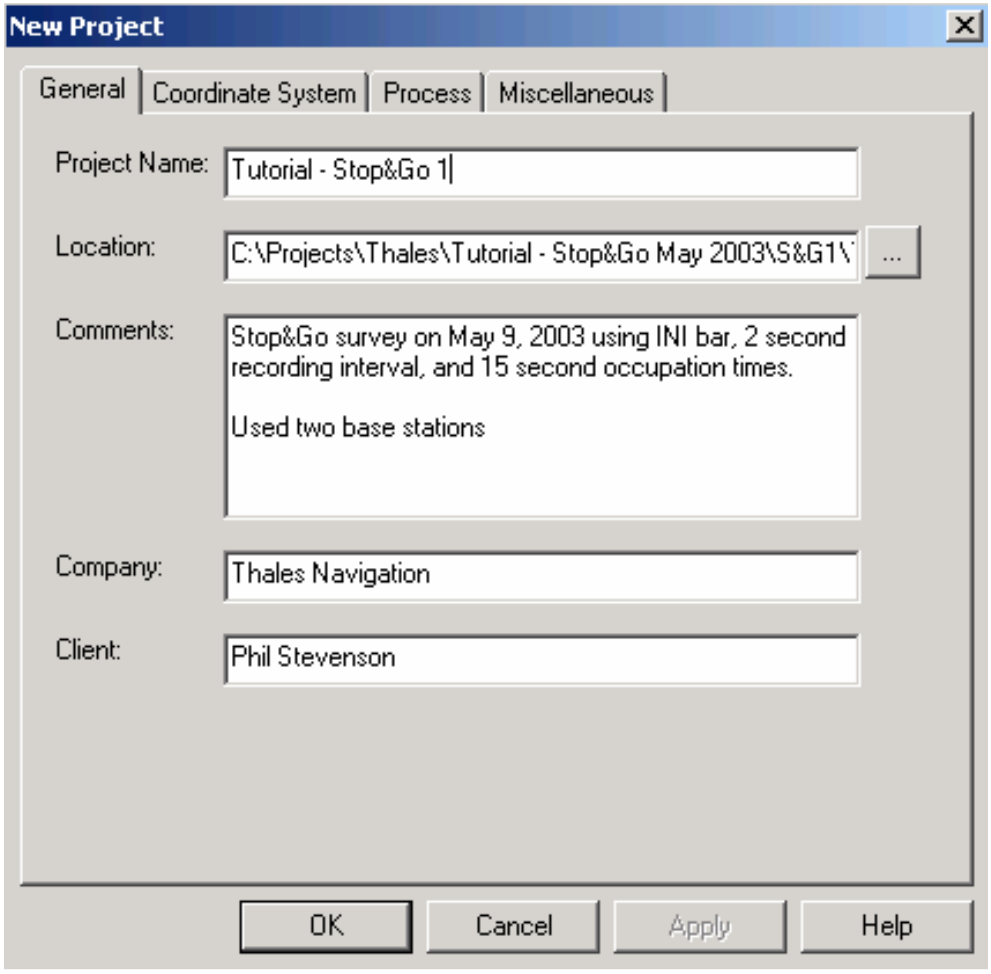
Ashtech Solutions: Data Cleaning Part Two

In part one of Data Cleaning the project was based on a static survey. It worked so well that not all of the data cleaning tools available in Ashtech Solutions were needed. In part two the field data will be processed using Stop&Go procedures, thus permitting the use of some additional tools.

Cleaning up field data

The static survey from part one can be used as a standard by which the Stop&Go can be evaluated. Recall in part one that the coordinates were converted to state plane coordinates in feet in order to make some comparisons with the Stop&Go work.



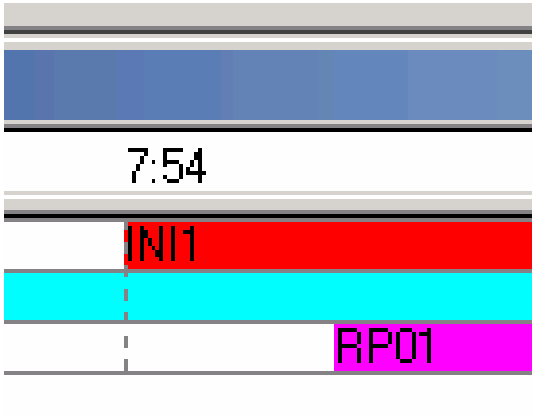
Screen shot 1

In part two the Stop&Go project is set up as an independent survey that will use the same control point. Notice in screen shot 1 that the new project name is Tutorial – Stop&Go 1. A new folder was also created for this specific project. A look at the comments will give some additional information about this project.

The coordinate systems and the units of measure will be the same initial settings used in the static survey project in part one – geodetic coordinates based on the North American Datum of 1983 and the Geoid99 model for the USA.

After loading the files into the project it is time to check for field blunders.

A glance at the time view in screen shot 2 reveals that the rover at INI1 was started before the base at RP01. It may only be 30 seconds of data but remember the initialization is only 300 seconds. Because of this, it is important to start the base first. Start the rover only after the base is running.



Screen shot 2

Fortunately Ashtech Solutions has the tools that may help to overcome this problem, providing of course that there is enough data to initialize the project. With this in mind an attempt will be made to process the data before throwing it out and starting the day over.

A look at the Antenna Height column in the Survey Project Manager – [Workbook] as illustrated in screen shot 3 shows the INI1 and the rover heights to be the same. While this may be possible it is not likely that they

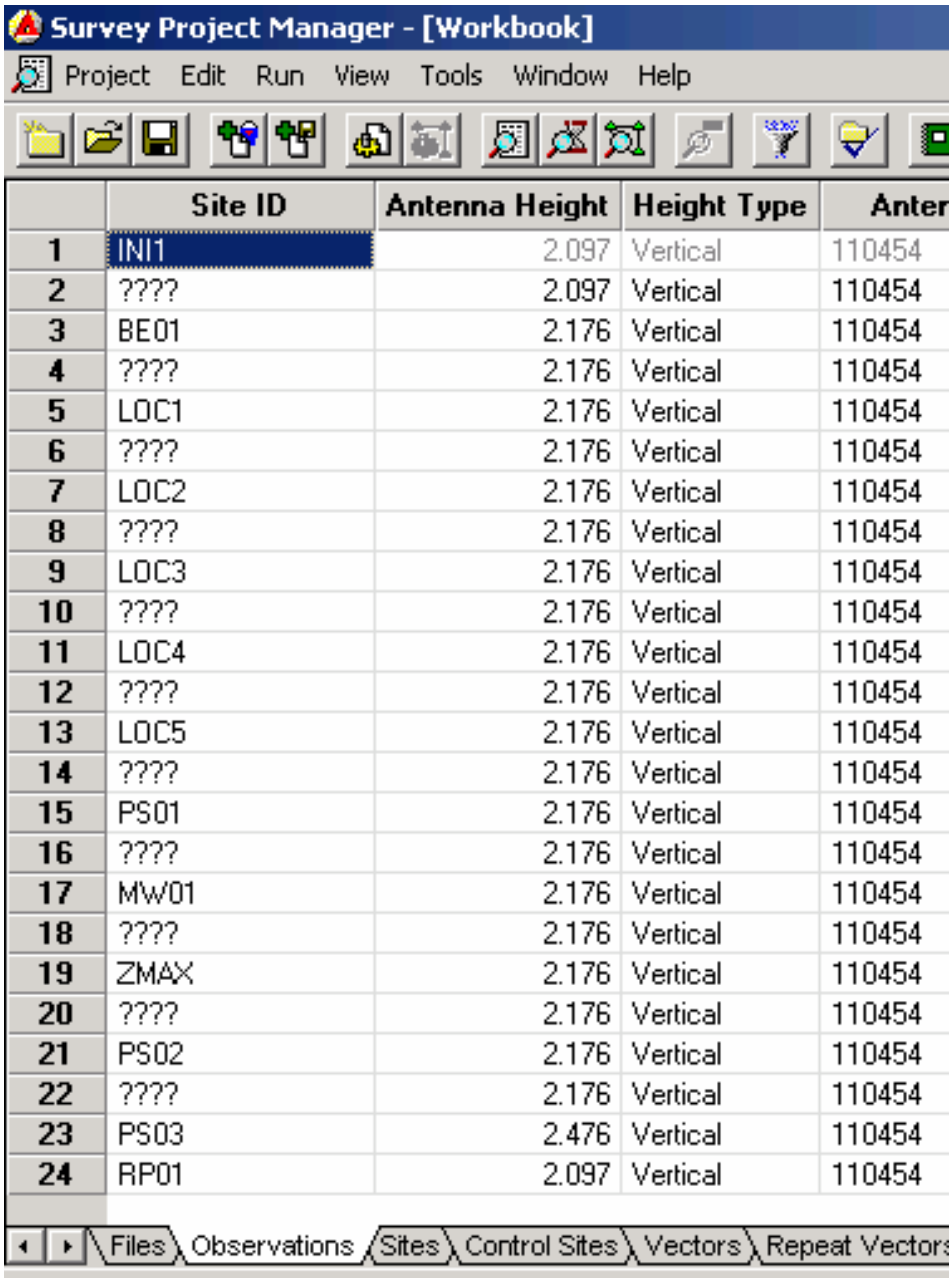
would be exactly the same. It's worth checking the vertical height to the ARP on the rover pole again. The height was entered incorrectly on this project so each rover shot will need to be edited to reflect the correct height of 2.091 meters. This can be done in either the time view or the observations tab in the workbook.

