VLBA clock estimates for EVN stations

Loránt Sjouwerman, JIVE

April 18, 2000

Abstract

This report summarizes my analysis of the clock searches of the feb00 Global session. It should reflect the feasibility of 'blind' or 'a-priori clock' correlation of EVN stations at the VLBA correlator. This report includes a postscript diagram of VLBA clock offset estimate versus time for 11 EVN stations and the VLA.

1 Introduction

NRAO, the VLBA correlator, has expressed the wish not to have to perform a clock search for every experiment involving EVN stations. By using GPS clock information, a clock search for VLBA stations is not needed (anymore), and the reasoning is that in principle it should not be necessary for EVN stations either once their particulars are known. I have been asked to investigate the behavior of EVN clock offsets derived from the estimates done by VLBA software, and I report on that here.

For the analysis I used the VLBA software clock estimates and the actual clock-searched offsets of the Global session in February 2000 (feb00). If found useful, and maybe after some adjustments, I could try to do similar analyses for sessions in the past and future. Please provide feedback to *sjouwerman@jive.nl*

I have excluded the experiments GS016A, EV009, and GM036; GS016A and EV009 are EVN only experiments and may float independently from the 'VLBA clock' (and they do as two people have been working on them), and GM036 has been omitted because the clocks for two different sets of filters differ enormously (up to $\sim$0.7 μs), and in the process of different correlator passes the original job scripts got overwritten. MC failed in GG042B, NT is missing from GG038B. Here, WB and Y are the phased arrays.

The rest of this report describes how the VLBA estimates the clocks, how these clocks are used in the scripts and how they compare to the clocks found. I will also point out some points of attention and summarize it all at the end.

2 VLBA clock estimates for EVN stations from GPS values

VLBA staff (the analysts) fetches the GPS clock information from the ftp-servers twice a week. When available, EVN clock estimates are updated at this rate. However, note that the 'Bologna rules' only require posting of the GPS values once a week, starting one week before and during the EVN/Global sessions, and once a month outside sessions.

The process of obtaining a clock estimate involves fitting the available GPS values around the date of observation with a straight line, starting from some reference date from when the fit is considered valid. The clock is then given by this reference date, a clock rate and a clock offset and,for a number of stations, *corrected for some empirical delay* (the 'clock' for short). Perhaps this clock fit can be further improved, but that is not relevant for the problem discussed here. For the VLBA stations all these empirical delays are known or calculable for each type of experiment (frequency, tape speed, filter); it is to be investigated here whether such an empirical delay can be derived (and is constant enough) for our EVN stations as well.

This fitted clock information gets stored in the clock section of the correlator job scripts; the correlator software then estimates the clock at the time of the experiment (at individual scan level) by adding the clock offset to the offset calculated from the rate as progressed since the reference date.

2.1 Delay search widow

The clock calculated by the method above is supposed to fall within the delay search window so that the correlated data does not suffer from decorrelation due to a faulty clock and such that the PI can obtain valid solutions in his/her fringe-fit. The total theoretical delay search window is set by the number of lags (FFT-size in VLBA jargon), the spectral averaging (SA) and the BBC bandwidth (BW) used in the experiment:

$$\text{total delay window} \;=\; \frac{\mathrm{FFT-size}}{2 \times \mathrm{SA} \times \mathrm{BW}}, \tag{1}$$

which is in μs if BW is specified in MHz.

However, at the edges of the total delay window, the sensitivity to the fringes degrades considerably, so that only half of the total delay window can effectively be used to find the fringes. Specifying the delay window symmetrically around the clock actually used, as is the default search window for AIPS, the clock estimate effectively should be as accurate as one quarter of the above:

$$\text{effective delay window} \;=\; \pm\frac{\mathrm{FFT-size}}{8 \times \mathrm{SA} \times \mathrm{BW}}. \tag{2}$$

By default, clock searches are done with SA=1, and FFT-size of 512 or 1024, depending on the number of BBC's used. FFT-size of 2048 cannot be used if any of the stations in the clock search job is barrel rolling, and for cross-polarizations 256 is the maximum. Because of the low SA in clock searches (and no cross-polarizations), a standard bandwidth of 8 MHz for a standard continuum project, and typically less for spectral line projects, allows for an effective search window of at least 8 μs on each side.

If clock searches are skipped, the user specified SA and polarization must be taken into account. Spectral line projects are probably okay anyway, but when assuming the default continuum, and default polarization observations, the correlator parameters will be FFT=512, SA=16, BW=8 and FFT=256, SA=8, BW=8 respectively.

