Robust Estimation of the Outliers in GPS Baseline Components

Yuksel BOZ and Ertan GOKALP, Turkey

Key Words: Outlier, Robust, Estimation, GPS

SUMMARY

Several definitions can be made for the term 'outlier' such as a single observation far away from the rest of the observation set. The outlier is the observation that not includes systematic and gross errors but random errors and differs from the other data with its distribution. It is assumed that geodetic observations free from systematic and gross errors always have random errors which are in normal distribution. If an observation is separated from the other data with respect to its random error which is not in normal distribution, that observation is named 'outlier'.

To date several approaches have been developed to identify the outliers more accurately. The first proposed approaches, conventional methods, used the statistical tests to detect outlying observations and removed them from the data set. The main drawback of such an application is the shape defect of the network. Unlike the conventional methods, the robust estimations do not remove any observation from the network. They reduce the effects of the possible outliers on the adjustment model as making the weights of the outlying observations smaller even zero. In this study, some robust estimations (Danish method, Huber's, Andrew's, and IGGIII M-Estimations) are applied to a 17-point GPS network in order to make decision about outliers and the baseline components ($\Delta X_{i-j}, \Delta Y_{i-j}, \Delta Z_{i-j}$) of that network are taken into account as measurements. The used GPS network consists of 34 baselines ranging from 302 meters to 1684 meters.

In robust estimations, it is crucial to define a proper critical value to form the weight functions. This critical value can be calculated or a constant value, which is defined by the a priori standard deviation of the unit weight, can be assigned to it. In this study, both choices are tested in examining the robust estimations. Danish method and Andrew's M-Estimation with constant critical values yield similar results by making zero the weights of the suspicious observations. All estimations with calculated critical values and IGGIII M-Estimation indicate same results by reducing the weights of the possible outlying observations.

Robust Estimation of the Outliers in GPS Baseline Components

Yuksel BOZ and Ertan GOKALP, Turkey

1. INTRODUCTION

The methods developed for outlier detection are classified in two models. One of these models is "The mean-shift model", which determines the outliers by means of statistical tests, and the other is "The variance-inflation model", which reduces the effects of the outlying observations on the adjustment model assigning to them small weights (Kubik and Wang, 1991). In this study, some robust estimations are applied to the GPS baseline components which are considered as measurements in order to determine the outlying measurements in a GPS network. The robust estimations belong to the second model and they use the iteratively reweighted least squares procedure.

Unlike conventional statistical test methods, there is not the drawback of network distortion in the robust methods. No measurement is removed from the network after the iterations, only the weights are changed. The residuals/standardized residuals are compared with a critical value and robust weights between 0-1 are assigned to them in order to obtain the final weights of the measurements. It is observed that the weights of the outlying measurements reduce, even become zero after the iteration steps. In this way, the robust methods decrease the influences of the outliers on the parameter estimation.

The robust estimations used in this study and the constant critical values connected with them are given below Table 1.

Table 1. The weight functions of the M-Estimations

M-Estimation	Weight Function	Critical Value (c)
Andrews	$w_i = \begin{cases} \sin(v_i/c)/(v_i/c) & \lvert v_i \rvert \le c\pi \\ 0 & \lvert v_i \rvert > c\pi \end{cases}$	$1.5s_0$ - $2s_0$
Huber (Gui and Zhang, 1998; Hekimoglu and Berber, 2003)	$w_i = \begin{cases} 1 & \lvert v_i \rvert \le c \\ \dfrac{c}{\lvert v_i \rvert} & \lvert v_i \rvert > c \end{cases}$	$1.5s_0$ - $2s_0$
Danish (Berberan, 1992; Berberan, 1995)	$w_i = \begin{cases} \exp\left(-\dfrac{v_i^2}{c^2}\right) & \lvert v_i \rvert > c \\ 1 & \lvert v_i \rvert \le c \end{cases}$	$1.5s_0$ - $2s_0$
IGGIII (Yang, et al., 2001)	$w_i = \begin{cases} 1 & \lvert \tilde{v}_i \rvert \le c_0 \\ \dfrac{c_0}{\lvert \tilde{v}_i \rvert} & c_0 < \lvert \tilde{v}_i \rvert \le c_1 \\ 0 & \lvert \tilde{v}_i \rvert > c_1 \end{cases}$	$c_0 = 2.0 \sim 3.0$ $c_1 = 4.5 \sim 8.5$

In Table 1, s_0 is the a priori standard deviation of the unit weight, $\tilde{v}_i$ is the normalized residual and is equal to v_i/s_{vi}. Since it is very crucial to define a proper the value to form the weight functions, the formulation related to median of the absolutes of the residuals with weights can be used to calculate s_0 (Yang et al., 1999).

$$s_0 = \text{med}\left\{\left|\sqrt{P_i}\, v_i\right|\right\}/0.6745 \qquad (1)$$

where "*med*" denotes median, P_i and v_i are the weight and residual of the observation ℓ_i, respectively. After s_0 is obtained, it should not be changed in the iteration steps of the robust estimation.