Antenna height errors are easy to make and easy to fix if good field notes are kept. Independent measurements of the antenna height in meters and feet can often save a project. The use of fixed height tripods and fixed height poles makes it easy to reconstruct a setup so the height can be measured again if needed.

Editing the heights of the dynamic data where the Site ID is ???? is not required, however for the sake of consistent looking data all of the rover data has been edited so the antenna heights match the height of the rover pole as illustrated in screen shot 4.

The height of the rover on the INI bar is measured as if it were set on the center point of the bar. In this case, with the base and rover antenna at the same height on the bar, the height of INI1 is equal to the height of the base point RP01

Once the rover heights have been corrected it's a good idea to check the Site ID's before proceeding.



	Site ID	Antenna Height	Height Type	Anter
1	INI1	2.097	Vertical	110454
2	????	2.097	Vertical	110454
3	BE01	2.176	Vertical	110454
4	????	2.176	Vertical	110454
5	LOC1	2.176	Vertical	110454
6	????	2.176	Vertical	110454
7	LOC2	2.176	Vertical	110454
8	????	2.176	Vertical	110454
9	LOC3	2.176	Vertical	110454
10	????	2.176	Vertical	110454
11	LOC4	2.176	Vertical	110454
12	????	2.176	Vertical	110454
13	LOC5	2.176	Vertical	110454
14	????	2.176	Vertical	110454
15	PS01	2.176	Vertical	110454
16	????	2.176	Vertical	110454
17	MW01	2.176	Vertical	110454
18	????	2.176	Vertical	110454
19	ZMAX	2.176	Vertical	110454
20	????	2.176	Vertical	110454
21	PS02	2.176	Vertical	110454
22	????	2.176	Vertical	110454
23	PS03	2.476	Vertical	110454
24	RP01	2.097	Vertical	110454

Screen shot 3

It is always a good habit to look at the project before and after editing.

Survey Project Manager - [Workbook]

Project Edit Run View Tools Window Help

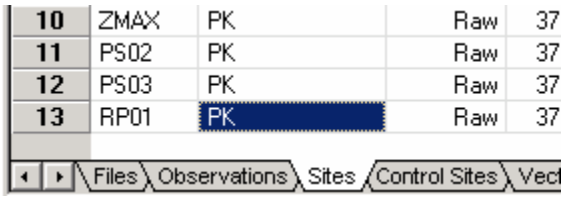
	Site ID	Antenna Height	Height Type	Anter
1	INI1	2.097	Vertical	110454
2	????	2.091	Vertical	110454
3	BE01	2.091	Vertical	110454
4	????	2.091	Vertical	110454
5	LOC1	2.091	Vertical	110454
6	????	2.091	Vertical	110454
7	LOC2	2.091	Vertical	110454
8	????	2.091	Vertical	110454
9	LOC3	2.091	Vertical	110454
10	????	2.091	Vertical	110454
11	LOC4	2.091	Vertical	110454
12	????	2.091	Vertical	110454
13	LOC5	2.091	Vertical	110454
14	????	2.091	Vertical	110454
15	PS01	2.091	Vertical	110454
16	????	2.091	Vertical	110454
17	MW01	2.091	Vertical	110454
18	????	2.091	Vertical	110454
19	ZMAX	2.091	Vertical	110454
20	????	2.091	Vertical	110454
21	PS02	2.091	Vertical	110454
22	????	2.091	Vertical	110454
23	PS03	2.476	Vertical	110454
24	RP01	2.097	Vertical	110454

Files Observations Sites Control Sites Vectors Repeat Vector

Screen shot 4

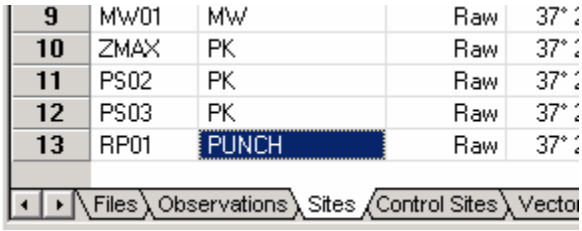
Sites is the next tab to check. On this project RP01 is a punch hole in a curb. A look at the Sites tab shows the Site Descriptor for RP01 as a PK. This is easily fixed by a click on the incorrect descriptor, PK, so it is highlighted, and then typing in the correct data, PUNCH.

Before:



Screen shot 5

After:



Screen shot 6

The control point for this project is RP01. Once again, as in part one, it is appropriate to compare the raw position shown in screen shot 7 with the official coordinates. In this project the official coordinates were obtained from OPUS.

Workbook

Site ID	Site Descriptor	Latitude	95% Err.	Longitude	95% Err.	Ortho. Ht.
RP01	PUNCH	37° 21' 4.90519" N	0.000	121° 56' 4.91510" W	0.000	25.539

Screen shot 7

Official coordinates from the OPUS report:

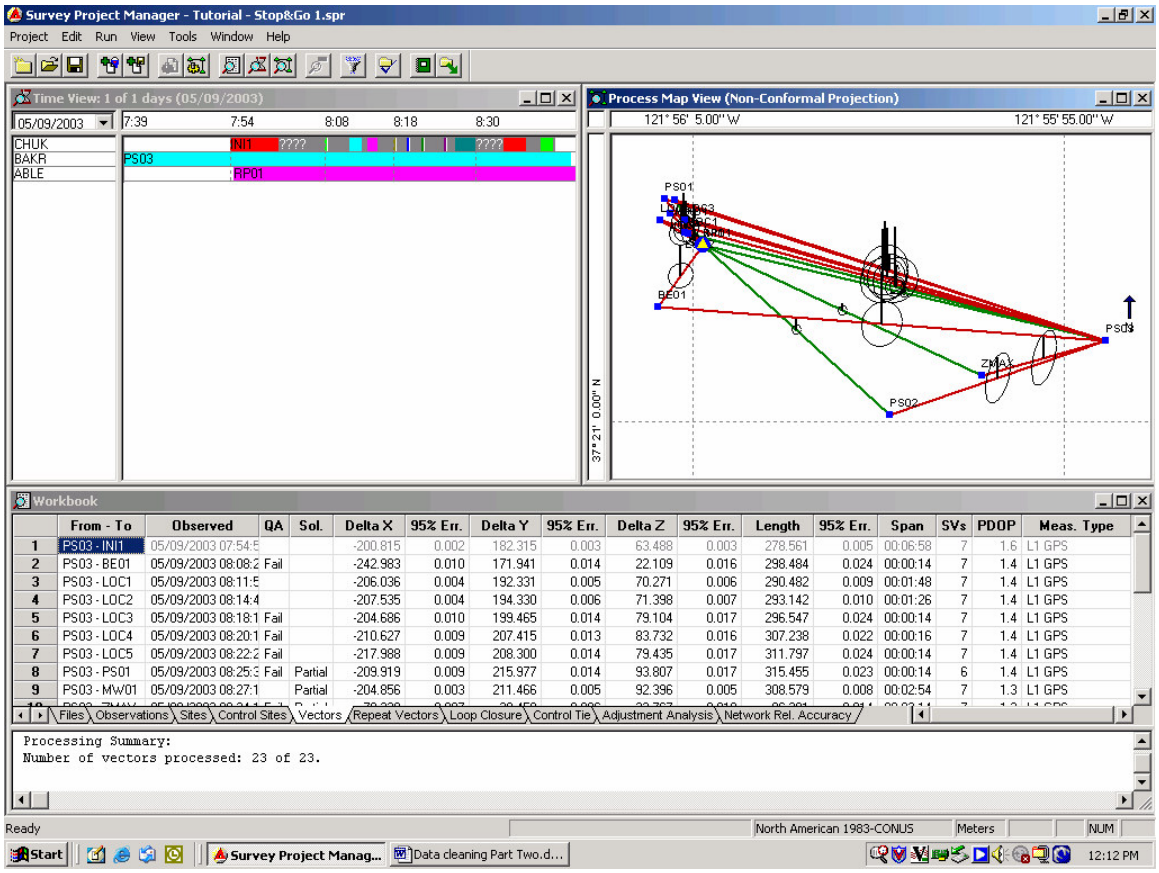
REF FRAME: NAD83(CORS96) (EPOCH:2002.0000)
LAT: 37°21'04.77521"
W LON: 121°56'04.71600"
EL HGT: -12.187(m)
ORTHO HGT: 20.426(m) [Geoid99 NAVD88]

Since the raw position is within 50 feet horizontal and 100 feet vertical of the OPUS coordinates they are acceptable for this project. In screen shot 8 the coordinates for RP01 have been edited to match the OPUS coordinates.

Site ID	Site Descriptor	Latitude	95% Err.	Longitude	95% Err.	Ortho. Ht.
RP01	PUNCH	37° 21' 4.77521" N	0.000	121° 56' 4.71600" W	0.000	20.426

Screen shot 8

Save the project and process the data.



Screen shot 9

Problems are quickly revealed by red vectors on the map view and in the Sol and the QA columns on the vectors tab in the workbook.

A click on the Sol column heading in the workbook brings four partial solutions to the top of the table. The first task is to clean up these four problem vectors if possible.

