

The case-deletion and mean-shift outlier models: equivalence and beyond

J. Guo

Received: 16 October 2012 / Accepted: 12 March 2013 / Published online: 10 April 2013
© Akadémiai Kiadó, Budapest, Hungary 2013

Abstract Deletion diagnostics have been widely adopted to evaluate the influence of one or more observations on the adjustment outputs. Both the case-deletion model and the mean-shift outlier model can be used to develop multiple case-deletion diagnostics for linear models. These two multiple outlier detection models are identical from the statistical point of view. However, the mean-shift outlier model, in which the underlying observations are implicitly deleted, outweighs the case-deletion model in term of computational efficiency. The influence of outliers on the adjustment outputs is also addressed. It reveals that the precision, minimal detectable bias (MDB) measure and dilution of precision metric (DOP) are all overestimated when outliers exist but were neglected under the assumption that *a priori* variance factor is known before.

Keywords Outlier · Case-deletion · Mean-shift model · Equivalence

1 Introduction

When outliers are present in a data set, a least-squares (LS) adjustment may not be possible or will produce poor or invalid results (Wolf and Ghilani 1997). Many approaches to mitigate or even eliminate the deteriorating effect of outlying observations on the parameter estimates have been developed (Cook 1977; Koch 1999; Monhor and Verö 2011), albeit there is no universally-accepted definition for an outlier (Barnett and Lewis 1994; Monhor and Takemoto 2005; Monhor and Verö 2011).

There are two essential approaches to control the corrupt effects of outliers: conventional outlier detection test procedures developed in geodetic literature (Baarda 1968; Pope 1976) and robust methods (Huber 1981; Hampel et al. 1986, Rousseeuw and Leroy 1987; Koch 1999; Yang 1999; Hekimoglu and Koch 2000; Xu 2005; Hekimoglu 2005). However, the conventional test procedures are only applicable under the assumption that no more than one outlier is present. In case of multiple outliers, the most practical strategy is to employ

J. Guo (✉)

Information Engineering Univ., 62 Kexuedaodao Road, P.O. Box 2201-160, Zhengzhou 450001, China
e-mail: jianfeng.guo@gmail.com

the iterative data snooping presented by Kok (1984), whilst procedures for detecting all outliers at once have also been proposed (Hadi and Simonoff 1993; Snow and Schaffrin 2003; Baselga 2011).

To evaluate the influence of one or more observations on the adjustment outputs, the deletion diagnostics have been extensively adopted (Cook 1977, 1979; Chatterjee and Hadi 1988). There are two methods to implement the diagnostics, namely, delete the underlying observation(s) explicitly or implicitly. The explicit one is case-deletion model and the other one is referred to as mean-shift outlier model (Hekimoglu et al. 2012). The aim of this contribution is twofold: first to prove the equivalence of these two methods; second to address influence of outlying observations on the quality measures.

The paper is organized as follows: the equivalence of two multiple outlier detection models is investigated, followed by the computational considerations in performing the mean-shift outlier model. Furthermore, theoretical analyses state that the precision, Minimal Detectable Bias (MDB) measure and Dilution of Precision (DOP) metric are all overoptimistic when the outlying observations should have been taken into account but were neglected.

2 Model description

Let us consider a linear Gauss-Markov model defined by Koch (1999)

$$E(L) = AX \quad \text{with} \quad \text{Cov}(L) = \sigma_0^2 P^{-1}, \quad (1)$$

where L is the $n \times 1$ vector of observations, A the $n \times u$ design matrix with full column rank, and X the $u \times 1$ vector of unknowns. σ_0^2 is the *a priori* variance factor of unit weight, and P the symmetric positive-definite weight matrix. Whenever necessary, the observations are supposed to be normally distributed.

Then, the (weighted) LS estimate of the unknowns in model Eq. (1) reads (Koch 1999)

$$\hat{X} = (A^T P A)^{-1} A^T P L \quad (2)$$

The corresponding residual vector is readily obtained as

$$V = L - A \hat{X} = R L \quad (3)$$

where $R = I_n - A(A^T P A)^{-1} A^T P$ maps the original observational vector onto the residual vector as a result of the LS adjustment (Schaffrin 1997; Guo et al. 2011). The matrix R plays an important role in linear adjustment techniques since it contains extremely useful information (Huber 1981; Guo et al. 2007, 2010). One can easily verify that R is idempotent and has the following useful properties

$$R^T P = P R = R^T P R, \quad R A = O, \quad A^T P R = O \quad (4)$$

the weighted sum of squares of the LS residuals reads

$$\Omega = V^T P V = L^T P R L \quad (5)$$

3 Multiple outlier detection models

As is known, LS method is very susceptible to outliers (Wolf and Ghilani 1997; Koch 1999; Guo et al. 2010). There are two procedures to implement the deletion diagnostics, namely, the case-deletion model and the mean-shift outlier model.

