Corrections to the pipeline

These notes are based on the short talk APJ presented to the SUSI meeting on 15/01/2007. APJ 22 Jan 2007.

Several small changes have been made to the pipeline and the modified idl files were uploaded to the ~susi/code/analysis directory on chef (in Sydney) at 1800hrs EASummerT on 070119 and on arthur (at Narrabri) at 1330hrs EASummerT on 070120.

The modified idl files are:

check_v2.script, extract_wobble.pro & calc_centroids.pro

I have created a new subdirectory, oldcode, and moved all the old code into it. Everything appears to work correctly, but if I've overlooked something you could always log in as user=susi and move the old files back, or better still copy the file you want into the directory where you opened IDL. IDL looks in this 'local' directory first for programs.

The changes are:

check_v2.script:

1. added stots keyword as default in call to extract_wobble.

This will force extract_wobble to correlate the fringe file scintillation with wobble image fluxes to align them in time.

extract_wobble.pro:

1. corrected type=1 error to type=cent_type in call to calc_centroids for s_ims.

This makes the calc_centroids loop iterate more than once.

2. Added keywords correlation, pltcorr & nc_width.

These are used to plot the correlation to screen and to investigate abnormal correlations. For normal data reduction you don't need to know about these.

calc_centroids.pro:

1. MJI moved the add_noise & random number generator bits out of the loop.
2. APJ increased the default number of iterations to 8 (from 3) to allow convergence of the wobble-image centroid values.

What do you need to do?

Run the pipeline as normal, these changes are largely internal.
You will notice the IDL window now displays the following lines for each run:

"Maximum correlation: 0.470423
Clock on gaheris leads by: -0.00600000 seconds..."

…and then pauses for a moment while it completes the extra iterations.
And then, for each run, a new plot window opens to display a plot of the correlation. Just note the appearance of the correlation plot. Obviously you want to see a significant peak. If so, all is OK. IDL continues processing your data.

If the correlation values or times change a lot over the night's data or you see double-peaked or otherwise odd correlation plots then there may be a synchronisation problem. See the document WobFr_Correlation.doc (or the SUSI web pages) for what to do if this is the case (this is where the keyword nc_width comes in).

I have seen occasions when a single correlation plot is double peaked but otherwise the night's data is OK. In this case perhaps it's simplest to flag the run as dubious (the run number appears on the plot to identify it) and exclude it from further processing.

NOTE: [added by APJ on 070205] The 'double-peaked' correlations are due to noise spikes. This is now solved. See WobFr_Correlation.doc for details.

What does the new pipeline do?

