

Minimal detectable outliers as measures of reliability

Karl-Rudolf Koch

Received: 26 August 2014 / Accepted: 29 January 2015 / Published online: 12 February 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract The concept of reliability was introduced into geodesy by Baarda (A testing procedure for use in geodetic networks. Publications on Geodesy, vol. 2. Netherlands Geodetic Commission, Delft, 1968). It gives a measure for the ability of a parameter estimation to detect outliers and leads in case of one outlier to the MDB, the minimal detectable bias or outlier. The MDB depends on the non-centrality parameter of the χ^2 -distribution, as the variance factor of the linear model is assumed to be known, on the size of the outlier test of an individual observation which is set to 0.001 and on the power of the test which is generally chosen to be 0.80. Starting from an estimated variance factor, the F -distribution is applied here. Furthermore, the size of the test of the individual observation is a function of the number of outliers to keep the size of the test of all observations constant, say 0.05. The power of the test is set to 0.80. The MDBs for multiple outliers are derived here under these assumptions. The method is applied to the reconstruction of a bell-shaped surface measured by a laser scanner. The MDBs are introduced as outliers for the alternative hypotheses of the outlier tests. A Monte Carlo method reveals that due to the way of introducing the outliers, the false null hypotheses cannot be rejected on the average with a power of 0.80 if the MDBs are not enlarged by a factor.

Keywords MDB · Hypothesis test · Power of a test · B-splines · Monte Carlo method

1 Introduction

Baarda (1968) by his pioneering work introduced outlier tests to geodesy. He assumed the variance factor, i.e. the variance of unit weight, of the linear model to be known. Thus, the test statistic of his data snooping follows by the normalized residual having the normal distribution. Pope (1976) started from an estimated variance factor and derived as test statistic the studentized residual with the τ -distribution. Koch (1983), (1999, p. 302) and Kok (1984) showed that the tests for multiple outliers can be derived by the mean-shift model, cf. Baarda (1967, p. 21), Beckman and Cook (1983). Applied to searching for one outlier, it leads to the data snooping or the τ -test.

Baarda (1968) also introduced the concept of reliability. He looked at the alternative hypothesis of his test that the outlier is not equal to zero in contrast to the null hypothesis that the outlier equals zero. Different choices for alternative hypotheses for outlier tests were discussed by Lehmann (2013).

The non-central χ^2 -distribution follows for the alternative hypothesis of Baarda's test. He determined the non-centrality parameter by setting the power of the test to a fixed value, generally 0.80, which means the probability of rejecting a false null hypothesis is 0.80. In case of one outlier, the expression for the non-centrality parameter can be solved for the detectable outlier. Thus, a measure of reliability is obtained.

This measure was called marginally detectable outlier by Kok (1982) and is now named the minimal detectable bias (MDB), cf. Teunissen (2000, p. 102). The MDB is a scalar for one outlier. It is generalized for multiple outliers to a unit vector with as many components as there are outliers to be searched for. The vector indicates the positions of the outliers within the observations, and it is multiplied by a scalar with the same expression as the MDB for one outlier, cf. Knight

K.-R. Koch (✉)
Institute of Geodesy and Geoinformation, Theoretical Geodesy,
University of Bonn, Nussallee 17, 53115 Bonn, Germany
e-mail: koch@geod.uni-bonn.de

et al. (2010), Teunissen and de Bakker (2013). The MDB is defined for all combinations of the observations with respect to the number of outliers. It depends on the submatrix of the matrix of reliabilities of the observations, which follows from the reliabilities of those observations to be assumed as falsified by outliers.

Bounds for the MDB were computed by solving eigenvalue problems, thus obtaining measures of reliability, cf. Teunissen (2000, p. 105), Knight et al. (2010). The non-centrality parameter of the χ^2 -distribution, on which the MDB depends, can be expressed by a quadratic form and therefore by a hyperellipsoid, the axes of which result from the eigenvalues (Teunissen 1986). The MDB is also a function of the size of the test, i.e. the level of significance, which for practical applications is set to 0.001, cf. Staff of the Geodetic Computing Center (1982).

Measures of reliability for multiple outliers shall be derived here. As known, multiple outlier testing is difficult to successfully apply. Furthermore, if a certain number of outliers is searched for and all combinations of this number within the observations are tested, different outliers may produce identical residuals and therefore identical test statistics (Baselga 2011). This does not cause a problem here as the magnitudes of the detectable outliers are important, their distributions among the observations are less of a concern.

