VLBI GEODESY: TECHNIQUES AND RECENT RESULTS [+]

David B. Shaffer
Interferometrics Inc.

INTRODUCTION

From its inception, it was realized that VLBI had the potential for very precise determinations of baselines between antennas. Since the first early, crude results (Cohen and Shaffer 1971), the data collection systems and analysis techniques have been continuously refined, and VLBI is now producing baseline measurements to accuracies of a few centimeters over trans- and inter-continental distances. Since the Mark III system and the basic techniques of delay and delay rate baseline determination have been previously described (Rogers et al. 1983; Shapiro 1976), and there have been several recent meetings that dealt exclusively with VLBI geodesy techniques (NASA 1980; NOAA 1982), I will concentrate in this review on recent developments in equipment and procedures, followed by presentation of new results. This review is based primarily on the efforts and results of the so-called "East Coast VLBI Group" comprised of staff and support contractors at NASA Goddard Space Flight Center (NASA/GSFC), the NEROC Haystack Observatory, MIT, and the Harvard-Smithsonian Center for Astrophysics. Additional input comes mostly from staff of the National Geodetic Survey (NGS), Onsala Space Observatory, and the Jet Propulsion Laboratory.

THE NASA CRUSTAL DYNAMICS PROJECT

During the 1970's, development of VLBI and laser techniques made it clear that distance measurements accurate to a few centimeters were feasible over very long distances. NASA undertook developments in both these areas, leading to the launch of the laser geodesy satellite (LAGEOS), improvement of laser ranging systems, and the Mark III VLBI system. This new equipment provided the expected precision, and NASA subsequently formed the Crustal Dynamics Project (CDP).

The goal of the CDP is to acquire data for, and to make available the results of laser and VLBI observations to, the geodesy and geophysics community. Measurements of continental drift, distortion and stability of individual crustal plates, and baseline changes that might precede or follow earthquakes in geologically active regions such as southern California are among the aims of the Project. To achieve scientifically useful results,

365

certain measurement accuracies must be attained during the 1980-1988 time
frame of the CDP. Table 1 shows the accuracies that the CDP has assured
will be available, and goals for further improvements.

Table 1
CRUSTAL DYNAMICS PROJECT MEASUREMENT REQUIREMENTS

	Commitment	Goal
	(all one sigma)	
GEODETIC PRODUCTS		
Intersite Velocity Accuracy	1 cm/year	0.4 cm/year
Baseline Length Precision	6 cm	2 cm
Earth's Pole Position Precision	10 cm	4 cm
INSTRUMENT PERFORMANCE		
Laser-Range/VLBI-Delay rms (3-minute normal point)	4 cm	1 cm

THE MARK III VLBI SYSTEM

The Mark III system can be regarded as a set of hardware and software
for VLBI geodesy. The hardware is not just the IF processor and tape
recorder, but includes hydrogen maser frequency and time standards,
dual-frequency, broad-band receivers, a tone injection system with cable
length calibration, meteorological sensors (temperature, pressure, and
humidity), and water vapor radiometers. The contribution of several of
these subsystems to the overall system accuracy is discussed below.

Broad-band Receivers

The precision of a group delay measurement is directly proportional
to the spanned bandwidth over which it is determined. At X-band, the CDP
receivers and Mark III terminal record signals in the range 8210 to 8570
MHz, using a bandwidth synthesis technique (Rogers 1970). With this
bandwidth, single measurement delay accuracies are better than about 100
picoseconds, or 3 cm equivalent range, even on the weakest detectable
sources. Strong sources have delay precisions of just a few picoseconds,
or only a millimeter or so of range. At S-band, the recorded band covers
2215 to 2305 MHz.

The delays and delay rates measured at X- and S-band are combined to
eliminate the $1/f^2$ dispersive effects of the ionosphere. For instance,
the true group delay is:

$$\tau_{true} = \frac{f_x^2}{f_x^2 - f_S^2} \tau_x - \frac{f_S^2}{f_x^2 - f_S^2} \tau_S ,$$

where f_S and f_x are the effective S- and X-band frequencies, and τ_S and τ_x are the group delays measured in the two bands. The S-band delay and its associated error come in only at about the 10% level because of the f-squared terms. Hence, the smaller spanned bandwidth at S-band does not degrade the final measurement.

