

# EVN Calibration Issues

## — Summary of to-dos

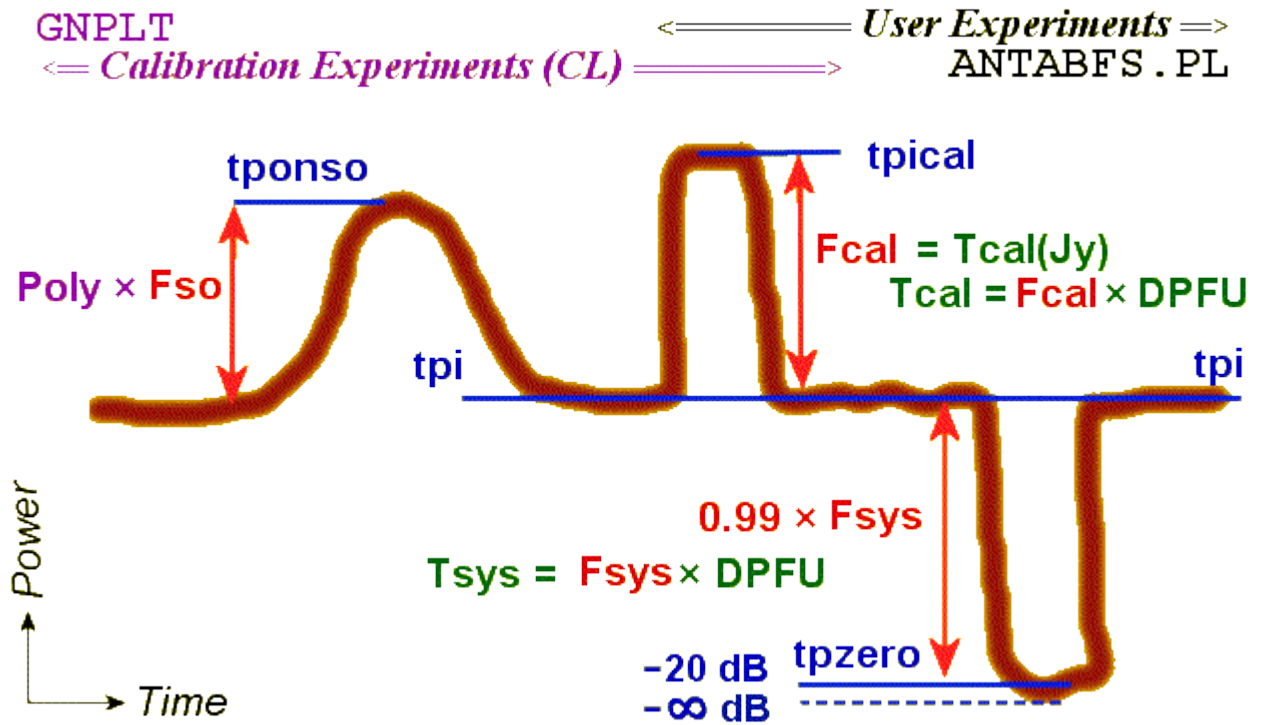
**Kaz Borkowski**

Torun Centre for Astronomy, NCU, Torun, Poland

This document is heavily based on earlier reports of studies carried out in 2007. Discussions over the internet and during the TOG Meeting at Yebes rectified and enriched my understanding of certain aspects, therefore I felt it necessary to collect ideas that are worth implementation in a new document. **I intend to maintain this page by adding information on possible reactions from VLBI community and on new ideas if and when they are put forward.** The mentioned reports are as follows:

- [\*Analysis of EVN Calibration Problems at Torun Station\*](#) – a web report of May 2007,
- [\*The DPFU Parameter Plays Havoc with EVN Calibration\*](#) – presented at the TRAO Seminar, 25 Oct. 2007,
- [\*The role the DPFU plays\*](#) – presented at the TOG Meeting, Yebes, 12 Nov. 2007.

