Outlier measurement analysis with the robust estimation

Article *in* Scientific Research and Essays · April 2010

1 author:

Yasemin Sisman
Ondokuz Mayıs Üniversitesi

Scientific Research and Essays Vol. 5 (6), pp. xxx-xxx, 18 March, 2010
Available online at http://www.academicjournals.org/SRE
ISSN 1992-2248 © 2010 Academic Journals

Full Length Research Paper

Outlier measurement analysis with the robust estimation

Yasemin Sisman

Department of Geomatics Engineering, Ondokuz Mayıs University, 55139 Kurupelit, Samsun, Turkey. E-mail: ysisman@omu.edu.tr.

Accepted 18 February, 2010

Consistent and outlier measurements coexist in the measurement group in the applicable sciences. The adjustment calculus, is made to obtain the nearest solution for real and is detached measurements as consistent or outlier. Various methods are being used in order to determine the outlier measurements in the measurement group. The conventional solution methods and making solutions in accordance with the Least Squares Method, determine the outlier measurements, however, they have some disadvantages like the fact that the corrections are much effected by the errors, they spread the measurement errors to other corrections of measurements, not more than one outlier measurement can be determined in each solution step and that they remove the measurement which is determined to be outlier from the measurement group depending on the structure of the aim function. These disadvantages of the conventional solution method have brought out the search for other methods in order to determine the outlier measurement groups. The robust estimation method presents a solution method compared to the aim function being less affected by the measurement errors. The robust estimation method makes an iterative solution by redetermining the measurement weights according to weight function gained from the aim function being less affected by measurement errors in the solution made in accordance with the Least Squares Method. A few different estimation methods are being used in redetermining the measurement weights. In this study, the methods used for redetermining the measurement weights are explained by expressing the reasons to use the robust estimation methods in the outlier measurements analysis. In order to match the theoretically explained methods, an application has been made by using the real network data and the applicability of these methods has been searched.

Key words: Outlier measurement, conventional methods, robust estimation, weight function.

INTRODUCTION

In the applicable sciences, increasing the accuracy of the measurements and the results of these measurements and making more measurements than necessary in order to increase the reliability is the principle. In order to get a significant result from the measurements made more than necessary, the adjustment calculation is done together with evaluating those measurements in accordance with an aim function. The selected aim function is usually the Least Squares Method (LSM). As LSM is simpler and does not require complex statistical information compared to other estimation methods and as it can be applied with its variance-covariance with the average of the measurements, it has the opportunity to be commonly used in the applications.

The measurements determineted in various ways are included with errors that can be classified as gross, systematic and random. Gross and systematic errors could be partially or completely purified from the measurement group, but it is too hard to determine the gross and systematic errors close to random errors in the measurement group and purify them from the measurement group. Random measurement errors do not highly affect the distribution of the measurements as the positives are equal to negatives, they adapt the normal distribution and the number of the ones with small value is greater than those with big value when the measurement number is absolute. As the random error-sized errors from the gross and systematic errors are unilateral, these measurements corrupt the coherence in the measurement group and cause outliers. The measurements that are in a different distribution than the distribution of the measurement groups are called the outlier measurement (Huber, 1964).

The mathematical model of the adjustment calculus consists of functional and stochastic models in such a way to reflect the relationship between the measurements and unknowns. The linearized functional model is given as:

$$\underline{v} = \underline{A}\underline{x} - \underline{\ell} \quad ; \quad \underline{x} = \left(\underline{A}^T \underline{Q}_{\ell\ell}^{-1} \underline{A}\right)^{-1} \underline{A}^T \underline{Q}_{\ell\ell}^{-1} \underline{\ell} \qquad (1)$$

In Equation 1, $\underline{x}$ is unknows, $\underline{v}$ is correction of measurements, $\underline{A}$ is the design matrix, $Q_{\ell\ell}$ is the inverse weights matrix of measurements. $\underline{v}$ corrections of measurements are formed by the errors of other measurements depending on the functional model as well as the random errors of the measurement and they show the differences with real value. In that sense, $\underline{v}$ vector is called adjustment residual instead of measurement corrections. Information about measurements could be gained and outlier measurements could be determined by analysing the $\underline{v}$ vector with special test methods (Thomson, 1976; Vanicek and Wells, 1972).

The validity of the adjustment calculation depends on performing the mathematical model accurately and completely. Whether the functional and stochastic models are applicable with the geometrical and physical relationships between the measurements and the unknowns or not and whether they reflect the sensible relationships between the observations accurately or not, they are examined by the test of the model hypothesis (Vanicek, 1972). In the application, the apriori variance σ_0^2 before the adjustment and the aposteriori variance s_0^2 after the adjustment are compared for the test of the model hypothesis which is also called the Global Test. In the comparison done with linear hypothesis test, the test value (T) is calculated in order to compare with the critical value (q) calculated from the related distribution.

The mathematical model set for adjustment calculation in the case of $T \leq q$ provides the geometrical and physical relationships between the measurements and the unknowns and the correlation between the sensibility of the measurements.

The mathematical model set for adjustment calculation in the case of $T > q$ is not valid. The invalidity of the mathematical model may be due to the gross errors in one or more measurement/s, the inefficient determination of the measurement weights (the wrongly-set stochastic model) or the inefficient determination of the geometrical and physical relationship between the measurements and unknowns (the wrongly-set functional model) (Baarda, 1968).

In the case where the mathematical model is not valid, in order to comprehend the existence of $\nabla\ell$ errors in the measurements, firstly the functional model is tested. The test of the functional model and the stochastic model is tested with the outlier measurements analysis in order to determine whether there is a gross error in the measurements used in the adjustment calculation or not, and whether the measurement weights are determined well or not. A few approaches have been used in order to determine the outlier measurements so far. The conventional outlier measurement test method based on LSM has commonly been used in geodetic studies for many years. The conventional solution method was explained by Baarda (1968) and then the iterative solution was developed (Koch, 1999). Due to some disadvantages of this method, the studies to determine the outlier measurements with the robust estimation method have started in recent years.