Let us assume the i_1 th, the i_2 th, ..., and the i_m th observations are to be deleted, while the i_{m+1} th, the i_{m+2} th, ..., and the i_n th observations are the remaining ones.

3.1 Mean-shift outlier model

For convenience we introduce the following notations,

$$\mathbf{H}_b = (\mathbf{h}_{i_1}, \mathbf{h}_{i_2}, \dots, \mathbf{h}_{i_m}), \quad \mathbf{H}_r = (\mathbf{h}_{i_{m+1}}, \mathbf{h}_{i_{m+2}}, \dots, \mathbf{h}_{i_n}) \tag{6}$$

where \mathbf{h}_i denotes the i th n -vector having a 1 as its i th entry and zeros otherwise. It can be seen $(\mathbf{H}_b, \mathbf{H}_r)$ is a permutation matrix (Strang and Borre 1997). Since a permutation matrix is orthogonal, one can obtain

$$(\mathbf{H}_b, \mathbf{H}_r)(\mathbf{H}_b, \mathbf{H}_r)^T = \mathbf{H}_b \mathbf{H}_b^T + \mathbf{H}_r \mathbf{H}_r^T = \mathbf{I}_n \tag{7}$$

and

$$(\mathbf{H}_b, \mathbf{H}_r)^T (\mathbf{H}_b, \mathbf{H}_r) = \begin{pmatrix} \mathbf{H}_b^T \mathbf{H}_b & \mathbf{H}_b^T \mathbf{H}_r \\ \mathbf{H}_r^T \mathbf{H}_b & \mathbf{H}_r^T \mathbf{H}_r \end{pmatrix} = \mathbf{I}_n$$

it follows immediately that

$$\mathbf{H}_b^T \mathbf{H}_b = \mathbf{I}_m, \quad \mathbf{H}_b^T \mathbf{H}_r = \mathbf{O}, \quad \mathbf{H}_r^T \mathbf{H}_r = \mathbf{I}_{n-m} \tag{8}$$

Accordingly, the corresponding mean-shift outlier model reads

$$E(\mathbf{L}) = \mathbf{A}\mathbf{X} + \mathbf{H}_b \nabla \quad \text{with } \text{Cov}(\mathbf{L}) = \sigma_0^2 \mathbf{P}^{-1}, \tag{9}$$

in which $(\mathbf{A}, \mathbf{H}_b)$ is of full column rank.

Based on the LS principle, one can obtain the following normal equation:

$$\begin{pmatrix} \mathbf{A}^T \mathbf{P} \mathbf{A} & \mathbf{A}^T \mathbf{P} \mathbf{H}_b \\ \mathbf{H}_b^T \mathbf{P} \mathbf{A} & \mathbf{H}_b^T \mathbf{P} \mathbf{H}_b \end{pmatrix} \begin{pmatrix} \hat{\mathbf{X}}_\nabla \\ \hat{\nabla} \end{pmatrix} = \begin{pmatrix} \mathbf{A}^T \mathbf{P} \mathbf{L} \\ \mathbf{H}_b^T \mathbf{P} \mathbf{L} \end{pmatrix} \tag{10}$$

with which and denoting

$$\mathbf{R}_{H_b} = \mathbf{I}_n - \mathbf{H}_b (\mathbf{H}_b^T \mathbf{P} \mathbf{H}_b)^{-1} \mathbf{H}_b^T \mathbf{P} \tag{11}$$

we have

$$\begin{cases} \hat{\mathbf{X}}_\nabla = (\mathbf{A}^T \cdot \mathbf{P} \mathbf{R}_{H_b} \cdot \mathbf{A})^{-1} \mathbf{A}^T \cdot \mathbf{P} \mathbf{R}_{H_b} \cdot \mathbf{L} \\ \hat{\nabla} = (\mathbf{H}_b^T \mathbf{P} \mathbf{H}_b)^{-1} \mathbf{H}_b^T \mathbf{P} (\mathbf{L} - \mathbf{A} \hat{\mathbf{X}}_\nabla) \end{cases} \tag{12}$$

