UNIT 6. Combining solutions from several surveys using *globk*

Ref. GLOBK manual

==========

1. OVERVIEW

We combine individual sessions (e.g., days) of observations (h-files) with SOPAC global h-files to obtain

a) Station coordinates averaged over a multi-day experiment.
b) Station coordinates uncertainties by scaling the covariance of the GAMIT solution (h-files).
c) Station coordinates reference frame

The *globk* combination of GPS data is a two step process:

a) Remove outliers from the corresponding h-files by renaming the outlier station in an *earthquake file*.
b) Run *globk* to combine daily h-files into a single h-file that represents the averaged station coordinates for the chosen time range (e.g., monthly averages). The scaling of the GAMIT solution covariance is done within the *globk* combination run.

Note: examples of SOPAC global solutions (h-files) include:

igs1, igs2 and igs3	IGS global network
eura	european stations (mostly EUREF)

Visit the SOPAC web site http://sopac.ucsd.edu/dataArchive/ to have a complete list of the global solutions (h-files) available.

2. REMOVING OUTLIERS

You can remove the outliers from the final combination using the *rename* command in the earthquake file (see GLOBK manual at pg. 38 – Defining earthquakes and renaming stations).

Do not jump right to the delete outlier stage before really assessing what might be causing the problem. Perhaps those outliers are caused by poor stabilization that can be fixed at the glorg stage. Perhaps there are antenna height problems in the station.info file. Perhaps there's some poor bias fixing that's degrading the east component. Do the log sheets indicate that there was a problem on that day - tripod not centered at the end of the session, etc?

Throwing out a little bit of data is not significant when you're working with continuous data, since there may be more than enough data to give a good velocity estimate and you can really pinpoint where problems arise. But when processing survey-mode data every little bit of data may be important.

==========

FIRST STEP – No need to create a new directory structure. You will employ the one already organized for the *glred/glorg* run

Expt					
expt/doy	expt/templates	expt/tables	expt/gsoln	expt/glbf	expt/gplot
h-files	command files, the itrf00.apr file, **earthquake file**		For running solutions (command files, lists of binary h-files, output files)	Binary h-files	For running plot of coordinate time series

==========

==========

SECOND STEP – Run *sh_plotcrd* with the options *–r –o 1* to plot the time series as residuals of a first order polynomial.

The file VAL.xxxx (where xxxx is the 4 char experiment name) has a list of residual and errors for every station/point in the time series. You can import this list in a spreadsheet and sort the data (by errors or by residuals) to identify outliers.

Example: (see Figure 1 after FOURTH STEP)
==========

==========

THIRD STEP – Make a backup copy the original MIT earthquake file available in the *expt/templates* directory
==========

==========

FOURTH STEP – Edit the earthquake file in the *expt/templates* directory.
Remove the outliers from the final combination using the *rename* command in the earthquake file (see GLOBK manual at pg. 38 – Defining earthquakes and renaming stations). Renaming a station to end in _xcl will cause it to be automatically removed.

rename <Orig site name> <New site name> [hfile code] [epoch range]

<Orig site name> site name that appears in the original binary h-file
<New site name>   site name used in the current solution

```
[hfile code]      name of the original binary h-file
[epoch range]     range of time over which the station will be removed

Example: (see Figure 1 below)
================================================================================
```

