Increasing the efficacy of the tests for outliers for geodetic networks

Article in Acta Geodaetica et Geophysica Hungarica · September 2011
DOI: 10.1556/AGeod.46.2011.3.2

4 authors:

Serif Hekimoglu
Yildiz Technical University

Bahattin Erdogan
Yildiz Technical University

Ramazan Cuneyt Erenoglu
Çanakkale Onsekiz Mart Üniversitesi

Ramazan Gürsel Hoşbaş
Yildiz Technical University

Acta Geod. Geoph. Hung., Vol. 46(3), pp. 291–308 (2011)
DOI: 10.1556/AGeod.46.2011.3.2

INCREASING THE EFFICACY OF THE TESTS FOR OUTLIERS FOR GEODETIC NETWORKS

S Hekimoglu, B Erdogan, R C Erenoglu, R G Hosbas

Department of Geomatic Engineering, Yildiz Technical University, Istanbul, Turkey,
e-mail: hekim@yildiz.edu.tr

[Manuscript received September 6, 2010; accepted March 10, 2011]

Outliers in geodetic networks badly affect all parameters and their variances estimated by least-squares. Tests for outliers (e.g. Baarda's and Pope's tests) are frequently used to detect outliers in geodetic networks. To measure the ability of these tests, the mean success rate (MSR) is proposed. Studies have shown that the MSRs of these tests in geodetic networks are low due to the smearing effect of the least-squares estimation even if there is only one outlier in the data set. In this paper, a new approach, for small outliers, is presented to increase the MSRs of the tests for outliers in geodetic networks. The main idea is that if the weight of one observation is increased, the corresponding studentized or normalized residuals are increased, too. This thesis is proved. Hence, the ability of the tests to detect outliers can be increased by appropriately increasing the weight of one observation at a time and repeating this for all observations. This approach is applied to three simulated geodetic networks. We show that the MSRs of the outlier tests are improved by approximately 5% if there is one small outlier in the data set. However, the improvements in the MSRs for more than one outlier are low.

Keywords: efficacy; geodetic networks; outliers; tests for outliers; weight

1. Introduction

Ordinary least-squares estimation (LSE) is commonly used for the adjustment of geodetic network and outlier detection. The LSE provides optimal results when random errors are normally distributed. If observations include outliers, the parameters and the variances based on LSE are strongly influenced. Therefore, these outliers must be detected. To do this, two approaches are widely used: the conventional tests for outliers (Baarda 1968, Pope 1976, Pelzer 1985, Koch 1999, Snow and Schaffrin 2003, Teunissen 2006, Baselga 2007) and the robust methods (Huber 1981, Fuchs 1982, Hampel et al. 1986, Rousseeuw and Leroy 1987, Harvey 1993, Benning 1995, Youcai 1995, Koch 1996, Wilcox 1997, Wicki 1999, Yang et al. 2001, Yang et al. 2002, Amiri-Simkooei 2003, Xu 2005).

The mean success rate (MSR) was introduced to measure the efficacy of the robust methods and of the tests for outliers (Hekimoglu and Koch 1999 and 2000).

The MSR for detecting outliers is globally the number of successful detections over the number of experiments. In earlier work different methods were compared using the MSR. For example, Hekimoglu (2005a) showed that the robust methods identify outliers more reliably than the conventional tests for outliers in simple regression problems. Moreover, similar results were obtained when analyzing geodetic networks (Hekimoglu and Erenoglu 2007).

In linear regression analysis, to detect more than one small outlier is getting difficult due to the masking and swamping effects of the LSE (Hadi and Siminoff 1993). According to the studies by Hekimoglu and Koch (2000), Hekimoglu (2005a, 2005b, 2005c), the MSRs are rather small. In order to increase the MSRs of Baarda's test, the weights of all observations are multiplied by the positive value k ($k = 1.5$) and Baarda's test was applied to the same sample with these new weights. Thus, the MSR of one small outlier is increased by 14% (Hekimoglu 2005b). Moreover, Hekimoglu (2005c) proposed another approach for the Baarda's and Pope's tests to increase the MSRs; such as: 1° determine the maximum number of detectable outliers ($m_{\max}$), 2° increase the weights of the randomly chosen $m_{\max}$ observations to the same given value, 3° apply these tests and detect suspicious observations, 4° repeat this procedure for a given number of times, 5° count the number (frequency) of arising suspicious observations, 6° compare the frequency of each observation with a given critical value. This algorithm was also applied to the M-estimators to increase their MSRs (Hekimoglu and Erenoglu 2009).

In geodetic networks, the MSRs of tests for outliers are also small even if a sample includes only one outlier (Hekimoglu and Erenoglu 2007). If we look at the simulation results by Hekimoglu and Erenoglu (2007), the MSRs for two outliers decrease drastically in respect to the ones for one outlier. These results verify the result by Baselga (2007) which the test may have its limitations even if there is only one gross error.

To increase the MSRs of Baarda's and Pope's tests, the approach given above has been adopted to the geodetic networks. Unfortunately, the simulation results showed that the increases of the MSRs are not significant for the Baarda's and Pope's test. The main motivation of this paper is the question: Is it possible to increase the ability of tests for outlier? Therefore a new algorithm that is based on increasing the normalized or the studentized residuals by increasing the weight of only one observation at a time is proposed. As increasing the only one observation weight, the chances of exceeding the critical value of outliers whose normalized or studentized residuals lie close to the critical value are decreased. This main thesis is proved here for the case when the weight of only one observation is increased. For a particular geodetic network, this approach is repeated for each observation in a sequence. Thus, the one outlier can be detected more successfully than before.

This paper is organized as follows: Section 2 presents linear models for geodetic networks. The outlier concept and the test for outliers will be briefly explained in Section 3, and our motivation is introduced in Section 4. Some useful theorems are given in Section 5. In Section 6, a new approach for increasing the MSRs will be presented. In Section 7, experimental results on simulated leveling, horizontal control and GPS networks are presented. The efficacy of the new approach is discussed in Section 8. Finally, in Section 9 we present our conclusions. Proofs for the three theorems are given in Appendices A and B.

2. Linear models

The Gauss-Markov linear model for the geodetic networks is given as follows:

$$\mathbf{v} = \mathbf{A}\mathbf{x} - \mathbf{l}\,, \tag{1}$$

$$\mathbf{C_{ll}} = \sigma_0^2 \mathbf{P}^{-1}\,, \tag{2}$$

where $\mathbf{l}$ is the $n \times 1$ vector of observations, $\mathbf{A}$ is the $n \times u$ design matrix which has full column rank, $\mathbf{x}$ is the $u \times 1$ vector of unknown parameters, $\mathbf{v}$ is the $n \times 1$ vector of residuals, σ_0^2 is the variance of unit weight, $\mathbf{C_{ll}}$ is the $n \times n$ covariance matrix of the observations, $\mathbf{P}$ is the $n \times n$ diagonal weight matrix of the observations, n is the number of observations and u is the number of unknown parameters.