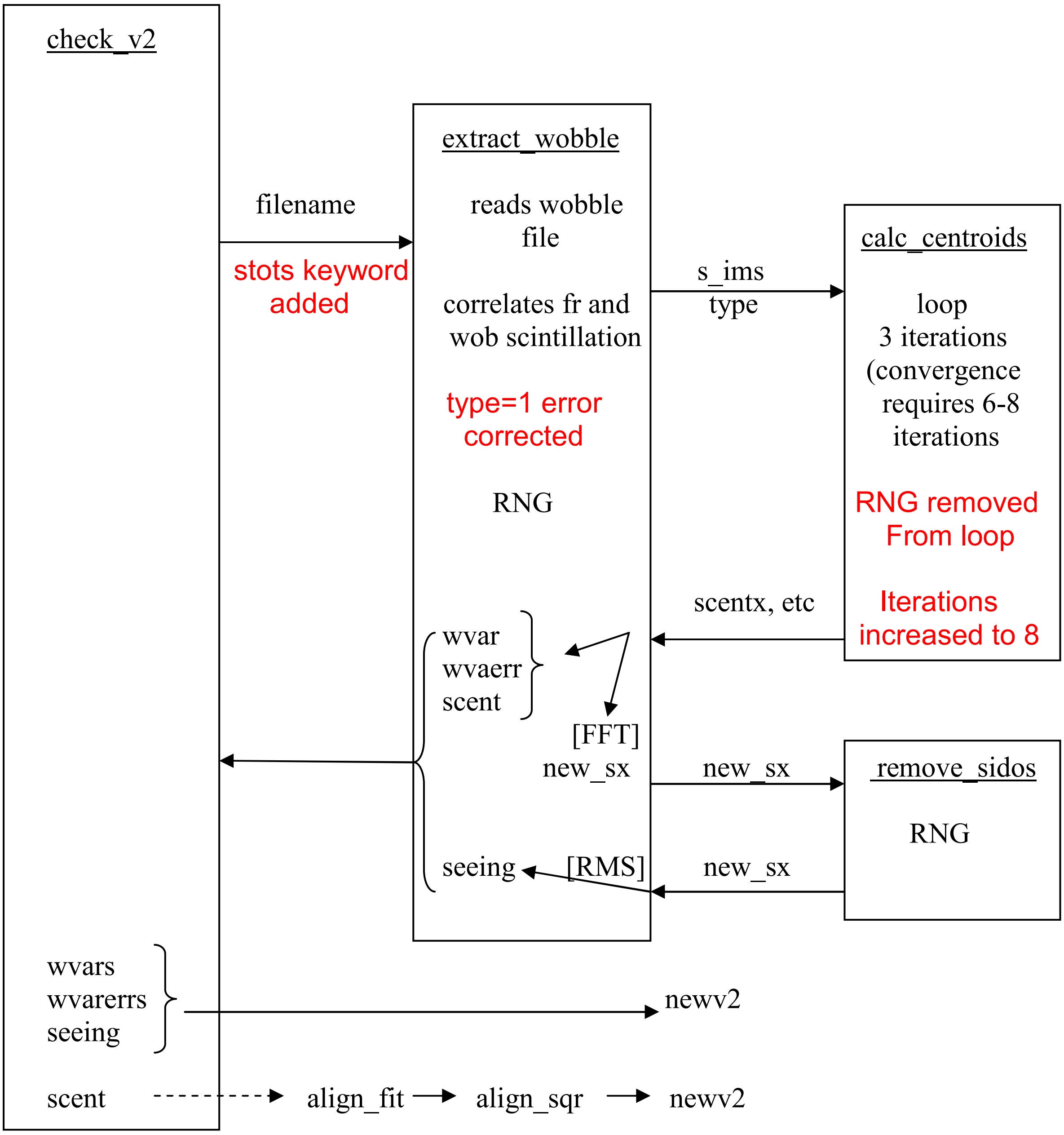

Fig.1 A rough schematic of relevant parts of the pipeline. Corrections are noted in red. RNG = random number generator. newv2 is the multi-parameter seeing corrected visibility output by the pipeline. This is the quantity usually used for further analysis.

The new pipeline works as follows:

The main check_v2 script calls extract_wobble with keywords stots. Extract_wobble reads the wobble file and correlates the fringe scintillation with the wobble image fluxes to synchronise the data. Extract_wobble calls calc_centroids (to calculate the centroids of all the wobble images) with the keywords s_ims (for south wobble images) type=cent_type (cent_type=2 for standard pipeline processing). North and dark wobble images are sent separately. Calc_centroids windows the image and iterates 8 times which allows convergence of the centroid value in most cases. If convergence hasn't happened the alignment is so bad you probably wouldn't want to use the data anyway. Calc_centroids returns north & south centroids in x & y for star & dark frames, i.e. 8 parameters are returned. Extract_wobble now finds the power spectrum of the centroid values, calls remove_sidos to remove siderostat oscillations, and calculates the 'seeing' from the normalised RMS power. Meanwhile, extract_wobble calculates the 'wobble variance', wvar, and its error, wvarerr. Now, wvar = sum(south x-centroid variance, south y-centroid variance, north x-centroid variance, north y-centroid variance). Wvar and seeing are passed back to check_v2 and directly into the calculation for newv2. Hence, because the centroid values are systematically increasing (see Fig. 2), wvar & seeing and therefore newv2 will systematically increase. Meanwhile, north & south centroid values are passed from extract_wobble to check_v2 where they contribute, in complicated spaghetti-like code, to an 'alignment fit'. This also is used in the calculation of newv2, but I can't say if there is any systematic effect on newv2.

Random Number Generators (RNG) and Fitting Procedures
There are a few random number generators and Monte Carlo-Markov Chain fitting procedures in the pipeline so the final newv2 values output are not identical on each invocation of the pipeline although the differences are buried in the errors, as they should be. The extract_wobble RNG simulates dark wobble images if they are not otherwise available. The scheduler should ensure they are. The calc_centroids RNG adds, e.g., photon noise effects to dark wobble frames. I don't know what the remove_sidos RNG does.

