

Rapid prototyping of GLOBK solutions

Lecture 11

Thomas Herring,
tah@mit.edu

Overview

- Processing roles:
 - GAMIT: Used to process phase data normally in 24-hour sessions. Station positions, atmospheric delay, phase ambiguity always estimated; often earth orientation and satellite orbit parameters estimated. Generally <50 stations in one analysis; up to 5550 parameters
 - GLOBK: Multiple roles in GPS (and other geodetic system) processing:
 - **Combination of subnets** from GAMIT for one-day (used when more than 50-stations processed). Fast and only done once.
 - **Time series analysis** in which many daily solutions are process individually to generate time series. Fast, often done many times. Can be run in parallel.
 - **Velocity solution combination** where many years of data are combined in a forward Kalman filter run. Often many thousands of parameters and many thousands of days of data. Can be very slow (24-48 hour runs) and prone to problems if there are “bad” data or un-modeled discontinuities. The prototyping is aimed at testing these types of solutions (and replace them in some cases).

GLOBK Velocity Solutions

- The aim of these solutions is to combined many years of data to generate position, velocity, offset, and postseismic parameter estimates. Not uncommon to have 10000 parameters in these solutions.
- Input requirements for these solutions:
 - Apriori coordinate and velocity file. Used as a check on positions in daily solutions (for editing of bad solutions) and adjustments are apriori values (apriori sigmas are for these values)
 - Earthquake file which specifies when earthquakes, discontinuities, and miss-named stations affect solution. Critical that this file correctly describe data.
 - Process noise parameters for each station. Critical for generating realistic standard deviations for the velocity estimates.

Velocity Solution Strategies

- In general careful setup (i.e., correct apriori coordinate, earthquake file and process noise files) is needed since each run that corrects a problem can take several days. In correct solutions may not complete correctly.
- Previous methods for constructing these solutions:
 - Define a core-set of sites (usually 20-200 sites) where the solution runs quickly. Test files on this solutions and use the coordinate/velocity estimates to form the reference frame for time series generation.
 - Time series using these reference frame sites and then test (RMS scatter, discontinuity tests) to form a more complete earthquake and apriori coordinate/velocity files.
 - Steps above are repeated, usually increasing number of stations until solution is complete. As new stations are added missed discontinuities and bad process noise models can cause problems.

Velocity strategies

- Other methods that are used in increase speed are:
 - Pre-combine daily solutions into weekly to monthly solutions and use these combined solutions in the velocity solutions. There are many advantages to this approach:
 - Runs are much faster. Each processing step takes about the same time with the monthly as a daily file but there are 30 fewer files so 30 times faster.
 - Numerical rounding errors are much better when monthlies are used
 - New MIDP output option refers the solutions to the middles of the month. (Earlier versions used last day of month as reference time, natural time for a sequential Kalman filter.
 - Random walk process noise models correct when velocity NOT estimated in combinations
 - Run decimated solutions (e.g., one day per week). Works fine and changing start day does not have large effect due to correlated noise models. Care needed when different start day results are combined to avoid white noise sigma reduction.

Prototyping tools

- There are two new programs that are used for prototyping solutions are:
 - **tscon** which converts a variety of data formats into the PBO .pos format while allowing a new reference frame realization using techniques similar to GLORG stabilization. Stabilization can be used to test selection of reference sites.
 - **tsfit** which fits time series with a variety of models some of which can be specified in a GLOBK .eq file format. tsfit also outputs a globk a priori coordinate files. Use of realistic sigma option here and sh_gen_stats allows process noise to be set for globk (site dependent random walk variances)
- There is also an additional program **xyzsave** that can be used to generate XYZ files for use in tscon when the pbo output option was not specified in the original globk runs. It is highly recommended that the pbo option be used in all output from globk and glred. The somewhat new program, tssum can be used to extract and append pbo time series files from globk and glred output files (normally .org files).

Prototyping concept

- The general idea of the solution prototyping is to generate an earthquake file and a list of stabilization sites that can be used in both velocity and time series analysis in GLOBK and GLRED runs. Tsfits can also be used to generate a priori coordinate files for use in tscon and globk/glred.
- Both tscon and tsfits can read standard globk earthquake and a priori coordinate files (include EXTENDED entries). The programs do not manipulate covariance matrices and so it is assumed that an initial time-series solution exists with stabilized coordinates (i.e., the output of a glred run with stabilization).