An alternative method in determining the outlier measurements is the robust statistics and the robust estimation method. The robust estimation is an approximate parametric method which is not affected by the small changes in the distribution functions of the measurements and the gross errors. Although the term “robust” was first used by Box (1953) in a study, its theory was then explained by Huber (1964) and developed for various cases (Andrews et al., 1972; Huber 1981). There are many studies in which the outlier measurement analysis is made in geodetic networks by applying this method (Krarup et al., 1980; Fuchs, 1982; Xu, 1989; Yang, 1991; Krauss, 1992; Harvey, 1993; Yang, 1999; Wieser and Brunner 2001, 2002; Hund et al., 2002; Hekimoglu and Erenoglu, 2007; Daszykowski et al., 2007; Erenoglu and Hekimoglu, 2009).

DETERMINING THE OUTLIER MEASUREMENTS

As a result of various errors in the measurements, outlier measurements could be formed. In evaluating the measurements which are done for geodetic studies, the determination of outlier measurements is significant in terms of reliability and quality. Not all of the outlier measurements are the bad measurements caused by the gross or systematic errors, in some cases, these measurements could be very important for the measurement group. The attitude to determine the outlier measurements reliably and rapidly is another problem. The frequency and the size of the gross or systematic errors could be evaluated from the information related to the reliabilty of the data. If the model is set finely and the majority tendency of the data is taken into consideration, the outlier measurements could directly be removed from the measurement group without an evaluation, however in this case, the information consisting of these measure-

Table 1. Conventional outlier measurements analysis methods.

Methods	Data-Snooping	Tau-Test	t-Test
Test Value	$W_i = \frac{\lvert v_i \rvert}{\sigma_0 \sqrt{Q_{v_i v_i}}}$	$T_i = \frac{\lvert v_i \rvert}{s_0 \sqrt{Q_{v_i v_i}}}$	$t_i = \frac{\lvert v_i \rvert}{s_{01} \sqrt{Q_{v_i v_i}}}$
Test Distribution	$\sim N(0, 1)$	$\sim \tau_{f,(1-\alpha_0/2)}$	$\sim t_{f-1,(1-\alpha_0/2)}$
Critical Value	$N_{(1-\alpha_0/2)}$	$\tau_{f,(1-\alpha_0/2)}$	$t_{f-1,(1-\alpha_0/2)}$

ments are abandoned (Hampel, 1986). Therefore the removal of the outlier measurement from the measurement group may not always be an appropriate solution. Two main approaches are used in determining the outlier measurements. The first approach includes taking the functional and stochastic models of the adjustment model as constant and determining the outlier measurements by comparing the test value determined by a result of the adjustment calculation with a critical value. This solution is called the conventional solution method (Baarda, 1968). The second approach in the outlier measurements test is changing the weights of the measurements (or stochastic model) by taking the functional model and the measurement group as constant. In this solution, the adjustment calculation is renewed by redetermining the measurement weights in accordance with an aim function. This solution is called the reweighted robust estimation solution (Simpson and Chang, 1997; Franklin and Brodeur, 1997; Wilcox, 1997; Koch, 1999).

Conventional solution methods

LSM method is a parametric method. Like all other applicable sciences, geodesy, too, is commonly used with LSM since the conventional solution methods are easily applied and the mathematical model remains the same till the end of the solution. When using LSM, it is accepted that only the measurement group consisting of the random errors complies with the normal distribution. It is very hard to get the condition that provides these acceptances from the real measurements. By LSM and the mathematical model set in the solution, the residuals determined by a result of the solution are affected by the errors of all measurements depending on the functional model. In this case, while a measurement which is not outlier with its own measurement error may be considered as outlier by being effected by the errors of other measurements, an outlier measurement may be considered as consistent. The parametric model, which is a close approach to real gives information about the bounds to be determined for the analysis of the data in the normal distribution and does not give information about how much the data is far from these bounds or the success of the estimates. As a result of the adjustment calculation, for the analysis of determining whether there is $\nabla\ell$ gross or systematic error in the measurements or not, a linear hypothesis test is carried out as:

$$H_0 : E\{\nabla\ell_i\} = 0; \qquad H_S : E\{\nabla\ell_i\} = \nabla\ell_i \neq 0 \tag{2}$$

by setting null and alternative hypothesis. In the outlier measurement analysis, test values for each measurement are calculated by first using $\underline{v}$ residuals of measurements. The test value is compared with the critical value determined from the table in which the distribution of the measurements complies. If there are test values over the critical value, the measurement with the highest test value is accepted to be outlier and is removed from the measurement group. With the newly formed measurement group, adjustment calculation and outlier measurement analysis proceedings are repeated. This proceeding is carried out until the entire test values are under the critical value. Three different approaches are used in conventional solution methods. These approaches are Data-Snooping (Baarda), Tau and t (student) test. All of these methods make solution in accordance with the same principles. The variance values they use in the solution and the distribution table of the measurements are different depending on these values. The test value, critical value and distribution information are given in Table 1, α_0 is the significance level, f is the freedom degree, s_{01}^2 is the posteriori variance eliminated from the model errors, and N, τ, t represent normal, tau and student distribution, respectively (Baarda, 1968; Koch, 1999; Valero and Moreno, 2005).

The robust estimation method and determining the outlier measurements

Robust statistic is a branch of science that deals with the estimates of assumptions like normality and linearity that are used commonly in statistic science. Robust statistic is a method that can efficiently be used in determining statistical outlier measurements. In the solution by the robust estimation method, the measurements are not effected by their own and other measurement errors, the corruptive effects of the measurement errors on the results are decreased and they are even destroyed (Huber, 1964). The classical statistical methods based on LSM that makes solutions with the acceptance that the measurement group is in normal distribution and is purified from the gross and systematic errors, the results is significant only when this approach is right. These models are quite weak especially against the little deviations of the measurement group (Hampel et al., 1986). Additionally, the statistical tests, being carried out according to the results of the adjustment calculation in which inevitable outlier measurements are made, are indirectly effected by these errors. While the robust estimation analysis the outlier measurements, it also decreases the effects of the outlier measurements on the results at the same time (Valero and Moreno, 2005).

