band group delay observables with respect to the X/S solution in that experiment showed differences in the range 2–7 mas. It is remarkable that the X/S position is away from both X- and S-band positions, although intuitively we can expect them to be between X and S positions. This can be explained if we surmise that the source J2020+2942 has two components, $0''.48$ apart, one of which is visible in the X band, but not visible in the original S-band image, and another which is visible in the S band, but not in the X band. An ionosphere-free linear combination of X- and S-band observables is used in the X/S solution: $(1 + \beta)\tau_{gx} - \beta\tau_{gs}$, where $\beta = 1/(\omega_x^2/\omega_s^2 - 1)$ as it follows from Equation (27). In the case where the position in the S band, \vec{k}_s , is shifted with respect to the X-band position vector, \vec{k}_x , the ionosphere free linear combination can be written as

$$\tau_{if} = (1 + \beta)\tau_{gx}(\vec{k}_x) - \beta\tau_{gs}(\vec{k}_s) - \beta \frac{\partial \tau}{\partial k}(\vec{k}_s - \vec{k}_x). \quad (32)$$

The first two terms correspond to a case with no offset between X- and S-band positions. Therefore, estimates of the source position from the ionosphere-free linear combination of observables will be shifted at $-\beta$ with respect to the offset $(\vec{k}_s - \vec{k}_x)$, i.e., in the opposite direction. Parameter β was 0.076153 in the VCS2 experiment. Therefore, if our hypothesis that the X-band and S-band observations detected emission from two components is true, then the shift of the X/S position with respect to the X-band position should be -1.96 mas in right ascension and -36.46 mas in declination, just within 0.3 mas from the reported S/X positions! The K-band position is within 1.5σ of the X-band position.

In order to check our hypothesis, we have re-imaged VCS2 observations of J2020+2942 in a wide field and have detected several previously unknown features “B,” “C,” and “D,” on a distance up to about 500 mas from the dominating extended structure “A” in the S band (Figure 18). The total flux density of the features A and D is 1.40 Jy and 0.09 Jy, respectively. Feature A is significantly more extended than D. We have fitted two circular Gaussian components to the uv -data in order to determine the positions of features A and D. We note that

the accuracy of position determination for component A from the image is very poor since its structure is extended over at least 40 mas and is not well represented by a single Gaussian component. The distance between Gaussian components A and D is 479 mas, while the positional difference in right ascension and declination is 25 and 479 mas, respectively—in very good agreement with independent astrometric measurements (Table 11).

X-band and K-band wide field imaging (Figure 19) did not reveal components with wide separation. We conclude from the astrometric analysis presented above that X- and K-band images represent the more compact component D. In this case we could also analyze its spectrum on the basis of simultaneous S/X-band observations. Its total flux density in the X-band is found to be 0.25 Jy which provides the 2.3–8.6 GHz spectral index estimate $\alpha = +0.8$ (flux density $\propto \nu^\alpha$)—an indication of synchrotron emission with significant self-absorption. The features A, B, and C become too weak and/or too resolved for us to detect in the snapshot VLBA images with a limited dynamic range and uv -coverage.

8.2. General Comparison: Uncertainties, Systematic K-S/X-band Difference, and the Core-shift Effect

Position differences for other objects do not show peculiarities. For instance, no declination-dependent systematic differences similar to those reported by Lanyi et al. (2010) were found. The wrms of the differences is 0.46 mas in right ascension scaled by $\cos \delta$ and 0.61 mas in declination. We have computed the normalized distances by dividing them by $\sqrt{e_k^2 + e_{xs}^2}$, where e_k is the projection of the error ellipse of the K-band position to the direction of position difference and e_{xs} is the similar projection of the error ellipse of the S/X position. In the case if position errors from K-band and S/X catalogs are independent and Gaussian with the variance equal to reported uncertainties, the distribution of normalized distances will be Rayleigh with $\sigma = 1$. The average of the normalized distances over all sources, except J0432+4138 and J2020+2942, is 1.276, only 2% greater than the mean of the Rayleigh distribution, $\sqrt{\pi}/2$. However, a close examination of the distribution (see Figure 20) reveals a slight deviation of its shape from the shape of the Rayleigh distribution. The Rayleigh distribution that best fits the distribution of normalized distances has $\sigma = 0.90$. This is an indication of a deviation of parent distributions from Gaussian.