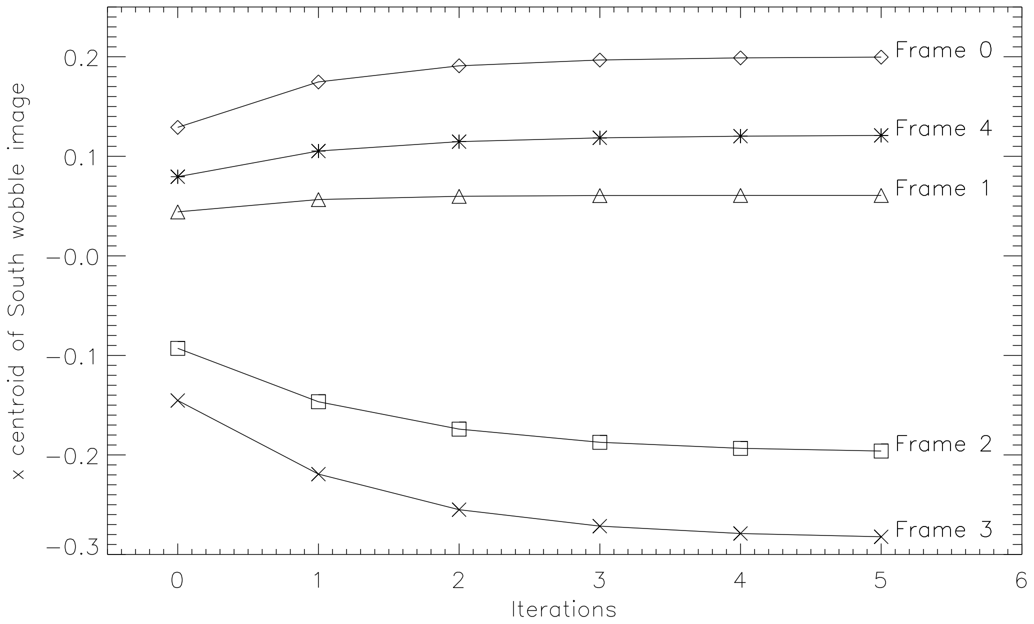

Fig. 2 Convergence of the calc_centroids.pro loop. This example is for the x centroid of the first 5 frames of the south wobble data stream. Note that abs(centroid) systematically increases with the number of iterations. The old pipeline iterated only 3 times (up to 2). The new pipeline iterates 8 times. Note also that (final_value/initial_value) increases for increasing abs(initial_value).

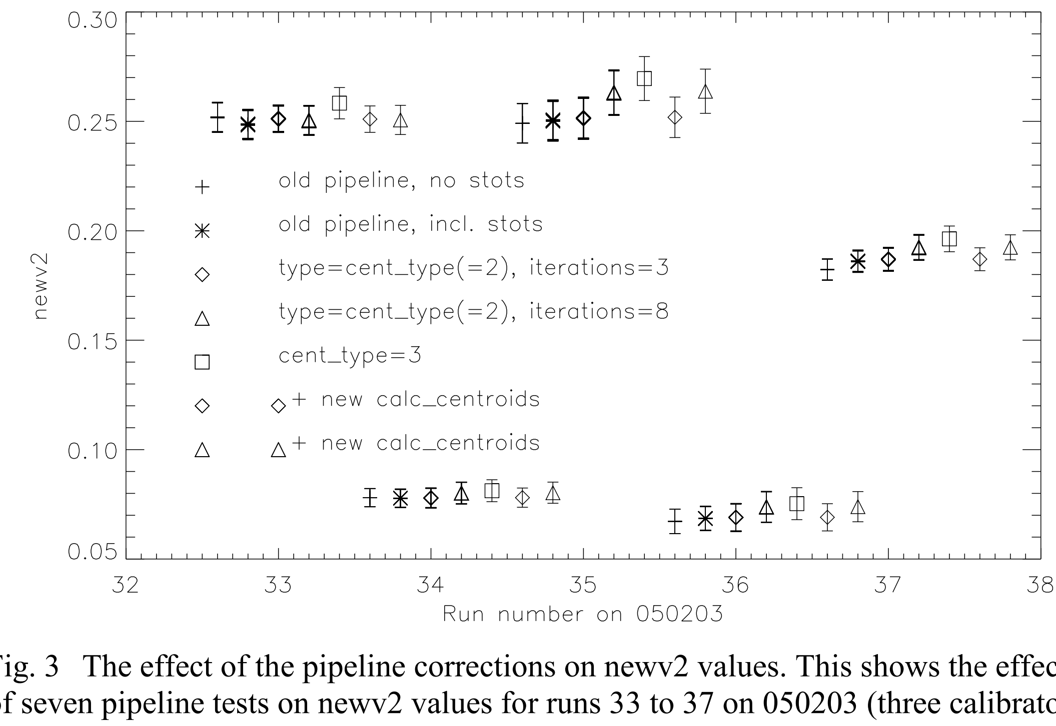

Fig. 3 The effect of the pipeline corrections on newv2 values. This shows the effect of seven pipeline tests on newv2 values for runs 33 to 37 on 050203 (three calibrator and two source observations).

