BLUNDER DETECTION TECHNIQUES IN ADJUSTMENT COMPUTATIONS

Mr. Bashar Saleem
Dept.of Surveying.
College of Engineering.
University of Baghdad

Mr. Hussein Alwan
Dept.of Surveying.
College of Engineering.
University of Baghdad

Mr. Jareer Jabir
Dept.of Surveying.
College of Engineering.
University of Baghdad

ABSTRACT

In adjustment of geodetic and photogrammetric networks, the surveying engineer faces many problems, such as errors of blunder nature in the observations (when comparing the homogeny of precision) make these observations odd from the result, and effecting directly on observation values after adjustment, and the statistical results after analysis as, adjusted coordinates of points, standard deviations, and ellipse of errors.

The research is deal with some of the most common usage of statistical methods to detect these odds observations to confirm which best method is, by studying the advantages and disadvantages of each method to geodetic network.

Three statistical methods will use in the analysis, these are:-

1. Standardized residuals method
2. F-t test
3. Robust estimation method

The adjustments were accomplished by preparing a Matlab program with the three blunder detection methods and the results were evaluated and some scientific conclusions were reached.

It was found that the robust estimation method represent the better blunder detection technique due to its ability in what is called (multi-blunder detection) , and the resulted higher accuracy indices.

الخلاصة

عند تصحيح الشبكات الجيوديسية وشبكات التثليث الجوي تعترض مهندس المساحة جملة من الأمور منها وجود أخطاء ذات طابع غلطي في بعض الارصادات (عند مقارنة تجانس الدقة) تجعل من هذه الارصادات شاذة عن البقية ,وتؤثر بصورة مباشرة على قيم الارصادات بعد التصحيح والنتائج الاحصائية بعد التحليل كالإحداثيات المصححه للنقاط والانحرافات المعيارية اضافة الى الشكل الاهليلجي للأخطاء الناتجة.

يتناول البحث مجموعة من الطرق الإحصائية الشائعة والمستخدمة للكشف عن هذه الرصدات الشاذة لاقرار الطريقة الافضل من خلال دراسة محاسن و مساوئ كل طريقة لشبكة جيوديسية .حيث اختيرت ثلاثة طرق احصائية متمثلة بالاتي:

1. طريقة البواقي القياسية (Standardized residuals method)
2. طريقة اختبار (F-t) test
3. طريقة التخمين المتين (Robust estimation method)

وباعداد برنامج بلغة (Matlab) تم اجراء التصحيح وتخمين النتائج للحصول على بعض الاستنتاجات العلمية والكشف عن الرصدات الشاذة بتطبيق الطرق الثلاث أعلاه.

وقد وجد ان طريقة التخمين المتين (Robust estimation method) تمثل أفضل تقنيه في الكشف عن الرصدات الشاذة استنادا إلى قابلية الطريقة في الكشف المتعدد للرصدات الشاذة (multi –blunder detection) وبسبب الدقة العالية التي توفرها هذه الطريقة مقارنة مع الطريقتين الأخريين.

KEY WORDS
Blunder detection, Geodetic Networks, Robust Estimation, Standardized Residual.

INTRODUCTION

When computing and adjusting geodetic networks, It is quiet important to ensure that the considered observations are free of blunders and systematic errors, and therefore the results are only affected by the random errors which presented in every measuring process.

Obviously, precaution in the observation process has to be the first step to avoid undesirable error appearance. Moreover, an adequate data filtering previous to the adjustment, checking reciprocal observations, and closures even with more detailed schemes is a suitable routine to be adopted in order to detect and eliminate the wrong observations.

Blunders or mistakes could be defined as obviously incorrect data points or results that are not reasonably close to the expected value. Some blunder or systematic errors may slip into the adjustment process. So it would be interested to deal with the subject that deals with the blunder detection.

BLUNDER DETECTION METHODS

The used methods for blunder detection will be explained theoretically and mathematically as follows:

Standardized Residual

The detection of blunder among the observations was treated using a technique pioneered by the geodesist Baarda (1968). In Baarda's method a statistical test which is called as *Standardized residuals* used to detect blunder [Baarda, 1968].

Blunders will affect the observations badly and produce incorrect estimations of the unknowns and their covariance matrix. If the blunders are detected by a statistical test, then those contaminated observations are removed, the network is re-adjusted, and we obtain the final results.