We can make several conclusions from this test. First, on average, reported formal uncertainties are correct within several percent. Second, the effect of the core-shift is too small to contribute significantly to results of single-epoch surveys. According to Kovalev et al. (2008), the typical apparent core-shift is expected to be 0.4 mas between the S and X bands. Porcas (2009) stressed that when the core-shift is proportional to f^{-1} , a source position derived from ionosphere-free linear combinations of X- and S-band group delays is not sensitive to the core-shift and corresponds to a true position of the jet base. The f^{-1} core-shift dependence is expected for a conical jet with synchrotron self-absorption in the regime of equipartition between the jet particle and magnetic field energy densities (Lobanov 1998). If we assume that this is indeed the case for the majority of the sources (see also Sokolovsky et al. 2011), the average core-shift between K-band and effective S/X positions is reduced to $0.4 \times \frac{f_x f_s}{f_x(f_x - f_s)} = 0.06$ mas. Our observations would allow detection of the core-shift between positions from S/X and K-band observables the 95% confidence level of a sample of 190

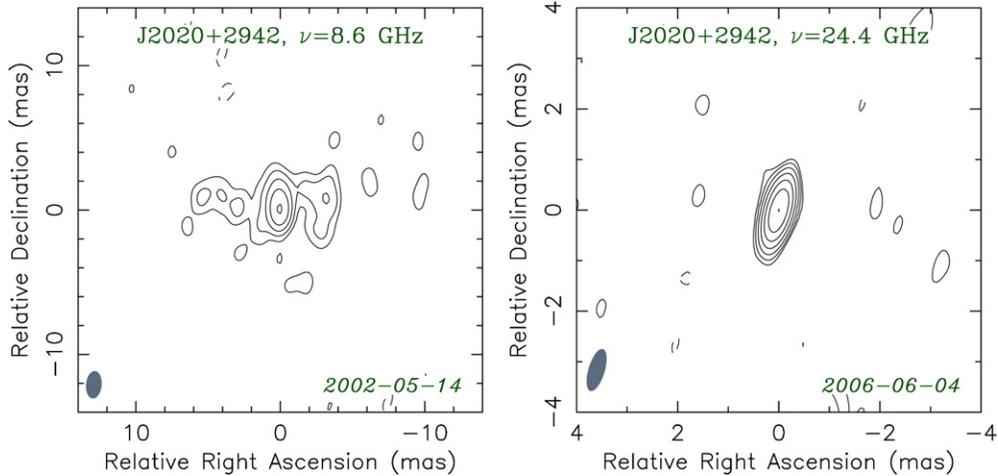


Figure 19. Naturally weighted X-band (left, VCS2 data) and K-band (right, this survey—VGaPS) CLEAN images of 3C410. The lowest contour in the X- and K-band is 3.3 and 2.8 mJy beam⁻¹, while the peak brightness is 56 and 178 mJy beam⁻¹, respectively. On the basis of our analysis, we identify the feature presented in this figure as feature “D” from Figure 18.

(A color version of this figure is available in the online journal.)

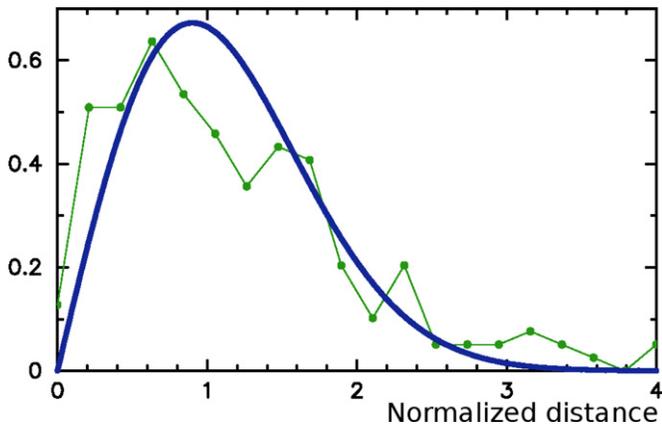


Figure 20. Empirical distribution of 190 normalized distances between K band position of target sources and their X/S positions (broken line) and the best-fit Rayleigh distribution with $\sigma = 0.90$ (solid thick line).

(A color version of this figure is available in the online journal.)

objects if the variance of the core-shift is greater than 1 mas. VGaPS observations set the upper limit of this variance to 1 mas, which does not contradict the result of core-shift measurements and predictions.

9. SUMMARY

In the VLBA Galactic Plane Survey we detected 327 compact radio sources not previously observed with VLBI at 24 GHz in absolute astrometry mode. Half of them are within 5° of the Galactic plane; 206 of them were also observed and detected within the VCS or RDV programs in the S/X bands in absolute astrometry mode. We determined K-band positions of all detected sources. The position uncertainties for all but one source are in the range from 0.2 to 20 mas with the median value of 0.9 mas. The quoted uncertainties account for various systematic effects and their validity within several percent was confirmed by comparison with independent S/X observations. The detection limit of our observations was in the range of 70–80 mJy. For the majority of detected sources, parsec-scale images were produced, and correlated parsec-scale flux densities

were estimated. These results are presented in the form of the position catalog, calibrated image and visibility data in FITS format, and visual plots.