The effect of the pipeline changes

Fig. 3 shows the effect of the pipeline changes on the newv2 values. Plus signs are for the old pipeline (and with the stots keyword turned off). The second diamonds are for newv2 with all errors corrected but still using too few iterations in calc_centroids. The second triangles are for the new pipeline, including the wobble-fringe correlation, all errors corrected and convergence in calc_centroids. Newv2 values from the new pipeline are systematically larger. The squares show results from an alternative, but very slow, centroiding algorithm. It uses a Monte Carlo, Markov Chain fitting algorithm and can take up to two minutes to process each fringe file.

While the newv2 values from the new pipeline are systematically higher the final calibrated visibilities and diameters are relatively unaffected because each newv2 is increased by the same amount, approximately 5%. Fig 4 shows UD diameters for ℓ Car on 050203 with the data processed in four different ways. Test 0 used the old pipeline. Test 2 used the new pipeline. For this data there is no significant difference in UD diameter although the errors are approximately 12% larger. Test 3 used the slow MCMC method and although its errors are slightly smaller here the extra time involved in processing may not be worth it.

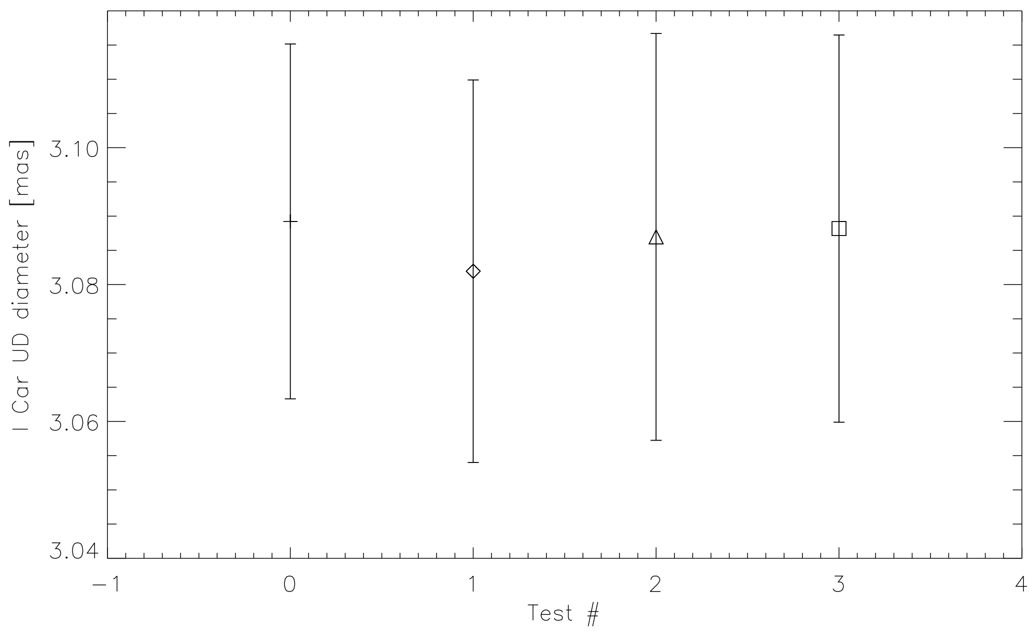

Fig 4 The effect of pipeline changes on UD diameters for runs 13 to 32 (includes 9 source runs) for ℓ Car on 050203.
Test 0: The old pipeline, corresponding to plus signs in Fig. 3.
Test 1: The new pipeline but with only 3 iterations in calc_centroids, corresponds to second diamonds in Fig. 3.
Test 2: as Test 1 but with 8 iterations, corresponds to second triangles in Fig.3, i.e. this is using the new pipeline.
Test 3: as Test 2 with cent_type=3, corresponds to second squares in Fig.3.

[Note: UD diameters in Fig. 4 are
0 3.089± 0.026
1 3.082± 0.028
2 3.087± 0.030
3 3.088± 0.028]