The effective delay search window, for reference, is ± 0.5 μs in either case.

2.2 Estimates versus offsets

To investigate whether the EVN clocks can be estimated sufficiently accurate, within this effective range of ± 0.5 μs, I have compared the estimates from the VLBA software with respect to the clock search results.

That is, I have extracted the reference date, the clock rate and the clock offset from the correlator jobs in the 'orig' directory - where a copy of the original jobs as made by the VLBA software and including the original clock estimates is kept - and the same information from the correlator jobs that were actually used after the clock search had been performed and incorporated (by editing or replacing the whole original clock section by the results from the search). Both the original and used clocks are progressed from the reference date to the start of the job by multiplying the time difference of the reference date and the date of observation by the rate and by adding the offset to this number.

The plot at the end of this report displays the clock used minus the estimated clock. For JB1, JB2, ON60 and WB, a constant ($+1$, $+1.5$, -2, and $+17$ μs respectively) was added to bring the clock in the window around zero. These numbers make good candidates to end up as the (corrected) *empirical delay* for these stations.

For each EVN station that produced more than one clock during the feb00 session (EB, HH, JB1, JB2, MC, NT, ON85, ON60, SH, TR, WB, and also Y; for CM, MH, YB, RO, and GO only one clock was available) we display the difference between the clocks as function of day of observation (in integers) as fat filled circles. Where multiple points blend together to an oval or a vertical stripe, the rate introduces non-constant differences between the clocks during the experiment with each correlator job. The data points are connected by solid lines to guide the eye, where a long-dashed horizontal line outlines the perfect clock estimate (the zero line). Short-dashed vertical lines divide the Day-of-Year axis in the different frequency sessions of the feb00 session (18 cm, 1.3 cm, 3.6 cm, and 6 cm respectively). Jodrell Bank Observatory and Onsala Space Observatory use different telescopes for different frequency sessions, which is apparent in the plot by assigning those telescopes each their own plot.

Before I go further into the details of the plot, a small aside section is in place.

3 Practice of clock searches at the VLBA correlator

It is probably a good idea to go deeper into the practice of clock searching at the VLBA correlator (and at JIVE too ?), as it is of influence on the clocks that get used; a perfect clock estimate can be ruined by even a successfully performed clock search.

As there is no significant back-log for the VLBA correlator, clock searches are started as soon as most of the tapes for a given experiment are in, usually about two weeks after the experiment. A clock search job of a few minutes on a strong calibrator is created, and submitted to the correlator queue, where it gets correlated with the highest priority. The correlator output then gets 'sniffed', and the clock for each of the stations in the clock search is checked. If a delay offset for one or more stations differs more than about 80 ns, a clock correction is necessary. Currently, this is *always* the case, because empirical delays are not applied to every station. Notably, already now improvement could be made at JB1, JB2, ON60, WB, and perhaps HH, ON85 and SH.

3.1 Copying and pasting of clock information

As soon as clock corrections should be made, the clock section of the job is copied to a file. This file finally contains the reference date, offset and rate for all stations in the experiment - not only the clocks that need to be updated. In this file the clock corrections are applied to the 'offset' value of the appropriate station. In case another clock search warrants it, additional offsets are added in this file. Depending on the person working on the experiment, clock offsets smaller than about 80 ns may be updated (but this is usually skipped by the analysts). *Clock rates are never updated.* Fringe-fitting by the PI removes the remaining clock offsets between the used and 'true' clocks, which thus in general are smaller than about 80 ns.

From this point, when a clock-correction file has been made, it is used to *fully replace* the clock section from all the next correlator jobs that are generated. This means that when stations post new GPS values in the mean time, the fit to the clock might be a better estimate to the real clock, but that *the updated estimate gets overwritten by the clock search clocks !* This limits the effectiveness of clock estimates when clock searches start before all GPS information is available.

I have seen many experiments in which the fit to the clock search job has a different reference date than the original production jobs. This indicates that meanwhile a new clock estimate in the original production job has rejected the old fit in the clock search job and now allows for a better rate and offset estimate. It may be disastrous to the experiment if the updated rate (e.g. for a non-linear maser such as EB) has an opposite sign to the rate used (from the old fit in the clock search job). Also, clocks found in one experiment are commonly copied over to use them as a first estimate for following experiments way down in the whole session, and in my view this practice should be abandoned for the reasons mentioned above.