In the robust estimation, efficient results could be obtained by the solution of the reweighted of the measurements that are the second approaches in determining the outlier measurements. In this solution, instead of LSM of an aim function, the minimum total of the residuals squares ($\underline{v}^T \underline{P}\underline{v}=$ min.) the aim function is less affected by the errors of the residuals are taken, and weight function is obtained by the robust estimation. If the weight function values that are gained by the aim function in the robust estimation are solved according to LSM, the solution will be made by reducing the robust estimation algorithm to LSM algorithm.

In the robust estimation, the derivative of $\rho(\underline{v})$, the aim function according to $\underline{v}$ determine the $\psi(\underline{v})$, the effect function; the derivative of $\psi(\underline{v})$ effect function according to $\underline{v}$, determines the $W(\underline{v})$, the weight function. In order to get the robust result, all of these functions shall be constant and their boundary shall be definite. Determination of only one of these functions is enough to determine others and for solution (Pilgrim, 1996; Yang, 1999). The robustness for the robust estimation functions is qualitative robustness. A qualitative robust estimation can be formed by choosing limited function of the aim function with a non-linear derivative. LSM method is not robust because the aim function is quadratic and its derivative is linear and limitless. The robust estimators can be classified as:

1. M-estimators or maximum likelihood estimators.
2. L-estimators or linear combination of order statistics estimators.
3. R-estimators or rank-test derived estimators.

Among these, M-estimators are the most flexible and the easiest to generalize to multiparameter cases. M-estimators which are also developed by Huber (1964) makes resolution in accordance with the principle of minimization of a function of the residuals. If the probability function of the measurement group which is a linear functional relationship between the measurements and the unknowns is taken as $F_{(\underline{x},\underline{\ell})}$, M-estimator is defined as unknown values $\underline{x}$ that maximize the multiplyings and are given in Equation (3).

$$L_{(\underline{x})}=\prod_{i=1}^{n} F_{(\underline{x},\underline{\ell})} \;;\; Log L_{(\underline{x})}=-\sum_{i=1}^{n} Log\, F_{(\underline{x},\underline{\ell})}=\sum_{i=1}^{n}\rho_{(\underline{x},\underline{\ell})} \qquad (3)$$

Here, the solution to make the total probability function maximum and to make the aim function minimum is sought. The generalized M-estimator can be written as:

$$\mathrm{M}=\sum_{\mathrm{i}=1}^{\mathrm{n}}\rho_{(\underline{\mathrm{x}},\underline{\ell})}=\sum_{\mathrm{i}=1}^{\mathrm{n}}\frac{\partial\rho(v_{\mathrm{i}})}{\partial v_{\mathrm{i}}}\frac{\partial v_{\mathrm{i}}}{\partial_{\mathrm{j}}}=\sum_{\mathrm{i}=1}^{\mathrm{n}}\psi(v_{\mathrm{i}})a_{ij}=0 \qquad (4)$$

by taking the function which is set in the solution with LSM into consideration according to Equation (1). When Equation (4) is solved,

$$\underline{A}^T\psi(\underline{v})=\underline{A}^T\psi(\underline{A}\underline{x}-\underline{\ell})=\underline{A}^T\underline{W}(\underline{A}\underline{x}-\underline{\ell})=0 \qquad (5)$$

Equation (5) can be written (Yang, 1999). The normal equation system of the M-estimation is non-linear. Therefore to solve this system iteratively, reweighted LSM estimation is used (Koch 1999).

$$\underline{x}=\left(\underline{A}^T\underline{W}\underline{A}\right)^{-1}\underline{A}^T\underline{W}\underline{\ell} \qquad (6)$$

By this way, iterative and reweighted solution can be made with the Equation (7) by LSM.

$$\underline{x}_t=\left(\underline{A}^T\underline{W}_t\underline{A}\right)^{-1}\underline{A}^T\underline{W}_t\underline{\ell} \quad ; \quad \underline{v}_t=\left[\underline{A}\left(\underline{A}^T\underline{W}_t\underline{A}\right)^{-1}\underline{A}^T\underline{W}_t-\underline{E}\right]\underline{\ell}$$
$$\underline{W}_t=\underline{P}\underline{W}_{(t-1)} \quad t=1,2,\ldots.. \quad ; \quad \underline{W}(v_0)=\underline{E} \qquad (7)$$

Here, the iteration number shows the weight function chosen as $\underline{W}$. For the beginning $\underline{W}=\underline{E}$ unit matrix and the solution can be summarized as defining the $\underline{W}$ weight matrix from $\underline{v}$ measurement residuals determined by solution LSM and robust estimation solution is iterative. The solution is made through providing the M-estimator condition which is given in the Equation of LSM and (4). In the Robust M-estimator whose solution has been found by reweighting LSM estimation in Equation (7), convenient weights for each measurement have been determined and a robust solution has been obtained.

As a result of such solution, it is seen that x_t unknowns and W_{t+1} weights of the consistent measurements given in Equation (7) are the same, W_{t+1} weights of the outlier measurements are gradually reduced and they even get closer to zero. In this case, the corruptive effects of the

Table 2. Weight functions in the robust estimation method.