It is important to note that this Stop&Go survey project is different from the static survey in that it is not a network. It is a radial survey from two base stations. A closer look at the Vectors in the workbook might help. But first a review of the field notes…

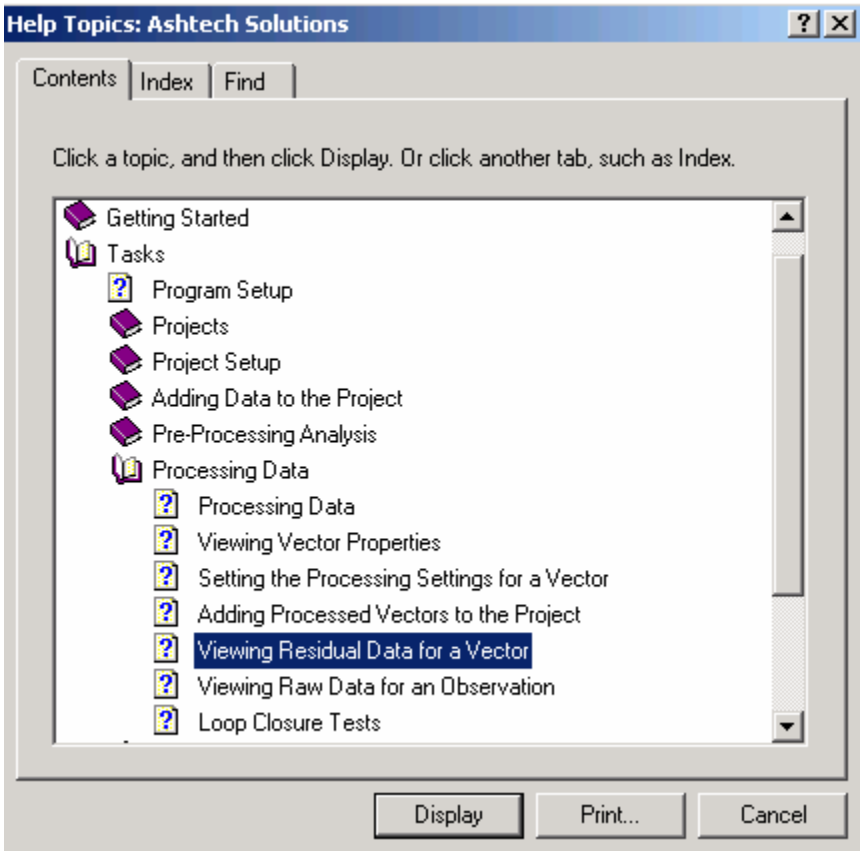
Screen shot 10

	From - To	Observed	QA	Sol.	Delta X	95% Err.	Delta Y	95% Err.	Delta Z	95% Err.	Length	95% Err.	Span	SVs	PDOP
1	PS03 - PS01	05/09/2003 08:25:34	Fail	Partial	-209.919	0.009	215.977	0.014	93.807	0.017	315.455	0.023	00:00:14	6	1.4
2	PS03 - PS02	05/09/2003 08:39:26	Fail	Partial	-140.846	0.007	44.155	0.006	-49.785	0.011	155.775	0.014	00:01:52	7	1.3
3	PS03 - ZMAX	05/09/2003 08:34:12	Fail	Partial	-78.228	0.007	28.458	0.006	-22.767	0.010	86.301	0.014	00:03:14	7	1.3
4	PS03 - MW01	05/09/2003 08:27:10		Partial	-204.856	0.003	211.466	0.005	92.396	0.005	308.579	0.008	00:02:54	7	1.3
5	PS03 - LOC2	05/09/2003 08:14:40			-207.535	0.004	194.330	0.006	71.398	0.007	293.142	0.010	00:01:26	7	1.4
6	PS03 - LOC5	05/09/2003 08:22:20	Fail		-217.988	0.009	208.300	0.014	79.435	0.017	311.797	0.024	00:00:14	7	1.4
7	PS03 - LOC1	05/09/2003 08:11:58			-206.036	0.004	192.331	0.005	70.271	0.006	290.482	0.009	00:01:48	7	1.4
8	PS03 - BE01	05/09/2003 08:08:24	Fail		-242.983	0.010	171.941	0.014	22.109	0.016	298.484	0.024	00:00:14	7	1.4
9	PS03 - LOC3	05/09/2003 08:18:16	Fail		-204.686	0.010	199.465	0.014	79.104	0.017	296.547	0.024	00:00:14	7	1.4
10	PS03 - LOC4	05/09/2003 08:20:14	Fail		-210.627	0.009	207.415	0.013	83.732	0.016	307.238	0.022	00:00:16	7	1.4
11	RP01 - PS03	05/09/2003 07:55:20			200.661	0.001	-182.323	0.001	-63.621	0.001	278.486	0.002	00:48:28	8	1.3
12	PS03 - INI1	05/09/2003 07:54:50			-200.815	0.002	182.315	0.003	63.488	0.003	278.561	0.005	00:06:58	7	1.6
13	RP01 - INI1	05/09/2003 07:55:20			-0.156	0.001	-0.008	0.002	-0.130	0.002	0.203	0.003	00:06:28	7	1.6
14	RP01 - BE01	05/09/2003 08:08:24	Fail		-42.319	0.006	-10.375	0.008	-41.513	0.009	60.182	0.014	00:00:14	7	1.4
15	RP01 - LOC1	05/09/2003 08:11:58			-5.383	0.002	9.999	0.003	6.650	0.004	13.159	0.005	00:01:48	7	1.4
16	RP01 - LOC2	05/09/2003 08:14:40			-6.871	0.002	12.009	0.003	7.778	0.004	15.872	0.006	00:01:26	7	1.4
17	RP01 - LOC3	05/09/2003 08:18:16	Fail		-4.025	0.005	17.146	0.008	15.477	0.009	23.446	0.013	00:00:14	7	1.4
18	RP01 - LOC4	05/09/2003 08:20:14	Fail		-9.964	0.005	25.095	0.007	20.110	0.009	33.667	0.013	00:00:16	7	1.4
19	RP01 - LOC5	05/09/2003 08:22:20	Fail		-17.328	0.005	25.965	0.008	15.814	0.009	34.993	0.013	00:00:14	7	1.4
20	RP01 - PS01	05/09/2003 08:25:34	Fail		-9.344	0.005	33.922	0.008	30.250	0.009	46.401	0.013	00:00:14	7	1.3
21	RP01 - MW01	05/09/2003 08:27:10			-4.291	0.002	29.405	0.002	28.843	0.003	41.413	0.004	00:02:54	7	1.3
22	RP01 - ZMAX	05/09/2003 08:34:12			122.365	0.002	-153.575	0.002	-86.317	0.003	214.497	0.004	00:03:14	7	1.4
23	RP01 - PS02	05/09/2003 08:39:26			59.740	0.002	-137.876	0.003	-113.331	0.003	188.209	0.005	00:01:52	7	1.3

Files | Observations | Sites | Control Sites | Vectors | Repeat Vectors | Loop Closure | Control Tie | Adjustment Analysis | Network Rel. Accuracy

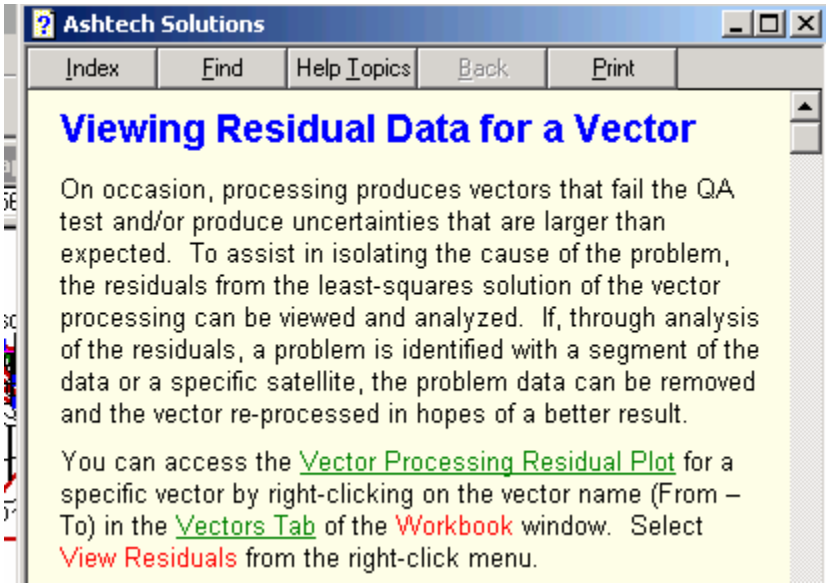
The rover never lost lock. During this entire session the rover maintained lock on at least five SV's. What can be the reason for problems with the data?

Ashtech Solutions includes tools that permit examination of the data. After a review of the workbook to see what needs work the help button will be the next click.



Screen shot 11

What can be learned from the on line help? There is a lot to read but if there is a way to save the data without another trip to the field it will be time well spent. By clicking on the green letters "Vector Processing Residual Plot" there is an opportunity to see an example with even more explanation.