The new wide-band fringe search baseline-oriented algorithm for processing correlator output has been developed and implemented in the software PIMA. This reduced the detection limit of the observations by a factor of \sqrt{N} , where N is the number of IFs, by determining group delays, fringe phases at the reference frequency, and phase delay rates from the coherent sum of the data from all IFs. The new algorithm increased the number of detected target sources by a factor of 2.4 for this survey of weak objects near the Galactic plane. We validated the new algorithm by parallel processing of 1080 hr, over 0.6 million observations, using both the traditional AIPS approach and the new approach. The differences between source position estimates processed with the wide-field and with the traditional AIPS algorithms do not exceed 0.15 mas, which is satisfactory for any practical application.

We investigated possible systematic errors caused by errors in the ionosphere path delay derived from GPS TEC maps. We derived an empirical model of the ionosphere-driven delay path errors. We found that for declinations $> -20^\circ$ for 90% of the sources, mismodeling path delay caused source position errors of less than 0.05 mas. At declinations below -20° these errors grow to 0.15 mas and for some sources may reach 0.5 mas.

Comparisons of new K-band VLBI positions with positions of 192 sources observed in S/X showed an agreement with the wrms of 0.46 and 0.6 mas in right ascensions and declination, respectively, within reported position uncertainties for all but two compact steep spectrum sources J0432+4138 and J2020+2942. For these two objects, positional differences are about 40 mas. We showed that the reason for these differences is that for sources with complex extended structures, positions referred to different source details. These two objects demonstrate the existence of an overlooked source of errors in VLBI position catalogs that will be studied in detail in the future. A 1 mas upper limit on an apparent core-shift effect between 8 and 24 GHz is found for the studied sample, in agreement with core-shift measurements and predictions by Kovalev et al. (2008), Sokolovskiy et al. (2011), and Porcas (2009).

We thank Leonid Kogan for fruitful discussions about hidden secrets of correlators. We used the data set MAI6NPANA provided by the NASA/Global Modeling and Assimilation Office (GMAO) in the framework of the MERRA atmospheric reanalysis project. This project was started when Y.Y.K. was working as a Jansky Fellow of the National Radio Astronomy Observatory in Green Bank. Y.Y.K. was supported in part by the JSPS Invitation Fellowship for Research in Japan (S-09143) and Russian Foundation for Basic Research (08-02-00545 and 11-02-00368). The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. The authors made use of the database CATS of the Special Astrophysical Observatory (Verkhodanov et al. 2005). This research has made use of the NASA/IPAC Extragalactic Database (NED; Eichhorn et al. 2002) which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the NASA.