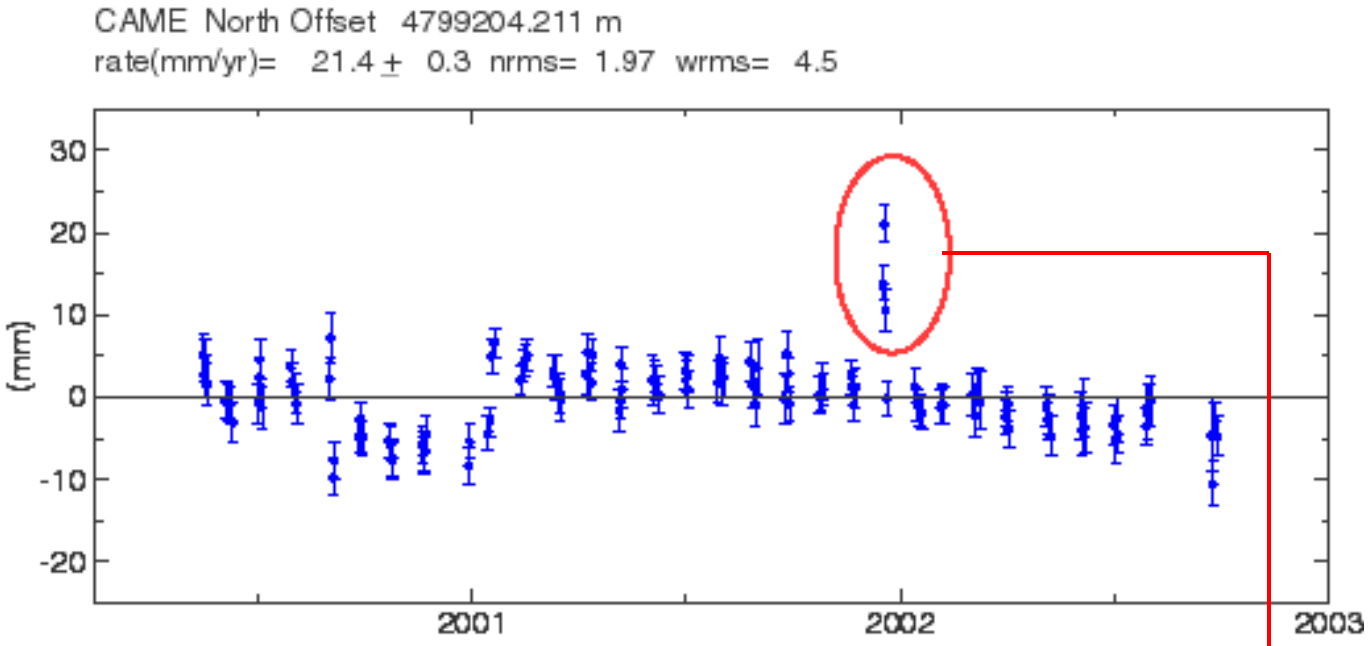

Figure 1a. Plot the time series using *sh_plotcrd*. Locate the outliers.

```
Combi    nat   ion          of    eura     igs1 igs2 igs3   networks
CAME_    GPS   to           N     Soluti   on   1

   2001    12   17   351   11     59    4799204.23     0.002    0.0139     0.002
   2001    12   18   352   11     59    4799204.24    0.0022     0.021    0.0022
   2001    12   19   353   11     59    4799204.23    0.0025    0.0105    0.0025
```

Figure 1b. Identify outliers in the file VAL.xxxx (where xxxx is the 4 char experiment name, the file used in Figure 1b is VAL.adri).

```
rename came_gps came_xcl h0112171200_adri.glx   01 12 17 0 0 01 12 17 24 0
rename came_gps came_xcl h0112181200_adri.glx   01 12 18 0 0 01 12 18 24 0
rename came_gps came_xcl h0112191200_adri.glx   01 12 19 0 0 01 12 19 24 0
```

Figure 1c. Remove the outliers from the final combination using the *rename* command in the earthquake file

3. COMBINING SOLUTIONG AND SCALING THE COVARIANCE

==========

FIRST STEP - Copy the SOPAC global h-files (e.g., igs1 & igs2 & igs3) in */glbf*

```
sh_get_hfiles
================================================================================
Script to get hfiles by day number from the scripps archive

Usage: sh_get_hfiles -yr <yr> -doy <doy> -ndays <num> -net <networks> -soln -ftp_prog
<ftp/ncftp>

<yr> 4 char year of hfile requested  [Required]
<doy> 3 char day of year of hfile requested [Required]
<num> Number of consecutive days of hfiles to retreive [Default 1]
<networks> List of networks to be retreived from the ftp archive [Default ALL]
         Network choices: igs1 igs2 bard cors noam dgga eura pgga net1 net2 net3 net4
-soln Check the solution archives rather than the h-file archives
<ftp/ncftp> choose the ftp program to be used [default is ftp]

 Examples: sh_get_hfiles -yr 1999 -doy 246
           sh_get_hfiles -yr 1999 -doy 246 -ndays 1 -net igs1 igs2 eura
```

==========

You can also download the global h-files directly from the SOPAC web site OR anonymous FTP server (update information available at http://sopac.ucsd.edu/dataArchive/)

==========

==========

SECOND STEP – Prepare the binary h-files in */glbf* using *htoglb*

htoglb converts ASCII h-files into bynary h-files readable by *globk*.
htoglb will generate 4 different binary files:
gcr and gcx (biases free and biases fixed tight constrained solutions)
glr and glx (biases free and biases fixed loose constrained solutions)

The binary file wanted by *globk* is the glx solution.

Runstring:

```
Usage: htoglb [glbf_dir] [ephemeris_file] <GAMIT h-file>
               output       output           input
where
[dir] is the directory for the output files.
[empheris file] Name of the file for output of the ephemeris for the satellites. Can
               then be used as input to GLOBK.
<input files  ... > is a list of input files

Example:
      htoglb    . ../tables/svs_myexp.svs ../008/hnbaya.01008
      htoglb    . ../tables/svs_myexp.svs higs?a.01008
```

==========

==========

THIRD STEP - Copy the *globk* (globk_comb.cmd) command file for combinations from */templates* to */gsoln*