As stated above, besides assigning constant values to the critical value, it can be calculated and the calculation procedure can be applied as follows:

$$c_i = s_0 \sqrt{Qvv_{ii}} \sqrt{P_{ii}}\, t_{f,1-\alpha_0/2} \qquad (2)$$

Here c represents the critical value, s_0 is the a priori standard deviation, Qvv is the cofactor matrix of the residuals, P is the weight matrix of the observations, f is the degree of freedom, α_0 is the significance level, and t represents the t table.
The desired critical value is calculated by taking the average of the critical values calculated for each observation:

$$c = \sum_{i=1}^{n} c_i / n \tag{3}$$

2. THE WRITTEN COMPUTER PROGRAM AND ITS EXECUTION USING THE DATA OF A GPS NETWORK

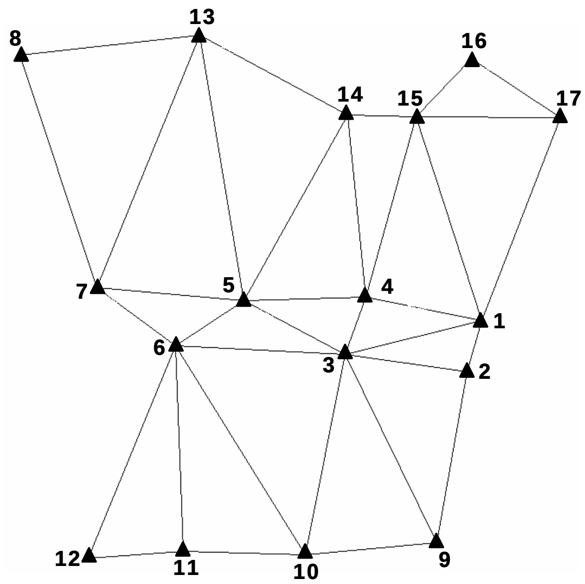

Figure 1. The 17-point GPS network

The GPS network used for outlier investigation consists of 17 points and 34 baselines ranging from 302 meters to 1684 meters (Figure 1). The baseline components of the network have been considered as measurements. In order to adjust the measurements free adjustment has been performed. The adjustment process and the estimation procedures have been written in MATLAB technical computing language (MATLAB, 2001). The input data such as approximate coordinates of the network points, the variance-covariance matrix of the measurements etc. are provided from the computer program GeoGenius (GeoGenius, 2000).

The a priori standard deviation of the unit weight s_0 is obtained as 4.05 mm. from equation (1). The constant critical values used in the robust estimations have been taken into account as $2s_0$, thereby they are obtained as 8.1 mm. The calculated critical value from equation (2) is equal to 6.9 mm.

2.1 The Measurements (Baseline Components) and the Numbers Assigned to Them

Table 2. The baseline components in the GPS network

Baselines		ΔX (m)	MN	ΔY (m)	MN	ΔZ (m)	MN
2	1	-233.7982	1	82.3269	2	172.6048	3
3	1	-166.2416	4	864.5049	5	131.2265	6
3	2	67.5540	7	782.1781	8	-41.3778	9
4	1	143.9372	10	737.2146	11	-90.8069	12
3	4	-310.1788	13	127.2902	14	222.0348	15
4	5	44.0973	16	-773.2000	17	-37.9013	18
3	5	-266.0820	19	-645.9108	20	184.1373	21
6	5	-219.1057	22	433.3547	23	160.2469	24
3	6	-46.9721	25	-1079.2660	26	23.8936	27
7	5	104.9636	28	936.4827	29	-77.6581	30
7	6	324.0713	31	503.1278	32	-237.9039	33
7	8	-1156.9079	34	-470.0131	35	837.9788	36
9	2	-843.1286	37	206.7179	38	623.6820	39
3	10	1033.6841	40	-286.6628	41	-760.4310	42
9	10	122.9529	43	-862.1080	44	-95.4264	45
9	3	-910.7314	46	-575.4424	47	665.0048	48
11	10	18.6660	49	796.8231	50	-10.0875	51
6	10	1080.6767	52	792.6002	53	-784.3027	54
6	11	1062.0076	55	-4.2209	56	-774.2169	57
11	12	46.9800	58	-593.6542	59	-34.1216	60
6	12	1108.9881	61	-597.8759	62	-808.3387	63
13	8	79.7232	64	-1121.1667	65	-61.2690	66
13	7	1236.6348	67	-651.1432	68	-899.2448	69
5	14	-1043.7486	70	674.1378	71	769.5501	72
13	14	297.8383	73	959.4782	74	-207.3410	75
13	5	1341.5867	76	285.3401	77	-976.8956	78
15	14	-54.2960	79	-415.9632	80	37.0028	81
4	14	-999.6738	82	-99.0542	83	731.6512	84
4	15	-945.3777	85	316.9089	86	694.6486	87
15	1	1089.3219	88	420.2947	89	-785.4428	90
17	1	1041.4928	91	-490.9504	92	-755.9327	93
17	15	-47.8309	94	-911.2453	95	29.5093	96
15	16	-282.8851	97	428.9426	98	219.5607	99
17	16	-330.7117	100	-482.3059	101	249.0796	102