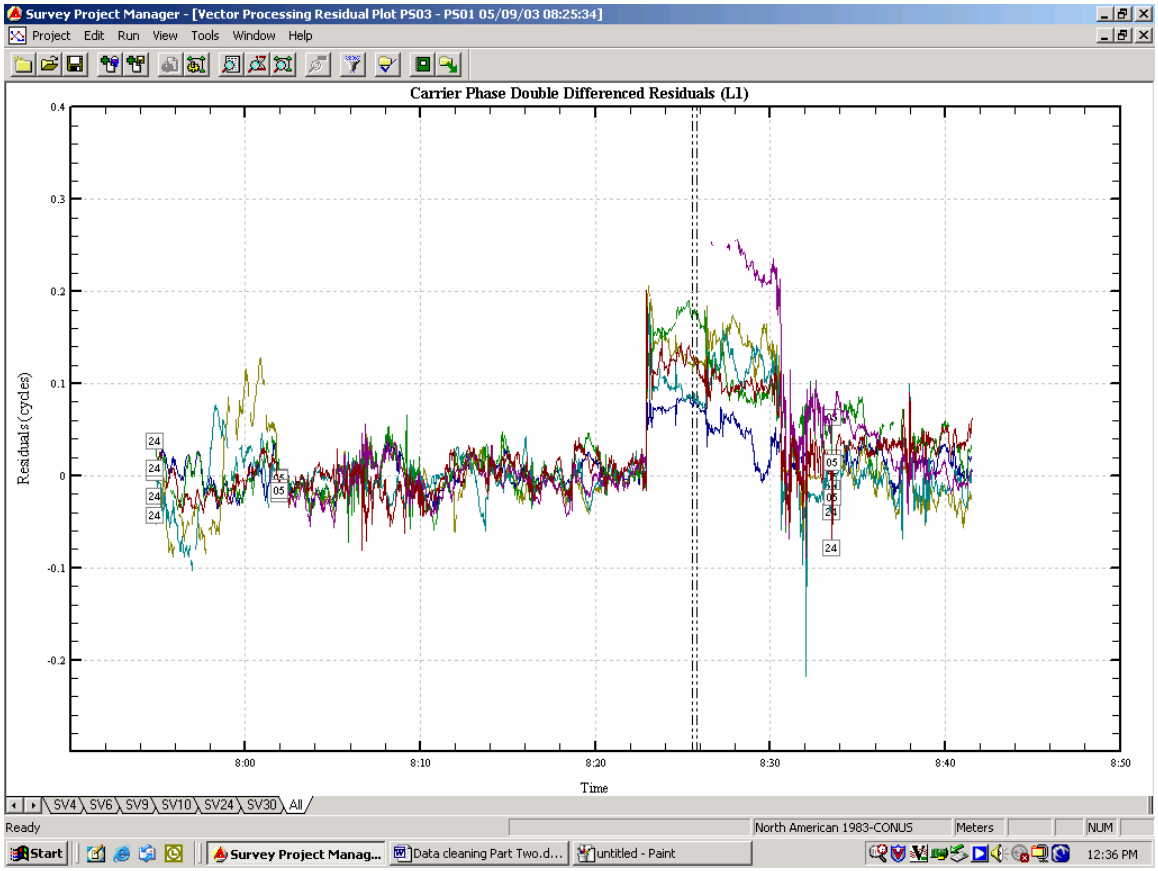


Screen shot 12

Following the instructions in the help file, a right click on the vector name PS03-PS01, and a left click on View Residuals, brings up the residual plot. A click on the ALL tab shows a graphic plot of all the SV's. For a better comprehension, read the Ashtech Solutions manual. It matters. More things are learned by working with the data. Whenever the topic is accuracy the question has to be, "Compared to what?" The residual plot shows the accuracy of each SV as compared to the reference SV. The residual plot shows the data that is common to both receivers for the duration of the session.

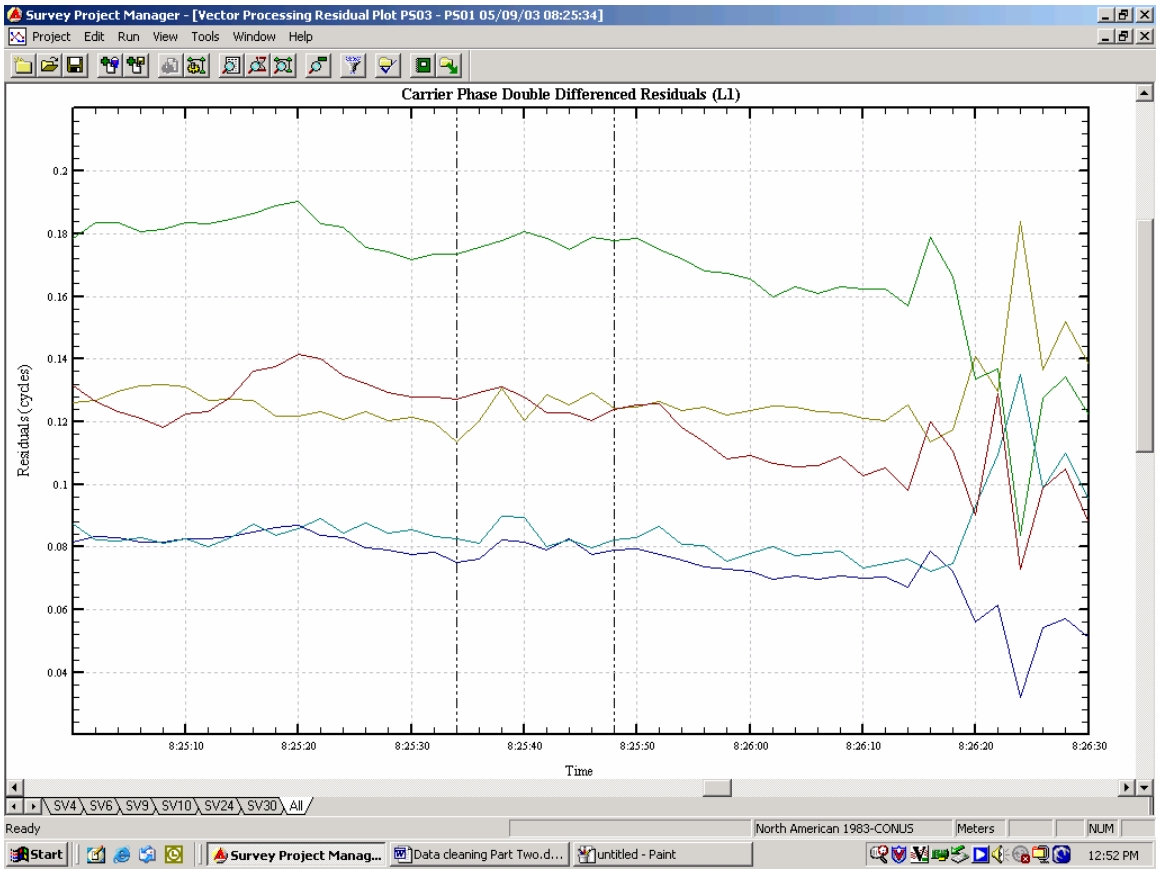
Ideally a tight group of lines that form a uniform pattern is what will be seen. A perfect measurement would be a horizontal line right on the zero line. However, most measurement experts know that there is no such thing as a perfect measurement so the goal is a tight graph with uniform data.

In screen shot 13 there are two bold vertical lines on the plot that show a particular shot taken on PS01.



Screen shot 13

Zooming in on the time of the two bold vertical lines gives a clue to what happened during that shot. Zoom in by dragging a box with the left mouse button.

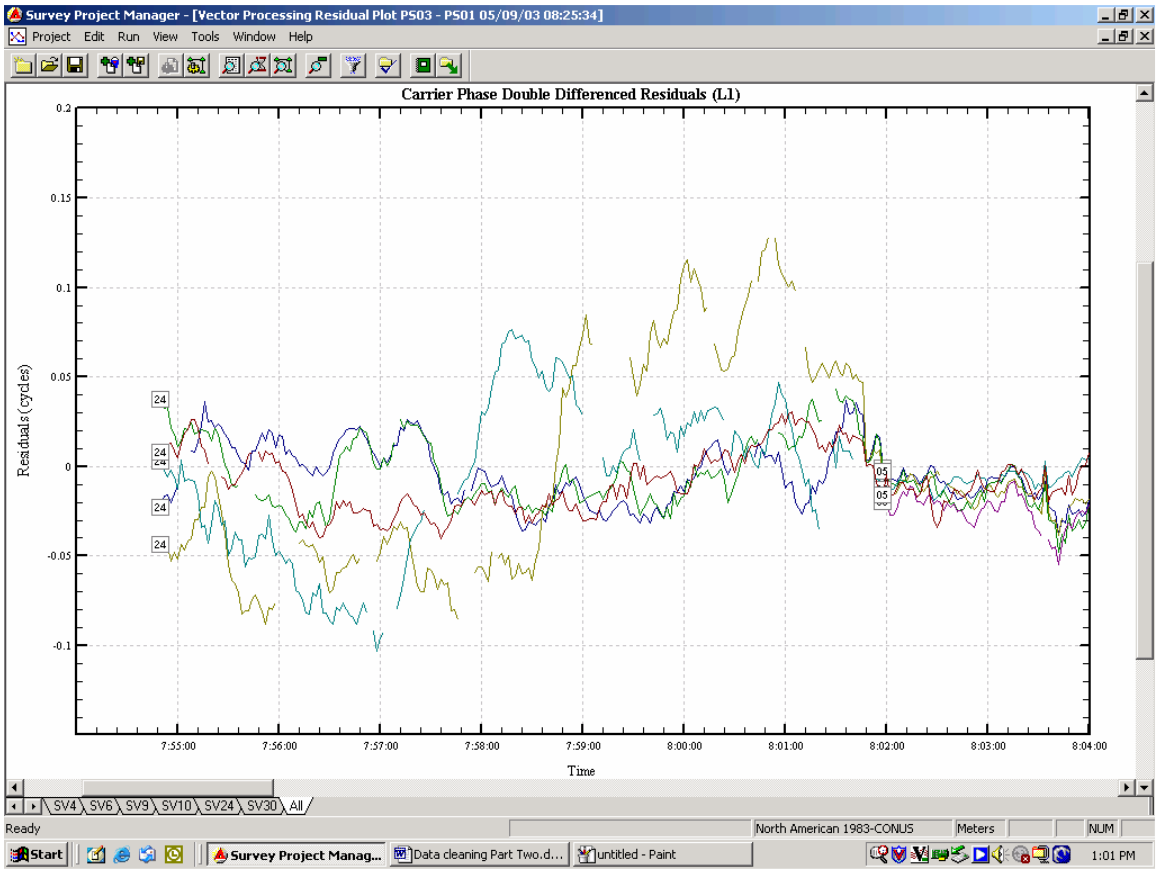


Screen shot 14

Count the squiggly lines to determine the number of SV's used and add one for the reference SV. Remember, the graph compares each SV to the reference SV. During the occupation of PS01 there were 6 SV's in use. So what is the problem?