Methods	Critical Value	Weight Function	Methods	Critical Value	Weight Function
Huber	$\|v\| \le c$	1	Yang-II	$\|v\| \le c_0$	1
	$\|v\| > c$	$c/\|v\|$		$c_0 < \|v\| \le c_1$	$\frac{c_0}{\|v\|}$
Andrews	$\|v\| \le c\Pi$	$(\|v\|/c)^{-1}\sin(\|v\|/c)$		$\|v\| > c_1$	0
	$\|v\| > c\Pi$	0	Denmark	$\|V\| \le c$	1
Beaton-Tukey	$\|v\| \le c$	$(1-(\|v\|/c)^2)^2$		$\|V\| > c$	$e^{\left(-\frac{v^2}{c^2}\right)}$
	$\|v\| > c$	0	LMS	None	1
L1-norm	None	$\frac{1}{\|v\|}$	LTS	None	1

outlier measurements on the unknowns are gradually reduced, too. This is one of the most significant features of the robust estimation especially for the analysis of the measures which can not be determined whether to be outlier or not (Caspary and Barutta, 1987; Pilgrim, 1996; Barberan, 1995; Yang, 1999; Valero and Moreno, 2005).

For the solution equations of the generalized M-estimator, several methods have been tried to determine the weight function compared to reweighted iterative LSM. The most favored methods: Huber, Andrews, Danish, Beaton-Tukey, Yang-I and Yang-II, Least Absolute Values(L1-norm), Least Median Squares (LMS), Least Trimmed Squares (LTS) in the application and the weight functions derived in accordance with the aim functions can be seen in Table 2 (Yang et al., 2001; Chen, 2002; Valero and Moreno, 2005; Moller et al., 2005; Hekimoğlu, 2007; Gökalp et al., 2008; Knight and Wang, 2009).

The robust weight factors are obtained by comparing the residuals with critical values derived from calculations or given constant values. In order to calculate the critical value, a procedure can be applied as follows:

$$c_i = s_0\sqrt{Q_{vv_{ii}}}\sqrt{P_{ii}}\ t_{f,1-\alpha_0/2} \qquad (8)$$

Here, c represents the critical value, Q_{vv} is the cofactor matrix of the residuals, P_{ii} is the weight matrix of the observations and t represents the t-distribution. The associate critical value is calculated by averaging the critical values calculated for each observation as:

$$c = \frac{\sum_{i=1}^{n} c_i}{n} \qquad (9)$$

In this study, the critical values of the estimations have been calculated by using Equation (9) (Gökalp et al., 2008).

NUMERAL APPLICATION

In the light of the theoric explanations made, a solution has been made by using the real data of a GPS network whose 20 bases are measured in order to determine the advantages and disadvantages of the methods for determining the outlier measurements (Figure 1). All the measurements of GPS network were performed in June 2008 using GPS receivers. Data reduction and post-processing were carried out using Leica LGO 2.0 software. Adjustment computation was performed accepting the coordinates of Turkish National GPS Network points as stable in the measurement epoch. As a result of the adjustment computation, the coordinates of geodetic points *(φ,λ,h)* and the standard deviations were determined as $(\sigma_\varphi, \sigma_\lambda, \sigma_h)$; $\sigma_\varphi, \sigma_\lambda \le \pm 3.0$ cm, $\sigma_h \le \pm 5.0$ cm, and the level of statistical confidence were accepted as $\beta = 0.95$. Then, coordinates of points *(X, Y, Z)* were calculated using coordinated of points *(φ, λ, h)*.

In determining the outlier measurements, the solution has been made through choosing the t-test from the conventional solution methods and L1-norm, Denmark, Yang-II, Huber, Beaton-Tukey and Andrews methods from the robust estimation methods for reweighted. When outlier measurements analysis was made by using the t-test from the methods for the conventional solution methods, the consistent measurement group was reached at the 10th iteration step and 9 measurements were respectively removed from the measurement group (Table 3).

The weights of the measurements have been redetermined by using the weight functions of L1-norm, Denmark, Yang-II, Huber, Beaton-Tukey and Andrews

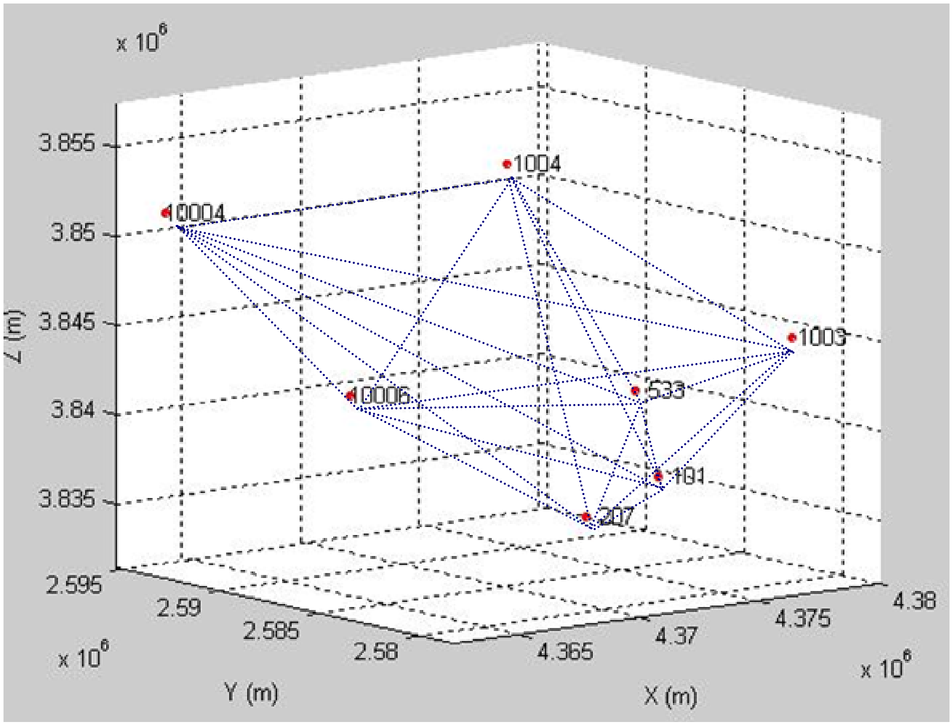

Figure 1. The network selected for the study.