The weights of the observations can be obtained as follows:

$$p_i = \frac{\sigma_0^2}{\sigma_i^2}\,, \tag{3}$$

where p_i belongs to the diagonal weight matrix $\mathbf{P}$ and σ_i^2 is the variance of ith observation.

3. Tests for outliers

Outlier detection procedures were proposed by Baarda (1968) and Pope (1976) for geodesy. In these outlier detection processes, "good" observations originate from the same distribution, which is generally a normal distribution $N(\mu, \sigma^2)$. It is assumed that the outliers are rare in the observations; they are called "bad" observations and their expectation value is larger than 3σ.

If an observation $\bar{l}_k$ has an outlier δl_k with $\bar{l}_k = l_k + \delta l_k$, the hypothesis

$$H_0: \ \delta l_k = 0 \quad \text{against} \quad H_1: \ \delta l_k \neq 0 \tag{4}$$

is tested. If the value of the variance of unit weight σ_0^2 is known, the residual v_i is normalized to obtain the test statistic as follows:

$$W_i = \frac{|v_i|}{\sigma_0 \sqrt{q_{vvi}}} = \frac{|v_i|}{\sigma_{vi}}\,, \tag{5}$$

where q_{vvi} is the ith diagonal element of $\mathbf{Q_{vv}}$. $\mathbf{Q_{vv}}$ is the cofactor matrix of the residuals ($\mathbf{Q_{vv}} = \mathbf{P}^{-1} - \mathbf{A}(\mathbf{A^T P A})^{-1}\mathbf{A^T}$), σ_{vi} is the standard deviation of the ith residual. If $w_i > z_{1-\alpha/2}$, then the ith observation is considered as a bad observation, where α generally is chosen as 0.001 for Baarda's method (Baarda 1968, pages 74 and 86).

If the variance σ_0^2 is not known, the residual v_i is studentized by using the estimated variance $\hat{\sigma}_0^2$. This gives the test statistic of the τ-test (Pope 1976),

$$\tau_i = \frac{|v_i|}{\hat{\sigma}_0 \sqrt{q_{vvi}}} = \frac{|v_i|}{\hat{\sigma}_{vi}} . \tag{6}$$

If the level of significance α corresponds to all observations, the level of significance for each observation must be α/n:

$$P(\tau_i < c_{1-\alpha;n,n-u}) = 1 - \alpha/n , \tag{7}$$

where

$$c_{1-\alpha;n,n-u} = \tau_{1-\alpha/n;1,n-u-1} . \tag{8}$$

If rank $\mathbf{A} = qu$ holds in the Gauss-Markov model, then:

$$c_{1-\alpha;n,n-q} = \tau_{1-\alpha/n;1,n-q-1} , \tag{9}$$

where α generally is chosen as 0.05 or 0.01 (Koch 1999).

If there is more than one outlier in the observations, conventional tests for outliers are iteratively used. Only the observation with the largest normalized or studentized residual is tested in one cycle of the iterations. If this observation is rejected, it is removed, and the remaining observations are adjusted again. This procedure is carried out until no more outliers are detected (Schwarz and Kok 1993).

4. Motivation

The success rates of the conventional tests for outliers in geodetic networks are quite low even for just one outlier (Hekimoglu and Erenoglu 2007, Baselga 2007). How can this be explained? Due to the smearing effect of the LSE, an outlier in the observation l_i does not have an impact on only the residual v_i, but also on all the related residuals (Yang et al. 1999). Moreover, the outlier δl_i in the observation l_i changes its residual v_i by decreasing approximately in the scale of partial redundancy number r_i i.e. $v_i \approx -r_i \delta l_i$. Furthermore, $|v_i| < |\delta l_i|$ since $0 < r_i < 1$. This explains why the tests cannot detect even one small outlier in the observations. For example, in the horizontal control networks described by Hekimoglu and Erenoglu (2007), the MSR of the Baarda's test is 36% for one small outlier that lies between 3σ and 6σ. For some cases, the impact of δl_i on the related normalized or studentized residual is simply not big enough to exceed the critical value; this is the main reason for failure of the methods.

Before the analysis, if we knew which was the contaminated observation and increased the weight of this observation, the corresponding normalized residuals w_i or studentized residual τ_i would be increased and easily exceed the critical value. But, we cannot know the contaminated observation in advance. Therefore, we propose a new approach based on applying the tests for outliers after increasing the weight of only one observation. This approach is repeated for each observation separately. The indexes of observations detected in each stage are stored in a null vector **f** with size $n \times 1$. If the ith observation is detected, the ith element of the vector **f** is assigned the value '1'. If the same observation is detected again, f_i is assigned the value 2, etc. To judge which observation is an outlier, the robust scale factor (S_n) is applied on frequencies and the likelihood that detection of the bad observations may improve. In such a case, the w_i (or τ_i) of a good observation whose w_i (or τ_i) is close to the critical value will also be larger and may exceed the critical value. Thus, a type I error may occur and it is the possibility of rejecting the H_o hypothesis although H_o is true.

5. Theorems

In this section some theorems which have been developed to improve the conventional tests for outliers will be presented.

5.1 Theorems for Baarda's test

Theorem 1: If the weight p_i of an observation l_i is changed by an amount Δp, the absolute values of all the new residuals from LSE are modified. If $\Delta p < 0$, the absolute value of the new corresponding residual v_i^c always becomes larger ($|v_i^c| > |v_i|$), and if $\Delta p > 0$, it always becomes smaller ($|v_i^c| < |v_i|$). The proof of Theorem 1 is given by Hekimoglu and Sanli (2003). The following theorem is developed in this respect.

Theorem 2: If the weight p_i of an observation l_i is changed by an amount Δp, the absolute values of all the new normalized residuals from LSE are also modified. If $\Delta p < 0$, the absolute value of the new corresponding normalized residual w_i^c always becomes smaller ($|w_i^c| < |w_i|$), and if $\Delta p > 0$, it always becomes larger ($|w_i^c| > |w_i|$). The proof of Theorem 2 is given in Appendix A.

5.2 Theorems for Pope's test

The theorem for Pope's test can be derived analogously to the theorems for Baarda's test.

Theorem 3: If the weight p_i of the observation l_i is increased by a positive amount Δp, the new *a posteriori* variance $(\hat{\sigma}_0^c)^2$ of unit weight in the Gauss-Markov model always becomes larger $\{\text{i.e., } (\hat{\sigma}_0^c)^2 > \hat{\sigma}_0^2\}$. If $\Delta p < 0$, it always becomes smaller $\{\text{i.e. } (\hat{\sigma}_0^c)^2 < \hat{\sigma}_0^2\}$.