But also clock search jobs and production jobs that use the same reference dates and values, are sometimes using an old fit due to relatively late posting of GPS data (or an early start of correlation). In my experience this error, which is completely unnecessarily introduced, is usually of the order of 0.2 μs or less. But it may also be a multiple of this if the GPS information is very old, which may be the case for stations that are not often included in the array (AR, CM, HH, MH, RO, SH, SM, UR, WZ, YB, ..).

4 Explaining the differences

The effect of an old fit mentioned in the previous paragraph can already be seen in the first part of the EB clocks. That part, the whole 18cm session, uses a fit valid from February 2 (day 33) in both the clock search as well as the production jobs (created one month later on March 2; day 62). The maser drift is assumed to be linear from that day, either because it is, or because there is no sufficiently new GPS information available to reject the fit in favor of a new one.

Apparently there is a correction needed in the 18cm part of the session, because the clock search indicates a correction to the offset that was initially estimated in the production jobs. In other words, the estimate deviates from the true (clock searched/found) value by about 150 to 200 ns. Assuming that the clock found in the first experiment is valid for the whole session, may or may not be justified depending on the reference date used, and the rate drift of the maser for that station.

Note however, that on the other hand the clock for TR, which uses a reference date of January 6 (day 6) in those jobs for day 42, is surprisingly still spot on ! (better than about 80 ns).

A new clock fit for EB became available for after day 52 (February 21); that is, the updated fit is considered to be valid from day 52 and further until a new fit overrides this fit. Clock searches from day 42 to 56 used a fit valid from February 2 (day 33). The clock is estimated by using the rate and offset starting at February 2. Production jobs generated for day 52 and later have an estimate from another rate and offset, considered to be valid around February 21 (day 52), thus implying that the fit from February 2 is not good enough for these experiments.

However, the corrections of the clock searches for days 52 to 56 do compare with the estimates from the new fit, telling us that the VLBA clock estimates with this new fit are correct (*but not a-priori applied !*). After the clock search the applied offset may be correct again, but because the clock sections are overwritten, the rate continues to be an old estimate.

The last two points for EB use the fit from February 21 for both the clock search as well as the production jobs. Differences between the recent estimate and the clock search are within the normal 80 ns.

4.1 Dependency on frequency, tape speed, filters

In the 1.3 and 3.6 cm sessions another effect is seen, but only for Mk4 formatters. The clock offsets alternate between two values, which differ about 400 ns and are consistently e.g. more negative for all the Mk4 stations in an experiment. These are experiments with a tape speedup factor of one (days 52, 54, and 55), whereas all other experiments have speedup factor two. The VLBA formatter apparently already corrects for this (EB, Y; no data for TR, and NT used the Mk4 during these experiments). Information on other speedup factors is not yet available. Maybe this can be tested and/or eliminated in the future. These numbers already should be incorporated in the VLBA clock estimates, as they are large compared to the window of 0.5 μs.

Apart from the usage of different telescopes and hence applying different empirical delay offsets at some sites (JB, ON, maybe RO, but also single dish or phased WB etc) that could already be done at the VLBA correlator, I see no systematic dependency of delay offsets on frequency. The small numbers available of course exclude a proper analysis, and perhaps we may investigate this further as a second order effect.

Another effect that I noticed is that there is a difference of about 500 ns between the VLBA formatter and the Mk4 formatter (NT switched formatters on day 47 during EV009), possibly due to the different electronics of the two systems.

Finally, it has been noticed by the analysts for GM036 (not included here) that Mk4 formatters also differ in delay offsets for different sets of (internal versus external ?) filters. GM036 used both wide band (8 MHz) as well as narrow band (0.5 MHz) filters and the maximum was a difference in clock offset of about 700 ns for MC. More data is needed to investigate this peculiarity. Whereas the narrow filters are probably okay, the clock estimate for the wide band filter still needs to be accurate.

4.2 Final clock offsets and rates

I have to emphasize that the offsets in this report are not the exact offset differences between the clocks estimated and the 'true' clocks (but the used clocks instead). The true clocks are found by scanning the final sniffer plots or fringe-fitting. Personally I try to always take out any residual delay while I am at it anyway, but as mentioned before, the analysts usually do not bother with small differences (up to about 80 ns). However, occasionally it also happens that a clock is further off (e.g. 150 ns for ON in GP025A, day 44; the second point for ON actually should be 150 ns 'higher' which puts it in line with all other points; another one is 150 ns for JB1 and −230 ns for MC in GK020, day 45, again both consistent with the general trend). In this report and the figure I have not tried to account for such offsets.