Table 3. Outlier measurements analysis with conventional solution methods

Iteration	**t-Test**				
	Number of Measurement		**Test Value**	**Critical Value**	**Decision**
1	1004-10004	ΔZ	2.710	2.39	☹
2	533-101	ΔZ	2.568	2.39	☹
3	1003-533	ΔX	2.476	2.39	☹
4	1004-101	ΔZ	3.407	2.39	☹
5	1004-10004	ΔY	2.562	2.39	☹
6	533-101	ΔY	2.655	2.40	☹
7	10004-533	ΔX	2.793	2.40	☹
8	1004-533	ΔX	2.726	2.40	☹
9	1003-533	ΔY	2.412	2.40	☹
10	207-1006	ΔX	2.143	2.40	☺

methods for outlier measurement anlysis for the Robust estimation methods. The solution with no significant l change in the measurement weights was reached at the 3rd iteration (Table 4).

DISCUSSION AND CONCLUSION

With the application made, an adjustment calculation has been made by using the baseline measurements in GPS network which has been chosen for the outlier measurement analysis. Then, the outlier measurements have been determined by using the t-test from the conventional methods by using the results of adjustment calculation. The consistent measurement group was reached at the 10th iteration with the conventional solution method and the 9 measurement was removed from the measurement group. Using the results of adjustment calculation, the outlier measurement group was determined by L1-norm, Denmark, Yang-II, Huber, Beaton-Tukey and Andrews estimation functions from the Robust estimation methods. While the weights were

Table 4. Outlier measurements analysis with the robust estimation methods.

Mesurement Number		L1-norm Method				Denmark Method				Yang-II Method				Huber Method				Beaon-Tukey Method				Andrews Method			
İteration		**I.**	**II.**	**III.**	**Decision**	**I.**	**II.**	**III.**	**Decision**	**I.**	**II.**	**III.**	**Decision**	**I.**	**II.**	**III.**	**Decision**	**I.**	**II.**	**III.**	**Decision**	**I.**	**II.**	**III.**	**Decision**
533-101	ΔX	0.217	0.211	0.207	☹	1.000	0.453	0.000	☹	0.217	0.245	0.302	☹	1.000	1.000	1.000	☺	0.120	0.001	0.000	☹	0.893	0.851	0.824	☺
	ΔY	0.150	0.134	0.132	☹	0.255	0.000	0.000	☹	0.000	0.000	0.000	☹	0.855	0.627	0.070	☹	0.000	0.001	0.000	☹	0.783	0.673	0.588	😐
	ΔZ	0.124	0.110	0.108	☹	0.136	1.000	0.004	☹	0.000	0.000	0.000	☹	0.708	0.381	0.065	☹	0.000	0.001	0.000	☹	0.693	0.522	0.380	☹
207-10006	ΔX	0.323	0.473	0.688	☺	1.000	1.000	1.000	☺	0.323	0.486	1.000	☺	1.000	1.000	1.000	☺	0.479	0.001	0.000	☹	0.951	0.934	0.922	☺
	ΔY	0.641	0.662	0.673	😐	1.000	0.825	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.850	0.001	0.000	☹	0.987	0.981	0.975	☺
	ΔZ	1.190	1.030	1.305	☺	1.000	0.917	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.955	0.214	0.000	☹	0.996	0.993	0.990	☺
207-10004	ΔX	0.204	0.182	0.167	☺	1.000	0.228	0.000	☹	0.204	0.178	0.170	☹	1.000	1.000	0.955	☺	0.053	0.001	0.000	☹	0.879	0.823	0.779	😐
	ΔY	1.292	28.170	214.434	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.962	0.985	0.696	😐	0.997	0.997	0.998	☺
	ΔZ	2.047	1.957	1.682	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.985	0.246	0.943	☺	0.999	0.995	0.987	☺
10006-10004	ΔX	1.707	4.510	29.479	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.979	0.898	0.923	☺	0.998	0.998	0.997	☺
	ΔY	1.400	39.693	109.822	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.969	0.900	0.191	☹	0.997	0.998	0.999	☺
	ΔZ	0.201	0.208	0.202	☹	1.000	1.000	1.000	☺	0.201	0.225	0.248	☹	1.000	1.000	1.000	☺	0.056	0.001	0.000	☹	0.875	0.850	0.848	☺
1004-10006	ΔX	0.492	0.671	0.827	☺	1.000	1.000	1.000	☺	0.492	1.000	1.000	☺	1.000	1.000	1.000	☺	0.752	0.001	0.000	☹	0.978	0.970	0.963	☺
	ΔY	0.490	1.260	1.666	☺	1.000	1.000	1.000	☺	0.490	1.000	1.000	☺	1.000	1.000	1.000	☺	0.750	0.816	0.000	☹	0.978	0.976	0.977	☺
	ΔZ	2.778	3.147	2.100	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.992	0.086	0.000	☹	0.999	1.000	0.995	☺
1004-10004	ΔX	2.229	3.724	5.709	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.987	0.905	0.000	☹	0.999	0.998	0.998	☺
	ΔY	0.146	0.114	0.111	☹	0.222	0.000	0.000	☹	0.000	0.000	0.000	☹	0.815	1.000	0.219	☹	0.000	0.001	0.000	☹	0.772	0.640	0.532	😐
	ΔZ	0.110	0.101	0.101	☹	0.071	0.000	0.000	☹	0.000	0.000	0.000	☹	0.615	1.000	0.096	☹	0.000	0.001	0.000	☹	0.621	0.388	0.190	☹
1003-207	ΔX	0.195	0.180	0.168	😐	1.000	0.337	0.000	☹	0.195	0.190	0.181	☹	1.000	1.000	1.000	☺	0.023	0.001	0.000	☹	0.867	0.816	0.780	😐
	ΔY	0.512	0.772	0.865	☺	1.000	0.827	1.000	☺	1.000	1.000	1.000	☺	1.000	0.543	1.000	☺	0.770	0.236	0.000	☹	0.980	0.976	0.973	☺
	ΔZ	0.380	0.344	0.310	😐	1.000	0.722	0.263	☹	0.380	0.324	0.343	☹	1.000	0.243	1.000	☺	0.604	0.001	0.000	☹	0.964	0.951	0.943	☺
1003-10006	ΔX	0.972	2.197	4.722	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.936	0.857	0.863	☺	0.994	0.995	0.995	☺
	ΔY	0.824	1.978	2.922	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.911	0.961	0.000	☹	0.992	0.991	0.990	☺
	ΔZ	1.892	8.415	203.027	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.983	0.934	0.825	☺	0.999	0.998	0.998	☺
1003-10004	ΔX	4.118	30.019	45.162	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.996	0.980	0.000	☹	1.000	1.000	1.000	☺
	ΔY	0.654	7.647	20.538	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.861	0.673	0.000	☹	0.988	0.989	0.991	☺
	ΔZ	0.329	0.303	0.283	😐	1.000	1.000	1.000	☺	0.329	0.333	0.443	☹	1.000	1.000	1.000	☺	0.510	0.001	0.000	☹	0.952	0.948	0.955	☺
1003-1004	ΔX	9.459	7.396	10.277	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.999	0.872	0.000	☹	1.000	0.999	0.999	☺
	ΔY	1.211	3.474	3.875	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.957	0.990	0.782	😐	0.996	0.997	0.997	☺
	ΔZ	0.592	1.664	3.043	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.825	0.979	0.537	😐	0.985	0.991	0.997	☺