Estimated variance factors are assumed here, the test statistics are therefore F -distributed. In addition, the size of the test is not fixed to 0.001 but depends on the number of observations out of the following reason. If one outlier is searched in the observations, all observations have to be tested. To keep the size of the overall test constant, for instance 0.05, the size of the test of the individual observation is depending on the number of measurements. This has been shown by Pope (1976). Instead of computing the size of the test for the individual observation, Lehmann (2012) applied Monte Carlo methods to compute percentage points for maximal normalized and studentized residuals.

In case of multiple outliers, all combinations of the observations with respect to the number of outliers have to be tested. Like the non-centrality parameter of the χ^2 -distribution, the non-centrality parameter of the F -distribution can be represented by a quadratic form. The eigenvector of the quadratic form which belongs to the smallest eigenvalue leads to the detectable outliers. The minimum of these values among all combinations gives the MDBs, the minimal detectable outliers. Instead of generalizing the MDB for one outlier, it is derived here for multiple outliers.

The method is applied to data of a laser scanner. A bell-shaped surface represented by the tensor product of B-splines is determined by laser scanning. MDBs are computed and introduced into the alternative hypotheses of the outlier tests. The powers of the tests are determined by a Monte Carlo method from 100,000 random variates of the observations.

The paper is organized as follows: Sect. 2 recalls the method of multiple outlier testing in a mean-shift model. Section 3 derives the minimal detectable outliers. Section 4 deals with fitting a B-spline surface to the measurements of a laser scanner. Minimal detectable outliers are computed in Sect. 5 for the surface fitting and Sect. 6 gives the conclusions.

2 Multiple outlier testing

Let the linear model

$$\begin{aligned} X\boldsymbol{\beta} &= E(\mathbf{y}) = \mathbf{y} + \mathbf{e} \quad \text{with} \quad E(\mathbf{e}) = \mathbf{0} \\ \text{and} \quad D(\mathbf{y}) &= D(\mathbf{e}) = \sigma^2 \mathbf{P}^{-1} \end{aligned} \quad (1)$$

be given, where X denotes the $n \times u$ matrix of known coefficients with rank $X = u$, $\boldsymbol{\beta}$ denotes the $u \times 1$ vector of unknown parameters, \mathbf{y} the $n \times 1$ vector of observations, \mathbf{e} the $n \times 1$ vector of errors, σ^2 the unknown variance factor and \mathbf{P} the $n \times n$ positive definite weight matrix of the observations. The unknown parameters $\boldsymbol{\beta}$ are estimated by $\hat{\boldsymbol{\beta}}$ with

$$\hat{\boldsymbol{\beta}} = (X'PX)^{-1}X'P\mathbf{y}, \quad (2)$$

the vector $\hat{\mathbf{e}}$ of the residuals follows with

$$\hat{\mathbf{e}} = -(\mathbf{I} - X(X'PX)^{-1}X'P)\mathbf{y}, \quad (3)$$

its covariance matrix $D(\hat{\mathbf{e}})$ with

$$D(\hat{\mathbf{e}}) = \sigma^2 \mathbf{Q}_e = \sigma^2 (\mathbf{P}^{-1} - X(X'PX)^{-1}X') \quad (4)$$

and the quadratic form Ω of the residuals with

$$\Omega = \mathbf{y}'(\mathbf{P} - PX(X'PX)^{-1}X'P)\mathbf{y}. \quad (5)$$

The mean-shift model results from augmenting (1) by the $r \times 1$ vector $\boldsymbol{\epsilon}$ of unknown outliers and the $n \times r$ matrix \mathbf{Z} of known coefficients, cf. Koch (1999, p. 208) for a linear model with a matrix X of coefficients not of full rank,

$$|X, \mathbf{Z}| \begin{vmatrix} \boldsymbol{\beta} \\ \boldsymbol{\epsilon} \end{vmatrix} = E(\mathbf{y}) \quad \text{with} \quad D(\mathbf{y}) = \sigma_\epsilon^2 \mathbf{P}^{-1} \quad (6)$$

where σ_ϵ^2 denotes the unknown variance factor of the mean-shift model. Let's assume that a test for r outliers has to be derived. The number r can be arbitrarily chosen but should reflect the number of outliers which is expected in the observations. The outliers are ordered in a sequence starting with $\mathbf{y} = (y_i)$ at the observation y_k and are called $\Delta y_k, \Delta y_{k+1}, \dots, \Delta y_{k+r-1}$. They are subtracted from the observations $y_k, y_{k+1}, \dots, y_{k+r-1}$. Thus,