5 Critical points

So far I have been commenting on the clock search results, but one can only do that for clocks that were *found*. A problem within the EVN is still, in my opinion, individual station complications. Although EV009 is not included here, it did suffer from problems probably justifying a clock search anyway (missing log for EB, switch of formatters for NT). Also there were individual problems for MC in GG042B (wrong LO frequency), NT in GM034 (crossed polarizations) and ON and NT in GG038B (wrong tape footages, and no fringes respectively). If you don't want to exclude fringes for these stations by default if no fringes are found without a clock search, you'll probably only find out the problem after some searches.

But I think that the biggest loss of efficiency comes from inclusion of stations that are not often used (CM, HH, MH, RO, SH, YB, etc). You want to do a clock search whenever (usually only) one of these stations is in the experiment concerned. Taking along the other stations in that experiment for a clock search is not much extra effort. At least 5 of the 11 experiments of this session would warrant a clock search on this argument already. Improvement could be made if we can blindly trust the clocks estimated for these stations, but that would require the effort of frequent posting of GPS clock information by these stations.

The lines above therefore implicitly indicate an effective way to proceed to limit the number of clock searches needed at the correlator; do a clock search with the experiments with non-default stations before anything else, and preferably take along the default EVN stations. Then compare the spectral line correlations, which probably don't need a clock search because their delay search window is large due to the narrow bandwidth used. As an extra check one may choose to do an independent clock search on an experiment which has the narrowest delay search window as specified by the PI's correlation parameters.

But there is more to be done for (semi-)blind correlation of EVN stations. I think it can be divided in the following (EVN) station and (VLBA) correlator items:

5.1 Station recommendations/responsibilities

I think that the most important, and probably most difficult task is for the stations to post the GPS information timely, and continuously. This will especially be a burden to stations that are not often used. Perhaps one should reconsider the 'Bologna rules' on this item, or maybe an automatic posting of this data could be sorted out.

Another item, in my view, is to try to inform the correlator as accurately as possible of errors or complications that might have occurred during the experiment, either in the log or in the feedback facility. Making mistakes is natural, but effort could be spent to try to avoid it. If something was noticeably wrong, a simple warning 'we did some things we cannot remember' would already be a valuable comment at the time of correlation. For example, it took me three clock search jobs to figure out that ON had reset the tape footage counter during the GG038B experiment. Knowing this beforehand would make the correlation more efficient.

Perhaps not really an item for the stations (but for JIVE or the EVN) is to find out the Mk4 particulars. That is, to test the clock dependencies on tape speed more precisely, as well as dependencies on filters and frequencies if apparent.

5.2 Recommendations for the correlation

This section separates in recommendations for a better clock estimate by the software, and suggestions for clock search practices. The first section is probably best when itemized:

- Apply/update empirical delays for all stations, in particular JB, ON, (phased) WB, and perhaps HH and SH.
- Implement different empirical delays for ON60 and ON85, JB1 and JB2, and single dish/phased WB (e.g. in *editobs*).
- Include the same reference date, offset and rate for CM as for JB (they use the same maser), but with an additional empirical offset of about −955 ns.
- Account for different clock offsets for Mk4 formatters as function of tape speedup factor.

When implemented, I think that these points already greatly reduce the need for clock searches for the regular EVN stations. However, I think the practice of clock searches should also be adjusted.