Table 4. Continued.

Mesurement Number		L1-norm Method				Denmark Method				Yang-II Method				Huber Method				Beaon-Tukey Method				Andrews Method			
İteration		I.	II.	III.	Decision	I.	II.	III.	Decision	I.	II.	III.	Decision	I.	II.	III.	Decision	I.	II.	III.	Decision	I.	II.	III.	Decision
207-533	ΔX	0.194	0.215	0.236	😐	1.000	0.331	0.050	☹	0.194	0.271	0.368	☹	1.000	1.000	0.997	☺	0.020	0.001	0.000	☹	0.866	0.822	0.796	☺
	ΔY	0.552	0.615	0.639	😐	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.800	0.001	0.000	☹	0.983	0.978	0.975	☺
	ΔZ	0.361	0.442	0.485	😐	1.000	1.000	1.000	☺	0.361	1.000	1.000	☺	1.000	1.000	1.000	☺	0.568	0.001	0.000	☹	0.960	0.951	0.947	☺
207-101	ΔX	0.566	0.907	1.624	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.810	0.497	0.961	☺	0.984	0.980	0.977	☺
	ΔY	0.699	1.277	1.274	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.873	0.687	0.000	☹	0.989	0.988	0.987	☺
	ΔZ	2.047	11.444	19.959	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.985	0.774	0.860	☺	0.999	1.000	1.000	☺
10006-533	ΔX	0.234	0.211	0.200	😐	1.000	1.000	1.000	☺	0.234	0.261	0.297	☹	1.000	1.000	1.000	☺	0.191	0.001	0.000	☹	0.907	0.872	0.849	☺
	ΔY	0.412	0.376	0.364	😐	1.000	1.000	1.000	☺	0.412	0.282	0.255	☹	1.000	1.000	1.000	☺	0.669	0.001	0.000	☹	0.969	0.953	0.939	☺
	ΔZ	1.591	139.31	171.20	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.976	0.831	0.991	☺	0.998	0.999	1.000	☺
10006-101	ΔX	0.405	0.358	0.338	😐	1.000	1.000	1.000	☺	0.405	0.393	0.351	☹	1.000	1.000	1.000	☺	0.659	0.001	0.000	☹	0.968	0.955	0.944	☺
	ΔY	0.598	0.932	0.910	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.835	0.586	0.000	☹	0.985	0.984	0.983	☺
	ΔZ	0.203	0.223	0.232	😐	1.000	1.000	0.128	☹	0.203	0.257	0.289	☹	1.000	1.000	1.000	☺	0.063	0.001	0.000	☹	0.877	0.843	0.825	☺
10004-533	ΔX	0.206	0.211	0.214	☹	1.000	0.124	0.000	☹	0.206	0.179	0.159	☹	1.000	1.000	0.883	☺	0.076	0.001	0.000	☹	0.881	0.823	0.774	😐
	ΔY	0.271	0.362	0.376	😐	1.000	1.000	1.000	☺	0.271	1.000	1.000	☺	1.000	1.000	1.000	☺	0.337	0.001	0.000	☹	0.930	0.913	0.906	☺
	ΔZ	0.538	0.728	0.797	☺	1.000	1.000	0.293	☹	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.798	0.001	0.259	☹	0.982	0.970	0.957	☺
10004-101	ΔX	17.500	82.786	310.208 7	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	0.975	0.515	😐	1.000	1.000	1.000	☺
	ΔY	7.000	19.099	94.132	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.999	0.990	0.000	☹	1.000	1.000	1.000	☺
	ΔZ	0.294	0.357	0.363	😐	1.000	1.000	1.000	☺	0.294	1.000	1.000	☺	1.000	1.000	1.000	☺	0.414	0.001	0.000	☹	0.940	0.946	0.961	☺
1004-533	ΔX	0.456	0.439	0.435	☹	1.000	1.000	0.209	☹	0.456	0.339	0.260	☹	1.000	1.000	1.000	☺	0.714	0.001	0.000	☹	0.975	0.960	0.946	☺
	ΔY	0.483	0.947	1.055	☺	1.000	1.000	1.000	☺	0.483	1.000	1.000	☺	1.000	1.000	1.000	☺	0.744	0.311	0.994	☺	0.978	0.974	0.972	☺
	ΔZ	0.218	0.305	0.321	☺	1.000	1.000	1.000	☺	0.218	0.455	1.000	☺	1.000	1.000	1.000	☺	0.105	0.001	0.000	☹	0.893	0.899	0.921	☺
1004-101	ΔX	2.536	1.614	1.374	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.990	0.628	0.000	☹	0.999	0.999	0.998	☺
	ΔY	0.523	15.084	9.877	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.779	0.069	0.000	☹	0.981	0.984	0.988	☺
	ΔZ	0.411	0.455	0.453	☹	1.000	1.000	1.000	☺	0.411	0.490	0.462	☹	1.000	1.000	1.000	☺	0.656	0.001	0.000	☹	0.969	0.966	0.968	☺
1003-533	ΔX	0.143	0.141	0.141	☹	0.220	0.000	0.000	☹	0.000	0.000	0.000	☹	0.813	0.550	0.246	☹	0.000	0.001	0.000	☹	0.762	0.644	0.550	😐
	ΔY	0.177	0.163	0.161	☹	1.000	0.251	0.000	☹	0.177	0.183	0.193	☹	1.000	0.965	0.852	☺	0.000	0.001	0.000	☹	0.840	0.769	0.720	😐
	ΔZ	0.556	0.494	0.523	😐	1.000	1.000	1.000	☺	1.000	0.443	0.299	☹	1.000	1.000	1.000	☺	0.810	0.001	0.000	☹	0.983	0.967	0.949	☺
1003-101	ΔX	1.429	2.245	2.669	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.970	0.997	0.000	☹	0.997	0.998	0.999	☺
	ΔY	0.560	2.090	2.935	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.813	0.907	0.000	☹	0.983	0.984	0.984	☺
	ΔZ	0.603	1.247	1.920	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	1.000	1.000	1.000	☺	0.838	0.457	0.748	😐	0.986	0.986	0.988	☺