The ionospheric correction is now routinely applied to all data that are collected and analyzed by the East Coast group. An experiment in September/October 1980 clearly showed the improvement provided by dual-frequency observing. These data, taken during a period of unusually high ionospheric total electron content, were analyzed with and without the correction. The uncorrected results gave baselines that were too long by 15 to 20 centimeters when compared to results obtained at times of lower electron content and the corrected results. (Error sources such as the ionosphere and the troposphere that cause extra path delay biases always have the effect of making the baselines seem too long. The antenna that is further from the source will be observing at a lower elevation angle, and thus a greater atmospheric/ionospheric path, with a concommitant larger delay bias. The parameter estimation program tries to match the excess delay by lengthening the baseline.) The uncorrected results also show much greater scatter in the residuals within a single day's analysis and in the day-to-day means. (Two seven-day sessions separated by two weeks comprise these observations, which were part of the MERIT preliminary campaign.)

The active ionosphere during the MERIT observations also pointed up a deficiency in the bandwidth synthesis frequency array that was in use at S-band: a 3-frequency minimum redundancy array that spanned 75 MHz with a 25 MHz unit spacing. The resultant delay ambiguities were 40 nanoseconds, far less than the day/night range of ionospheric delays (almost 100 nanoseconds). Determination of the correct delay ambiguity was rather difficult. This difficulty was the incentive for several improvements to observing and analysis techniques. The previous 7-frequency Mark III mode ("Mode B", with 4 frequencies at X-band and 3 frequencies at S-band) was replaced by a 14-frequency mode ("Mode C", with 8 frequencies at X-band and 6 frequencies at S-band). These frequencies are used in arrays derived by D. S. Robertson of the NGS that have zero redundancy but some missing spacings, moderate sidelobes (50 to 60%), and small unit spacings (5 MHz at S-band and 10 MHz at X-band). The corresponding ambiguity spacings are 200 nanoseconds at S-band and 100 nanoseconds at X-band. The widely separated ambiguity lobes greatly simplify ambiguity resolution. The baseline analysis program was modified to check closure delays on arbitrary triads of stations, sc that incorrect ambiguity resolution can be spotted quickly.

The multiple frequency channels are calibrated by a 1 MHz rail that is injected at the receiver, ahead of the low noise amplifiers. These "phase cal" tones are extracted by the correlator and the resultant phases

are used to correct for dispersion in the RF to digital path and to coherently combine the channels. The phase calibration rail also provides the fundamental time reference for the observations. (It can be thought of as injecting one microsecond ticks into the receiver.) The system is driven by the station maser, and cable length variations between the maser and the front end are monitored by a reflectometer length measuring system (the "cable cal" system).

The current CDP receivers are uncooled parametric amplifiers and have rather high system temperatures: about 100 K at S-band and 200 K at X-band. Efforts are now underway to upgrade the receivers by replacing the paramps with cooled GaAs-FET amplifiers. The first of these systems has been installed at the new Mojave base station at Goldstone. The new system temperatures are about 70 K at both S- and X-band. (The S-band temperature includes some 30 to 40 K of spillover that results from rather poor illumination provided by the dual-frequency feed.)

Meteorological Sensors and Water Vapor Radiometers

Propagation delays which arise in the troposphere currently set the limits on the accuracy of VLBI baseline determinations. These delays are caused by the dry and wet components of the atmosphere.

The dry component of the atmosphere causes a zenith path delay of about 7 nanoseconds, or about 2 meters equivalent length. The exact value depends on the altitude of the station and the local temperature and pressure, which determine the mass of the air overhead. Temperature and pressure are recorded for every VLBI scan, and an atmospheric model is used to calculate a correction for all the observations.