This method of standardized residuals, detect one blunder in every iteration and remove it from the observations, and readjust the network with the original observations, the iterations continued until remove all blunders and until the chi square test (χ^2) passed depend upon the significance level, variance and the degree of freedom. The method also depends on computing the variance-covariance matrix of residuals Σ_{vv}.

The standardized residuals are computed as:

$$\bar{v}_i = \frac{v_i}{\sqrt{q_{ii}}} \qquad (1)$$

Where $\bar{v}_i$ is the standardized residual,

v_i is the residual, and

q_{ii} is the diagonal element of the Σ_{vv} matrix.

Since a computed parameter divided by its standard deviation which is a random normal variable, we can compute a (t) value as:

$$t_i = \frac{v_i}{S_0\sqrt{q_{ii}}} = \frac{v_i}{S_{v_i}} = \frac{\bar{v}_i}{S_0} \tag{2}$$

However, this equation should be based on "good "value for the reference variance since a blunder automatically affects the value of S_0, the method summarized as follows:

1) Locate all measurements that qualify for rejection.
2) Reject the single observation with the largest standardized residuals.
3) Repeat adjustment 1& 2 until all observations qualify for rejection are rejected, and until (goodness of fit) pass.

Robust estimation method

The method is the latest method for blunders detection, that was obtained by scientist Huber in 1981 [Huber, 1981] and developed in 1990's [Francis, 2004]. The basics of this method are depending upon the original weights of observations.

Ordinary Least square adjustment is not sensitive to the blunder in the observation. From the Figure (1) below, it could be noticed how the regression line of least square dropped to the blunder points ,and affect the regression line by this amount ,but if we use the method of robust estimation that resistant to blunder the problem will be very different .

The regression line by both ordinary least square method, and robust line fit shown, and it could be noticed how the blunder point could not affect the regression line of the robust estimation method, and how affects on the least square line regression.

Robust estimation method will treat the blunder from a new view of point, which will compute a weight through a special function upon the scientific function which detect blunder and give a correct result, then treat the blunder.

One can notice how the blunder points dropped the least square line towards the blunder points, how the regression line of robust estimation is canceled the blunder point and how resistant the dropping of the blunder point and did not let it to affect the line of regression.

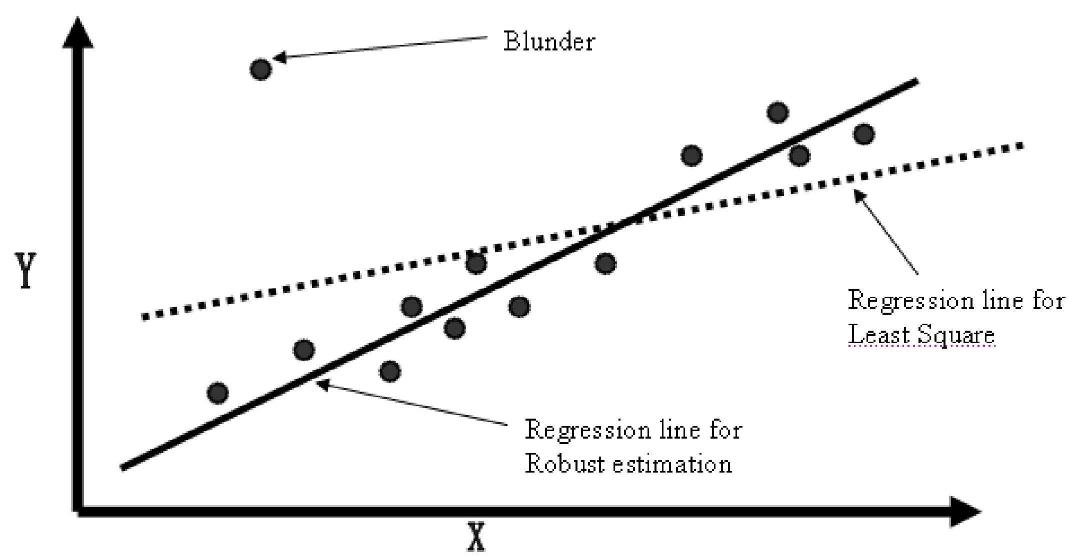


Fig (1). robust regression line and ordinary least square regression line (with one blunder).

Specifies a weight function, a tuning constant, and the presence or absence of a constant term. The weight function can be any of the names listed in the following Table (1).