Table 5. The results of the outlier measurements analysis methods.

Method	Number of Outlier Measurement	%	Number of Suspicious Measurement	%	Number of Consistent Measurement	%
t-test	9	15.0	0	0	51	85.0
L1-norm	11	18.3	14	23.3	35	58.4
Denmark	15	25.0	0	0	45	75.0
Yang-II	21	35.0	0	0	39	65.0
Huber	5	8.3	0	0	55	91.7
Beaton-Tukey	47	78.3	5	8.3	8	13.3
Andrews	2	3.3	8	13.3	50	83.3

Table 6. Number of common outlier measurements.

Method	Number of Common Outlier Measurements						
	t-test	L1 norm	Denmark	Yang-II	Huber	Beaton-Tukey	Andrews
t-test	9	9	8	9	5	9	2
L1 norm	9	11	9	11	5	11	2
Denmark	8	9	15	14	5	15	2
Yang-II	9	11	14	21	5	21	2
Huber	5	5	5	5	5	5	2
Beaton-Tukey	9	11	15	21	5	47	2
Andrews	2	2	2	2	2	2	2

redetermined by the estimation function in the robust estimation method, the outlier measurements were not removed from the measurement group but their effects on the results were reduced.

In the solution which was made for the robust estimation, the weights of the measurements which were found outlier in Denmark, Yang-II, Huber, Beaton-Tukey and Andrews methods were reduced and got closer to zero. In L1-norm method, however, the weights of the outlier measurements remained the same and the weights of other measurements increased. The results that were acquired by these methods are shown in Table 4. The main difference between the conventional solution methods and the robust estimation methods in determining the outlier measurements is that the measurements in the robust estimation can be determined as suspicious as well as consistent and outlier according to the weight values. The robust estimation method can determine the measurements with a changing measurement weights in some amount, but not being 0 or 1, as "suspicious". In this study, the weight values have been taken as "consistent" for the measurements weight greater than 0.8, "suspicious" for the measurements weight between 0.5 - 0.8 and "outlier" for the measurements weight smaller than 0.5. All the measure-ments which were found outlier with the conventional solution methods could be determined as outlier also in the robust estimation methods, except for Huber and Andrews methods. The results of the methods which were carried out in order to determine the outlier measurements are given in Table 5.

Also, to be able to examine the achievements of the methods in determining the outlier measurements for any outlier measurements analysis method, the determination number from outlier measurements of the other methods has been calculated. The results have been given in Table 6.

Examining Table 6, the convenience of the method has been researched. While deciding on the convenience of a method, the methods determining the outlier measurements determined by other methods and the fact that the measurements determined to be outlier would be different from the outlier measurements of other methods have been taken into consideration. In such an examination:

1. The t-test method from the conventional solution methods has been removed from the measurement group after determining 9 measurements as outlier. In the solution which is done in accordance with the robust estimation method, L1-norm, Yang-II, Beaton-Tukey method has determined those 9 measurements as outlier and the Denmark method has determined 8 measurements as outlier. Huber and Andrews methods have determined 5 and 2 measurements as outlier.

2. L1-norm method, which is not defined as completely

robust since it makes solutions that do not approach the weights to zero in iterative solution, has determined 11 measurements as outlier. While the solution which is done by this method determines all the outliers in t-test, Huber and Andrews methods, it has determined 6 outlier measurements in the Denmark method, 10 outlier measurements in Yang-II method and 36 outlier measurements in Beaton-Tukey method as consistent.

3. While the Denmark method determines 15 points as outlier and all outliers in Huber and Andrews methods, it has determined 1 measurement from the outlier measurements in the t-test, 2 measurements from the outlier measurements in L1-norm method, 7 measurements from the Yang-II method and 32 measurements from the Beaton-Tukey method as consistent.

4. While Yang-II method determines 21 of the measurements as outlier, it has been able to determine all the outlier measurements defined by t-test, L1-norm, Denmark, Huber, Andrews methods; Beaton-Tukey has determined 26 measurements as consistent.

5. Beaton-Tukey Method is a method that determines the outlier measurements most. This method has determined 47 measurements as outlier. This method has been able to determine all the outlier measurements determined by t-test, Huber and Andrews methods. This method, determining several outlier measurements has given rise to the thought that c critical value is not appropriate for this method.

It is seen that Huber and Andrews methods have determined the least outlier measurement with respectively 5 and 2 outlier measurements. The measurements have also been determined by all other methods and that these two methods can not suggest a different solution for this application.