# EVN Calibration scheme



**Fig. 1: The idea and practice of EVN calibration.** In this diagram, the thick curve represents the power of signal at BBC or VC outputs as a function of time. The three prominent features correspond to a source of known flux density  $F_{so}$  (left), calibration signal from noise diode  $F_{cal}$  (middle), and a depression due to 20 dB (factor of 100) attenuated background  $0.99 F_{sys}$  (right). The original Field System readings  $tp...$  as recorded in the experiment logs, i.e. counts as read from total power integrators, are marked in blue. The broken line just below  $tpzero$  shows the level expected of an infinite attenuation, the small difference being equal to  $F_{sys} / 100$ . Calibration experiments serve to determine the gain curve  $Poly$  and the equivalent flux density of diode signal  $F_{cal}$ , or its temperature equivalent  $T_{cal}$ , at number of frequencies using the GNPLT application. The results are kept in special rxg files. Continuous monitoring during user experiments relies on so obtained  $F_{cal}$  to determine  $SEFD = F_{sys} / Poly$ , required by the final user to scale his maps. In practice, this is accomplished with the ANTABFS.PL script which uses  $tp...$  measurements (performed while observing target sources and collected in the log files) along with  $T_{cal}$  and  $Poly$  (taken from the rxg files) to produce tables of  $T_{sys}$ , accompanied by  $DPFU$  and  $Poly$  data.

# Methodological issues

1. The DPFU parameter must not be updated basing on the CL experiments, simply because there is not any justification for such an operation. Unfortunately however, the GNPLT does allow for this (by offering options for adjusting DPFU alone or fitting DPFU along with Poly). Thus, **GNPLT should be amended so as to prevent users from fitting the DPFU parameter and to fit the gain curve alone** (Poly; normalized at its maximum).
2. **While calculating Tcal or Tsys, the count difference of  $t_{pi} - t_{pzero}$  should be equated with  $0.99 F_{sys}$  or  $0.99 T_{sys}$**  (and not with  $F_{sys}$  or  $T_{sys}$ ) to account for the power at  $t_{pzero}$  being not equal to zero but about 1/100 (-20 dB) of that at  $t_{pi}$  (see [Fig. 1](#)). Otherwise we will continue to introduce this small bias in  $T_{sys}$  or SEFD (systematic underestimation by 1 %) which gets transferred right onto the final users maps.
3. **Onoff measurements scheme could be somewhat optimized to**
  - (A) make data acquisition of more important quantities, these that are to be later differenced ( $t_{ponso}$ ,  $t_{pi}$  and  $t_{pical}$ ), to take place immediately one after the other. This would minimize the effect of slow variations in receiver gains. While doing this we might
  - (B) assign the lowest priority to the gain compression (onsource  $t_{pical}$ ) and  $T_{sys}$  measurements ( $t_{pzero}$ ), which do not directly contribute to accuracy of the final results ( $T_{cal}$  and Poly determinations). A minimal version of basic sequence could be e.g.  
`tpical(onsource), { tponso, tpi, tpical(offsource) }, tpzero`with the part enclosed in braces optionally repeated a few times.
4. The  $t_{pgain}$  used as a proxy for VLBA racks while computing  $T_{sys}$  by the ANTABFS may not be the best choice. It seems better to **'continually' record  $t_{pical}$  and  $t_{pi}$  instead of  $t_{pgain}$  and to measure zero level ( $t_{pzero}$ ) only once in a while**. Since gain variations appear in  $t_{pzero}$  100 times suppressed, the ratio  $(t_{pi} - t_{pzero}) / (t_{pical} - t_{pi})$  should remain quite gain insensitive. Also,  $t_{pzero}$  could be tested in ANTABFS for abnormal values (in magnitude and/or standard deviation) and in case of problems could be automatically replaced by the last good value or even by  $t_{pi}/100$ . Such replacement would prevent spoiling of all the  $T_{sys}$  measurements between successive  $t_{pzero}$  measurements by one wrong readout. This scheme of monitoring should also eliminate the familiar steps in continuous  $T_{sys}$  that are due to the quantized nature of the BBC AGC ( $t_{pgain}$ ).
5.  $T_{cal}$  and  $T_{sys}$  (or their flux density counterparts) computations should take into account measurement errors. *As [shown elsewhere](#), weighting by the reciprocals of variances may constitute an excellent tool in future dealing with the ever worsening problem of RFI.* For computing of  $F_{cal}$  or  $T_{cal}$  we already have the necessary error data in the form of the onoff measurements recorded in the logs of CL-type experiments (the lines marked #onoff#ONSO, ONSC, OFFS, OFFC and ZERO). To be able to include the weighting (or other type of dicrimination) of measurements for  $T_{sys}$  in the user experiments it would require some **modification within FS, so that such readings like  $t_{pi}$ ,  $t_{pical}$  and  $t_{pzero}$  are all accompanied by respective standard deviations or variances.**