Table (1). Weight function used in robust estimation function**.**

Tune constant	Meaning	Weight function
1.339	$W = (abs(r)) \times \sin(r)/r$	Andrew
4.685	$W = (abs(r)) \times (1 - r^2)^2$	Bisquare
2.385	$W = 1/(1 + r^2)$	Cauchy
1.400	$W = 1/(1 + abs(r))$	Fair
1.345	$W = 1/(abs(r))$	Huber
1.205	$W = \tanh(r)/r$	Logistics
2.795	$W = 1/(abs(r))$	Talwar
2.985	$W = \exp(r^{-2})$	Welsech

Procedure for robust estimation method could be summarized as follow:

1) Solve as ordinary least square method

2) Solve for hat matrix $\hat{H}$ (covariance of observations)

Where:

$$H = B \times (B^t \times B)^{-1} \times B^t \tag{3}$$

3) Compute r

$$r_i = \frac{V_i}{tune \times s \times \sqrt{(1 - h_{ii})}} \tag{4}$$

Where:

V_i : the residuals of observation i,

tune: the tuning constant from table (1),

s: an estimate of the standard deviation of the error term. s = MAD/0.6745, and

h_{ii} : the vector of leverage values (diagonal element of the hat matrix H).

The quantity MAD is the median absolute deviation of the residuals. The constant 0.6745 makes the estimate unbiased for the normal distribution.

4) If the value of r_i is greater than rejection level, then this observation is considered as a blunder, and it must be down weighted to a new weight according to the weight function in the table above.

5) Readjust according to the new weight.

F-T TEST METHOD

It is one of the traditional methods for detecting blunders; it was presented by the researcher Xu Pieliang. The basic of this method depends upon the tests *F-t*. Initially *F* test used for the global test and then *t* test to check for each blunder immediately after an *F* test. The *U* statistics here is actually the χ^2 statistics given by Stefanovic [Xu Pieliang, 1987].

Compared with an *F* test or (χ^2) test, this method is convenient for testing each blunder after the global test, and makes it possible to discuss the relationship of significance levels between the global test and the test of each blunder.

Let the mathematical model be:

$$L = B\Delta + V \qquad (5)$$

Where

B: the design matrix,
L: the observational vector with the weight matrix W,
Δ: the vector with *u* unknown parameters, and
V: the normally distributed error vector.

If there are blunder in (l), eq. (5) is rewritten as:

$$l_1 = B_1\Delta + V_1 \qquad (6)$$

$$l_2 = B_2\Delta + \Delta l_2 + V_2 \qquad (7)$$

Where:

B_1 : a matrix with full column rank,
l_1: a no-blunder observational error,
l_2: may be considered to be a vector containing some blunders (Δl_2) after the initial identification, and
V_1 and V_2 are the normally distributed error vectors of (l_1) and (l_2), respectively.

The ordinary least square solutions are:

$$\Delta = (B_1^t W_1 B_1)^{-1}(B_1^t W_1 l_1) \qquad (8)$$

$$\Sigma_{xx} = (B_1^t W_1 B_1)^{-1}\ \sigma^2 \qquad (9)$$

$$\hat{\sigma}^2 = V_1^t W_1 V_1 / (n - m - u) \qquad (10)$$

$$V_1 = B_1\Delta + l_1 \qquad (11)$$

Where

W_1: the weight matrix of (l_1), and

n: total number of observations,

The predicted vector ($\hat{l}_2$) of (l_2) is:

$$\hat{l}_2 = B_2.\hat{\Delta} = B_2(B_1^t.W_1.B_1)^{-1}B_1^tW_1l_1 \quad (12)$$

For each element ($\hat{l}_2$) of (l_2) (i > *n-m*):

$$\hat{l}_i = B_i(B_1^tW_1B_1)^{-1}B_1^tW_1l_1 \quad (13)$$

Where

B_i: a row vector of B_2.