Process

- Basic processing ordering:
 - First run glred to generate time series with the pbo output option set. This solution might for example use ITRF05 sites for stabilization, or for more regionally focused networks, globk might be used for a velocity solution and the good sites from this analysis used as the stabilization sites in the glred run.
 - (There is a "catch-22" here in that knowing which sites are well behaved requires generating time series first and so these approaches tend to be iterative with the list of good sites being determined from their behavior in different analyses.)
 - Once the initial time-series are generated, **tscon** can be used to generate new time-series with different stabilization sites and with different apriori coordinate models than those used in the original run.
 - Analyses of these time series can be carried out using **tsfit** to estimate new apriori coordinate models and additional parameters associated with seasonal variations, earthquake post-seismic deformations and jumps in the time series due to antenna and the instrument changes and earthquakes.

Basic Processing (cont.)

- The statistics of the fits to the time series are generated by `tsfit` and these can be used to judge the quality of the analyses. The summary file output by `tsfit` can be used in the version of `sh_gen_stats` with the `-ts` option.
- Removal of outlier data using an n-sigma condition can also be performed by `tfsit` with the output in standard eq-file format.
- The new coordinate apriori files from `tsfit` can be used in a new reference frame realization using `tscon`. The newly generated time series can be used to refine the analysis more using `tsfit`. Iterating the reference frame in this manner could lead to some systematic behaviors and it is ideally best to generate the reference frame with a `globk` solution.

Prototyping output

- At the completion of the tscon/tsfit process, there should be available an earthquake file that contains earthquakes, renames for offsets and for time series editing (renames to XPS names), and an apriori coordinate file with optional EXTENDED entries that should provide a good match to the behavior of the time series.
- A refined list of reference frame sites and process noise models may also have been generated (sh_gen_stats).
- The earthquake and apriori file and other information can be used in an updated globk velocity solution or in glred repeatability time series run. These final globk and glred analyses should run with no major problems and would be used to generate final results.

tsfit

- **tsfit** is a program to fit PBO-formatted times series using a globk earthquake file input and other optional parameters (such as periodic signals). PBO format time series are generated using the pbo output option in glorg and program tssum to extract the time series. **tssum** allows incremental updates of time series rather the full re-generation used by ensum and multibase.
- For the prototyping role, the most important commands are eq_file (input) and out_aprf and rep_edits (outputs).
- The command line for tsfit is:
 - tsfit <command file> <summary file> <list of files/file containing list>

tsfit commands

- EQ_FILE <File Name>
 - Name of standard globk earthquake file. Command may be used multiple times as in the latest version of globk.
- OUT_APRF <file name>
 - Specifies name of a globk apriori coordinate file to be generated from the fits. This file contains EXTENDED entries if needed and can be used directly in globk or tscon.
- REP_EDITS <rename file>
 - Set to report edits to file <rename file>. Edit lines start with R. The rename file if given will contain globk rename to _XPS lines.
- REAL_SIGMA
 - Apply the tsview/ensum realistic sigma algorithm to generate sigmas that account for temporal correlations in the data. This option is needed to use sh_gen_stats

Other tsfit commands

- PERIODIC <Period (days)>
 - Estimates Cosine and Sine terms with Period. This command may be issued multiple times to estimate signals with different periods.
- DETROOT <det_root>
 - String to be used at the start of the site dependent parameter estimate files. Each site generates its own file. Default is ts_. NONE generates no files
- VELFILE <vel file name>
 - Name of the output file containing velocity estimates in the standard globk velocity file format.
- NSIGMA <nsigma limit>
 - Edit time series based on a n-sigma condition.

Other tsfit commands

- `MAX_SIGMA <Sig N> <Sig E> <Sig U> meters`
 - Allows limit to be set on sigma of data included in the solutions.
 - Default values are 0.1 meters in all three coordinates.
- `TIME_RANGE <Start Date> <End Date>`
 - Allows time range of data to be processed to be specified. Dates are Year Mon Day Hr Min. End date is optional.
- `OUT_EQROOT <root for Earthquake files> <out days>`
 - Specifies the root part of the name for earthquake estimates outputs. The outputs are in globk .vel file format and so can be used with `sh_plotvel` and `velview`. The outputs are coseismic offset and log and exponential coefficient estimates. If the `<out days>` argument is included the total post-seismic motion is computed that many days after each of the earthquakes. If exponential and log terms are estimated for the same event (same `eq_def` code) then they are summed and correlations accounted for in computing the sigmas of the total motion. Output file format is .vel file format.