The wet component of the atmosphere is less amenable to modelling based on ground-level humidity readings, since the water vapor is not well mixed. However, the 22.2 GHz emission from water shows promise of allowing a direct measurement of the path delay caused by water vapor. The brightness temperature in the water line is roughly proportional to the excess path length. Existing water vapor radiometers (WVR's) are two-channel devices, with one channel at about 21 GHz and the other near 31 GHz. The theory and operation of two WVR's are described by Elgered and by Resch and Miller (both in NOAA 1982).

Several WVR's of the Resch design are in use by the CDP, but because of reliability problems, no convincing demonstration of their ability to improve VLBI results has yet been made. In tests at the Very Large Array (VLA) in New Mexico, Resch et al. (1983) have shown that the use of WVR's did improve the phase stability, sometimes dramatically. In the VLA tests, the WVR's used the VLA cassegrain optics, and sampled the same air column as was being used for the VLA measurements. As used at most CDP VLBI sites, however, the WVR's are often at a distance of 100 meters or more from the antennas which are tracking the radio sources. This offset and the 7-degree beam of the WVR's means that the radiometers are not necessarily sampling the same air column as the VLBI observations. On very cloudy and/or rainy

days, the differences could be significant. I believe that future WVR improvements should include mounting them on the VLBI antennas, with modest size dishes (1 to 2 meters) of their own, so that the water vapor delay is measured more accurately.

Hydrogen Masers

Masers are currently the best readily available time and frequency standards. Recent experience has shown that the masers are rather sensitive to temperature changes. This fact was not properly appreciated in the past, and the masers were often operated in telescope control rooms with poor temperature regulation. The high sensitivity of the Mark III revealed daily, cyclical variations of a nanosecond or more in the residuals from some experiments in which the masers had inadequate temperature control. Most of the CDP stations are now outfitted with very well insulated and temperature-controlled "isolation boxes" for the masers. These boxes regulate the maser temperature to a few hundredths of a degree Celsius.

The best VLBI baseline solutions have an rms scatter of group delay residuals of 50 to 70 picoseconds about the best-fit solution for 24 hours of data. This scatter is comparable to the Allan variance expected for two masers over a 24-hour period. Thus, after the wet component of the troposphere, maser behavior is a significant contribution to the quality of VLBI baseline determination. A possible way to overcome the limitations of the masers, and perhaps even allow the use of less stable clocks, is the use of a "clock source" at a declination greater than about 70 degrees that is visible throughout the day for all (Northern!) stations. The clock source would be observed every other scan, and all other observations would be referenced to it. By analyzing only differences with respect to the clock source, variations of the time standards with periods longer than the clock source observation interval would not affect the data. Use of this technique has been hampered by the lack of suitably strong sources and the long slew times of most antennas. The new generation of receivers will overcome the source strength problem, and the implementation of faster slewing antennas (Mojave and Wettzell slew at 1 degree per second or more) will allow this technique to be more thoroughly tested.

Radio Source Structure

The ideal geodetic source is strong and unresolved on all baselines. As shown by nearly everyone at this Symposium, many compact sources are far from simple in structure. The brightness temperature limit of about 10^{12} K (Kellermann and Pauliny-Toth 1969), set by inverse Compton scattering, means that the strongest sources will not be the most compact. Structure in a source causes phase gradients in the (u,v)-plane. These gradients cause phase slopes over the width of a bandwidth synthesis frequency range since the individual frequency channels sweep out slightly different paths in the (u,v)-plane. These phase slopes are indistinguishable from the phase slope ($d\phi/d\omega$) that determines group delay. Hence, source structure can be a cause of error in the VLBI measurements.

Errors due to structure can be greatly reduced by correcting for the phase effects as predicted by a map of the source. The strongest phase gradients occur near the minima of visibility functions, so large errors can be avoided by not observing at low visibilities (Thomas 1980). The easiest way to avoid structure problems is to observe point sources. The East Coast group has followed this approach, and nearly eliminated sources with significant structure from their set of reference sources. Table 2 shows the set of sources used for trans- and inter-continental VLBI by the East Coast group.