It can be verified that \mathbf{R}_{H_b} is idempotent and has the following useful properties

$$\mathbf{R}_{H_b}^T \mathbf{P} \mathbf{R}_{H_b} = \mathbf{P} \mathbf{R}_{H_b} = \mathbf{R}_{H_b}^T \mathbf{P}, \quad \mathbf{R}_{H_b} \mathbf{H}_b = \mathbf{O}, \quad \mathbf{H}_b^T \mathbf{P} \mathbf{R}_{H_b} = \mathbf{O} \tag{13}$$

The corresponding residual vector is

$$\begin{aligned} \mathbf{V}_\nabla &= (\mathbf{L} - \mathbf{A} \hat{\mathbf{X}}_\nabla) - \mathbf{H}_b \hat{\nabla} \\ &= (\mathbf{I}_n - \mathbf{H}_b (\mathbf{H}_b^T \mathbf{P} \mathbf{H}_b)^{-1} \mathbf{H}_b^T \mathbf{P}) (\mathbf{L} - \mathbf{A} \hat{\mathbf{X}}_\nabla) \\ &= \mathbf{R}_{H_b} (\mathbf{L} - \mathbf{A} \hat{\mathbf{X}}_\nabla) \end{aligned} \tag{14}$$

and thus

$$\hat{\sigma}_\nabla^2 = \frac{\Omega_\nabla}{n - (m + u)} \tag{15}$$

with

$$\Omega_\nabla = \mathbf{V}_\nabla^T \mathbf{P} \mathbf{V}_\nabla = (\mathbf{L} - \mathbf{A} \hat{\mathbf{X}}_\nabla)^T \mathbf{P} \mathbf{R}_{H_b} (\mathbf{L} - \mathbf{A} \hat{\mathbf{X}}_\nabla) \tag{16}$$

3.2 Multiple case-deletion model

Under the same condition, the multiple case-deletion model reads

$$E(\mathbf{H}_r^T \mathbf{L}) = \mathbf{H}_r^T \mathbf{A} \mathbf{X} \quad \text{with} \quad \text{Cov}(\mathbf{H}_r^T \mathbf{L}) = \sigma_0^2 \mathbf{H}_r^T \mathbf{P}^{-1} \mathbf{H}_r, \tag{17}$$

with which one can obtain the LS estimator as follows

$$\hat{\mathbf{X}}_r = (\mathbf{A}^T \mathbf{H}_r \cdot \mathbf{P}_r \cdot \mathbf{H}_r^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{H}_r \cdot \mathbf{P}_r \cdot \mathbf{H}_r^T \mathbf{L} \tag{18}$$

where

$$\mathbf{P}_r = (\mathbf{H}_r^T \mathbf{P}^{-1} \mathbf{H}_r)^{-1} \tag{19}$$

The permutation matrix $(\mathbf{H}_b, \mathbf{H}_r)$ is invertible. Therefore, one can obtain

$$\mathbf{P}^{-1} = (\mathbf{H}_b, \mathbf{H}_r) [(\mathbf{H}_b, \mathbf{H}_r)^T \mathbf{P} (\mathbf{H}_b, \mathbf{H}_r)]^{-1} (\mathbf{H}_b, \mathbf{H}_r)^T \tag{20}$$

which in combination with Eq. (8) yields

$$\begin{aligned} \mathbf{H}_r^T \mathbf{P}^{-1} \mathbf{H}_r &= (\mathbf{O}, \mathbf{I}_{n-m}) [(\mathbf{H}_b, \mathbf{H}_r)^T \mathbf{P} (\mathbf{H}_b, \mathbf{H}_r)]^{-1} (\mathbf{O}, \mathbf{I}_{n-m})^T \\ &= [\mathbf{H}_r^T \mathbf{P} \mathbf{H}_r - \mathbf{H}_r^T \mathbf{P} \mathbf{H}_b (\mathbf{H}_b^T \mathbf{P} \mathbf{H}_b)^{-1} \mathbf{H}_b^T \mathbf{P} \mathbf{H}_r]^{-1} \\ &= (\mathbf{H}_r^T \mathbf{P} \mathbf{R}_{\mathbf{H}_b} \mathbf{H}_r)^{-1} \end{aligned} \tag{21}$$