tscon

- The program `tscon` converts timeseries from Reason/JPL/SIO XYZ files and SCEC CSV format to PBO time series format and optionally re-realizes the reference frame used to generate the time series for the format above and standard PBO time series files generated with `tssum`.
- The program assumes that the position time series are reported at a regular 1-day interval. This is the normal timing used in `gamit` for 24-hr sessions of data.
- The command line for `tscon` is:
 - `tscon <dir> <prod_id> <cmd file> <XYZ/PBO files/file with list>`

tscon commands

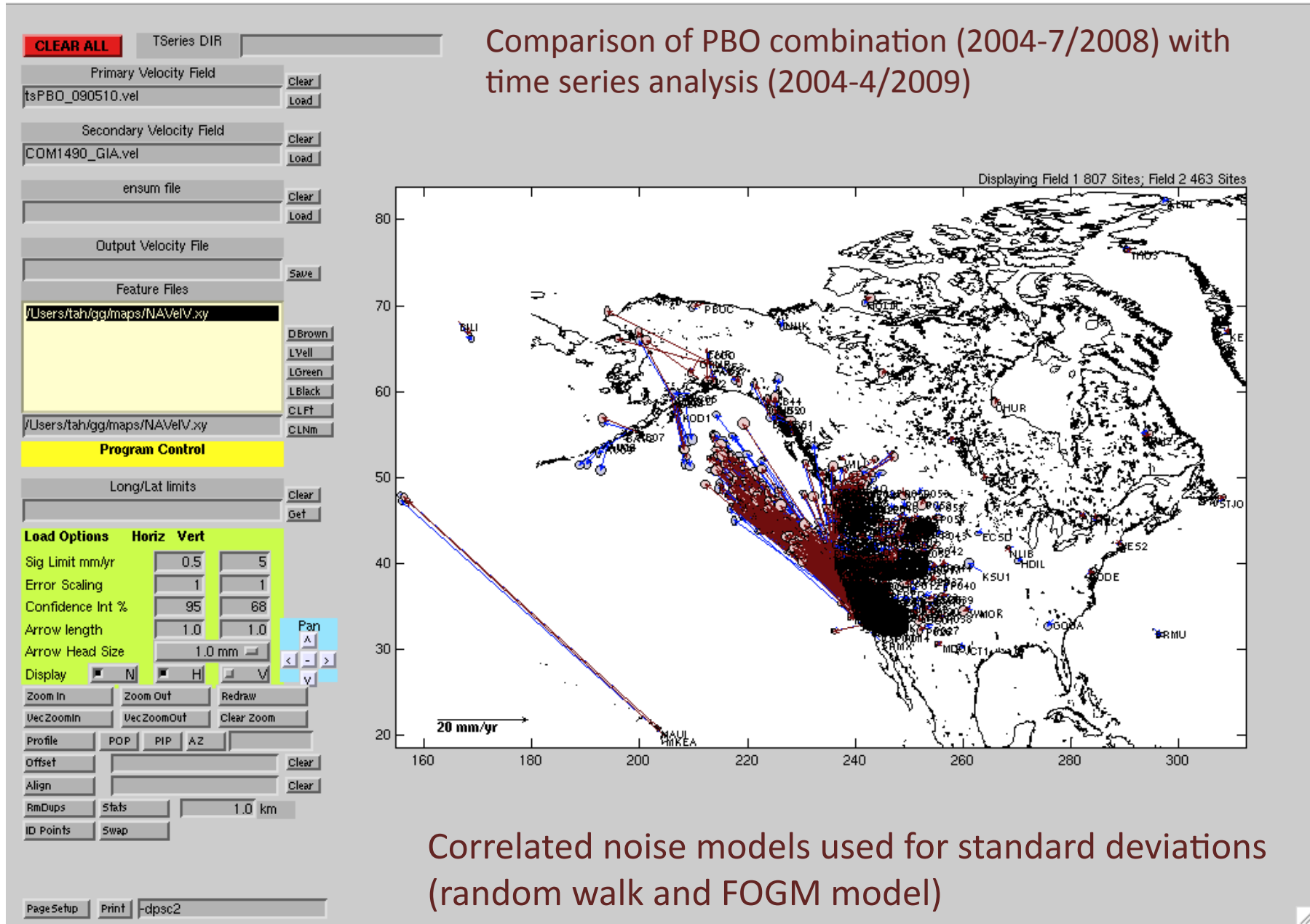
- Summary of commands are:
 - eq_file <file name> (maybe issued multiple times)
 - apr_file <apriori coordinate file> (may be issued multiple times)
 - stab_site <list of stabilization sites> (multiple times)
 - pos_org <xtran> <ytran> <ztran> <xrot> <yrot> <zrot> <scale>
 - stab_ite [# iterations] [Site Relative weight] [n-sigma]
 - stab_min [dHsig min pos] [dNEsig min pos]
 - cnd_hgtv [Height variance] [Sigma ratio]
 - time_range [Start YY,MM,DD,HR,MIN] [End YY,MM,DD,HR,MIN]
- These commands mimic the glorg equivalent commands and operate in a very similar way. There are some small differences because tscon starts with frame realized time series.