Table 2
VLBI REFERENCE SOURCES

0106+013		1226+023	(3C273)*
0234+285		1404+286	(OQ208)
0300+471		1641+399	(3C345)**
0528+134		1741-038	
0552+398		2216-038	
0851+202	(OJ287)	2251+158	(3C454.3)
0923+392	(4C39.25)*		

* Significant structure
** Slight structure

Unfortunately, 3C273 has been chosen to define the right ascension zero point. This source has complex, time variable structure (Pearson et al. 1981) and in retrospect is a poor choice for this purpose. Better choices would be 0106+013 or 0851+202, which have simple structure, lie near the ecliptic, and are visible from the southern hemisphere.

Table 3 shows the expected contribution of the various error sources to the error budget for measurements of intermediate and intercontinental baselines. These estimates were arrived at after two days of discussion in October 1982 by most of the USA practitioners of geodetic VLBI. Some of these estimates are pessimistic, since they were enlarged to include worst cases. For instance, VLBI instrumentation has been used to measure baselines of 600 and 1200 meters, to accuracies of millimeters (Lundqvist in NOAA 1982; Rogers et al. 1978). Source structure was found to cause errors of less than 1 mm on a Haystack-NRAO-HRAS-OVRO set of baselines (Shaffer 1982).

Table 3 indicates that VLBI is meeting the CDP commitment of 6 cm for baseline length accuracy, and is well on the way to the 2 cm goal shown in Table 1. Baseline orientation, especially in the vertical location of the stations, is more difficult to determine than baseline length, since length is invariant to orientation. Height errors in a station location arise from atmospheric/ionospheric propagation errors. Excess path lengths mimic the signature of a lowered elevation. This effect is more severe for short

baselines where the propagation delays are highly correlated for the several stations and do not give rise to strong station-dependent signatures.

Table 3
VLBI ERROR SOURCE SUMMARY

Error Source	400 km Baseline			4000 km Baseline Length
	Vertical	Horizontal		
		Transverse	Length	
Instrumentation	1.4 cm	0.9 cm	0.9 cm	1.0 cm
Ionosphere	0.2	0.1	0.1	0.2
Dry troposphere	1.5	0.6	0.8	0.5 - 1.0
Wet troposphere*	4.0/<9.0	0.5/1.0	0.6/1.2	1.5/<3.0
UT1/Polar Motion+	0.6/1.3	0.6/1.3	≅ 0	≅ 0
Source Structure	0.3	0.3	0.3	0.3
Aggregate rss Estimate (cm)	4.6-9.3	1.4-2.0	1.4-1.7	2.0-3.3

* First number assumes WVR, second assumes the use of modeling or surface meteorology data.

+ Assumes 10-20 cm a priori data, such as from POLARIS, or satellite or lunar laser ranging.

RECENT RESULTS

The densest data set of baseline length determinations comes from the National Geodetic Survey POLARIS project (Robertson and Carter, in NOAA 1982). This program makes one measurement per week between the Westford and GRAS (ex-HRAS) antennas. The goal of POLARIS is high precision determination of polar motion and UT1 variations, with baseline lengths as a by-product. Figure 1 shows the results of two and a half years of measurements between Haystack and GRAS, including POLARIS results. [The Westford-GRAS baseline has been referenced to Haystack by adding the Haystack-Westford baseline determined by Rogers et al. (1978).] Typical errors on a length measurement are a few centimeters (see Figure 2). The rms scatter of all the baseline length measurements is about 2 cm on this 3136 km baseline. The various experiments were analyzed by the NASA/GSFC group with a set of source coordinates in the J2000 system (Kaplan 1981). (This set is designated SSC(GSFC)82 R 01 in the annual report of the BIH

for 1982.) Slightly different analysis procedures applied to these data by the NGS investigators result in a comparably small scatter. However, there is a small offset of about 2 cm between the mean baseline lengths determined by the two groups, probably caused by different source catalogs. This offset highlights the fact that VLBI is somewhat better suited to measuring changes in baseline length rather than absolute baseline lengths. Length variations are nearly unaffected by any biases as long as the analysis procedures are consistent from experiment to experiment.