Part of the problem is the desire to focus on a few seconds of data for a particular observation. In a Stop&Go survey it is not just about the STOP it is also about the GO. A good solution must be maintained from the initialization through each successive shot. An observation of the entire residual plot is needed to understand what went wrong.

It is not hard to understand that at the beginning of the session things will be a little scattered out as illustrated in screen shot 15, but as the initialization comes together so does the residual plot.



Screen shot 15

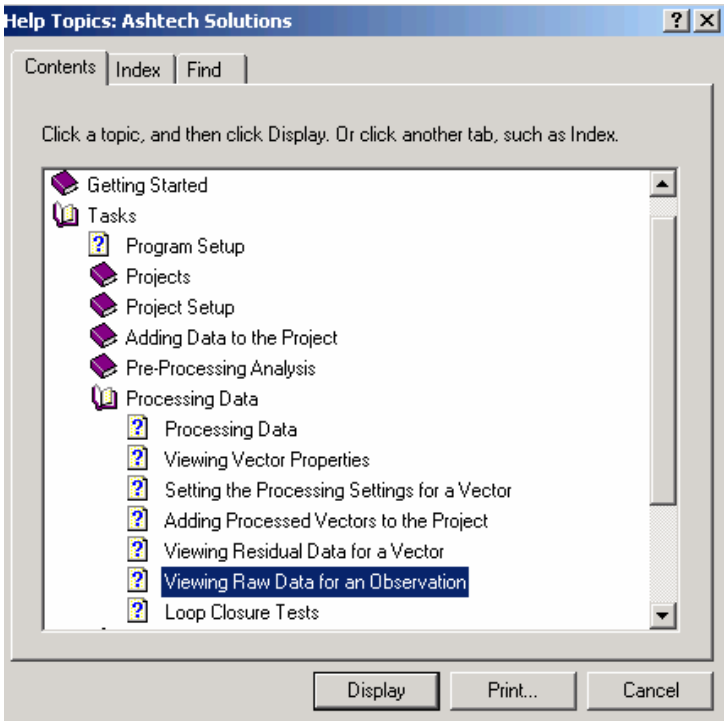
Apparently something went wrong at about 8:23 that morning and the indication of the problem is obvious in the examination of the residual plot as shown in screen shot 15.

Close this plot and go back to the workbook. Sorting the vectors by Observed time indicates that the vector from PS03-LOC5 taken at 08:22:20 obtained a fixed solution but everything after that failed.

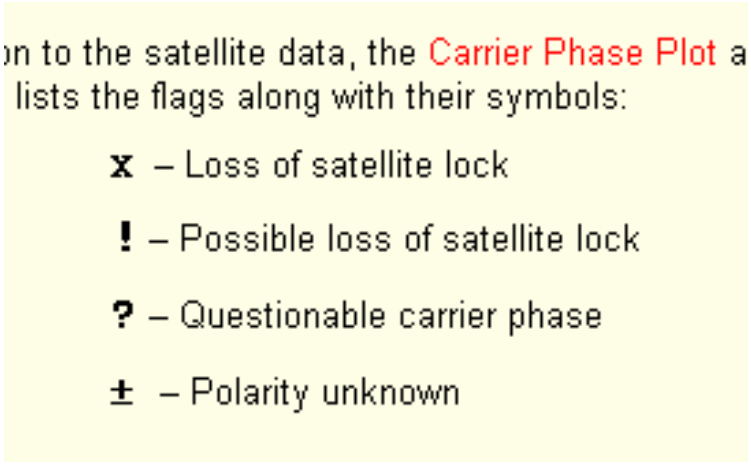
Does this mean a trip back to the field? Perhaps not, there may be another solution. A look at the raw data may indicate a way to solve the problem.

Screen shot 16

Before a look at the raw data take another look at the help file to see what can be learned about raw data plots.



Screen shot 17



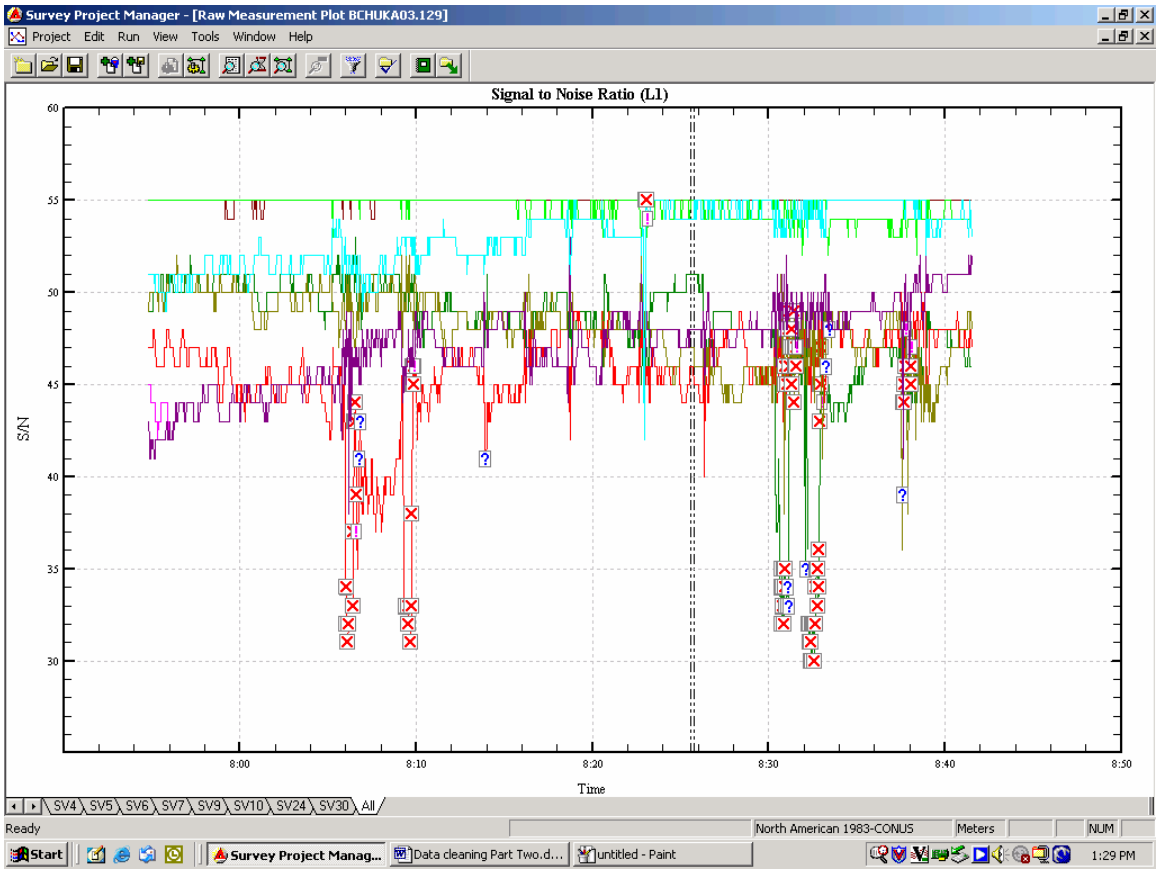
Screen shot 18

The help file regarding raw data contains important, pertinent information to keep in mind when examining raw data plots. The icons shown in screen shot 18 in a raw data plot are bad news.

Now, back to the raw data plot in screen shot 19.

The problem data is from 8:23. So the next task is to find problem data at 8:23, or thereabouts.