Theorem 4: If the weight p_i of any observation l_i is changed by an amount Δp, the new studentized residual τ_i^c is also changed. If $\Delta p > 0$, the new related studentized residual τ_i^c always becomes larger ($|\tau_i^c| > |\tau_i|$) and if $\Delta p < 0$, it always becomes smaller ($|\tau_i^c| < |\tau_i|$).

The proofs are given in Appendix B.

6. The new approach for increasing MSR of the tests for outliers

In this study, we try to verify that the ability to detect outliers in observations can be improved by increasing the weight of only one observation. If the weight of a single observation is increased, the related studentized and normalized residuals are also increased. The greater the studentized and normalized residuals are, the higher the chance of exceeding the critical value will be. Thus the abilities of the tests for outliers can be improved. To increase the MSRs of the tests for outliers; the following new approach is applied:

- Give index (i) to each observation l_i, ($i = 1, 2, \ldots, n$).
- Increase only the weight of the 1st observation p_1 to p'_1, (i.e. $p'_1 = p_1 + \Delta p_1$).
- Apply the conventional tests for outliers to the sample considering the new weight p'_1.
- Detect one outlier with the largest test statistic and remove it, repeat this test until no additional outlier is detected.
- Store the index or indexes of these outliers in a null vector $\mathbf{f}$. If the ith observation is detected as a possible outlier, f_i is given the value "1". This is the first step.
- Increase the next observation weight, (i.e. p_2) by Δp_2 (i.e. $p'_2 = p_2 + \Delta p_2$) and apply the conventional tests for outliers. If the ith observation is identified as a possible outlier f_i is increased by "1", i.e. $f_i = 2$.
- Proceed until all n observations have been tested in this way.
- The ith element f_i of the vector $\mathbf{f}$ is named "frequency".
- Compute the robust scale factor S_n of these frequencies.
- Identify the observation whose frequency is greater than three times S_n as an outlier.

In this new approach, the frequencies may vary from 0 to n (number of observations). The S_n can easily be applied to the stored frequencies as follows:

$$S_n = 1.4826 \text{ median} \left| \mathbf{f} \right| , \tag{10}$$

where 1.4826 ($1/\Phi^{-1}$ ($z = 0.75$)) is a usual correction vector, that $\Phi(z)$ is the cumulative distribution function for the standard normal distribution. As a special case of Eq. (10), if median $|\mathbf{f}| = 0$ (Rousseeuw and Leroy 1987, p. 160), we can compute the S_n as:

$$S_n = 1.2533 \frac{1}{n} \sum_{i=1}^{n} \left| f_i \right| , \tag{11}$$

where $1.2533(\sqrt{\pi/2})$ is an other correction factor.

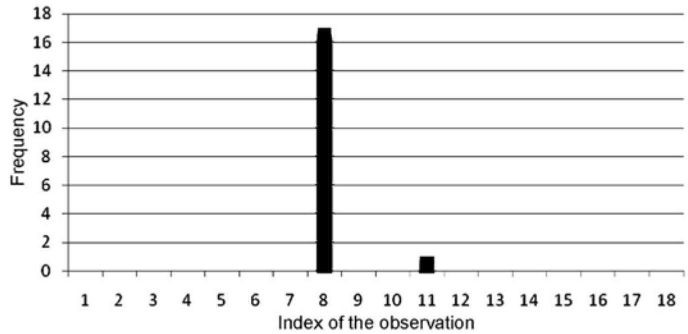

Fig. 1. Frequencies for levelling network given in Fig. 3 contaminated by one outlier ($\bar{l}_8$)

In summary, during each adjustment one observation weight is changed; each observation which is rejected gets a value '1' as a frequency. The next observation weight is changed and the process is repeated. The 'frequency' of the rejected observations is increased by 1 each time. This approach is applied until all n observations have been dealt with. Any observation with a frequency larger than $3S_n$ is considered as an outlier.

An example is given for a leveling network where only the 8th observation is contaminated in Fig. 1.

For the sample given in Fig. 1, the median of the frequency vector is 0, and the critical value is estimated from Eq. (11) as being 3.76. Finally, the frequency of the 8th observation is greater than the critical value and this observation is detected as being an outlier.

7. Monte Carlo simulation on geodetic networks

7.1 Simulation

In this study, a Monte-Carlo simulation technique has been used. To apply the new approach for detecting outliers, three geodetic networks have been simulated. They are a horizontal control network, a leveling network and a GPS network. Figures 2, 3 and 4 present the positions of the points and observations for the horizontal control, the leveling and the GPS networks, respectively. The MSRs for the conventional tests for outliers and of the new approach are presented for the same networks. All of the results have been calculated using MATLAB version R2006a.

The random errors e_i are generated from a normal distribution $N(\mu, \sigma^2)$ with mean zero and variance σ^2 by using the random number generator of MATLAB. The progress of simulating the good observations l_i' and bad observations $\bar{l}_i$ are obtained as the same as Erenoglu and Hekimoglu (2010).

For the horizontal control network, the observations, such as direction measurements l_{1i} and distance measurements l_{2j} are computed from the coordinates of the points. They are free of random errors. The random errors are generated from a normal distribution as follows: for direction measurements $e_{l1} \sim N(0, \sigma_{l1}^2$, where the precision is $\sigma_{l1} = 0.3$ mgon, and for the distance measurements $e_{l2j} \sim N(0, \sigma^2_{l2j})$, where the precision is $\sigma_{l2j} = \pm(5 \text{ mm} + 5 \times 10^{-6} S_{ij})$, S_{ij} is the distance between ith and jth points. σ_0 is taken as 0.3 mgon when forming weight matrix.

For the leveling and GPS networks, the height differences Δh_k and the baseline components ΔX_k, ΔY_k, ΔZ_k are computed from the fixed points. They are free of random errors. Similarly, the random errors are generated from a normal distribution. They are added to the height differences and baseline components. The precision is taken as $\sigma_h = \sigma_0 \sqrt{S} (\sigma_0 = 1 \text{ mm}/\sqrt{1 \text{ km}})$ for the leveling network where S is the length of the leveling line in km; for the GPS network $\sigma_{\Delta x} = \pm(5 + 10^{-6}\Delta X)$, $\sigma_{\Delta y} = \pm(5 + 10^{-6}\Delta Y)$, $\sigma_{\Delta z} = \pm(3(5 + 10^{-6}\Delta Z))$. Then, the random error e_i is replaced by the outlier δl_i in the related observation l_i, thus the observations are contaminated.