$$\begin{aligned} \boldsymbol{\epsilon} &= |\Delta y_k, \Delta y_{k+1}, \dots, \Delta y_{k+r-1}|' \quad \text{and} \\ \mathbf{Z} &= |\mathbf{e}_k, \mathbf{e}_{k+1}, \dots, \mathbf{e}_{k+r-1}| \end{aligned} \quad (7)$$

with

$$\mathbf{e}_l = |0, \dots, 0, 1, 0, \dots, 0|' \quad (8)$$

where the l th component with $l \in \{k, k + 1, \dots, k + r - 1\}$ gets the value 1. The estimate $\hat{\epsilon}$ of ϵ follows by

$$\hat{\epsilon} = \mathbf{S}\mathbf{Z}'\mathbf{P}\mathbf{Q}_e\mathbf{P}\mathbf{y} = -\mathbf{S}\mathbf{Z}'\mathbf{P}\hat{\epsilon} \tag{9}$$

with

$$\mathbf{S} = (\mathbf{Z}'\mathbf{P}\mathbf{Q}_e\mathbf{P}\mathbf{Z})^{-1}. \tag{10}$$

The matrix $\mathbf{Q}_e\mathbf{P}$ is called redundancy matrix, cf. Koch (1999, p. 305), Guo et al. (2011), because the i th diagonal element of $\mathbf{Q}_e\mathbf{P}$ gives the contribution of the observation y_i to the redundancy $n - u$

$$\text{tr}(\mathbf{Q}_e\mathbf{P}) = \text{tr}(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{P}\mathbf{X})^{-1}\mathbf{X}'\mathbf{P}) = n - u. \tag{11}$$

The multiple outlier test is defined by the null hypothesis H_0 that all outliers ϵ are equal to zero against the alternative hypothesis H_1 that outliers $\bar{\epsilon}$ are present

$$H_0 : \epsilon = \mathbf{0} \text{ versus } H_1 : \epsilon = \bar{\epsilon} \neq \mathbf{0}. \tag{12}$$

This hypothesis is introduced into the linear model (6) as special case of a general linear hypothesis. It gives the test statistic T by, cf. Koch (1999, p. 302),

$$T = \frac{R/r}{(\Omega - R)/(n - u - r)} \tag{13}$$

with

$$R = \hat{\epsilon}'\mathbf{S}^{-1}\hat{\epsilon} = \hat{\epsilon}'\mathbf{P}\mathbf{Z}\mathbf{S}\mathbf{Z}'\mathbf{P}\hat{\epsilon}. \tag{14}$$

In case the null hypothesis is true, T has the F -distribution

$$T \sim F(r, n - u - r). \tag{15}$$

The null hypothesis is rejected if

$$T > F_{1-\alpha; r, n-u-r} \tag{16}$$

where α denotes the size of the test and $F_{1-\alpha; r, n-u-r}$ the upper α -percentage point of the F -distribution. It is defined by

$$\int_0^{F_{1-\alpha; r, n-u-r}} F(r, n - u - r) dT = 1 - \alpha \tag{17}$$

where $F(r, n - u - r)$ now denotes the density function of the F -distribution. Thus,

$$\int_{F_{1-\alpha; r, n-u-r}}^\infty F(r, n - u - r) dT = \alpha \tag{18}$$

gives the probability α of the Type I error.

In case the alternative hypothesis is true, T has the non-central F -distribution

$$T \sim F'(r, n - u - r, \lambda) \tag{19}$$

with the non-centrality parameter λ

$$\lambda = \bar{\epsilon}'\mathbf{S}^{-1}\bar{\epsilon}/\sigma_\epsilon^2 \tag{20}$$

and

$$\sigma_\epsilon^2 = (\Omega - R)/(n - u - r). \tag{21}$$

The probability of the Type II error of the test is computed by

$$P(\text{Type II error}) = \int_0^{F_{1-\alpha; r, n-u-r}} F'(r, n - u - r, \lambda) dT = 1 - \beta \tag{22}$$

where $F'(r, n - u - r, \lambda)$ now denotes the density of the non-central F -distribution. The probability of rejecting a false null hypothesis is given by the power β of the test. The test by (13) is a uniformly most powerful invariant (UMPI) test, cf. Arnold (1981, p. 108), Kargoll (2008, p. 37). It maximizes the power within the class of unbiased, invariant tests.