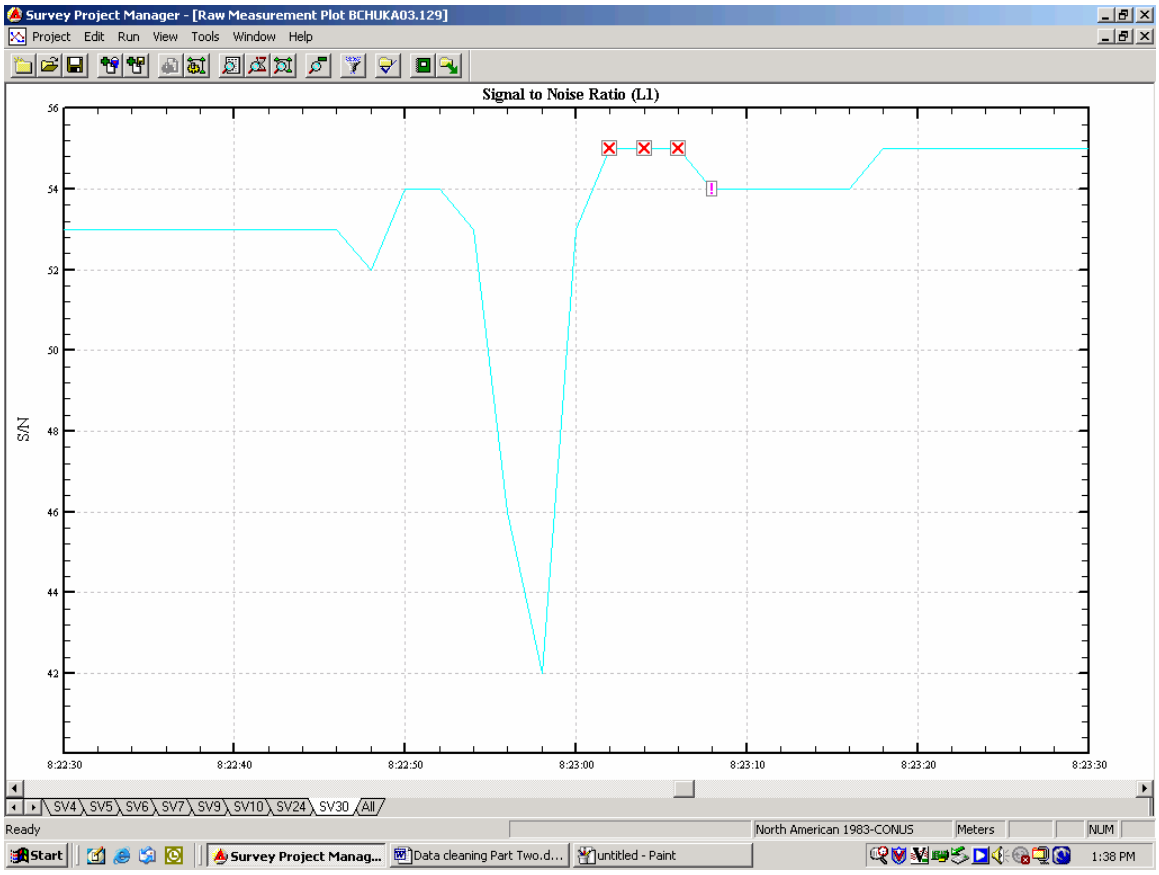
The raw data plot differs from the residual plot. Instead of viewing both receivers in one plot they are viewed individually. Also, instead of viewing one SV compared to the reference SV the data from each SV is viewed on an individual tab on the plot. A look at all of the SV's raw data together may provide surprises. When looking at the data from this project keep in mind that there was never a loss of lock alarm in the field.



Screen shot 19

Focus on the time of 8:23 is important since this is where the residual plot shows that things went wrong. A problem at that time resulted in the first partial solution. Zoom in on the time for a closer examination.

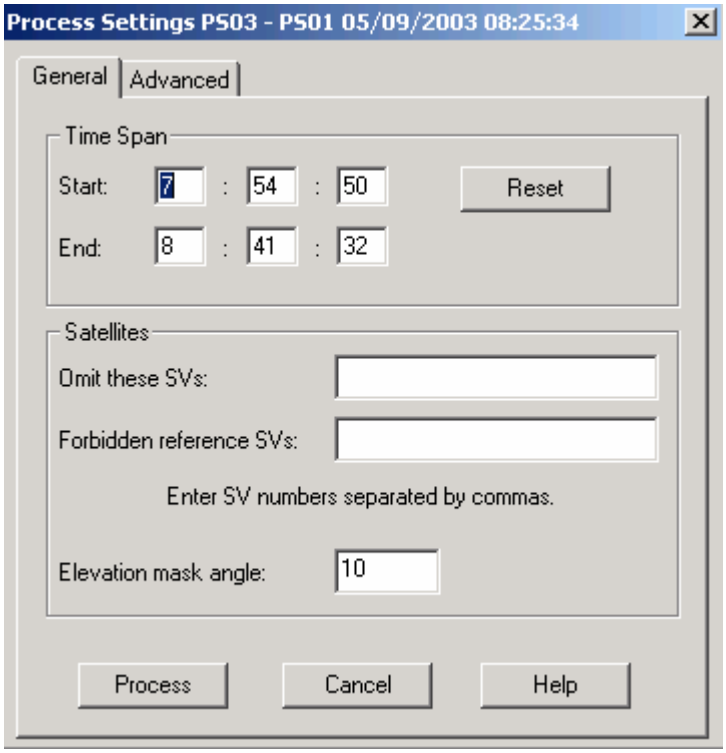
By zooming in on that time frame it is shown that the problem is specific to SV30. A click on the SV30 tab isolates the problem more as shown in screen shot 20. The big dip in the raw data plot just before the loss of lock icons may indicate multipath. Refer to the Ashtech Solutions manual and the information contained in the raw data help file. The loss of lock is easier to understand and is most likely the source of the problem with the vector from PS03 to PS01. Trimming out the data from SV30 between 8:22:50 and 8:23:10 could make this vector work without a return to the field.



Screen shot 20

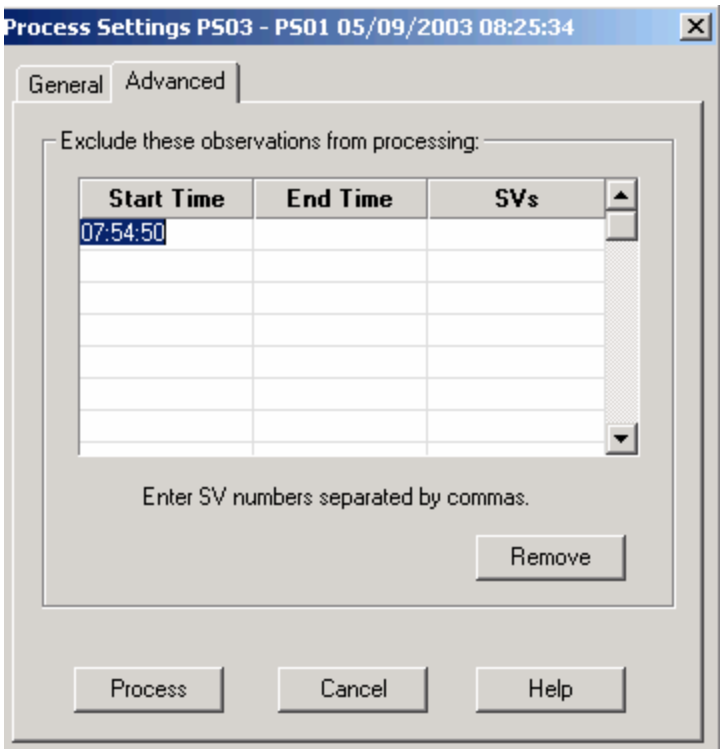
Use extreme care when trimming out data. With Stop&Go survey work it is not just one vector that is affected by reprocessing. When Stop&Go data is reprocessed the entire session is reprocessed.

A right click on the vector name from PS03-PS01 will pull up a menu. Then, a left click on process will pull up the Process Settings box shown in screen shot 21.



Screen shot 21

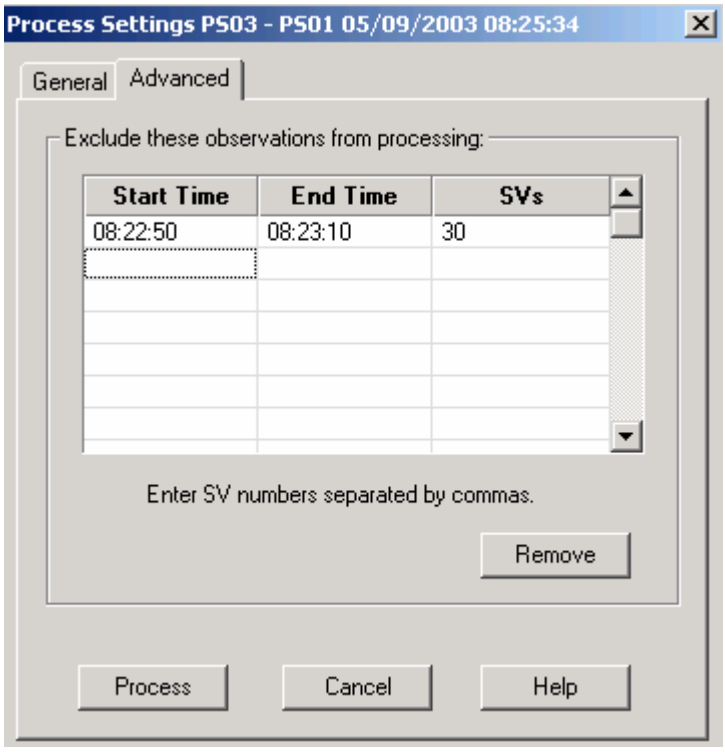
Had this session been a static survey one possible solution would be to just enter 30 in the Omit these SVs: box. Since this is Stop&Go it would be better to be more selective and exclude the defective observation time. Click on the advanced tab.



Screen shot 22

A double click on the blank line under Start Time fills in the start time of this session. Using the arrow keys to maintain the format, the start time can be changed by entering the start time that we want to begin the trim. In this case the time needs to be changed to 08:22:50. The same procedure is used to change the end time to 08:23:10. The SVs column gets 30 since that is the space vehicle that lost lock.

With the data to exclude all filled in click on Process to execute reprocessing of the Stop&Go session.