A hundred different contaminated samples of $\bar{\mathbf{l}}$ are simulated for each of the data sets $\mathbf{l}'$. Thus, 10 000 different contaminated samples of $\bar{\mathbf{l}}$ are obtained separately. In this study, α is chosen to be 0.001 for Baarda's test and 0.05 Pope's test. Note that, the fourth columns in Tables I, II and III are obtained from the conventional tests for outliers. We present these results for a comparison between the conventional tests and new approach described here. We take $\Delta p = 0.25 p_i$ that is obtained from many experiments. The new approach is performed for each observation in a network. Thus, it is computed n times for a network.

To measure the reliabilities of the conventional tests for outliers, the MSR criterion is used. Since a simulation is used to generate the outliers, it is possible to know exactly whether an observation is contaminated or not, in advance of carrying out the analysis. After applying the outlier detection method, if the observation identified as an outlier corresponds to the truly contaminated observation, the method is considered as being successful. If the method fails, it is considered unsuccessful (Hekimoglu and Koch 1999, 2000, Erenoglu and Hekimoglu 2010).

7.2 Example 1 – Horizontal control network

The simulated horizontal control network consists of 7 points where $n = 51$, $u = 21$ and degrees of freedom $f = 33$. Figure 2 presents the position of the points as well as the distances and directions.

The magnitudes of the outliers for small outliers (whose magnitudes lie between 3σ and 6σ), and for large outliers (whose magnitudes lie between 6σ and 12σ), are generated separately. An initial approximation of the orientation is estimated by using the arithmetic mean. Also, conventional tests for outliers are iteratively used. Only the observation with the largest normalized or studentized residual is tested and if it is rejected, it is removed, and the remaining observations are adjusted again. But, in this case, a geometric defect of the network may occur. To prevent such a geometric defect, the detected observation is not removed; the related weight p_i of the observation l_i is made smaller for the next iteration step. This situation is applied for all types of network.

Table I, column (A) includes the MSRs of both Baarda's and Pope's tests. To judge the ability of the new approach in detecting outliers, it is also applied to the

Table I. The MSRs of the test methods applied to the horizontal control network

Tests for outliers	Magnitude of outliers	Number of outliers	(A)* %	(B) %
Baarda	–	0	02	06
Pope			00	01
Baarda	$3\sigma - 6\sigma$	1	42	53
Pope			28	40
Baarda	$6\sigma - 12\sigma$	1	89	89
Pope			89	93

*(A): Conventional tests for outliers, (B): The new approach

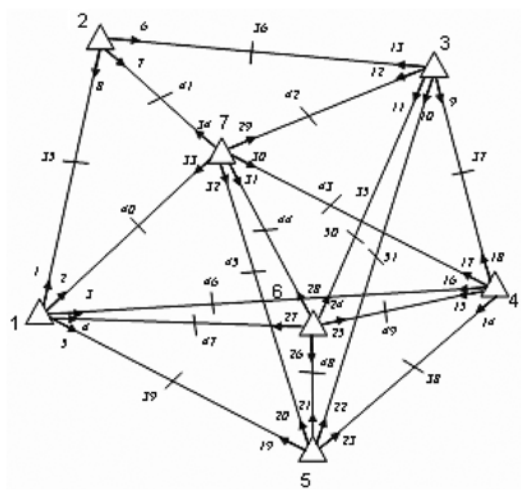

Fig. 2. The simulated horizontal control network

same samples. The MSRs for the new approach are shown in Table I, column (B). The MSRs are increased by 11% (53%–42%) and 12% (40%–28%) for the Baarda's and Pope's tests, respectively, for one small outlier.

In the case of one large outlier in the horizontal control network, the MSR of the new approach based on Pope's test is also increased. However, the MSR of the new approach based on Baarda's test is not increased significantly.

If the weight of a single observation is increased, the newly corresponding normalized and studentized residuals clearly exceed the critical value. However, in this case, one or more of the good observations may also be detected as being outliers. Thus, type I errors occur, which are 2% for Baarda's test and 0% for Pope's test. The increases in the MSRs of type I error are 4% (6%–2%) for Baarda's test and 1% (1%–0%) for Pope's test by using new approach. Thus, the increase in the MSR of type I errors by using the new approach is much smaller than the improvement of the MSR for one small outlier. The advantages of the new approach are 7% (11%–4%) for Baarda's test and 11% (12%–1%) for Pope's test.

7.3 Example 2 – Leveling network

In this subsection, we simulated a leveling network given in Fig. 3. The network consists of 9 benchmarks with 18 height differences. The degrees of freedom are $f = 10$. Table II, column (A) includes the MSRs of both Baarda's and Pope's tests, separately. For both Baarda's and Pope's tests the new approach is applied to the same samples.

The MSRs are increased at the rate of 7% (61%–54%) for Baarda's test and 7% (43%–36%) for Pope's test in the case of one small outlier. When there is one large outlier, the new approach has about the same MSR as the conventional Baarda's and Pope's tests.

Similarly with the horizontal control network, if the weight of a single observation whose normalized and studentized residuals closed to the critical value is increased, the new normalized and studentized residuals may clearly exceed the critical value. In this case, one or more good observations may also be detected as an outlier. Thus, a type I error occurs. The type I errors are 0% and 2% for Baarda's test and Pope's test, respectively. The increases in the MSRs of type I error are 1% (1%–0%) for Baarda's test and 1% (3%–2%) for Pope's test by using the new

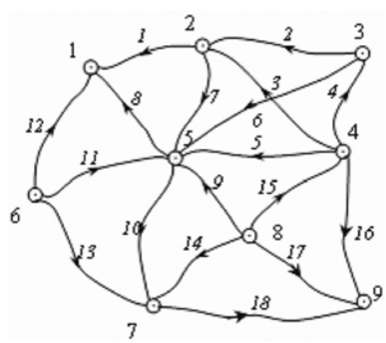

Fig. 3. The simulated leveling network

Table II. The MSRs of the test methods applied to the leveling network

Tests for outliers	Magnitude of outliers	Number of outliers	(A)* %	(B) %
Baarda Pope	–	0	00 02	01 03
Baarda Pope	$3\sigma - 6\sigma$	1	54 36	61 43
Baarda Pope	$6\sigma - 12\sigma$	1	98 91	98 92

*(A): Conventional tests for outliers, (B): The new approach

approach. Therefore, the increase in the MSR of type I errors by using new approach is much smaller than the increase in the MSR for one small outlier. The advantages of the new approach are 6% (7%–1%) for Baarda's test and 6% (7%–1%) for Pope's test.