If r outliers are searched in n observations, q test statistics have to be computed by the combination of r th order

$$\binom{n}{r} = q. \tag{23}$$

Let α be the size of testing the q combinations. The probability of the type I error of this test then results with (16) from

$$P(T_1 > F_{1-\alpha; r, n-u-r} \cup T_2 > F_{1-\alpha; r, n-u-r} \cup \dots \cup T_q > F_{1-\alpha; r, n-u-r}) = \alpha. \tag{24}$$

By applying the Bonferroni inequality, cf. Arnold (1981, p. 195), we get

$$P(T_1 > F_{1-\alpha; r, n-u-r}) + P(T_2 > F_{1-\alpha; r, n-u-r}) + \dots + P(T_q > F_{1-\alpha; r, n-u-r}) \geq \alpha \tag{25}$$

because of the relation which holds for $l \neq m$ and $l, m \in \{1, 2, \dots, q\}$, cf. Koch (2007, p. 7),

$$P(T_l > F_{1-\alpha; r, n-u-r} \cup T_m > F_{1-\alpha; r, n-u-r}) = P(T_l > F_{1-\alpha; r, n-u-r}) + P(T_m > F_{1-\alpha; r, n-u-r}) - P(T_l > F_{1-\alpha; r, n-u-r} \cap T_m > F_{1-\alpha; r, n-u-r}).$$

It follows approximately from (25)

$$P(T_i < F_{1-\alpha; r, n-u-r}) = 1 - \alpha/q. \tag{26}$$

Thus, the i th among the q combinations has to be tested with the size α/q to obtain α as size of the test for all q combinations. For such a test, the upper percentage point of the F -distribution in (17) needs to be replaced by $F_{1-\alpha/q; r, n-u-r}$ so that instead of (18) the Type I error follows from

$$\int_{F_{1-\alpha/q; r, n-u-r}}^\infty F(r, n - u - r) dT = \alpha/q. \tag{27}$$

By applying (26), the null hypothesis is accepted more often than by the inequality (25) as the upper percentage point is greater than the percentage point which results from the inequality.

3 Minimal detectable outliers

The ability of a parameter estimation to detect outliers has been called reliability by [Baarda \(1968\)](#). He started from the alternative hypothesis of his outlier test and determined the non-centrality parameter of the non-central χ^2 -distribution by assuming a fixed value for the power of the test, 0.80 is generally taken. The alternative hypothesis H_1 with the non-central F -distribution is used here. Furthermore, one of the q combinations is tested with the size α/q so that the upper percentage point $F_{1-\alpha/q;r,n-u-r}$ in (27) is applied, which follows from (17) by

$$\int_0^{F_{1-\alpha/q;r,n-u-r}} F(r, n - u - r) dT = 1 - \alpha/q. \tag{28}$$

The non-centrality parameter λ in (20) is determined by λ_0 with fixing the power β of the test to $\beta = 0.80$

$$\lambda_0 = \lambda(1 - \alpha/q; r, n - u - r, \beta = 0.80). \tag{29}$$

The power β follows from (22) by

$$\int_0^{F_{1-\alpha/q;r,n-u-r}} F'(r, n - u - r, \lambda_0) dT = 1 - \beta \tag{30}$$

so that λ_0 is computed as the $(1 - \beta)$ -percentage point $F'_{1-\beta;r,n-u-r,\lambda_0}$ of the non-central F -distribution. The integral in (28) and the integral in (30) with λ_0 being known are integrated over the identical domain $[0, F_{1-\alpha/q;r,n-u-r}]$, therefore

$$F'_{1-\beta;r,n-u-r,\lambda_0} = F_{1-\alpha/q;r,n-u-r}. \tag{31}$$

The cumulative distribution functions of the F -distribution and of the non-central F -distribution are computed by the series expansion of the incomplete beta function, cf. [Koch \(1999, p. 116\)](#). To obtain the percentage points, approximate values are chosen and then iteratively corrected in nested intervals ([Gaida and Koch 1985](#)), also called binary search ([Schader and Schmid 1986](#)).