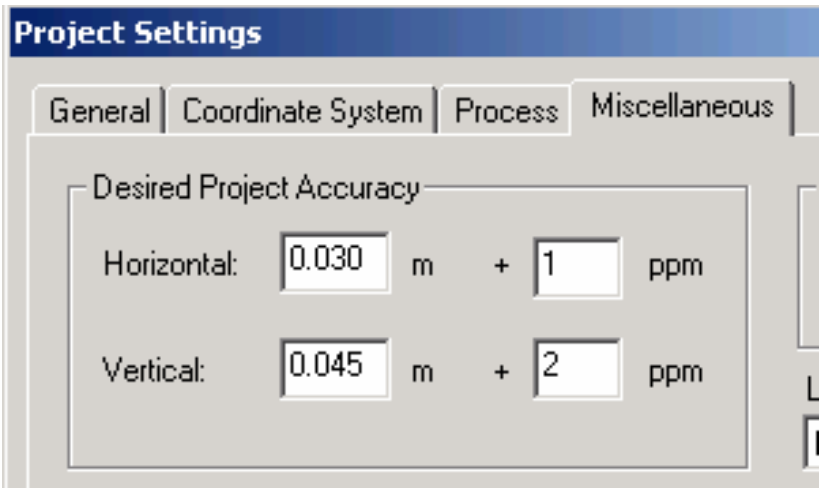
Screen shot 23

Now Ashtech Solutions reports that not only does the vector from PS03-PS01 provide a fixed solution but ALL four Partial solutions have been fixed. This just illustrates what twenty seconds of bad data can do to a project.

Here is a tip for removing data from a vector. Look at the 95% error to the right of length. Every step that makes that error smaller was the right move. When the 95% error gets bigger it is time to back up and do something different.

Data cleaning can be a tricky process. Take it slow and eliminate one bad thing at a time. Haste can mean wasting good data and too much scrubbing just washes everything away, leaving nothing to process.

What about those QA failures? Remember that the QA pass/fail is based on the desired project accuracy. The question to ask is whether enough field work is going to be done to satisfy the desired project accuracy. This project was done using short occupation times on an obstructed project site. Is it reasonable to expect horizontal and vertical accuracy at 0.009 meters? Perhaps loosening of that standard is appropriate for this project.



Screen shot 24

With less pressure on the accuracy expectations the QA failures are eliminated. When it comes to QA pass or fail the question is about whether enough work is going to be done to meet the accuracy expectation. As always, the choices are to do the work or lower the expectations.

	From - To	Observed	QA	Sol.
1	PS03 - INI1	05/09/2003 07:54:50		
2	RP01 - PS03	05/09/2003 07:55:20		
3	RP01 - INI1	05/09/2003 07:55:20		
4	PS03 - BE01	05/09/2003 08:08:24		
5	RP01 - BE01	05/09/2003 08:08:24		
6	PS03 - LOC1	05/09/2003 08:11:58		
7	RP01 - LOC1	05/09/2003 08:11:58		
8	PS03 - LOC2	05/09/2003 08:14:40		
9	RP01 - LOC2	05/09/2003 08:14:40		
10	RP01 - LOC3	05/09/2003 08:18:16		
11	PS03 - LOC3	05/09/2003 08:18:16		
12	RP01 - LOC4	05/09/2003 08:20:14		
13	PS03 - LOC4	05/09/2003 08:20:14		
14	PS03 - LOC5	05/09/2003 08:22:20		
15	RP01 - LOC5	05/09/2003 08:22:20		
16	PS03 - PS01	05/09/2003 08:25:34		
17	RP01 - PS01	05/09/2003 08:25:34		
18	PS03 - MW01	05/09/2003 08:27:10		
19	RP01 - MW01	05/09/2003 08:27:10		
20	PS03 - ZMAX	05/09/2003 08:34:12		
21	RP01 - ZMAX	05/09/2003 08:34:12		
22	RP01 - PS02	05/09/2003 08:39:26		
23	PS03 - PS02	05/09/2003 08:39:26		

Files | Observations | Sites | Control Sites | Vectors

Screen shot 25

With nice clean vectors a network adjustment can be made – just as soon as the project is saved once again.

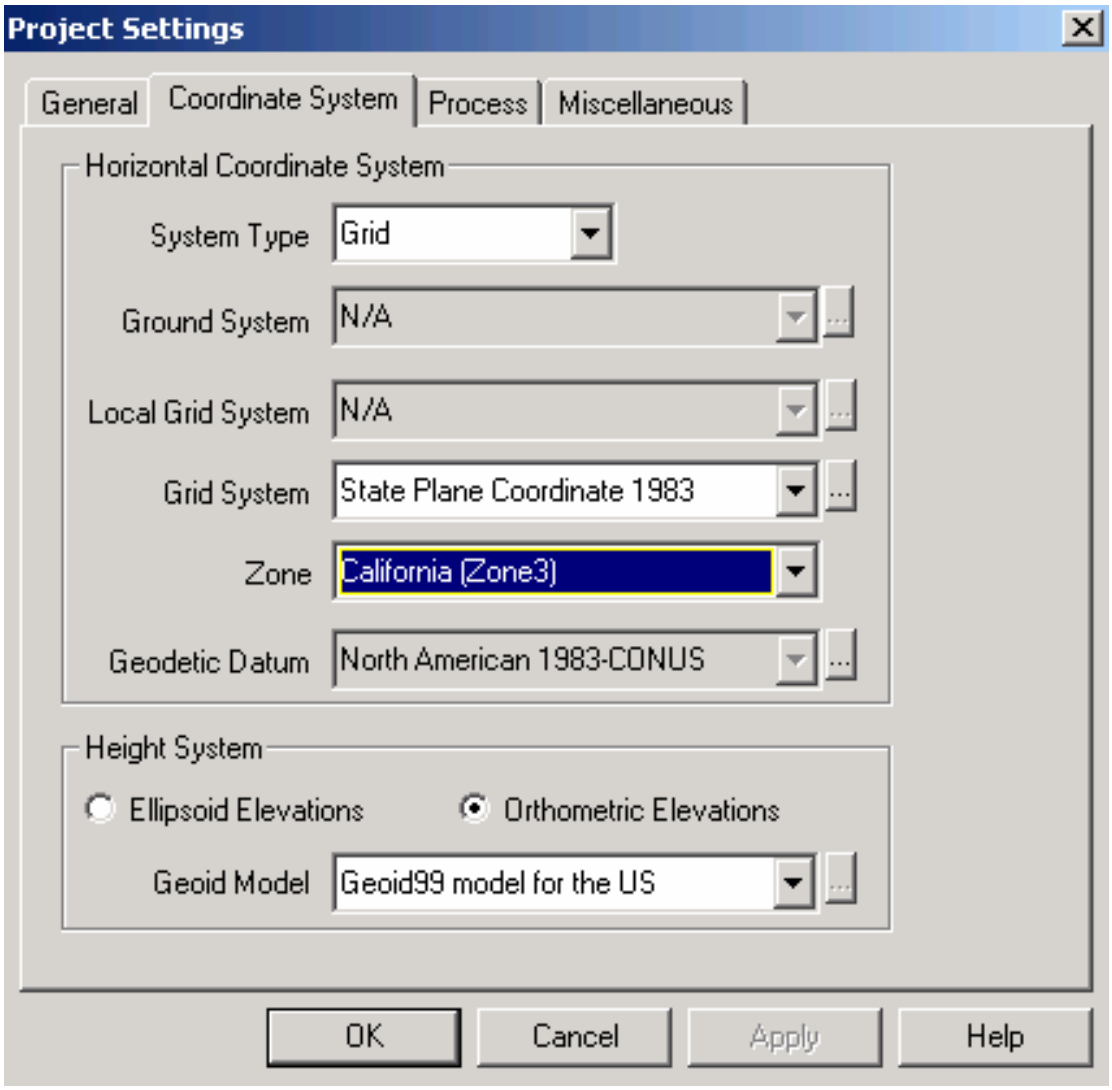
Screen shot 26

	Site Pair	QA	Horz. Rel. Error	Vert. Rel. Error	Horz. Rel. Accuracy	Vert. Rel. Accuracy	Distance
1	PS03 - LOC3		0.006	0.013	1:46891	1:22811	296.548
2	RP01 - LOC3		0.006	0.013	1:3709	1:1804	23.446
3	RP01 - BE01		0.006	0.013	1:9525	1:4629	60.183
4	PS03 - BE01		0.006	0.013	1:47240	1:22960	298.486
5	RP01 - LOC5		0.006	0.013	1:5572	1:2692	34.997
6	PS03 - LOC5		0.006	0.013	1:49638	1:23984	311.793
7	RP01 - PS01		0.006	0.013	1:7411	1:3569	46.402
8	PS03 - PS01		0.006	0.013	1:50423	1:24286	315.715
9	RP01 - LOC4		0.006	0.012	1:5552	1:2805	33.666
10	PS03 - LOC4		0.006	0.012	1:50671	1:25603	307.239
11	PS03 - LOC2		0.003	0.006	1:92538	1:48857	293.142
12	RP01 - LOC2		0.003	0.006	1:5012	1:2645	15.871
13	RP01 - ZMAX		0.002	0.004	1:100669	1:53624	214.497
14	RP01 - PS03		0.002	0.003	1:130815	1:92829	278.486
15	PS03 - ZMAX		0.002	0.004	1:40694	1:21610	86.440
16	RP01 - MW01		0.002	0.004	1:19628	1:10353	41.414
17	RP01 - LOC1		0.002	0.005	1:6237	1:2632	13.161
18	PS03 - LOC1		0.002	0.005	1:137733	1:58096	290.482
19	PS03 - MW01		0.002	0.004	1:146504	1:77211	308.843
20	PS03 - PS02		0.002	0.005	1:75605	1:31180	155.901
21	RP01 - PS02		0.002	0.005	1:91317	1:37642	188.211

Files | Observations | Sites | Control Sites | Vectors | Repeat Vectors | Loop Closure | Control Tie | Adjustment Analysis | Network Rel.