7.4 Example 3 – GPS network

In the following part, a real GPS network based on the International GNSS Service (IGS) stations is used. GPS data on May 20, 2007 was downloaded from the Scripps Orbit and Permanent Array Center (SOPAC). 24 hours of GPS data was processed using Bernese Software version 5.0. During the processing, the IGS final SP3 product is used, and thus the coordinates of the stations are determined. The network consists of 23 baselines and 10 stations. The GPS network is given in Fig. 4 where $n = 69$, $u = 30$ and $f = 42$. To obtain good baseline components of $\Delta X'_k$, $\Delta Y'_k$ $\Delta Z'_k$ the random errors of $e_{\Delta X_k}$, $e_{\Delta Y_k}$, $e_{\Delta Z_k}$ are added to ΔX_k, ΔY_k, ΔZ_k, respectively. The magnitudes of outliers are taken to be between 3σ and 6σ for small outliers, and 6σ and 12σ for large outliers.

These observations are contaminated by using the previously described procedure which generates the bad observations. The MSRs of the Baarda's test, Pope's test and the new approach are given in Table III, columns (A) and (B), respectively.

The Baarda's test, Pope's test and the new approach are all applied to the same samples. The MSRs for one small outlier are increased at the rate of 6% (60%–54%) and 8% (50%–42%) for Baarda's and Pope's tests, respectively. Further, the increases in the type I errors for the new approach are 1% (8%–7%) and 2% (8%–6%) for Baarda's test and Pope's test, respectively. The increase in the MSR of type I errors when using the new approach is much smaller than the increase in the MSR for one small outlier. The improvements of MSRs of the new approach are 5% (6%–1%) for Baarda's test and 6% (8%–2%) for Pope's test. However, for one large outlier, the MSRs for Baarda's and Pope's tests are not increased.

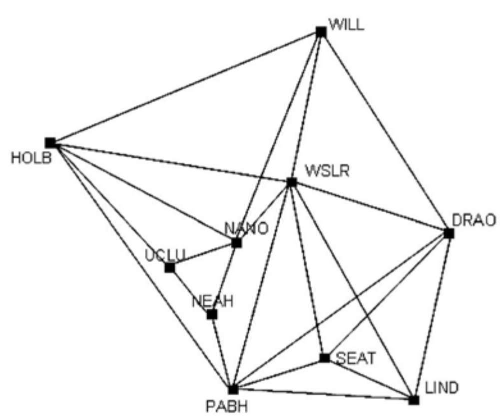

Fig. 4. The simulated GPS network

Table III. The MSRs of the test methods applied to the GPS network

Tests for outliers	Magnitude of outliers	Number of outliers	(A)* %	(B) %
Baarda	–	0	07	08
Pope			06	08
Baarda	$3\sigma - 6\sigma$	1	54	60
Pope			42	50
Baarda	$6\sigma - 12\sigma$	1	97	95
Pope			95	95

*(A): Conventional tests for outliers, (B): The new approach

8. Discussion

With our approach two choices have to be made: 1. How many observation weights will be increased in the first stage? and 2. What weight should be assigned to improve the MSRs of the conventional tests for outliers?

It is obvious that, if we change more than one observation weight, the MSRs obtained will be greater than before. If network is free of outliers, the number of good observations detected as outlier becomes higher, too. We could not find any optimal solution; therefore, we decide to change the weight of only one observation in each stage.

To determine a suitable weight increase, we do a lot of experiments, simulating 10 000 different working samples for each network and then changing the weight of a single observation. Tables IV, V and VI show the results of MSRs with one small outlier.

As seen in Tables IV, V and VI, the MSRs of the new approach increase when the weight of one observation is increased. Unfortunately, the MSRs of the new approach also increase when an outlier is not present. Consequently, a good observation whose normalized or studentized residual lies close to the critical value may be detected as an outlier. This is a type I error and it is obviously undesirable. For this reason it is necessary that the increase in the MSR for the case of no outlier must be smaller than the increase in the MSR when the sample includes one outlier. To find out the appropriate weight increase (Δp) 10 000 experiments have simulated. The differences between the increases in the MSRs in case of no outlier and the increases in the MSRs in case of no outlier are equal or greater than approximately 5% when Δp is chosen between $0.15p_i$ and $0.4p_i$. In this study Δp is chosen as $0.25p_i$.

Table IV. The variations in MSRs of the test methods when changing Δp for the horizontal control network

	Baarda's test				Pope's test			
Amount of weight increase	New approach %		New approach (no outliers) %		New approach %		New approach (no outliers) %	
$\Delta_p = 0.15p_i$	49		04		35		00	
$\Delta_p = 0.20p_i$	51		05		37		01	
$\mathbf{\Delta_p = 0.25p_i}$	**53**	**(42)***	**06**	**(02)***	**40**	**(28)****	**01**	**(00)****
$\Delta_p = 0.30p_i$	54		09		42		02	
$\Delta_p = 0.35p_i$	54		11		44		02	
$\Delta_p = 0.40p_i$	54		16		46		07	
$\Delta_p = 0.50p_i$	53		30		48		11	

* the MSRs of the conventional Baarda's test
** the MSRs of the conventional Pope's test

9. Conclusions

The main benefit of the new approach developed here is to increase the ability of the conventional tests to detect small outliers in geodetic networks. It is based on increasing the weight of one observation at a time and repeating this for all observations. If the weight p_i of an observation l_i is increased as Δp, the related absolute value of the normalized or studentized residuals are increased, too. This thesis has been proved in this paper. The MSRs of Baarda's test can be increased by using the new approach at the rate of 11% for the horizontal control network, at the rate of 7% for the leveling network and at the rate of 6% for the GPS network for one small outlier. In addition, the MSRs of Pope's test can also be increased by using the new approach at the rate of 12% for the horizontal control network, 7% for the leveling network and 8% for the GPS network for one small outlier. But, both tests and the new approach detect outliers in cases where there is no outlier. This is a type I error. However, the increase in the MSR of the new approach in case of one small outlier is much greater than the improvement in the MSR in case there is no outlier.

Finally, the advantage of the new approach when Δp is chosen between $0.15p_i$ and $0.4p_i$ is approximately 5% at the MSR for one small outlier in all geodetic networks. Therefore, it should be used to detect outliers in geodetic networks.