As a result of this solution, it is seen that the robust estimation methods are very successful in determining the outlier measurements. It is also seen that while this method destroys the effects of the measurements which are not removed, the measurement group on the solution by setting the weights of the measurements to zero, scales the weights of other measurements. Leaning on these results, it is seen that the use of only conventional methods in the outlier measurements analysis is not a right approach and that besides this method, the robust estimation method has to be used as supportive method. According to the examination which is done between the robust estimation methods, it is concluded that Denmark and Yang-II methods, are more successful in determining the outlier measurements. It is also found that although L1-norm method is succesfull in determining outlier measurements, this method could not provide a distinctive scaling in the weights of measurement. Because Beaton-Tukey method was determined as outlier measurements of 78% of all measurements, this method is not suitable for determining outlier measurement for this application. Huber and Andrews methods have determined the results that were determined by all other methods. Therefore, these methods are found more unsuccessful when compared with the other outlier detection methods.

REFERENCES

Andrews DF, Bickel PJ, Hampel FR, Huber PJ, Rogers WH (1972). Robust estimates of location, Princeton University Press, Princeton, p.372.

Berberan A (1995). Multiple outlier detection: a real case study, Surv. Rev., 33 (255): 41-49.

Berné Valero JL, Baselga Moreno S (2005). Robust estimation in geodetic networks, Física de la Tierra, 17: 7-22.

Box GEP (1953). Non-normality and tests on variances, Biometrika, 40: 318-355.

Caspary W, Barutta H (1987). Robust estimation deformation models, Survey Rewiev, 29 (2333); 29-45.

Chen C (2002), Robust regression and outlie: detection with therobustreg procedure, SAS User Group International (SUGI)-27Proceedings, Orlando, FL, pp. 265-327.

Daszykowski M, Kaczmarek K, Vander Heyden Y, Walczak B (2007). Robust statistics in data analysis - A review basic concepts, Chemometrics and Intelligent Laboratory Systems 85: 203-219.

Erenoglu RC, Hekimoglu S (2009). An investigation into robust estimation applied to correlated GPSnetworks , M.G. Sideris (ed.), Observing our Changing Earth, International Association of Geodesy Symposia 133: 639-644.

Franklin S, Brodeur M (1997). A practical application of a robust multivaiiate outlier detection method, Proceedings of the Survey Research Methods Section, ASA, pp.186-191.

Frosch Moller S, Von Frese J, Bro R (2005). Robust methods for multivariate data analysis, J. Chemometr., 19: 549-563.

Fuchs H (1982). Contribution to the adjustment by minimizing the sum of absolute residuals, Manuscripta Geodaetica, 7: 151-207.

Gökalp E, Güngör O, Boz Y (2008). Evaluation of different outlier detection methods for GPS networks, Sensors, 8(11): 7344-7358.

Hampel F, Ronchetti EM, Rousseeuw PJ, Stahel WA (1986). Robust Statistics The Approach Based on Influence Functions, A Wiley-Interscience Publication John Wiley & Sons. New York.

Harvey BR (1993). Survey network adjustments by the L1 method, Aust. J. Geodesy, Photogram. Surv., 59: 39-52.

Hekimoglu S, Erenoglu RC (2007). Effect of heteroscedasticity and heterogeneousness on outlier detection for geodetic networks, J. Geod., 81: 137-148.

Huber PJ (1964). Robust estimation of a location parameter, The Annals of Mathematical Statistics, 35(1): 73-101.

Huber PJ (1981). Robust statistics, John Wiley and Sons Inc., New York., USA, pp.1-17, 43-68.

Hund E, Massart DL, Smeyers-Verbeke J (2002). Robust regression and outlier detection in the evaluation of robustness tests with different experimental designs, Analytica Chimica Acta 463: 53-73.

Knight NL, Wang J (2009). A comparison of outlier detection procedures and robust estimation methods in GPS positioning, The J. Navigation 62 (04): 699-709.

Koch KR (1999). Parameter estimation and hypothesis testing in linear models, Springer, Berlin-Heidelberg, NewYork, USA.

Krarup T, Juhl J, Kubik K (1980). Götterdämmerung over least squares adjustment, In: Proc. 14th Congress of the International Society of Photogrammetry, B3: 369-378.

Krauss K (1992). Photogrammetry, Volume 2, Advanced Methods and Applications, Ümmler Verlag, Berlin.

Pilgrim L (1996). Robust estimation applied to surface matching, ISPRS Journal of Photogrammetry And Remote Sensing, 51: 243-257.

Simpson DG, Chang YI (1997), Reweighting approximate GM estimators: asymptotics and residual-based graphics, J. Stat. Plan. Inference 57: 273-293.

Thomson DB (1976). Combination of geodetic networks, Department of Surveying Engineering, University of New Brunswick, Canada.

Vanicek P, Wells DE (1972). The least squares approximation, Department of Geodesy and Geomatics Engineering University of

New Brunswick, Canada.
Wieser A, Brunner FK (2001). Robust estimation applied to correlated GPS phase observations, In. Carosio A. & H. Kutterer (Eds.) Proc, First Int. Symposium on Robust Statistics and Fuzzy Techniques in Geodesy and GIS, ETH Zürich, pp. 193-198.
Wieser A, Brunner FK (2002), Short static GPS sessions: robust estimation, GPS Solutions. 5 (3): 70-79.
Wilcox RR (1997). Introduction to Robust Estimation and Hypothesis Testing, Academic Press, New York.
Xu PL (1989). On robust estimation with correlated observations, Bull. Geod., 63: 237-252.
Yang Y (1991). Robust bayesian estimation, Bull. Geod., 65: 145-150.
Yang Y (1999). Robust estimation of geodetic datum transformation, J. Geod., 73: 268-274.
Yang Y, Song L, Xu T (2001). Robust estimator for the adjustment of correlated gps networks, In. Carosio A. & H. Kutterer (Eds.) Proc, First International Symposium on Robust Statistics and Fuzzy Techniques in Geodesy and GIS. Zurich. Switzerland.