The Haystack-GRAS measurements show no significant changes in baseline length over the two and a half year period shown in Figure 1. This result is not surprising, since both stations are on the North American plate; and there are no areas of recent geological activity between the two stations, although the baseline does cross the New Madrid, Missouri area, site of large earthquakes in the early 1800's. The formal velocity error is comparable to the CDP goal given in Table 1 and will improve as the data span grows over the next few years.

Figure 2 shows baseline length measurements of the Haystack-OVRO baseline for a six year period. The errors on each solution are shown, and are typically a few centimeters. The average error is somewhat reduced from 1979 onward, when the Mark III went into service. The fact that the errors did not decrease dramatically, despite the high precision of the Mark III, shows the predominant effect of external error sources on the VLBI reductions, mostly from the troposphere.

The Haystack-OVRO baseline also shows no change in length, with a formal "speed limit" that is better than the CDP goals and shows the quality of data that can be obtained during a moderately long observing campaign. Although both stations are again on the North American plate, it is not so clear what level of stability should be expected on this baseline. OVRO is located in a geologically active area. It is only a few kilometers from a large, active fault that runs the length of the Owens Valley at the base of the Sierra Nevada mountains. The baseline crosses the Basin and Range province, an area of expected extension. Sites in Nevada, Utah, and Colorado will be visited by mobile VLBI systems during the life of the CDP in order to determine whether the Massachusetts-California baseline is truly invariant, or if the apparent stability is a coincidence caused by offsetting motions.