Comparison of tsfit and globk

- tsfit uses time series analysis and is very fast. In many cases, the results obtained with method are very similar to the more rigorous globk results. Works well when there is a high quality reference network of sites.
- The comparison here used a rigorous combination using a subset of “good” sites. The reference frame of this solution is established and these sites are then used for frame realization of the entire network of sites in time series analysis mode.
- The time series of each site is then analyzed separately and can have very complex process noise models. Station analyses can be done parallel processing mode.
- Example: PBO analysis. 230 sites used as frame definition sites. Translation, rotation and scale rate used to align to SNARF 1.0. (230 site analysis is fairly rapid -- runtime cube of number of number of sites.)

Velocity comparison

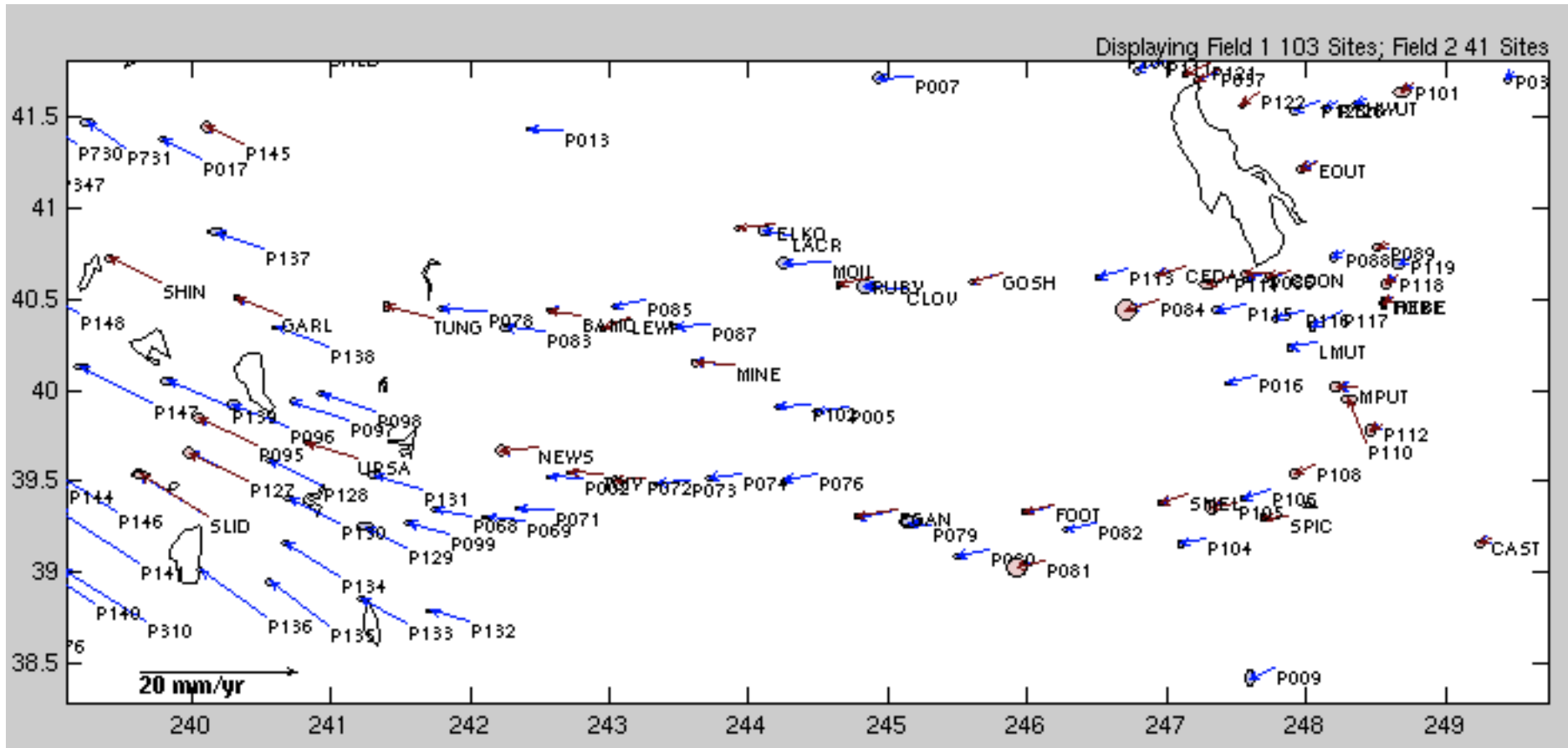
Comparison of PBO combination (2004-7/2008) with time series analysis (2004-4/2009)