The predicted residuals vector is therefore (denoted by $\hat{V}_2$)

$$\hat{V}_2 = l_2 - \hat{l}_2 = l_2 - B_2\Delta \quad (14)$$

For each element of $\hat{V}_2$

$$\hat{V}_i = l_i - B_i\Delta \quad (15)$$

If (l_1) and (l_2) are unrelated, (l_2) is unrelated to Δ, thus we can get from eq. (14) and (15)

$$\Sigma_{vv} = D(l_2) + D(\hat{l}_2)$$

$$\sum_{vv} = \left[W_2^{-1} + B_2(B_1^tW_1B_1)^{-1}B_2^t\right]\sigma^2 \quad (16)$$

$$\Sigma_{vv} = [W_i^{-1} + B_i(B_1^tW_1B_1)^{-1}B_i^t]\sigma^2 \quad (17)$$

To test if blunders exist in $\hat{V}_2$ we test the zero hypotheses:

$$H_o = \Delta L_2 = 0 \quad (18)$$

When a priori value of the unit weight variance is known, the quadratic form:

$$U = \hat{V}_2^t\Sigma_{vv}^{-1}\hat{V}_2 \sim \chi^2(\delta) \quad (19)$$

If eq. (18) is correct the noncentral parameter (*δ)* of *U* is equal to zero; otherwise,

$$\delta = \Delta L_2^t D^{-1}(\hat{V}_2)\Delta L_2. \quad (20)$$

In fact statistics U is Stefanovic's χ^2 tests.

Now we further establish two statistics, denoted by *F* and *t*, respectively [Xu Pieliang, 1987]:

$$F = U/m/\hat{\sigma}^2/\sigma^2$$

$$F = \hat{V}_2{}^t[W_2{}^{-1} + B_2(B^t{}_1W_1B_1)\ {}^1B_2{}^t]\ {}^1\hat{V}_2/m\hat{\sigma}^2 \quad (21)$$

~F (m, n-m-t)

Using eq. (21) we can conveniently test whether (l_2) contain some blunders. When the zero hypotheses are rejected, further testing is needed to determine which (l_2) is responsible.

Therefore establish the zero hypotheses for each elements of (l_2):

$$H_o = \Delta L_2 = 0 \qquad (22)$$

It is clear that when eq. (21) is correct,

$$t = \hat{v}\sqrt{\Sigma_{vv}} \,/\, \sqrt{\hat{\sigma}^2 / \sigma^2} \qquad (23)$$

~t (n-m-t)

After the *F* test, the *t* test can therefore be used to detect blunder in (l_2). It is not difficult to see that a (*t*) test makes it easy to test each element (l_i) of (l_2).

The steps according to this method may be summarized as follow:

1) Initial identification of blunders.

2) Division of model in eq. (5) into two parts–the adjustment model in eq.(6) and the prediction model in eq.(7) according to the initial identification and then solving for Δ in eq. (8), the predicted residuals vector $\hat{V}_2$ from eq. (16) and its variance-covariance matrix.

3) Use of the *F* test for the global test. If the zero hypotheses are rejected, further investigation (a *t* test) is needed to determine which (l_i) is responsible.

APPLICATION TO ACTUAL NETWORK:

For mathematical verification, a combined (Hybrid) geodetic net were chosen, which was adjusted by [Ghilani, 1994]. The systematic errors are corrected and the observed distances are reduced to mean see level. Goodness of fit test for this net after ordinary least square adjustment was failed which means that blunders exist, so the net was suitable to check our methods for blunders detection.

The specifications of this network were as follows [Ghilani, 1994]; (see appendix A)

-Two control points (2000, 2001).

-11 unknown points.

-19 distance observations, and 17 angles. Fig (2) shows the configuration of the net.