We now look at the q combinations with respect to the r outliers. Let $\bar{\epsilon}_j$ with $j \in \{k, k + 1, \dots, k + r - 1\}$ be one $r \times 1$ vector of outliers. The non-centrality parameter λ_0 can then be expressed with (10), (12) and (20) by the quadratic form

$$\lambda_0 = \bar{\epsilon}'_j \mathbf{Z}' \mathbf{P} \mathbf{Q}_e \mathbf{P} \mathbf{Z} \bar{\epsilon}_j / \sigma_\epsilon^2. \tag{32}$$

It shall be minimized subject to the condition $\bar{\epsilon}'_j \bar{\epsilon}_j = 1$, cf. [Golub and van Loan \(1984, p. 267\)](#),

$$\bar{\epsilon}'_j \mathbf{Z}' \mathbf{P} \mathbf{Q}_e \mathbf{P} \mathbf{Z} \bar{\epsilon}_j / \sigma_\epsilon^2 \rightarrow \min \text{ for } \bar{\epsilon}'_j \bar{\epsilon}_j = 1. \tag{33}$$

Let the eigenvector $\bar{\epsilon}_{jm}$ belonging to the smallest eigenvalue of the symmetric matrix $\mathbf{Z}' \mathbf{P} \mathbf{Q}_e \mathbf{P} \mathbf{Z}$ give the minimum

$$\bar{\epsilon}'_{jm} \mathbf{Z}' \mathbf{P} \mathbf{Q}_e \mathbf{P} \mathbf{Z} \bar{\epsilon}_{jm} / \sigma_\epsilon^2 = \min \text{ for } \bar{\epsilon}'_{jm} \bar{\epsilon}_{jm} = 1. \tag{34}$$

The smallest value of $\bar{\epsilon}'_{jm} \mathbf{Z}' \mathbf{P} \mathbf{Q}_e \mathbf{P} \mathbf{Z} \bar{\epsilon}_{jm}$ is then determined among all q combinations. To find it, we get by the factor f since the quadratic form (34) has to be equal to λ_0 from (32)

$$\lambda_0 = f \bar{\epsilon}'_{jm} \mathbf{Z}' \mathbf{P} \mathbf{Q}_e \mathbf{P} \mathbf{Z} \bar{\epsilon}_{jm} / \sigma_\epsilon^2. \tag{35}$$

Solving for f , we obtain the expression which has to be minimized

$$f = \sqrt{\frac{\lambda_0 \sigma_\epsilon^2}{\bar{\epsilon}'_{jm} \mathbf{Z}' \mathbf{P} \mathbf{Q}_e \mathbf{P} \mathbf{Z} \bar{\epsilon}_{jm}}}. \tag{36}$$

Let the minimum be attained with $\bar{\epsilon}_{lm}$, we then find the MDBs, the minimal detectable biases or the minimal detectable outliers by, cf. [Knight et al. \(2010\)](#), [Teunissen and de Bakker \(2013\)](#),

$$MDB = f \bar{\epsilon}_{lm} = \sqrt{\frac{\lambda_0 \sigma_\epsilon^2}{\bar{\epsilon}'_{lm} \mathbf{Z}' \mathbf{P} \mathbf{Q}_e \mathbf{P} \mathbf{Z} \bar{\epsilon}_{lm}}} \bar{\epsilon}_{lm} \text{ for } \bar{\epsilon}'_{lm} \bar{\epsilon}_{lm} = 1. \tag{37}$$

They give a measure for the reliability of the parameter estimation.

The matrix $\mathbf{P} \mathbf{Q}_e \mathbf{P}$ is therefore called the reliability matrix \mathbf{R} of the observations

$$\mathbf{R} = \mathbf{P} \mathbf{Q}_e \mathbf{P} = (r_{ij}), \tag{38}$$

although [Proszynski \(2010\)](#) denotes the redundancy matrix $\mathbf{Q}_e \mathbf{P}$ in (11) as reliability matrix. The $r \times r$ submatrix $\mathbf{Z}' \mathbf{P} \mathbf{Q}_e \mathbf{P} \mathbf{Z}$ in (37) is found with \mathbf{Z} from (7) by

$$\mathbf{Z}' \mathbf{P} \mathbf{Q}_e \mathbf{P} \mathbf{Z} = \begin{vmatrix} r_{kk} & r_{k,k+1} & \dots & r_{k,k+r-1} \\ \dots & \dots & \dots & \dots \\ r_{k+r-1,k} & r_{k+r-1,k+1} & \dots & r_{k+r-1,k+r-1} \end{vmatrix} \tag{39}$$

The MDB for $r = 1$ follows from (37) with

$$MDB = \sqrt{\frac{\lambda_0 \sigma_\epsilon^2}{r_{mm}}}. \tag{40}$$

where r_{mm} denotes the minimal diagonal element of the reliability matrix $\mathbf{P} \mathbf{Q}_e \mathbf{P}$. The non-centrality parameter λ_0 for the non-central χ^2 -distribution with $\alpha = 0.001$ and $1 - \beta = 0.20$ is obtained from, cf. [Baarda \(1968\)](#), [Koch \(1983\)](#),

$$\lambda_0 = 4.13^2 = 17.05. \tag{41}$$

If in addition to the MDBs, maximal detectable outliers shall be computed, the eigenvector belonging to the largest eigenvalue is then used to find the maximum of (36) among the q combinations. The maximal detectable outliers have not been determined here.