By virtue of Eqs. (7), (13), (19) and (21), we have

$$\begin{aligned} \mathbf{H}_r \cdot \mathbf{P}_r \cdot \mathbf{H}_r^T &= \mathbf{H}_r \cdot \mathbf{H}_r^T \mathbf{P} \mathbf{R}_{\mathbf{H}_b} \mathbf{H}_r \cdot \mathbf{H}_r^T \\ &= (\mathbf{H}_b \mathbf{H}_b^T + \mathbf{H}_r \mathbf{H}_r^T) \mathbf{P} \mathbf{R}_{\mathbf{H}_b} (\mathbf{H}_b \mathbf{H}_b^T + \mathbf{H}_r \mathbf{H}_r^T) \\ &= \mathbf{P} \mathbf{R}_{\mathbf{H}_b} \end{aligned} \tag{22}$$

It follows that

$$\hat{\mathbf{X}}_r = \hat{\mathbf{X}}_{\nabla} \tag{23}$$

The weighted sum of squares of the LS residuals in this multiple case-deletion model reads

$$\begin{aligned} \Omega_r &= (\mathbf{H}_r^T \mathbf{L} - \mathbf{H}_r^T \mathbf{A} \hat{\mathbf{X}}_r)^T \cdot \mathbf{P}_r \cdot (\mathbf{H}_r^T \mathbf{L} - \mathbf{H}_r^T \mathbf{A} \hat{\mathbf{X}}_r) \\ &= (\mathbf{L} - \mathbf{A} \hat{\mathbf{X}}_{\nabla})^T \mathbf{P} \mathbf{R}_{\mathbf{H}_b} (\mathbf{L} - \mathbf{A} \hat{\mathbf{X}}_{\nabla}) = \Omega_{\nabla} \end{aligned}$$

and thus

$$\hat{\sigma}_r^2 = \frac{\Omega_r}{(n - m) - u} = \hat{\sigma}_{\nabla}^2 \tag{24}$$

It can be seen from Eqs. (23) and (24) that the mean-shift outlier model is equivalent to the multiple case-deletion model, no matter whether the deleted observations are correlated with the remaining or not.

According to the above discussions, one can conclude that the adjustment outputs are equal to each other no matter whether the (potential) outliers are deleted explicitly or implicitly, even though the removed observations are correlated with the remaining ones.

3.3 Computational consideration

With Eq. (12), one has to deal with the two matrix inversions with orders u and m , as opposed to the two matrix inversions with orders u and $n - m$ in Eq. (18). Therefore, Eq. (12) outperforms Eq. (18) in term of computational efficiency for in most applications the number of outliers m is small relative to the number of the original observations n .

However, the computational burden can be further reduced by taking the partitioned structure of the normal matrix in Eq. (10) into account. In fact, the normal equation (10) can also be solved as

$$\begin{cases} \hat{\nabla} = (\mathbf{H}_b^T \mathbf{P} \mathbf{R} \mathbf{H}_b)^{-1} \mathbf{H}_b^T \mathbf{P} \mathbf{R} \mathbf{L} \\ \hat{\mathbf{X}}_{\nabla} = (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{P} (\mathbf{L} - \mathbf{H}_b \hat{\nabla}) \end{cases} \quad (25)$$

or in more explicit form

$$\begin{cases} \hat{\nabla} = (\mathbf{H}_b^T \mathbf{P} \mathbf{R} \mathbf{H}_b)^{-1} \mathbf{H}_b^T \mathbf{P} \mathbf{V} \\ \hat{\mathbf{X}}_{\nabla} = \hat{\mathbf{X}} - (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{P} \mathbf{H}_b \hat{\nabla} \end{cases} \quad (26)$$

with which we obtain

$$\mathbf{V}_{\nabla} = \mathbf{R} (\mathbf{L} - \mathbf{H}_b \hat{\nabla}) \quad (27)$$