MN: Measurement **N**umber

2.2 Results with Constant Critical Value

Danish method and Andrews's M-Estimation tend to produce many outliers making zero the weights of the measurements in the iteration steps of the adjustment. In Danish method, the measurements with the numbers 9, 33, 34, 40, 42, 58, 59, 60, 73, 90 and in Andrews's M-Estimation, the measurements 7, 9, 18, 33, 34, 40, 42, 58, 59, 60, 73, and 90 got zero weight after 7 iterations. Huber's M-Estimation caused small variations in almost all measurements after 5 iterations.

2.3 Results with Calculated Critical Value

The weights of the measurements with the numbers 58 and 60 became smaller after 3 iterations in Danish method and after 4 iterations in Huber's M-Estimation. In addition to these measurements, 3rd and 9th measurements got smaller weights after 3 iterations in Andrews's M-Estimation. In IGGIII M-Estimation, only the weight of the 60th measurement became smaller after 2 iterations.

3. CONCLUSIONS AND RECOMMENDATIONS

The behaviors of Danish method and Andrews's M-Estimation tend to reduce the weights of outlying measurements to zero in case of using constant critical values. In this condition, it is a difficult process to obtain the unknown parameters of the adjustment and sometimes it is also impossible to reach a result. All of the estimations with calculated critical value and also IGGIII M-Estimation reduce the weights of suspicious measurements throughout the iterations and indicate almost the same measurements as outliers. Considering this result, it is the better choice to use the robust estimations with calculated critical values in detection of the outliers.

REFERENCES

Berberan, A., 1992. Outlier Detection and Heterogeneous Observations a simulation case study. *Australian Journal of Geodesy, Photogrammetry, and Surveying* **56**, 49-61.

Berberan, A., 1995. Multiple Outlier Detection. A Real Case Study. *Survey Review* **33**, 255, 41-49.

GeoGenius, 2000, Spectra Precision Terrasat GmbH.

Gui, Q., Zhang, J., 1998. Robust Biased Estimation and Its Applications in Geodetic Adjustments. *Journal of Geodesy* **72**, 430-435.

Hekimoglu, S., Berber, M., 2003. Effectiveness of Robust Methods in Heterogeneous Linear Models. *Journal of Geodesy* **76**, 706-713.

Kubik, K., Wang, Y., 1991. Comparison of Different Principles for Outlier Detection. *Australian Journal of Geodesy, Photogrammetry, and Surveying* **54**, 67-80.

MATLAB, 2001, The Language of Technical Computing, Version 6.1.0.450 Release 12.1.

Yang, Y., Cheng, M. K., Shum, C. K., and Tapley, B. D., 1999. Robust Estimation of Systematic Errors of Satellite Laser Range *Journal of Geodesy* **73**, 345-349.

Yang, Y., Song, L., and Xu, T., 2001. Robust Estimator for the Adjustment of Correlated GPS Networks. *First International Symposium on Robust Statistics and Fuzzy Techniques in Geodesy and GIS* March 12-16, Zurich / Switzerland.

BIOGRAPHICAL NOTES

Yuksel Boz is a Ph.D. student at Karadeniz Technical University (KTU), Turkey. He graduated from the Department of Geodesy and Photogrammetry Engineering at Karadeniz Technical University (KTU) in 2002. He got his M.Sc. degree from the same department in 2005. His interest areas are Satellite Geodesy and GPS. He is a member of Chamber of Surveying Engineers.

Ertan Gokalp is an associate professor at Karadeniz Technical University (KTU), Turkey. He graduated from the Department of Geodesy and Photogrammetry Engineering at KTU in 1986. He got his M.Eng. degree from the Department of Surveying Engineering at University of New Brunswick (UNB), Fredericton, Canada in 1991. He got his Ph.D. degree from the Department of Geodesy and Photogrammetry Engineering at KTU in 1995. He is currently working at the Department of Geodesy and Photogrammetry Engineering at KTU. His interest areas are GPS (Global Positioning System), Engineering Surveying, and Satellite Geodesy. He is a member of Chamber of Surveying Engineers.

CONTACTS

Yuksel Boz
Karadeniz Technical University
Department of Geodesy and Photogrammetry Engineering
61080 Trabzon
TURKEY.
Tel. + 90 462 3772760
Fax: +90 462 3280918
E-mail: yboz@ktu.edu.tr

Ertan GOKALP
Karadeniz Technical University
Department of Geodesy and Photogrammetry Engineering
61080 Trabzon
TURKEY.
Tel. + 90 462 3772770
Fax: +90 462 3280918
E-mail: ertan@ktu.edu.tr