Table V. The variations in MSRs of the test methods when changing Δp for the leveling network

	Baarda's test				Pope's test			
Amount of weight increase	New approach %		New approach (no outliers) %		New approach %		New approach (no outliers) %	
$\Delta_p = 0.15p_i$	58		01		40		02	
$\Delta_p = 0.20p_i$	59		01		42		02	
$\mathbf{\Delta_p = 0.25p_i}$	**61**	**(54)***	**01**	**(00)***	**43**	**(36)****	**03**	**(02)****
$\Delta_p = 0.30p_i$	61		01		44		03	
$\Delta_p = 0.35p_i$	62		01		45		03	
$\Delta_p = 0.40p_i$	63		01		46		04	
$\Delta_p = 0.50p_i$	64		04		47		05	

* the MSRs of the conventional Baarda's test
** the MSRs of the conventional Pope's test

Table VI. The variations in MSRs of the test methods when changing Δp for the GPS network

	Baarda's test				Pope's test			
Amount of weight increase	New approach %		New approach (no outliers) %		New approach %		New approach (no outliers) %	
$\Delta_p = 0.15p_i$	58		08		47		07	
$\Delta_p = 0.20p_i$	59		08		49		07	
$\mathbf{\Delta_p = 0.25p_i}$	**60**	**(54)***	**08**	**(07)***	**50**	**(42)****	**08**	**(06)****
$\Delta_p = 0.30p_i$	61		10		51		09	
$\Delta_p = 0.35p_i$	61		12		52		09	
$\Delta_p = 0.40p_i$	62		14		53		11	
$\Delta_p = 0.50p_i$	63		15		55		15	

* the MSRs of the conventional Baarda's test
** the MSRs of the conventional Pope's test

Acknowledgements

The authors would like to thank Dr. Simon Butterworth for English corrections and the author B Erdogan thanks The Scientific and Technological Research Council of Turkey for its valuable grant.

Appendix A

Proof of Theorem 2: Firstly, the observation equations ($\mathbf{l}+\mathbf{v} = \mathbf{Ax}$ with $\mathbf{C_{ll}} = \sigma_0^2\mathbf{P}^{-1}$) is homogenized by transformation with;

$$\mathbf{P^T}(\mathbf{l}+\mathbf{v}) = \mathbf{P^T}\mathbf{Ax}\,,$$

where $\mathbf{P} = diag\left(\sqrt{p_1}, \sqrt{p_2}, \ldots, \sqrt{p_n}\right)$.

Thus, a simple model is obtained:

$$\bar{\mathbf{l}} + \overline{\mathbf{v}} = \overline{\mathbf{Ax}}, \quad \overline{\mathbf{C}}_{ll} = \sigma_0^2\overline{\mathbf{P}} = \sigma_0^2\mathbf{I}\,, \tag{1}$$

where $\bar{\mathbf{l}} = \mathbf{P^T}\mathbf{l}$, $\overline{\mathbf{v}} = \mathbf{P^T}\mathbf{v}$, $\overline{\mathbf{A}} = \mathbf{P^T}\mathbf{A}$ and $\overline{\mathbf{P}} = diag\ (1, 1, \ldots, 1)$.

If only the weight $\overline{p}_i$ of an observation $\overline{l}_i$ is changed by adding the amount of Δp, i.e. $\overline{p}_i^c = \overline{p}_i + \Delta p = 1 + \Delta p$, the normalized residual $\overline{w}_i$ changes to $\overline{w}_i^c$:

$$\overline{w}_i^c = \frac{|\overline{v}_i^c|}{\sigma_0\sqrt{\overline{q}_{\overline{v}_i^c\overline{v}_i^c}}} = \frac{|\overline{v}_i^c|}{\sigma_0\sqrt{\overline{r}_i^c}}\sqrt{\overline{p}_i^c}\,, \tag{2}$$

where $\overline{r}_i^c(\overline{\mathbf{R}}^{\mathbf{c}})_{ii}$ ($\overline{r}_i^c$ is partial redundancy), $\overline{\mathbf{R}}^{\mathbf{c}} = \mathbf{I} - \overline{\mathbf{H}}^{\mathbf{c}}$, ($\overline{\mathbf{H}}^{\mathbf{c}}$ denotes hat matrix) $p^c = I + \Delta p$, $\overline{\mathbf{H}}^{\mathbf{c}} = \overline{\mathbf{A}}(\overline{\mathbf{A}}^{\mathbf{T}}\mathbf{P^c}\overline{\mathbf{A}})^+\overline{\mathbf{A}}^{\mathbf{T}}\mathbf{P^c}$, $\overline{q}_{\overline{v}_i^c\overline{v}_i^c} = \frac{\overline{r}_i^c}{\overline{p}_i^c}$, $\overline{q}_{v_iv_i} = \frac{\overline{r}_i}{\overline{p}_i}$.

The residual $\overline{v}_i^c$ and the partial redundancy $\overline{r}_i^c$ are given as follows:

$$\overline{v}_i^c = \frac{\overline{v}_i}{1 + \Delta p\overline{h}_{ii}} \quad \text{(Hekimoglu and Sanli 2003)}\,, \tag{3}$$

$$\overline{h}_{ii}^c = \frac{\overline{h}_{ii}(1+\Delta p)}{1 + \Delta p\overline{h}_{ii}^c} \quad \text{(Even-Tzur 1999)}\,, \tag{4}$$

and

$$\overline{r}_i^c = 1 - \overline{h}_{ii}^c = \frac{\overline{r}_i}{1 + \Delta p\overline{h}_{ii}}\,, \tag{5}$$

where $\overline{h}_{ii}^c = (\overline{\mathbf{H}}^{\mathbf{c}})_{ii}$, $\overline{h}_{ii} = (\overline{\mathbf{h}})_{ii}$, $\overline{\mathbf{H}} = \overline{\mathbf{A}}(\overline{\mathbf{A}}^{\mathbf{T}}\overline{\mathbf{P}}\mathbf{A}^+\overline{\mathbf{A}}^{\mathbf{T}}\overline{\mathbf{P}}$.

If the relations are taken into account in Eq. (2), the normalized residual $\overline{w}_i^c$ can be written,

$$\overline{w}_i^c = \overline{w}_i\sqrt{\frac{1+\Delta p}{1 + \Delta p\overline{h}_{ii}}}\,. \tag{6}$$

If $\Delta p > 0$, always $1 + \Delta p > 1 + \Delta p\overline{h}_{ii}$ because $0 < \overline{h}_{ii} \leq 1$. Hence, $\overline{w}_i^c > \overline{w}_i$. If $\Delta p < 0$, $\overline{w}_i^c < \overline{w}_i$.