4 B-spline surface

A B-spline surface in three-dimensional space is a function of two parameters which are called ξ_1 and ξ_2 . It is obtained

Let the $n_p \times 1$ vector \hat{e} collect the residuals for the coordinates y_i from $\text{vec} \hat{E}$. The estimate $\hat{\sigma}^2$ of the variance factor σ^2 in (1) then follows by

$$\hat{\sigma}^2 = \hat{e}'\hat{e}/(e_1e_2 - I_1I_2). \tag{55}$$

The lofting method (Koch 2009) estimates the unknown control points more efficiently than (53). However, it cannot be applied because of the weight matrix P .

5 Numerical example

The rectangular coordinates x_i and z_i with $i \in \{1, \dots, n_p\}$ of a grid of 26 rows and 28 columns with $n_p = 26 \times 28 = 728$ points together with the coordinates y_i , which represent a bell-shaped surface, were measured by the laser scanner HDS 3000, see Fig. 1. The blue colors of the points indicate high intensities of the reflected laser beams. The coordinates refer to the local coordinate system of the instrument with the x -axis lying horizontally, the y -axis coinciding with the center of the lines of sight and the z axis pointing to the zenith. The scans start at the lower left corner of the grid from negative to positive z values with increasing x values and end at the upper right corner. The distances between the points on the surface are about 1 cm, the shortest y coordinate is 501 cm.

To determine the standard deviations of the observed coordinates, the points of the grid were measured in addition with $n_w = 15$ repetitions. With the mean \bar{x}_i of the repeatedly measured coordinates x_{ij} for $i \in \{1, \dots, n_p\}$, $j \in \{1, \dots, n_w\}$ from

$$\bar{x}_i = \frac{1}{n_w} \sum_{j=1}^{n_w} x_{ij}, \tag{56}$$

we compute the variance $\sigma_{x_i}^2$ of x_i by

$$\sigma_{x_i}^2 = \frac{1}{n_w - 1} \sum_{j=1}^{n_w} (x_{ij} - \bar{x}_i)^2 \tag{57}$$

and accordingly $\sigma_{y_i}^2$ and $\sigma_{z_i}^2$ as well as the standard deviations σ_{x_i} , σ_{y_i} , σ_{z_i} by taking the square roots. The mean values $\bar{\sigma}_{x_i}$, $\bar{\sigma}_{y_i}$ and $\bar{\sigma}_{z_i}$ of the n_w standard deviations are obtained with

$$\bar{\sigma}_{x_i} = 0.016 \text{ cm}, \bar{\sigma}_{y_i} = 0.097 \text{ cm}, \bar{\sigma}_{z_i} = 0.003 \text{ cm}. \tag{58}$$

As mentioned in Sect. 4, the B-spline surface (42) is fitted by a least squares adjustment to the y coordinates. The more knots are introduced, which means the more polynomial segments and the more unknown control points because of (45) and (46), the better the fit. The number of knots is chosen such that the variances for the coordinates y are not degraded. This is accomplished if the variance factor σ^2 is estimated from (55) by $\sqrt{\hat{\sigma}^2} \approx 1$. With 13 knots in the direction of z , 14 knots in the direction of x and $q_1 = q_2 = 3$, the number of unknown control points is 90 and the estimate from (55)

$$\sqrt{\hat{\sigma}^2} = 1.080. \tag{59}$$

Due to the local property of the B-splines, the more polynomial segments are chosen, the more difficult becomes the detection of outliers as the coordinates y do not control each other. On the other hand because of the local property, evenly distributed outliers do not much affect the estimation of the control points.