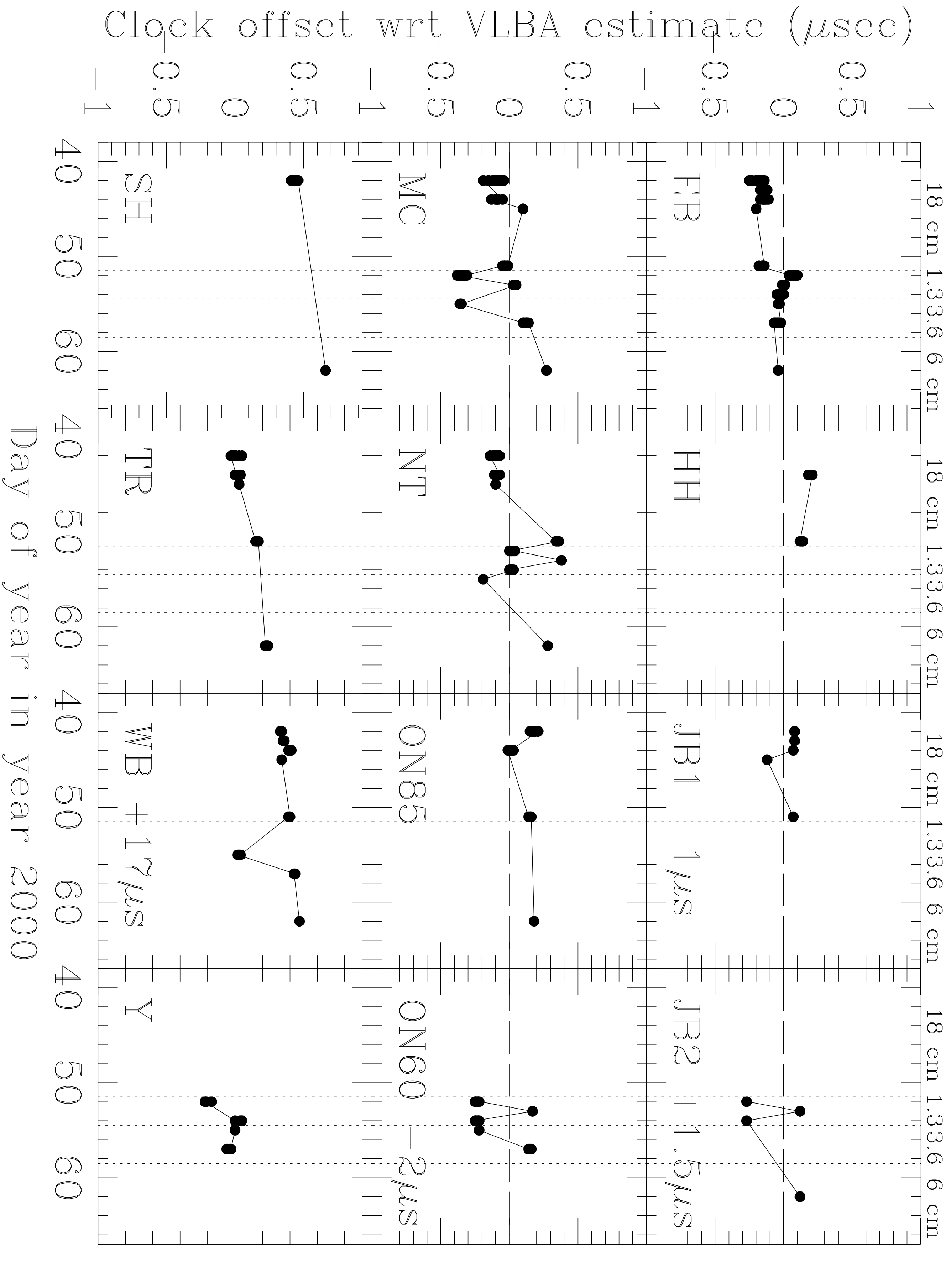
Apparently it takes about 6-8 weeks after the experiment to obtain a good fit to the GPS data with the current practice of GPS postings (you want to have clock fits with both reference dates before and after the experiment). Of course the availability of thin tapes requests rapid processing, but I would recommend to wait starting with clock searches at least until 6 weeks after the experiment. Otherwise, copying and pasting of the clock search job clocks remains an additional weak point. Also, the assumption that a clock found can be used for the rest of the (frequency) session is not generally justified, in particular for stations with high rates (e.g. EB).

Experiments with stations that are not often used need a clock search anyway. I would suggest to take along all EVN stations in this clock search to check the clocks. I think thorough checking and verification of the clocks for all stations in each experiment is needed until we are confident that we understand the particulars of each station. Recording the findings for all the stations, say for the next one or two years, may result in not having to do clock searches for any of the stations in the future.

Finally, attention remains to be paid to experiments with stations that experienced problems during the recording. The suggested new practice of skipping clock searches should not result in discarding data from those stations without further investigation.

6 Short summary/conclusion

Altogether, I think we are moving towards a better understanding of the problems encountered in the clock search procedures for EVN stations, and probably are able to blindly correlate EVN stations in due time (1-2 years). Not everything is sorted out yet, but already now much improvement can be made. Some points need continuous attention (posting of GPS information), others only incidentally (mostly software, updates). The coming sessions will, after the software for clock estimates includes the items above, show how to proceed and what needs to be investigated further. The stations that are not often used, and individual stations with problems would still need a clock search, which is best started after 6 to 8 weeks have passed since the experiment.



Clock offset wrt VLBA estimate (μsec)
Day of year in year 2000
18 cm 1.3 3.6 6 cm
EB
HH
JB1 +1μs
JB2 +1.5μs
MC
NT
ON85
ON60 −2μs
SH
TR
WB +17μs
Y