Correlated noise models used for standard deviations (random walk and FOGM model)

Velocity Comparison and Zoom

Standard deviations < 0.5 mm/yr shown

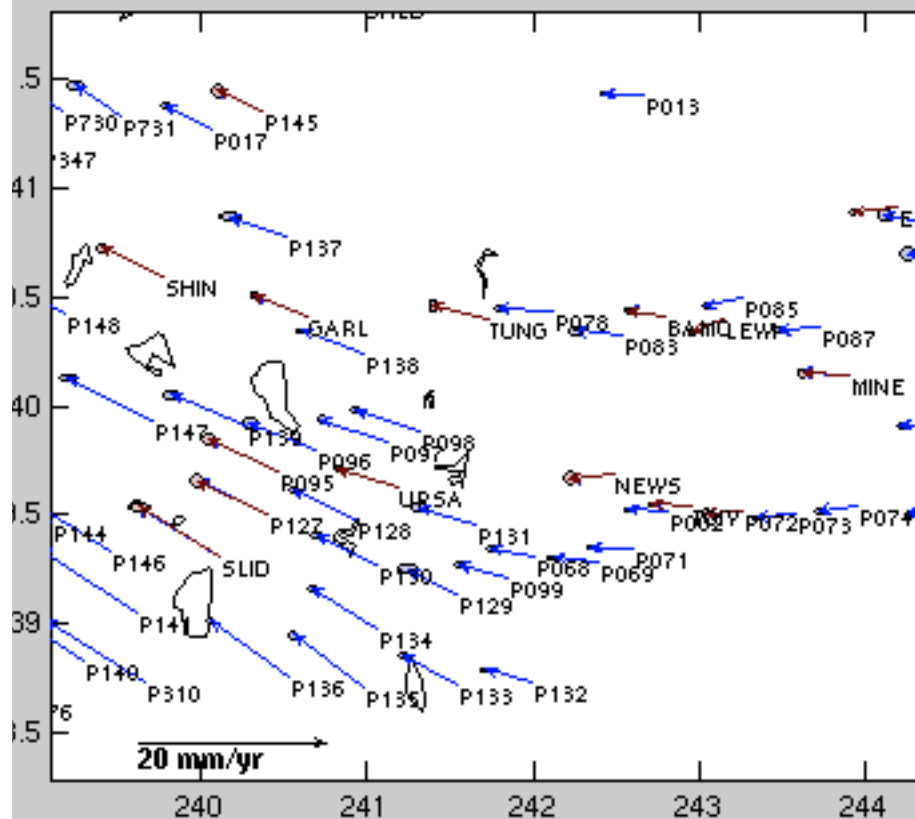


Since time series solution was more recently constructed (getting ready for next rigorous combination), it has more stations.

Alignment and RMS difference between the field.