and

$$\begin{aligned} \Omega_{\nabla} &= \mathbf{V}_{\nabla}^T \mathbf{P} \mathbf{V}_{\nabla} \\ &= (\mathbf{L} - \mathbf{H}_b \hat{\nabla})^T \mathbf{P} \mathbf{R} (\mathbf{L} - \mathbf{H}_b \hat{\nabla}) \\ &= \Omega - \hat{\nabla}^T \mathbf{H}_b^T \mathbf{P} \mathbf{R} \mathbf{H}_b \hat{\nabla} \\ &= \mathbf{L}^T [\mathbf{P} \mathbf{R} - \mathbf{P} \mathbf{R} \mathbf{H}_b (\mathbf{H}_b^T \mathbf{P} \mathbf{R} \mathbf{H}_b)^{-1} \mathbf{H}_b^T \mathbf{P} \mathbf{R}] \mathbf{L} \end{aligned} \quad (28)$$

Apparently, in this situation it only requires extra calculation of the inverse of the $m \times m$ normal matrix $\mathbf{H}_b^T \mathbf{P} \mathbf{R} \mathbf{H}_b$. As a by-product, the estimate of the vector of the disturbance parameters ∇ can also be obtained with Eq. (26). This is a sufficient reason for choosing the mean-shift outlier model over the case-deletion model from the computational point of view.

4 Quality Assessment of outlying observations

With Sherman-Morrison-Woodbury-Schur formula (Strang and Borre 1997), we have

$$\begin{aligned} (\mathbf{A}^T \mathbf{P} \mathbf{R}_{H_b} \mathbf{A})^{-1} &= (\mathbf{A}^T \mathbf{P} \mathbf{A} - \mathbf{A}^T \mathbf{P} \mathbf{H}_b (\mathbf{H}_b^T \mathbf{P} \mathbf{R} \mathbf{H}_b)^{-1} \mathbf{H}_b^T \mathbf{P} \mathbf{A})^{-1} \\ &= (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} + (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{P} \mathbf{H}_b [\mathbf{H}_b^T \mathbf{P} \mathbf{R} \mathbf{H}_b \\ &\quad - \mathbf{H}_b^T \mathbf{P} \mathbf{A} (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{P} \mathbf{H}_b]^{-1} \mathbf{H}_b^T \mathbf{P} \mathbf{A} (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \\ &= (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} + (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{P} \mathbf{H}_b (\mathbf{H}_b^T \mathbf{P} \mathbf{R} \mathbf{H}_b)^{-1} \mathbf{H}_b^T \mathbf{P} \mathbf{A} (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \end{aligned} \quad (29)$$

This formula states the apparent increase in precision when the outlying observations should have been taken into account but were neglected under the assumption that *a priori* variance factor is known before (Schaffrin 1997).

The second term of Eq. (29) has a quadratic form, it follows that

$$[(A^T P R H_b A)^{-1}]_{ii} \geq [(A^T P A)^{-1}]_{ii}, \quad i = 1, 2, \dots, u \tag{30}$$

This inequality shows that all types of DOP metrics (Strang and Borre 1997) will be over-optimistic if the outliers were ignored, even though the outlying observations are correlated with the remaining ones.

After some matrix manipulation, it follows that

$$\Omega_r = L^T H_r \cdot P_r R_r \cdot H_r^T L \tag{31}$$

where $R_r = I_r - H_r^T A \cdot (A^T H_r P_r H_r^T A)^{-1} \cdot A^T H_r \cdot P_r$.

By virtue of Eqs. (28) and (31) and since the two quadratic forms, Ω_{∇} and Ω_r , are equivalent for any realization of the random observational vector L , we have

$$H_r \cdot P_r R_r \cdot H_r^T = P R - P R H_b (H_b^T P R H_b)^{-1} H_b^T P R \tag{32}$$

Obviously, the k th observation in the multiple case-deletion model is just the i_{m+k} th one in the original linear Gauss–Markov model. Consequently, we get

$$H_r^T h_{i_{m+k}} = \tilde{h}_k \tag{33}$$

where \tilde{h}_k denotes the k th $(n - m)$ -dimensional canonical unit vector with 1 as its i th entry.