Appendix B

Proof of Theorem 3: After a simple model was obtained by homogenizing the linear model given in Eq. (1), the new weight matrix $\overline{\mathbf{P}}^{\mathbf{c}}$ can be written after changing the weight $\overline{p}_i$ of the observation l_i.

$$\overline{\mathbf{P}}^{\mathbf{c}} = \overline{\mathbf{P}} + \mathbf{\Delta P} = \mathbf{I\Delta P}\,, \tag{7}$$

with $\overline{p}_i^c = \overline{p}_i + p_i = 1 + \Delta p$, $i = (1, 2, \ldots, n)$,

$$(\overline{\mathbf{P}})_{jj} = 0, \quad j = (1, 2, \ldots, n), \quad i \neq j,$$
$$(\overline{\mathbf{P}})_{ii} = \Delta p\,.$$

The new residual vector $\overline{v}^c$ is given,

$$\overline{\mathbf{v}}^c = (\overline{\mathbf{H}}^{\mathbf{c}} - \mathbf{I})\overline{\mathbf{l}} = \overline{\mathbf{v}} + \boldsymbol{\Delta}\mathbf{v}\,, \tag{8}$$

where $\overline{\mathbf{H}}^{\mathbf{c}} = \overline{\mathbf{H}} + \overline{\mathbf{H}}\boldsymbol{\Delta}\mathbf{P} - \overline{\mathbf{H}}(\mathbf{I} + \boldsymbol{\Delta}\mathbf{P}\overline{\mathbf{H}})^{-1}\boldsymbol{\Delta}\mathbf{P}\overline{\mathbf{H}}\overline{\mathbf{P}}^{\mathbf{c}}$ (Hekimoglu and Sanli 2003),

$$\mathbf{v} = (\overline{\mathbf{H}} - \mathbf{I})\overline{\mathbf{l}}\,. \tag{9}$$

If only the weight $\overline{p}_i$ of the observation l_i is changed by adding the amount of Δp (i.e. $\overline{p}_i^c = 1 + \Delta p$), the corresponding residual $\overline{v}_i^c$ is given:

$$\overline{v}_i^c = \overline{v}_i - \frac{\Delta p \overline{h}_{ii} \overline{v}_i}{1 + \Delta p \overline{h}_{ii}} = \frac{\overline{v}_i}{1 + \Delta p \overline{h}_{ii}} \quad \text{(Hekimoglu and Sanli 2003)}\,, \tag{10}$$

and the other residual $\overline{v}_j^c$,

$$\overline{v}_j^c = \overline{v}_j - \frac{\Delta p \overline{h}_{ji} \overline{v}_i}{1 + \Delta p \overline{h}_{ii}} \quad \text{(Hekimoglu and Sanli 2003)}\,. \tag{11}$$

Considering Eq. (9) it can be written;

$$\begin{aligned}(\overline{\mathbf{v}}^{\mathbf{c}})^{\mathbf{T}}\overline{\mathbf{P}}^{\mathbf{c}}\overline{\mathbf{v}}^{\mathbf{c}} &= (\overline{\mathbf{v}} + \boldsymbol{\Delta}\mathbf{v})^{\mathbf{T}}(\mathbf{I} + \boldsymbol{\Delta}\mathbf{P})(\overline{\mathbf{v}} + \boldsymbol{\Delta}\mathbf{v}),\\ (\overline{\mathbf{v}}^{\mathbf{c}})^{\mathbf{T}}\overline{\mathbf{P}}^{\mathbf{c}}\overline{\mathbf{v}}^{\mathbf{c}} &= \overline{\mathbf{v}}^{\mathbf{T}}\overline{\mathbf{v}} + \overline{\mathbf{v}}^{\mathbf{T}}\boldsymbol{\Delta}\mathbf{P}\overline{\mathbf{v}} + 2\overline{\mathbf{v}}^{\mathbf{T}}\boldsymbol{\Delta}\mathbf{v} + 2\overline{\mathbf{v}}^{\mathbf{T}}\boldsymbol{\Delta}\mathbf{P}\boldsymbol{\Delta}\mathbf{v} +\\ &\quad + \boldsymbol{\Delta}\mathbf{v}^{\mathbf{T}}\boldsymbol{\Delta}\mathbf{v} + \boldsymbol{\Delta}\mathbf{v}^{\mathbf{T}}\boldsymbol{\Delta}\mathbf{P}\boldsymbol{\Delta}\mathbf{v}\,.\end{aligned} \tag{12}$$

We can find the following relations easily;

$$\overline{\mathbf{v}}^{\mathbf{T}}\boldsymbol{\Delta}\mathbf{P}\overline{\mathbf{v}} = \Delta p \overline{v}_i^2\,, \tag{13}$$

$$\overline{\mathbf{H}}\overline{\mathbf{v}} = \mathbf{0}\,, \tag{14}$$

$$\boldsymbol{\Delta}\mathbf{v}^{\mathbf{T}} = -\frac{\Delta p \overline{v}_i}{1 + \Delta p \overline{h}_{ii}}[\overline{h}_{1i}\overline{h}_{2i} \ldots \overline{h}_{ii} \ldots \overline{h}_{ji} \ldots \overline{h}_{ni}]\,, \tag{15}$$

$$\overline{\mathbf{v}}^{\mathbf{T}}\boldsymbol{\Delta}\mathbf{v} = \mathbf{0}\,, \tag{16}$$

$$\mathbf{v}^{\mathbf{T}}\boldsymbol{\Delta}\mathbf{P}\boldsymbol{\Delta}\mathbf{v} = -\frac{\Delta p^2 \overline{v}_i^2}{1 + \Delta p \overline{h}_{ii}}\,, \tag{17}$$

$$\overline{\mathbf{H}}\overline{\mathbf{H}} = \overline{\mathbf{H}} \quad \text{and} \quad \overline{\mathbf{H}}^{\mathbf{T}} = \overline{\mathbf{H}}\,, \tag{18}$$

$$\boldsymbol{\Delta}\mathbf{v}^{\mathbf{T}}\boldsymbol{\Delta}\mathbf{v} = \frac{\Delta p^2 \overline{v}_i^2 \overline{h}_{ii}}{(1 + \Delta p \overline{h}_{ii})^2}\,, \tag{19}$$

and

$$\mathbf{\Delta v^T \Delta P \Delta v} = \frac{\Delta p^3 \overline{v}_i^2 \overline{h}_{ii}^2}{(1+\Delta p \overline{h}_{ii})^2} \,. \tag{20}$$

As result we can write;

$$(\overline{\mathbf{v}}^\mathbf{c})^\mathbf{T} \overline{\mathbf{P}}^\mathbf{c} \overline{\mathbf{v}}^\mathbf{c} = \overline{\mathbf{v}}^\mathbf{T} \overline{\mathbf{v}} + \frac{\Delta p \overline{v}_i^2}{1+\Delta p \overline{h}_{ii}} \,, \tag{21}$$

$$(\hat{\sigma}_0^c)^2 = \hat{\sigma}_0^2 + \frac{\Delta p \overline{v}_i^2}{(n-u)(1+\Delta p \overline{h}_{ii})}, \qquad \text{where} \quad \hat{\sigma}_0^2 = \frac{\overline{v}^T \overline{v}}{n-u} \,. \tag{22}$$

Consequently, if $\Delta p > 0$, $(\hat{\sigma}_0^c)^2 > \hat{\sigma}_0^2$. If $\Delta p < 0$, $(\hat{\sigma}_0^c)^2 < \hat{\sigma}_0^2$.