Facilities: VLBA (project code BP125), RATAN

REFERENCES

- Abdo, A. A., et al. 2010a, *ApJS*, **188**, 405
 Abdo, A. A., et al. 2010b, *ApJ*, **715**, 429
 Agnew, D. C. 1997, *J. Geophys. Res.*, **102**, 5109
 Beasley, A. J., Gordon, D., Peck, A. B., Petrov, L., MacMillan, D. S., Fomalont, E. B., & Ma, C. 2002, *ApJS*, **141**, 13
 Eichhorn, G., et al. 2002, *Ap&SS*, **282**, 299
 Fomalont, E., Petrov, L., McMillan, D. S., Gordon, D., & Ma, C. 2003, *AJ*, **126**, 2562
 Greisen, E. W. 1988, in *Acquisition, Processing and Archiving of Astronomical Images*, ed. G. Longo & G. Sedmak (Napoli: Osservatorio Astronomico di Capodimonte), 125
 Honma, M., et al. 2003, *PASJ*, **55**, L57
 Kalnay, E. M., et al. 1996, *Bull. Am. Meteorol. Soc.*, **77**, 437
 Kogan, L. 1993, NRAO VLBI Scientific Memo, 6, <http://www.vlba.nrao.edu/memos/sci/sci06memo.pdf>
 Kogan, L. 1995, NRAO VLBI Scientific Memo, 9, <http://www.vlba.nrao.edu/memos/sci/sci09memo.pdf>
 Kogan, L. 1998, *Radio Sci.*, **33**, 1289
 Kopeikin, S. M., & Schäfer, G. 1999, *Phys. Rev. D*, **60**, 124002
 Kovalev, Y. Y. 2009, *ApJ*, **707**, L56
 Kovalev, Y. Y., Kovalev, Yu. A., Nizhelsky, N. A., & Bogdantsov, A. B. 2002, *PASA*, **19**, 83
 Kovalev, Y. Y., Lobanov, A. P., Pushkarev, A. B., & Zensus, J. A. 2008, *A&A*, **483**, 759
 Kovalev, Y. Y., Nizhelsky, N. A., Kovalev, Yu. A., Berlin, A. B., Zhekanis, G. V., Mingaliev, M. G., & Bogdantsov, A. V. 1999, *A&AS*, **139**, 545
 Kovalev, Y. Y., Petrov, L., Fomalont, E., & Gordon, D. 2007, *AJ*, **133**, 1236
 Lanyi, G. E., et al. 2010, *AJ*, **139**, 1695
 Lobanov, A. P. 1998, *A&A*, **330**, 79
 Ma, C., et al. 1998, *AJ*, **116**, 516
 Mantovani, F., Rossetti, A., Junor, W., Saikia, D. J., & Salter, C. J. 2010, *A&A*, **518**, A33
 Mathews, P. M. 2001, *J. Geod. Soc. Japan*, **47**, 231
 Matsumoto, K., Takanezawa, T., & Ooe, M. 2000, *J. Oceanogr.*, **56**, 567
 Matveenko, L. I., Kardashev, N. S., & Sholomitskii, G. B. 1965, *Izv. Vyssh. Uchebn. Zaved. Radiofiz.*, **8**, 651 (English transl. *Sov. Radiophys.*, **8**, 461)
 Murphy, T., et al. 2010, *MNRAS*, **420**, 2403
 Pearson, T. J., Shepherd, M. C., Taylor, G. B., & Myers, S. T. 1994, *BAAS*, **185**, 0808
 Petrov, L. 2007, *A&A*, **467**, 359
 Petrov, L., & Boy, J.-P. 2004, *J. Geophys. Res.*, **109**, B03405
 Petrov, L., Gordon, D., Gipson, J., MacMillan, D., Ma, C., Fomalont, E., Walker, R. C., & Carabajal, C. 2009, *J. Geod.*, **83**, 859
 Petrov, L., Hirota, T., Honma, M., Shibata, S. M., Jike, T., & Kobayashi, H. 2007a, *AJ*, **133**, 2487
 Petrov, L., Kovalev, Y. Y., Fomalont, E., & Gordon, D. 2005, *AJ*, **129**, 1163
 Petrov, L., Kovalev, Y. Y., Fomalont, E., & Gordon, D. 2006, *AJ*, **131**, 1872
 Petrov, L., Kovalev, Y. Y., Fomalont, E., & Gordon, D. 2007b, *AJ*, **136**, 580
 Petrov, L., & Ma, C. 2003, *J. Geophys. Res.*, **108**, 2190
 Petrov, L., Phillips, C., Bertarini, A., Murphy, T., & Sadler, E. M. 2011, *MNRAS*, in press (arXiv:1012.2607)
 Porcas, R. W. 2009, *A&A*, **505**, L1
 Ray, R. D. 1999, NASA/TM-1999-209478 (Greenbelt, MD: Goddard Space Flight Center/NASA)
 Ros, E., Marcaide, J. M., Guirado, J. C., Sardón, E., & Shapiro, I. I. 2000, *A&A*, **356**, 357
 Ryan, J. W., & Clark, T. A. 1988, in *IAU Symp. 129, The Impact of VLBI on Astrophysics and Geophysics*, ed. M. J. Reid & J. M. Moran (Dordrecht: Kluwer), 339
 Schaer, S. 1998, PhD thesis, Univ. Bern; accessible at <ftp://ftp.unibe.ch/aiub/papers/ionodiss.ps.gz>
 Schubert, S., et al. 2008, in *Proc. 3rd WCRP Int. Conf. on Reanalysis*, Tokyo, V1-104 (http://wcrp.ipsl.jussieu.fr/Workshops/Reanalysis2008/Documents/V1-104_ea.pdf)
 Sekido, M., Kondo, T., Kawai, E., & Imae, M. 2003, *Radio Sci.*, **38**, 8
 Shepherd, M. C. 1997, in *ASP Conf. Ser. 125, Astronomical Data Analysis Software and Systems VI*, ed. G. Hunt & H. E. Payne (San Francisco, CA: ASP), 77
 Sokolovsky, K. V., Kovalev, Y. Y., Pushkarev, A. B., & Lobanov, A. P. 2011, *A&A*, in press (arXiv:1103.6032)
 Sovers, O. J., Fanelow, J. L., & Jacobs, C. S. 1998, *Rev. Mod. Phys.*, **70**, 1393
 Thompson, A. R., Moran, J. M., & Swenson, G. W., Jr. 2001, *Interferometry and Synthesis in Radio Astronomy* (New York: Wiley)
 Verkhodanov, O. V., Trushkin, S. A., Andernach, H., & Chernenkov, V. N. 2005, *Bull. SAO*, **58**, 118
 Wrobel, J. 2009, *NRAO eNews*, **2**, 6
 Wrobel, J. M., Patnaik, A. R., Browne, I. W. A., & Wilkinson, P. N. 1998, *BAAS*, **193**, 4004