6. **Field System and its auxiliary programs** while measuring the signal power (for tp... commands) should also **do some internal weighting** before storing the results in logs. This would efficiently diminish the impact of the short-time system instabilities and RFI. A simple algorithm could be based on splitting of integration period just into two halves and calculating an average waighted by the reciprocals of variances in the two halves. Mathematical formulae for weighted average and its error in a more general case of splittning the period into N equal-length (0.5 s comes to mind) sections are given in the following box.

Let the i-th of N sections of averaged data have the mean value  $x_i$  and the variance  $v_i$  (in practice to be calculated as the RMS squared), then weighted means  $x$  stored in logs could be calculated as

$$x = (\sum x_i/v_i) / \sum 1/v_i, \quad (1)$$

where the summations are carried over all the N sections. The  $x$  quantity represents an unbiased estimate of the mean value in the entire integration period. So obtained value has in general an expected error (standard deviation or RMS value) different than the one for the normal (arithmetic) full period averaging, which would be just  $\sigma = (\sum v_i)^{1/2} / N$ . [It is proposed](#) that for the suggested kind of weighting the error of  $x$  be estimated from

$$\sigma' = 1 / (\sum 1/v_i)^{1/2}. \quad (2)$$

Note, however, that when all  $v_i$  happen to assume the same value,  $v_i = N \sigma^2$ , these formulae get equivalent to the case of arithmetic averaging. Setting  $N = 1$  reduces them similarly.

Being relatively simple, the extra computations according to Eqs. (1) and (2) above could possibly be implemented without excessive FS changes. They, of course, would not require any other procedure to be modified. However, such tools as GNPLT and ANTABFS should independently implement, wherever possible, a weighting of processed data using the errors (or standard deviations) stored in logs by the FS. In this case, combating of interference on a longer time-scale would be aimed at.

## ***GNPLT wishlist***

1. Names of currently used rxg file and analysed log file could be displayed somewhere within the GUI.
2. Such plots like [Plotting TCal\(K\) vs. Frequency](#) have only three widely separated ticks with labels on their scales. More ticks (with or without labels) would be welcome.
3. A new option to display error bars would also be useful in discerning bad points.
4. A new option to toggle inclusion/removal of IFa and IFb (wide bandwidth) points. Presently they are included automatically, if present in the log file, which may lead to false Tcal if the user does not explicitly delete them. A CL experiment may be scheduled not to record these data, but they are in fact useful in case of problems.

## ***ANTABFS wishlist***

1. When the final plot with all channels is displayed (as in [this example](#)) it is difficult to identify in it a particular BBC/channel. Therefore, it would be helpful to have displayed also a color-channel legend.
2. For each frequency channel the ANTABFS lets the user to choose from keyboard an option to edit or go ahead to the next channel. When upon inspection of the displayed measurements he decides to edit, the display is replotted and only then the editing is enabled. If observations are made in the presence of RFI (which happens almost always at L-band), virtually every channel (BBC and sideband) data have to be edited during the ANTABFS session. Life of the user would thus be more pleasant if he got the editing enabled already in the first popup of displayed data and had there one or two more options (something like 'proceed to the next channel' in addition to 'end editing and replot') for picking with the mouse, besides these keyboard options. The idea is to free the user of excessive switching between windows (the x-window where ANTABFS runs, and the graphical display of output data).