The Haystack to Onsala, Sweden baseline has been measured some 40 times from mid-1980 through early 1983 by the NASA/GSFC and NGS groups. The formal baseline length results show a possible lengthening of the baseline at a rate of 1.8 ± 0.6 cm/year. This change is in the direction and about the right size predicted by continental drift theory, but the experimenters are reluctant to claim a true detection of plate motion. The random and systematic errors are somewhat larger on this 5600 km baseline than on the intra-USA baselines, and a larger set of data is required to establish the USA to Europe velocity.
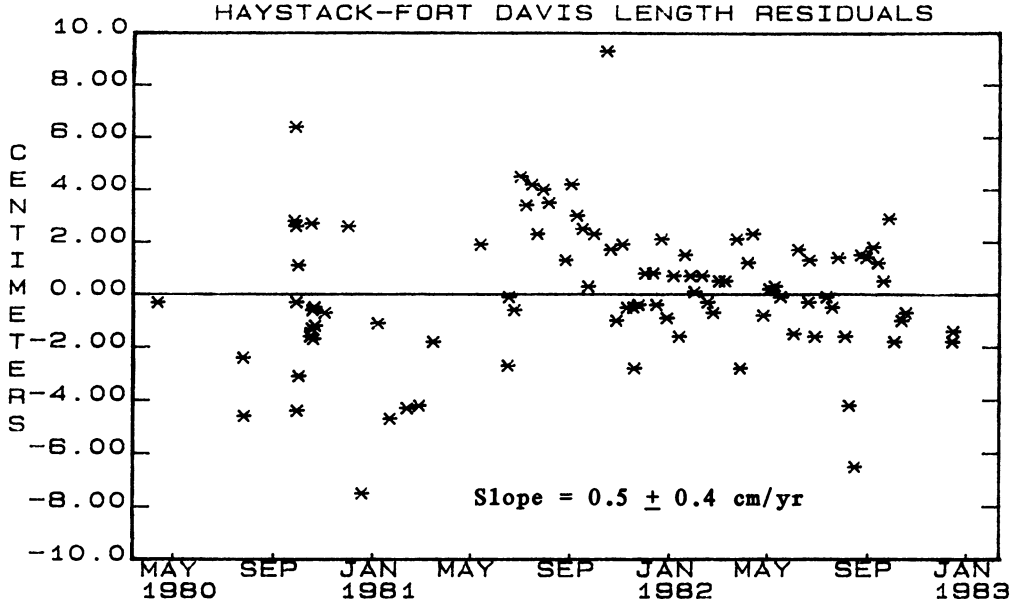
Figure 1.   Haystack to GRAS length residuals
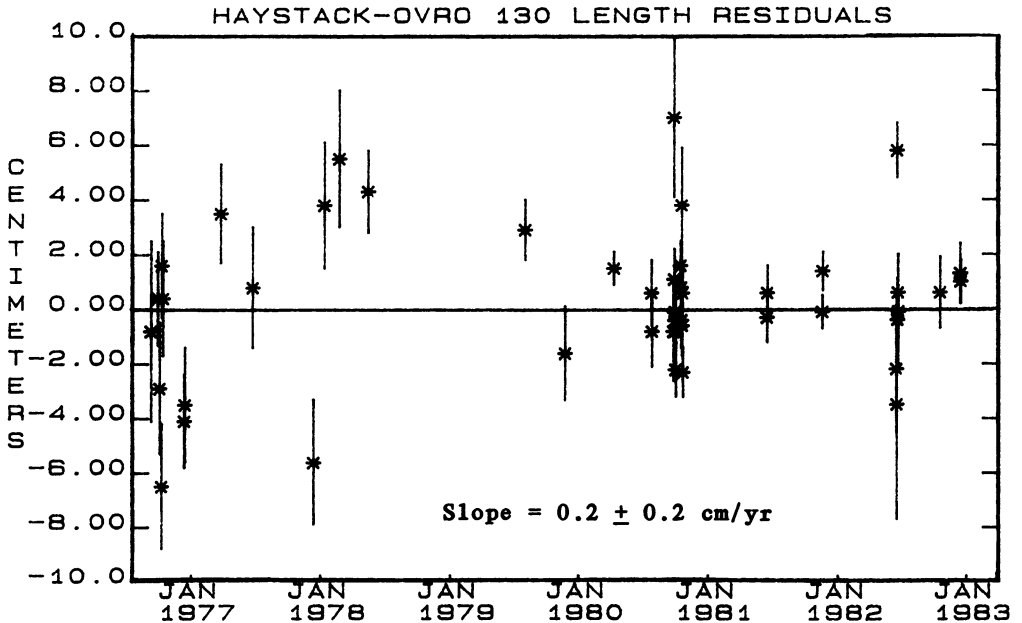   about 3 135 841.02 meters.

Figure 2.   Haystack to OVRO length residuals
            about 3 928 881.63 meters.

I thank my colleagues for discussions and advice in the preparation
of this paper, especially Dr. C. Ma of NASA/GSFC.

REFERENCES

Cohen, M. H. and Shaffer, D. B.: 1971, A. J. 76, pp. 91-100.
Kaplan, G. H. (editor): 1981, U. S. Naval Observatory Circular No. 163.
Kellermann, K. I. and Pauliny-Toth, I. I. K.: 1969, Ap. J. 155,
    pp. L71-L78.
NASA: 1980, Conference Publication 2115, "Radio Interferometry Techniques
    for Geodesy".
NOAA: 1982, Technical Report NOS 95 NGS 24, "Geodetic Applications of
    Radio Interferometry".
Pearson, T. J. et al.: 1981, Nature 290, pp. 365-368.
Resch, G. M., Hogg, D. E., and Napier, P. J.: 1983, Radio Science,
    (submitted).
Rogers, A. E. E.: 1970, Radio Science 5, pp. 1239-1247.
Rogers, A. E. E. et al.: 1978, J.G.R. 83, pp. 325-334.
Rogers, A. E. E. et al.: 1983, Science 219, pp. 51-54.
Shaffer, D. B.: 1982, paper at the Crustal Dynamics Project investigators
    meeting.
Shapiro, I. I.: 1976, Methods of Experimental Physics, M. L. Meeks,
    editor (Academic Press, New York), Vol. 12, part C, pp. 261-276.
Thomas, J. B.: 1980, JPL Publication 80-84.