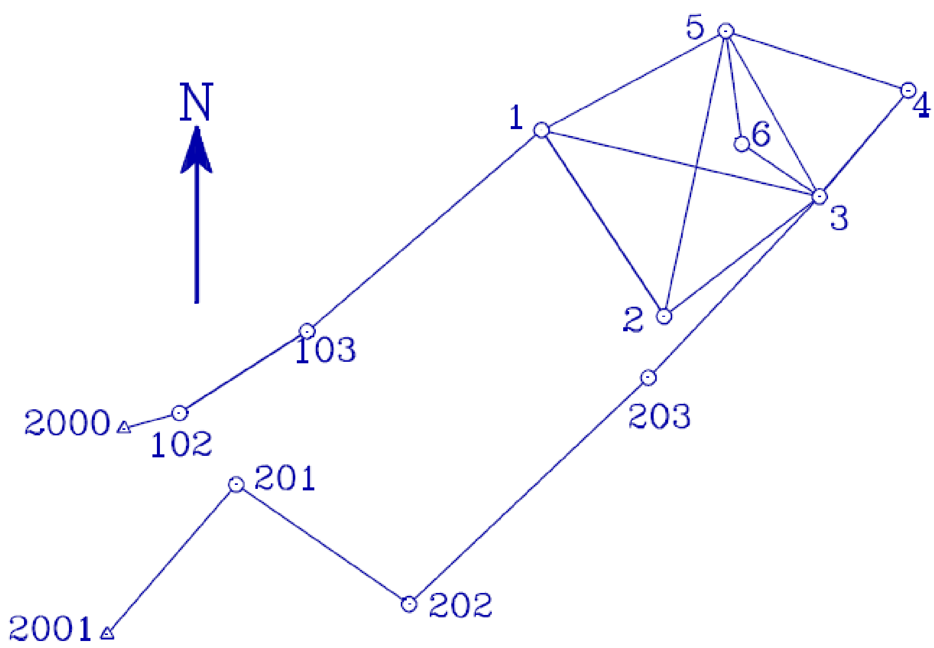


Fig (1). Configuration of the geodetic net

RESULTS AND DISCUSSIONS

After performing the least square adjustment, each method was checked for blunders detection in the observation. The results had been summarized in the Table (2)

Table (2) Comparison between blunder detection methods used for adjusting the geodetic network.

Method	Iteration	χ^2 test	±σ	Blundered observations	Redundancy
Standardized residuals method	3	Pass	1.162	2	12
Robust estimation method	1 (One step)	Pass	0.983	2	14
F-*t* test method	1 (Two steps)	Pass	1.165	2	12

The adjustment of the above network by using 1st method (Standardized residuals) was done by [Ghilani, 1994], while this research adjust the same network by using the two other methods.

It was found that the robust estimation method represent the better blunder detection technique due to its ability in what is called (multi-blunder detection), and the resulted higher accuracy indices, this is obvious from the comparing the values of ($\pm\sigma$).

CONCLUSIONS

It could be concluded the following:

1. It is clear that standardized residuals method, Robust estimation method and F-t test method are all effective in blunder detection.
2. A disadvantage on standardized residuals method was noticed, in that it is an iterative procedure or a single blunder detection technique. Robust estimation and F-t test has the advantage in that it is a multiple blunder detection techniques.
3. From the two applications discussed, it is quite evident that Robust estimation technique gives a lower value for the final variance of unit weight ($\hat{\sigma}_0^2$) and also for the standard deviations of the adjusted unknown, which indicates that Robust estimation is the best blunder detection method.

   This lower value of variance could be explained in that the redundancy (n-u) remain fixed while in the standardized residuals method and F-t test method the redundancy decreased each time a blunder detected and removed

4. Among all the weight functions used in robust estimation method it is highly recommended to use the (Cauchy) weight function.

REFERENCES

Baarda, W. (1968), A Testing Procedure for Use in Geodetic Networks, Netherlands Geodetic Commission Publ. On Geodesy, New series 2, No. 5, Delft 1968

Francis, W.O.Adoul (2004), Robust Geodetic Parameter Estimation under Least Square through Weighting on the Basis of the Mean Square Error, Netherlands Geodetic Commission, publication on Geodesy ,New Series ,Vol.2,No.5,Delft, 2004.

Huber, P.J., (1981), Robust Statistics, John Wiley& Sons, New York.

Wolf & Ghilani (1994), Adjustment Computations, John Wily & Sons, New York.

Wolf, Paul R.(1982), Elements of Photogrammetry, 2nd edition, John Wily & Sons, New York

Xu Pieliang (1987) A test method for many outliers, ITC journal, No.4, 1987.

APPENDIX (A)

Blunder detection example
===========================

Number of Control Stations = 2
Number of Unknown Stations =11
Number of Distance observations = 19
Number of Angle observations = 17

Initial approximations for unknown stations

Station	X	Y
1	2 477 233.88	420 353.62
2	2 477 497.99	419 951.98
3	2 477 832.67	420 210.17
4	2 478 023.86	420 438.88
5	2 477 630.64	420 567.44
6	2 477 665.36	420 323.31
102	2 476 454.17	419 743.39
103	2 476 728.88	419 919.69
201	2 476 576.25	419 589.24
202	2 476 948.76	419 331.29
203	2 477 463.90	419 819.56