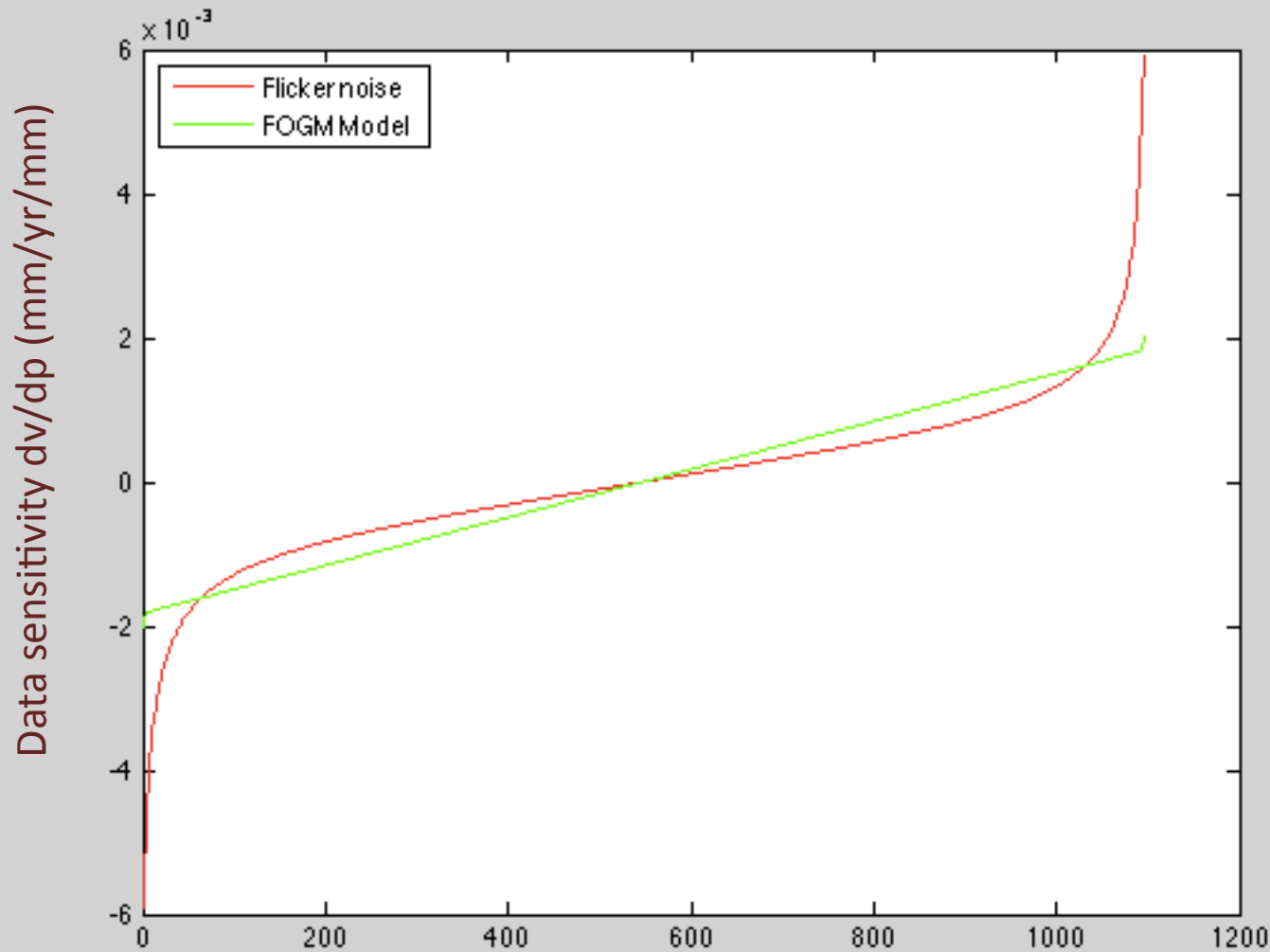
Alignment Results from 36 sites (right-click to remove)