Screen shot 26 shows that Ashtech Solutions estimates the worst horizontal error as 0.006 meters and the worst vertical error is 0.013 meters.

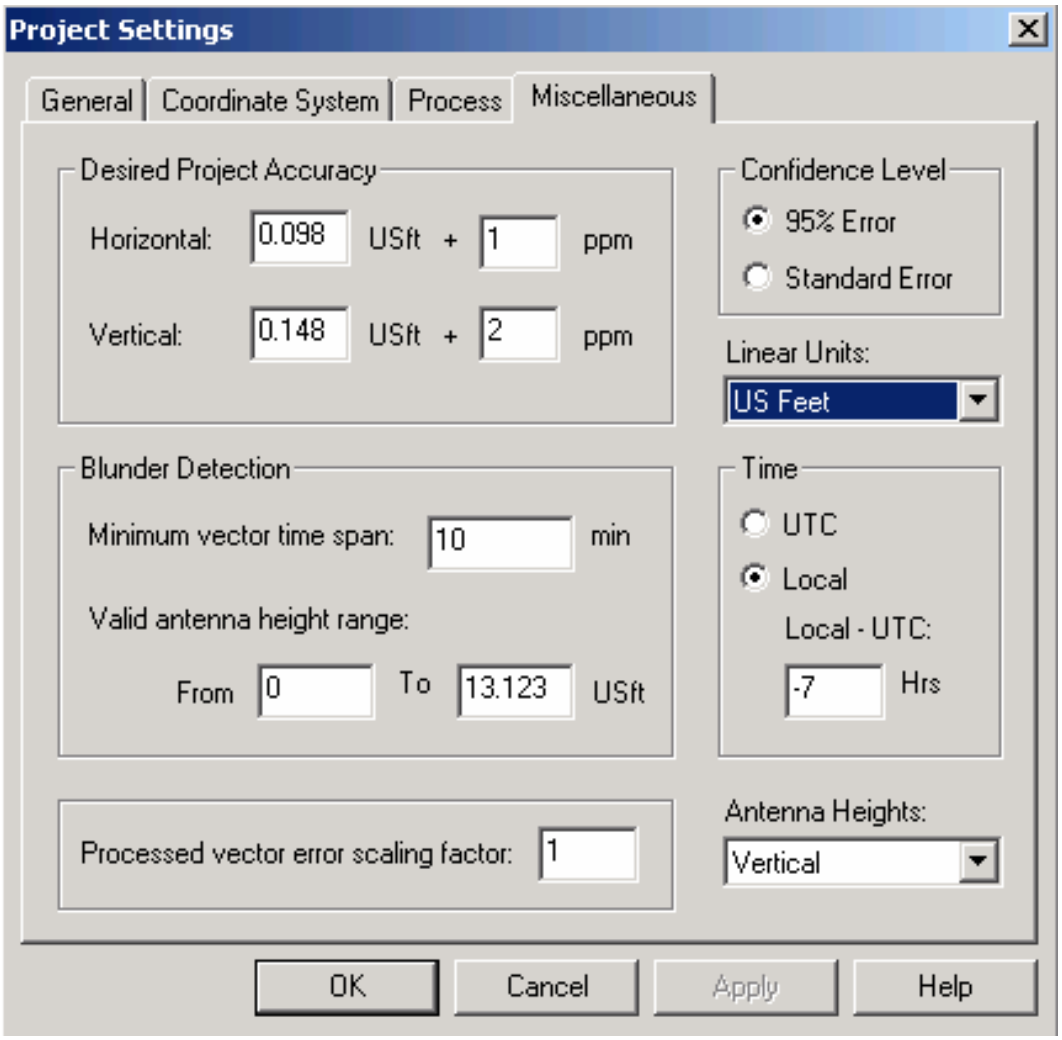
Since the data has been processed and is acceptable the project settings can be changed to State Plane Coordinates in US Feet for comparison with the results of the static survey done in part one.



Screen shot 27

The coordinate system is set to match the area where the project is located

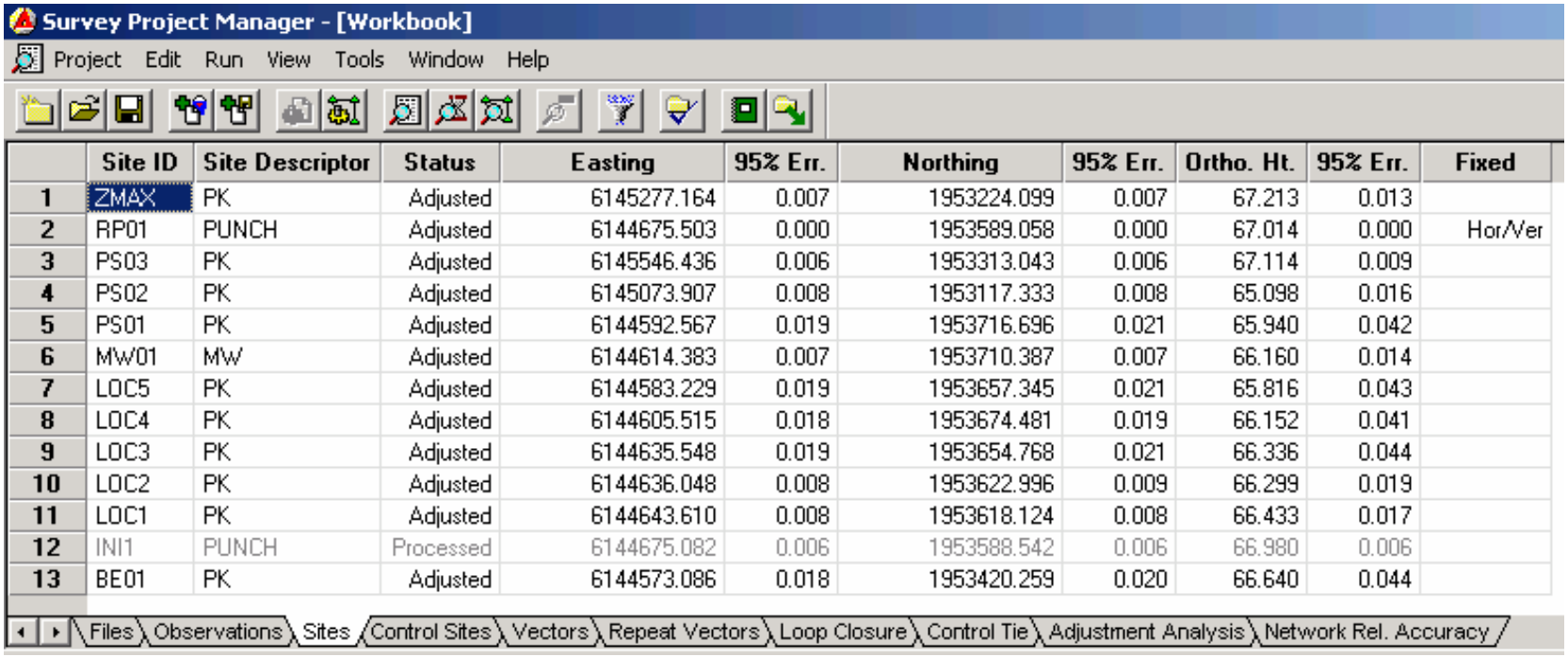
and the linear units is set to US Feet under the Miscellaneous tab.



Screen shot 28

After a final adjustment the Sites tab provides the coordinates as shown in screen shot 29 for a comparison with the static data from part one.

Screen shot 29



	Site ID	Site Descriptor	Status	Easting	95% Err.	Northing	95% Err.	Ortho. Ht.	95% Err.	Fixed
1	ZMAX	PK	Adjusted	6145277.164	0.007	1953224.099	0.007	67.213	0.013	
2	RP01	PUNCH	Adjusted	6144675.503	0.000	1953589.058	0.000	67.014	0.000	Hor/Ver
3	PS03	PK	Adjusted	6145546.436	0.006	1953313.043	0.006	67.114	0.009	
4	PS02	PK	Adjusted	6145073.907	0.008	1953117.333	0.008	65.098	0.016	
5	PS01	PK	Adjusted	6144592.567	0.019	1953716.696	0.021	65.940	0.042	
6	MW01	MW	Adjusted	6144614.383	0.007	1953710.387	0.007	66.160	0.014	
7	LOC5	PK	Adjusted	6144583.229	0.019	1953657.345	0.021	65.816	0.043	
8	LOC4	PK	Adjusted	6144605.515	0.018	1953674.481	0.019	66.152	0.041	
9	LOC3	PK	Adjusted	6144635.548	0.019	1953654.768	0.021	66.336	0.044	
10	LOC2	PK	Adjusted	6144636.048	0.008	1953622.996	0.009	66.299	0.019	
11	LOC1	PK	Adjusted	6144643.610	0.008	1953618.124	0.008	66.433	0.017	
12	INI1	PUNCH	Processed	6144675.082	0.006	1953588.542	0.006	66.980	0.006	
13	BE01	PK	Adjusted	6144573.086	0.018	1953420.259	0.020	66.640	0.044	