To determine by (37) the MDBs, one, two and three outliers have been assumed, i.e. $r \in \{1, 2, 3\}$, see first column of Table 1. Depending on q in (23), the submatrix (39) following from the reliability matrix (38) has to be determined. The variance σ_ϵ^2 is also needed in (37), it follows with (5) and (14) from (21). The non-centrality parameter λ_0 in (37) is computed with $\beta = 0.80$ as the $(1-\beta)$ -percentage point of the non-central F -distribution from (30). It follows from the upper percentage point (28) of the F -distribution with size α/q for the test of the individual observation. The size of the outlier test (12) is chosen as $\alpha = 0.05$. The numbers q of combinations to be tested are obtained from (23) with $n = n_p = 728$ for $r \in \{1, 2, 3\}$. They are given in the second column of Table 1 together with the values for the size $0.05/q$. The upper percentage points (28) and the non-centrality parameters λ_0 follow in the third column of Table 1.

The quadratic form (32) is minimized with computing the eigenvalues and eigenvectors by the Jacobi method, cf. Golub and van Loan (1984, p. 295). The smallest value of (36) is

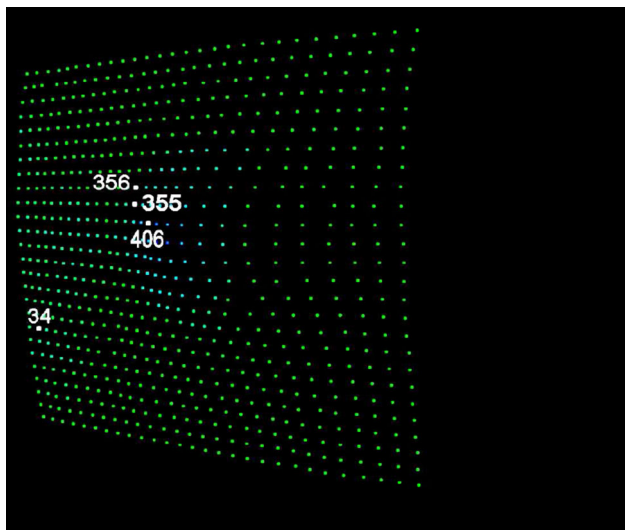


Fig. 1 Points representing the bell-shaped surface from coordinates x , y , z measured by the laser scanner. One minimal detectable outlier lies at point 355, two at points 356 and 406, three at 34, 355 and 406

Table 1 Number r of outliers, combinations q and size $0.05/q$ of the tests, upper percentage points and non-centrality parameter λ_0 , minimal detectable outliers with magnitudes and points, standard deviation

$\sqrt{\hat{\sigma}^2}$, test statistic T , power β from 100,000 generated coordinates y_i , factor for minimal detectable outliers, test statistic T and power β

| r | q Size $0.05/q$ | Percent. p. λ_0 | Min. detect. outl. magn. points (cm) | | $\sqrt{\hat{\sigma}^2}$ | | | |
|-----|----------------------|----------------------------|--------------------------------------|-----|-------------------------|--------------------------|-------|-------|
| | | | | | T β | Factor T β | | |
| 1 | 728 | 16.06 | 0.262 | 355 | 1.102 | 1.32 | | |
| | 0.686e-06 | 23.55 | | | 25.78 | 43.94 | | |
| | | | | | 0.268 | 0.799 | | |
| 2 | 264,628 | 15.86 | 0.122 | 356 | 1.097 | 1.94 | | |
| | 0.189e-08 | 23.31 | | | 0.238 | 406 | 10.15 | 41.13 |
| | | | | | | | 0.000 | 0.805 |
| 3 | 64,039,976 | 15.66 | 0.030 | 34 | 1.100 | 2.22 | | |
| | 0.781e-11 | 23.06 | | | 0.136 | 355 | 8.651 | 40.12 |
| | | | | | 0.234 | 406 | 0.000 | 0.805 |

determined among the q combinations. As can be recognized in Table 1, the number q of combinations increases rapidly with the number r of outliers. Instead of systematically searching the minimum of (36) for large numbers of q , one could randomly seek for r outliers among the n_p coordinates y_i to which points the minimum is moving. The search can then be restricted to the vicinity of these points. The minimum leads to the MDBs from (37). Their magnitudes and the numbers of the points where they occur are presented in the following two columns of Table 1. The points are also shown in Fig. 1. The numbering of the points follows the sequence of the scans from the lower left corner of the grid to the upper right corner.

The MDB from (40) with λ_0 from (41) of Baarda (1968) equals 0.223 cm. The difference against 0.262 in Table 1 is caused by applying the F -distribution instead of the χ^2 -distribution and by the size $0.05/728=0.686e-06$ of the test instead of 0.001.