Control Stations

Station	X	Y
2000	2 476 334.60	419 710.09
2001	2 476 297.98	419 266.82

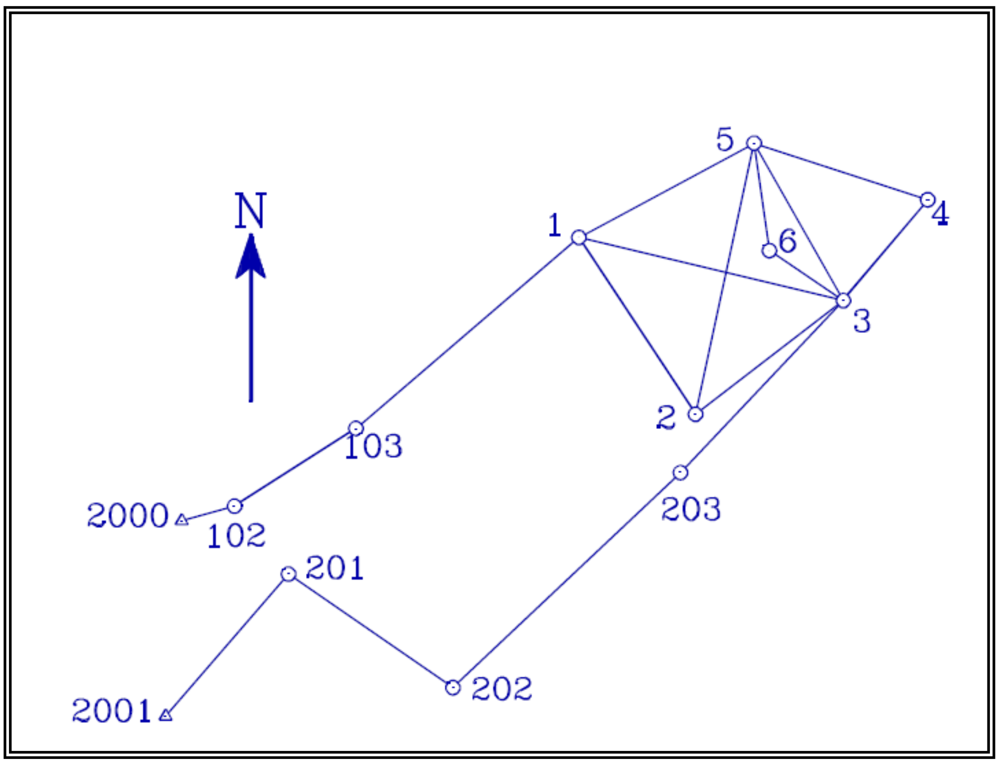
Distance Observations

Station Occupied	Station Sighted	Distance	S
1	3	615.74	0.02
1	2	480.71	0.02
3	1	615.74	0.02
3	4	298.10	0.02
3	6	201.98	0.02
3	5	410.44	0.02
3	2	422.70	0.02
5	2	629.58	0.02
5	1	450.67	0.02
5	6	246.61	0.02
5	4	397.89	0.02
5	3	410.46	0.02
102	103	327.37	0.02
103	1	665.79	0.02
201	202	453.10	0.02
202	203	709.78	0.02
203	3	537.18	0.02
2000	102	125.24	0.02

2001	201	425.90	0.02

Angle Observations

Station Back sighted	Station Occupied	Station Fore sighted	Angle	S
2	1	3	316° 48' 00.5"	6.3"
2	3	4	167° 32' 28.0"	14.5"
2	3	6	71° 42' 51.5"	15.1"
2	3	5	98° 09' 36.5"	10.3"
2	3	1	51° 07' 11.0"	7.2"
203	3	2	8° 59' 56.0"	6.5"
2	5	3	318° 20' 54.5"	7.0"
1	5	3	268° 49' 32.5"	9.8"
6	5	3	338° 36' 38.5"	10.7"
3	5	4	324° 17' 44.0"	8.1"
2000	102	103	162° 58' 16.0"	28.9"
102	103	1	172° 01' 43.0"	11.8"
2001	201	202	263° 54' 18.7"	9.7"
201	202	203	101° 49' 55.0"	8.1"
202	203	3	176° 49' 10.0"	8.4"
102	2000	2001	109° 10' 54.0"	25.5"
2000	2001	201	36° 04' 26.2"	7.4"