Proof of Theorem 4: If only the weight $\overline{p}_i$ of an observation $\overline{l}_i$ is changed by adding Δp, (i.e. $\overline{p}_i^c = 1 + \Delta p$), the studentized residual $\overline{\tau}_i^c$ is given,

$$\overline{\tau}_i^c = \frac{|\overline{v}_i^c|}{\hat{\sigma}_0^c \sqrt{\overline{q}_{\overline{v}_i^c \overline{v}_i^c}}} \,. \tag{23}$$

Considering (3), (4), (6) and (22), $\overline{\tau}_i^c$ can be written as:

$$\overline{\tau}_i^c = \frac{|\overline{v}_i|}{(1+\Delta p \overline{h}_{ii})} \frac{\sqrt{(1+\Delta p \overline{h}_{ii})(1+\Delta p)}}{\sqrt{\overline{r}_i}} \frac{1}{\hat{\sigma}_0^2 + \frac{\Delta p \overline{v}_i^2}{(n-u)(1+\Delta p \overline{h}_{ii})}} \,, \tag{24}$$

or

$$\overline{\tau}_i^c = \tau_i \frac{\sqrt{1+\Delta p}}{\sqrt{(1+\Delta p \overline{h}_{ii}) + \frac{\Delta p \overline{v}_i^2}{(n-u)\hat{\sigma}_0^2}}} = \tau_i \frac{\sqrt{1+\Delta p}}{\sqrt{1+\Delta p \left(\overline{h}_{ii} + \frac{\overline{v}_i^2}{(n-u)\hat{\sigma}_0^2} \right)}} \,. \tag{25}$$

Considering the limit for the studentized residual $\overline{\tau}_i^c$

$0 < \overline{\tau}_i^c \le \sqrt{n-u}$ (Baselga 2007), we argue that $\overline{h}_{ii} + \frac{\overline{v}_i^2}{(n-u)\hat{\sigma}_0^2} < 1$.

Let $\overline{h}_{ii} + \frac{\overline{v}_i^2}{(n-u)\hat{\sigma}_0^2} = 1$. In this case, we can write $\frac{\overline{v}_i^2}{\hat{\sigma}_0^2 \overline{r}_i} = n-u$; and hence $\overline{\tau}_i = \sqrt{n-u}$.

Let $\overline{h}_{ii} + \frac{\overline{v}_i^2}{(n-u)\hat{\sigma}_0^2} > 1$. Thus, we can find $\overline{\tau}_i^2 > n-u$ and hence $\overline{\tau}_i > \sqrt{n-u}$, this is impossible. Therefore, $\overline{h}_{ii} + \frac{\overline{v}_i^2}{(n-u)\hat{\sigma}_0^2} < 1$. If $\Delta p > 0$, $\overline{\tau}_i^c > \tau_i$. If $\Delta p < 0$, $\overline{\tau}_i^c < \tau_i$.

References

Amiri-Simkooei A R 2003: *J. Surv. Eng.*, 129, 37–43.
Baarda W 1968: A testing procedure for use in geodetic Networks. Publication on Geodesy, New Series 2, No. 5., Netherlands Geodetic Commission, Delft
Baselga S 2007: *J. Surv. Eng.*, 133, 52–55.
Benning W 1995: *zfv*, 12, 606–617.
Erenoglu R C, Hekimoglu S 2010: *Acta Geod. Geoph. Hung.*, 45, 426–439.
Even-Tzur G 1999: *zfv*, 124, 128–134.
Fuchs H 1982: *Manuscr. Geod.*, 7, 151–207.
Hadi A S, Siminoff J S 1993: *J. American Statistical Ass.*, 88, 65–72.
Hampel F, Ronchetti E, Rousseeuw P, Stahel W 1986: Robust statistics: the approach based on influence functions. John Wiley and Sons, New York
Harvey P R 1993: *Aust. J. Geod. Photogram. Surv.*,
Hekimoglu S 2005a: *zfv*, 3, 174–180.
Hekimoglu S 2005b: *AVN*, 112, 7–12.
Hekimoglu S 2005c: *Survey Review*, 38, 274–285.59, 39–52.
Hekimoglu S, Erenoglu R C 2007: *J. Geodesy*, 81, 137–148.
Hekimoglu S, Erenoglu R C 2009: *J. Surv. Eng.*, 135, 1–5.
Hekimoglu S, Koch K R 1999: In: Proc. Third Turkish-German Joint Geodetic Days. M O Altan, L Gründig eds, Istanbul, 1, 179–196.
Hekimoglu S, Koch K R 2000: *AVN*, 107, 247–254.
Hekimoglu S, Sanli D U 2003: *zfv*, 128, Heft 4, 271.
Huber P J 1981: Robust statistics. John Wiley and Sons. Inc., New York
Koch K R 1996: *AVN*, 103, 1–18.
Koch K R 1999: Parameter estimation and hypothesis testing in linear models. 2nd Ed. Springer-Verlag, Berlin
Pelzer H 1985: Überprüfung von Ausgleichungsmodellen. Geodätische Netze in Landes- und Ingenieurvermessung II. Verlag Konrad Witter, Stuttgart
Pope A J 1976: The statistics of residuals and the outlier detection of outliers. NOAA Technical Report, NOS 65, NGS 1, Rockville, MD
Rousseeuw P J, Leroy A M 1987: Robust regression and outlier detection. John Wiley and Sons, Inc., New York.
Schwarz C R, Kok J J 1993: *J. Surv. Eng.*, 119, 128–136.
Snow K, Schaffrin B 2003: *GPS Solutions*, 7, 130–139.
Teunissen P J G 2006: Testing theory, an introduction. 2nd edition, Delft University Press, Delft
Wicki F 1999: Robuste Schätzverfahren für die Parameterschätzung in geodätischen Netzen. Institut für Geodäsie und Photogrammetrie an der ETH, Zürich, Mitt. No. 67.
Wilcox R R 1997: Introduction to robust estimation and hypothesis testing, Academic Press. San Diego
Xu P L 2005: *J. Geodesy*, 79, 146–159.
Yang Y, Cheng M K, Shum C K, Tapley B D 1999: *J. Geodesy*, 73, 345–349.
Yang Y, He H, Xu G 2001: *J. Geodesy*, 75, 109–116.
Yang Y, Song L, Xu T 2002: *J. Geodesy*, 76, 353–358.
Youcai H 1995: *Bull. Geod.*, 69, 292–299.