| Comp | Wmean (mm) | WRMS | NRMS | | | | |
|-------------------|--------------|---------|---------|------|-------|------|--|
| North | 0.02 +- 0.02 | 0.13 | 0.686 | | | | |
| East | 0.02 +- 0.02 | 0.11 | 0.543 | | | | |
| Height | 0.17 +- 0.12 | 0.70 | 0.996 | | | | |
| Sites | NRes +- | ERes +- | URes +- | | | | |
| BAMO_GPS-BAMO_GPS | -0.06 | 0.17 | -0.03 | 0.19 | 0.15 | 0.39 | |
| CEDA_GPS-CEDA_GPS | -0.00 | 0.16 | 0.10 | 0.15 | -0.50 | 0.78 | |
| COOH_GPS-COOH_GPS | 0.35 | 0.22 | 0.20 | 0.22 | -0.43 | 1.04 | |
| EGAN_GPS-EGAN_GPS | 0.15 | 0.16 | -0.02 | 0.14 | -0.45 | 0.68 | |
| ELKO_GPS-ELKO_GPS | 0.03 | 0.14 | -0.02 | 0.17 | -0.71 | 0.81 | |
| EOUT_GPS-EOUT_GPS | -0.28 | 0.16 | -0.07 | 0.27 | 0.34 | 0.61 | |
| FOOT_GPS-FOOT_GPS | 0.06 | 0.14 | 0.08 | 0.17 | -0.29 | 0.76 | |
| HEBE_GPS-FORE_GPS | -0.17 | 0.22 | 0.03 | 0.25 | 4.09 | 0.96 | |
| GARL_GPS-GARL_GPS | 0.05 | 0.13 | -0.07 | 0.17 | -1.02 | 1.17 | |
| GOSH_GPS-GOSH_GPS | 0.04 | 0.16 | 0.04 | 0.14 | -0.25 | 0.81 | |
| HEBE_GPS-HEBE_GPS | -0.03 | 0.32 | 0.01 | 0.22 | 1.11 | 2.51 | |
| LEWI_GPS-LEWI_GPS | 0.06 | 0.16 | -0.03 | 0.16 | 0.37 | 1.11 | |
| MINE_GPS-MINE_GPS | -0.13 | 0.17 | 0.19 | 0.22 | 0.46 | 0.75 | |
| MPUT_GPS-MPUT_GPS | -0.01 | 0.26 | -0.44 | 0.38 | 0.82 | 1.13 | |
| NEWS_GPS-NEWS_GPS | 0.15 | 0.32 | -0.04 | 0.31 | 0.82 | 1.21 | |
| P072_GPS-P072_GPS | 0.01 | 0.22 | 0.01 | 0.21 | 0.18 | 0.51 | |
| P081_GPS-P081_GPS | -0.12 | 0.50 | 0.36 | 0.52 | 0.21 | 0.69 | |
| P084_GPS-P084_GPS | -0.12 | 0.51 | -0.44 | 0.51 | 0.77 | 0.70 | |
| P086_GPS-P086_GPS | 0.49 | 0.30 | -0.08 | 0.26 | 0.05 | 0.59 | |
| P089_GPS-P089_GPS | 0.17 | 0.19 | -0.21 | 0.24 | 0.01 | 0.60 | |
| P095_GPS-P095_GPS | -0.03 | 0.27 | -0.02 | 0.45 | 0.02 | 0.96 | |
| P105_GPS-P105_GPS | 0.18 | 0.46 | -0.13 | 0.33 | 3.01 | 2.26 | |
| P108_GPS-P108_GPS | 0.02 | 0.27 | -0.05 | 0.27 | -0.66 | 0.70 | |
| P112_GPS-P112_GPS | 0.13 | 0.29 | -0.26 | 0.25 | -0.87 | 0.78 | |
| P114_GPS-P114_GPS | 0.13 | 0.22 | 0.07 | 0.45 | -0.07 | 0.79 | |
| P118_GPS-P118_GPS | -0.23 | 0.32 | -0.21 | 0.28 | -0.71 | 0.71 | |
| P127_GPS-P127_GPS | -0.29 | 0.30 | 0.03 | 0.33 | -0.44 | 0.89 | |
| P145_GPS-P145_GPS | -0.10 | 0.26 | 0.19 | 0.31 | 0.27 | 0.69 | |
| RUBY_GPS-RUBY_GPS | 0.09 | 0.19 | 0.06 | 0.17 | -0.97 | 0.76 | |
| SHLN_GPS-SHLN_GPS | -0.09 | 0.17 | -0.06 | 0.18 | -0.72 | 0.84 | |
| SLID_GPS-SLID_GPS | -0.06 | 0.27 | 0.34 | 0.57 | -0.80 | 0.77 | |
| SMEL_GPS-SMEL_GPS | 0.10 | 0.16 | 0.07 | 0.14 | -0.22 | 0.70 | |
| SPIC_GPS-SPIC_GPS | 0.01 | 0.14 | 0.10 | 0.17 | -0.05 | 0.47 | |
| TOIY_GPS-TOIY_GPS | -0.12 | 0.13 | -0.11 | 0.17 | 0.35 | 0.36 | |
| TUNG_GPS-TUNG_GPS | -0.07 | 0.25 | -0.08 | 0.16 | -0.69 | 0.99 | |
| UPSA_GPS-UPSA_GPS | -0.03 | 0.14 | -0.04 | 0.17 | -0.18 | 0.61 | |



Noise model effect

Red: Flicker noise 1 mm, White noise 2mm; Green First Gauss Markov 1 day correlation time