SOLUTION

Adjusted stations

Error ellipse confidence level at 0.950

Station	X	Y	Sx	Sy	Su	Sv	t
1	2,477,236.78	420,351.57	26.138	27.582	95.745	40.382	137.21°
2	2,477,500.02	419,949.06	18.862	32.883	95.031	41.414	158.93°
3	2,477,835.61	420,206.18	22.972	41.726	124.575	38.036	155.07°
4	2,478,007.59	420,410.17	28.839	47.264	145.573	41.617	151.30°
5	2,477,631.63	420,566.15	32.636	36.320	127.266	40.399	138.75°
6	2,477,667.20	420,320.89	26.259	37.620	118.995	39.741	147.78°
102	2,476,455.42	419,742.35	9.931	6.023	27.726	15.494	75.95°
103	2,476,731.25	419,918.44	14.445	13.310	40.980	34.721	120.10°
201	2,476,576.61	419,589.00	8.329	9.077	27.541	19.397	37.62°
202	2,476,948.76	419,330.06	12.410	16.494	46.100	32.569	16.98°
203	2,477,465.47	419,816.79	15.649	30.287	85.777	36.510	163.29°

Adjusted Distance Observations

Station Occupied	Station Sighted	Distance	V	S	Std.Res.	Red.#
1	3	616.23	0.495	5.356	26.05	0.746
1	2	480.95	0.243	5.921	13.29	0.689
3	1	616.23	0.495	5.356	26.05	0.746
3	4	266.81	-31.287	6.525	-1802.59	0.622*
3	6	203.77	1.789	6.889	106.83	0.579
3	5	413.76	3.317	5.059	171.47	0.773
3	2	422.77	0.069	5.459	3.67	0.736
5	2	630.97	1.394	6.012	76.87	0.679
5	1	449.39	-1.281	7.192	-79.16	0.541
5	6	247.83	1.225	7.100	74.85	0.553
5	4	407.04	9.153	7.819	614.84	0.458
5	3	413.76	3.297	5.059	170.43	0.773
102	103	327.25	-0.121	10.056	-17.16	0.103
103	1	665.70	-0.087	10.048	-12.18	0.105
201	202	453.37	0.268	10.526	92.01	0.017
202	203	709.85	0.075	10.051	10.51	0.104
203	3	537.24	0.060	10.056	8.53	0.103
2000	102	125.05	-0.188	10.138	-28.72	0.089
2001	201	425.95	0.048	10.063	6.82	0.102

Adjusted Angle Observations

Station Back Sighted	Station Occupied	Station Fore Sighted	Angle	V	S	Std. Res.	Red. #
2	1	3	316°49'53.8"	113.32"	2334.36	28.07	0.411
2	3	4	167°35'29.5"	181.54"	5764.13	22.07	0.322
2	3	6	71°43'01.5"	10.01"	5660.17	1.05	0.397
2	3	5	97°55'08.5"	-868.02"	2793.81	-101.88	0.684
2	3	1	51°06'18.6"	-52.42"	2462.81	-10.32	0.498
203	3	2	8°59'35.6"	-20.36"	3049.99	-13.36	0.055
2	5	3	318°25'19.8"	265.26"	1871.66	45.52	0.693
1	5	3	268°58'58.7"	566.22"	3246.9	79.45	0.529
6	5	3	338°42'49.1"	370.61"	4309.5	62.85	0.304
3	5	4	322°04'20.0"	-8004.04"	3222.81	-1745.41	0.321*
2000	102	103	162°23'46.9"	-2069.09"	10624.86	-110.49	0.420
102	103	1	171°57'47.3"	-235.73"	5605.07	-112.46	0.032
2001	201	202	263°58'29.6"	250.90"	4536.27	104.48	0.061
201	202	203	101°52'55.9"	180.90"	3608.54	58.03	0.148
202	203	3	176°50'15.3"	65.26"	3819.12	23.14	0.113
102	2000	2001	109°40'20.8"	1766.77"	9348.51	106.51	0.423
2000	2001	201	36°07'53.9"	207.69"	3440.52	104.46	0.072

Adjustment Statistics

Iterations = 4
Redundancies = 14

Reference Variance = 232,981.728
Reference So = ±482.7

Failed to pass X^2 test at 95.0% significance level!
X^2 lower value = 5.63
X^2 upper value = 26.12
Possible blunder in observations with Std.Res. > 1,588
Convergence!

Then the problem be as following:-

Number of Control Stations = 2
Number of Unknown Stations = 11
Number of Distance observations = **18** (It was 19)
Number of Angle observations = 17

After that readjustment will be done.