==========

FOURTH STEP - Edit globk_rep.cmd (directory */gsoln*) -= see GLOBK manual for commands explanation

```
*
* Globk file for combination of daily h-files
*
  eq_file ../templates/scec_eq.v2.1
  make_svs ../tables/sat1.apr
  com_file globk_comb.com
  srt_file glb.srt
  sol_file globk_comb.sol

#  earth-rotation values
x  (not needed if pmu free in final combination; pmu.bull_b begins only at 1 Jan 1992)
 in_pmu  ../tables/pmu.bull_b

# apr site file(s)
  apr_file ../templates/itrf00.apr

# (1) Max chi**2, (2) Max prefit diff, (3) Max rotation; defaults are 100 10000 10000
  max_chi 30 50 2000.0

* Apply the pole tide whenever not applied in GAMIT
  app_ptid ALL

# Allow the network to be loose since using glorg for stabilization
  apr_neu  all  10 10 10  0 0 0

# Satellites are loose if combining with global SOPAC H-files
#               X       Y      Z    XDOT  YDOT  ZDOT  DRAD YRAD ZRAD BRAD XRAD DCOS DSIN YCOS
YSIN BCOS BSIN
  apr_svs all   100     100    100   10     10    10    1     1  .02  .02  .02 . 02  .02 . 02
. 02  .02 . 02
#  tight if not combining with global data
x  apr_svs all .05 .05 .05 .005 .005 .005 .01 .01 F F F F F F F F F F F F F F

# Keep EOP loose
  apr_wob  100.  100.  10. 10.0  0.0  0.0  0.0  0.0
  apr_ut1  100.  10.  0.0  0.0  0.0  0.0
  mar_wob  36500 36500 365 365 0 0 0 0
  mar_ut1  36500 365 0 0 0 0
#  unless not using global data
x  apr_wob  .25 .25  .001 .001 0 0 0 0
x  apr_ut1  .25 .25 .001 .001 0 0
x  mar_wob    1 1  .001 .001
x  mar_ut1    1 .001

* Estimate translation   -   .0005 m**2/yr = 15 mm/half-yr
 apr_tran .005 .005 .005 0 0 0
 mar_tran .0025 .0025 .0025 0 0 0

# Set minimal globk print options since using glorg output
   prt_opt cdms gldf

# Invoke glorg for stabilization
(NOT USING GLORG, THE OUTPUT WILL BE A LOOSELY CONSTRAINED SOLUTION)
x org_cmd glorg_comb.cmd
x org_opt cmds psum gdlf
x omit this for naming with sh_glred:
x org_out globk_comb.org

# Write out an h-file if needed for future combinations
  out_glb EMED97.GLX
```

==========

FIFTH STEP – Find the scaling for the global h-files (igs1, igs2, igs3)

a) Generate (directory */gsoln)* the input file list (*.gdl) for all the global h-file in the time span of your combination. A quite straigthforward approach is to use the ls command

Example: ls -1 ../glbf/*igs*.glx > igs_glx.gdl

b) Run *globk* (the scaling parameter is the nrms from the *.prt file)

```
 Usage: globk  crt prt log input_list markov_file
   crt = 6
   prt = globk_rep.prt
   log = globk_rep.log
   input_list = gdl file (e.g., nbay01_glx.gdl)
   markov_file = globk_rep.cmd

 Example: globk 6 globk_igs.prt globk_igs.log igs_glx.gdl globk_comb.cmd
```

c) Get the scaling parameters for the covariance (formal errors) of the igs global h-files (the scaling parameter is the *prefit chi**2* from the *.prt file)

grep “prefit chi” globk_igs.prt

```
The  prefit chi**2 for   22758 input parameters is     1.030
```

==========

The *prefit chi**2* provides a measure of the misfit of data relative to their formal uncertainties. By adjusting the scaling on the hfiles such that chi**2 ~ 1, the resulting one-sigma data uncertainties then equal the average scatter of the data residuals about the model.

The data scatter provides an unbiased estimator of the uncertainties if the error spectrum is white and the data are spatially and temporally sampled homogeneously. More complex procedures should be considered when mixing daily continuous data from IGS networks with infrequently sampled survey-mode data, and when trying to account for correlated noise due, for example, to monument wander or seasonal effects. See McClusky et al., JGR, 5695-5719, 2000, for further discussion.

==========

SIXTH STEP – Repeat FIFTH STEP to find the scaling for other global h-files (e.g., eura) and your survey

a) Generate the input file list (*.gdl) for the h-files
b) Run *globk* (the scaling parameter is the nrms from the *.prt file)
c) Get the scaling parameters for the covariance

==========

==========

SEVENTH STEP – Run the combination, estimate the uncertainties

At this point, two (or more) covariance scaling parameters should be available (let’s say that the scaling for the igs global h-files is 0.4, while the scaling for your h-files is 1.0). To combine the files, you will repeat mostly the procedure of the FIFTH STEP (note difference in the .gdl file).

a) Generate the input file list (*.gdl) for the h-files

```
../glbf/h0101081200_igs1.glx 0.400
../glbf/h0101081200_igs2.glx 0.400
```

```
../glbf/h0101081200_igs3.glx 0.400
../glbf/h0101081200_nbay.glx 1.000
```

b) Run *globk*
c) Get the combination *prefit chi**2* from the *.prt file. If the chi**2 is equal (or very close) to 1, you have the right scaling for the covariances and the uncertainties of your solution are a good estimate of the “real” ones
d) If the chi**2 is much different from 1, then repeat points a), b) and c) tuning (increasing or decresing) the scaling for your h-files until you get a chi**2 close to 1.