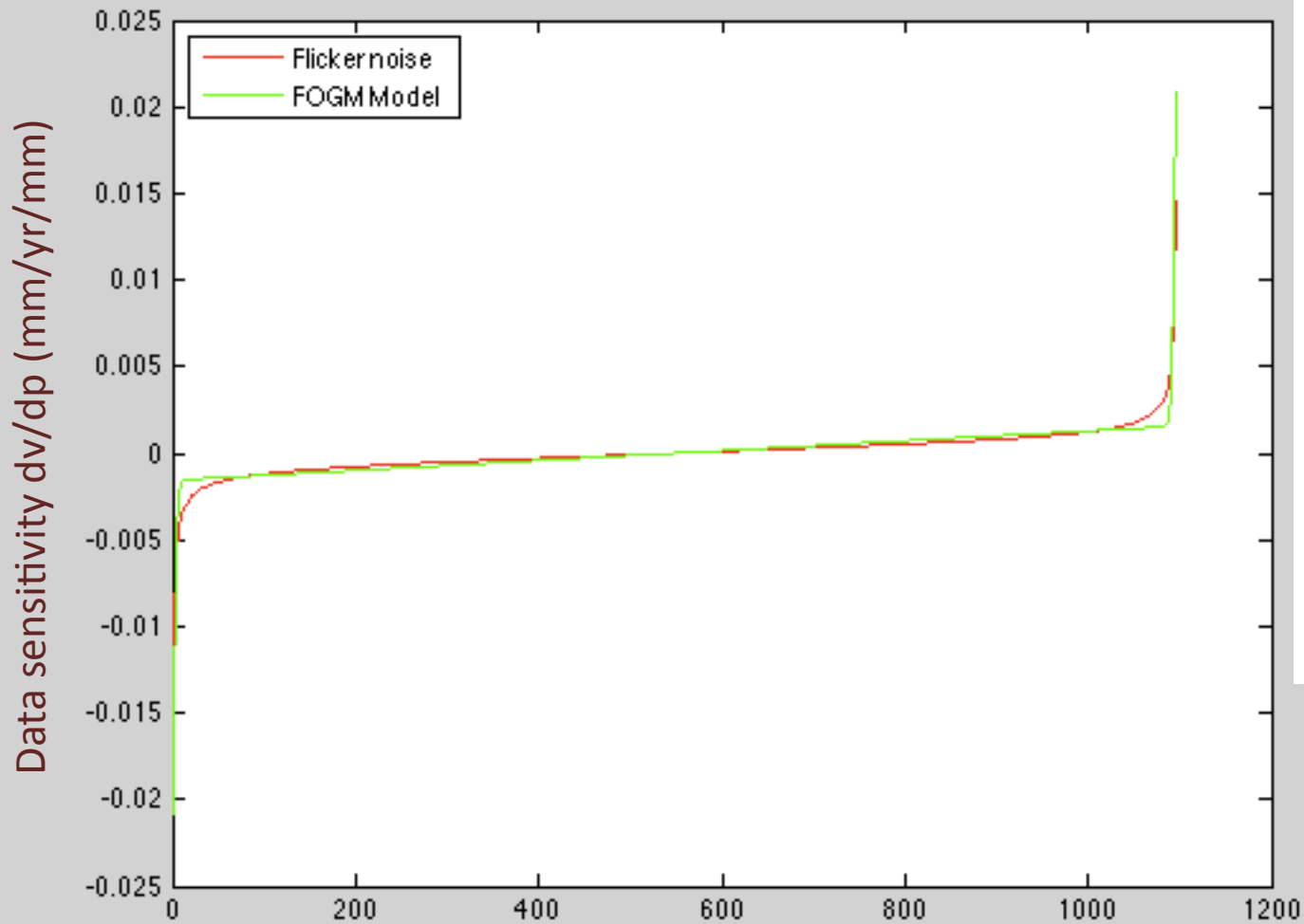
Flicker Noise Sig 1.00
WN 2.00 mm
Offset and Rate Sig
0.51 mm
0.27 mm/yr

FOGM Noise Sig 1.00
tau 1.0 WN 2.00
mm
Offset and Rate Sig
0.15 mm
0.09 mm/yr

3-years of data

Different process noise

Flicker noise 2mm, White noise 1 mm; First order Gauss Markov correlation time 32 days



Flicker Noise Sig 2.00 WN
1.00 mm

Offset and Rate Sig 0.99
mm
0.51 mm/yr

FOGM Noise Sig 1.00 tau
1.0 WN 2.00

Offset and Rate Sig 0.91
mm
0.51 mm/yr

3-years of data

Conclusion

- Most critical steps to creating velocity field:
 - Correct frame realization -- probably best done with a small group of well characterized sites
 - Time series generated from this frame to allow process noise models to be developed and anomalous site identified.
 - Determining the correct standard deviations for the velocities is the most difficult and important step.
 - With modern data and a well defined frame, the method of the constructing the velocity is probably not that critical i.e., tsfit and globk generate very similar results when globk is used to generate the reference frame solution.
- The tsfit and tscon combination allows apriori coordinate and earthquake files to be tested before running large multi-hour globk runs.
- tsview and velview are interactive matlab based tools that allow viewing of timeseries and velocity fields
<http://geoweb.mit.edu/~tah/GGMatlab>