The k th Baarda’s w -test in the multiple case-deletion model reads (Baarda 1968)

$$\tilde{w}_k = \frac{\tilde{h}_k^T P_r R_r H_r^T L}{\sigma_0 \sqrt{\tilde{h}_k^T P_r R_r \tilde{h}_k}} \sim \mathcal{N}(0, 1) \tag{34}$$

The corresponding MDB measure is given by

$$\sigma_0 \sqrt{\frac{\lambda_0}{\tilde{h}_k^T P_r R_r \tilde{h}_k}} = \sigma_0 \sqrt{\frac{\lambda_0}{h_{i_{m+k}}^T H_r P_r R_r H_r^T h_{i_{m+k}}}} \tag{35}$$

which in combination with Eq. (32) yields

$$\sigma_0 \sqrt{\frac{\lambda_0}{\tilde{h}_k^T P_r R_r \tilde{h}_k}} \geq \sigma_0 \sqrt{\frac{\lambda_0}{h_{i_{m+k}}^T P R h_{i_{m+k}}}} \tag{36}$$

This indicates that all the MDB measures of the remaining observations will become larger.

5 Conclusions

Both the case-deletion model and the mean-shift outlier model can be employed to perform multiple case-deletion diagnostics for linear models. The advantage of the case-deletion model is its intuitive appeal, for the suspicious observations are removed explicitly. The mean-shift outlier model, in which the underlying observations are implicitly deleted, has found wider acceptance because of its computational simplicity. However, these two models are equivalent from the mathematical point of view. Under the assumption that *a priori* variance factor is known before, theoretical analyses indicate that the precision, MDB measure and all kinds of DOP metrics are all over-optimistic when outliers were neglected.

Acknowledgements This research was sponsored by National Key Basic Research Program of China (2012CB825604), and the Natural Science Foundation of China (Grant No. 40874007). The author is also supported by the China Scholarship Council (File No. 2011317045).

References

- Baarda W (1968) A testing procedure for use in geodetic networks. Publication on geodesy, vol 2(5). Netherlands Geodetic Commission, Delft
- Barnett V, Lewis T (1994) Outliers in statistical data, 3rd edn. Wiley, New York
- Baselga S (2011) Acta Geod Geophys Hung 46(4):401–416
- Chatterjee S, Hadi AS (1988) Sensitivity analysis in linear regression. Wiley, New York
- Cook RD (1977) Technometrics 19(1):15–18
- Cook RD (1979) J Am Stat Assoc 74(365):169–174
- Guo J, Ou J, Wang H (2007) J Surv Eng 133(3):129–133
- Guo J, Ou J, Wang H (2010) J Geod 84(4):243–250
- Guo J, Ou J, Yuan Y (2011) J Surv Eng 137(1):9–13
- Hadi AS, Simonoff JS (1993) J Am Stat Assoc 88(424):1264–1272
- Hampel FR, Ronchetti EM, Rousseeuw PJ, Stahel WA (1986) Robust statistics: the approach based on influence functions. Wiley, New York
- Hekimoglu S (2005) ZFV, Z Vermess.wes 130(3):174–180
- Hekimoglu S, Koch KR (2000) AVN 107(7):247–253
- Hekimoglu S, Erdogan B, Erenoglu RC (2012) Exp Tech. doi:10.1111/j.1747-1567.2012.00876.x
- Huber PJ (1981) Robust statistics. Wiley, New York
- Kern M, Preimesberger T, Allesch M, Pail R, Bouman J, Koop R (2010) J Geod 78(9):509–519
- Koch KR (1999) Parameter estimation and hypothesis testing in linear models, 2nd edn. Springer, Berlin
- Kok JJ (1984) On data snooping and multiple outlier testing. NOAA technical report, NOS NGS 30, Rockville, MD
- Monhor D, Takemoto S (2005) Earth Planets Space 57(11):1009–1018
- Monhor D, Verö J (2011) Acta Geod Geophys Hung 46(1):84–92
- Pope AJ (1976) The statistics of residuals and the detection of outliers. NOAA technical report, NOS 65, NGS 1 Rockville, MD
- Rousseeuw PJ, Leroy AM (1987) Robust regression and outlier detection. Wiley, New York
- Schaffrin B (1997) J Surv Eng 123(3):126–137
- Snow KB, Schaffrin B (2003) GPS Solut 7(2):130–139
- Strang G, Borre K (1997) Linear algebra, geodesy, and GPS. Wellesley-Cambridge Press, Wellesley
- Wolf PR, Ghilani CD (1997) Adjustment computations: statistics and least squares in surveying and GIS, 3rd edn. Wiley, New York
- Xu P (2005) J Geod 79(1–3):146–159
- Yang Y (1999) J Geod 73(5):268–274