The MDBs are introduced as outliers with positive signs for the coordinates y_i at the points where they occur. One could have selected the sign such that the standard deviations $\sqrt{\hat{\sigma}^2}$ for the surface fit increase. This has not been tried to avoid the impression of manipulating the data. The sixth column in Table 1 shows $\sqrt{\hat{\sigma}^2}$.

The MDBs give the outliers for the alternative hypothesis H_1 in (12). To find out whether the false null hypothesis H_0 for such a case is on the average rejected with a power of 0.80, independent normally distributed errors with zero expectations and standard deviations σ_{y_i} from (57) times $\sqrt{\hat{\sigma}^2}$ from Table 1 are added to the coordinates y_i plus outliers, cf. Koch (2007, p. 197), to obtain 100,000 random variates of the coordinates y_i and to compute the test statistics T from (13). The average percent of rejecting the false null hypothesis, i.e. the power β , is thus determined.

By introducing in case of $r = 1$ the MDB for point 355 as outlier, $\sqrt{\hat{\sigma}^2}$ increases from 1.080 of (59) to 1.102. The test statistic $T = 25.78$ given in the sixth column of Table 1 is larger than the upper percentage point of 16.06 of the third column of Table 1. The null hypothesis that the outlier is equal to zero is therefore rejected. However, a power $\beta = 0.268$ is only reached for the minimal detectable outlier of point 355 as shown in Table 1. The reason might be the local property of the B-splines as mentioned after (59). Furthermore, the MDB = 0.262 cm is quite optimistic. The standard deviation from (58) times $\sqrt{\hat{\sigma}^2}$ of Table 1 is $\bar{\sigma}_{y_i} \sqrt{\hat{\sigma}^2} = 0.107$ cm. This standard deviation has to be multiplied by only 2.45 to obtain the MDB = 0.262. The factor of 3.00 according to the 3 σ -rule would have allowed a better detection of the outlier.

In case of $r = 2$ outliers, the standard deviation $\sqrt{\hat{\sigma}^2}$ in Table 1 drops from 1.102 to 1.097. As the MDBs for points 356 and 406 are introduced as outliers with positive signs, they improve the surface fit by reducing the residuals of the coordinates y_i of these points. The test statistic $T = 10.15$ is smaller than the percentage point 15.86 and the null hypothesis is accepted. The power $\beta = 0.000$ therefore follows. By introducing three outliers, $\sqrt{\hat{\sigma}^2}$ in Table 1 increases to 1.100 but it is still smaller than 1.102. The test statistic $T = 8.651$ is smaller than the percentage point 15.66. The null hypothesis is accepted and the power is $\beta = 0.000$.

The seventh column of Table 1 shows together with the test statistic T the factors by which the MDBs have to be enlarged in order to obtain the power $\beta = 0.80$ for 100,000 random variates for the coordinates y_i . In case of two and three outliers, the effects of improving the surface fits by adding positive MDBs have to be compensated by enlarging the MDBs so that the outliers act as outliers. The necessary factors of 1.94 and 2.22 approximate the factor of 2.45 given above to obtain the MDB = 0.262 cm from the standard

deviation. This explains the increase of the MDBs to obtain $\beta = 0.80$.

6 Conclusions

Starting from a linear model with an unknown variance factor, the mean-shift model is presented for multiple outlier testing. The test statistic, which has the F -distribution, and the non-centrality parameter, in case the alternative hypothesis is true, are given. Outliers are searched for all combinations of the observations with respect to the number of outliers. To keep the size of the outlier test constant, the test of the individual observations has to depend on the number of combinations. The power of the test is set to 0.80.

The non-centrality parameter is computed as the $(1-0.80)$ -percentage point of the non-central F -distribution. It is a function of a quadratic form, which is minimized by computing the eigenvalues and eigenvectors. The smallest minimum among all combinations gives the MDBs, the minimal detectable outliers.

The method is applied to the reconstruction of a bell-shaped surface measured by a laser scanner. The MDBs of one, two and three outliers are determined together with the points where they occur. The MDBs are introduced as given outliers into the alternative hypothesis of the outlier test so that the null hypothesis is false. A Monte Carlo method checks whether on the average the false null hypotheses are rejected with a power of 0.80. It turns out that this is not the case. The MDBs have to be multiplied by 1.32 for one, 1.94 for two and 2.22 for three outliers. The reasons for enlarging the MDBs are outlined.

Acknowledgments The author is indebted to Boris Kargoll for his valuable comments and to Ernst-Martin Blome for assistance with the measurements and the graphical depiction of the results.

References

- Arnold SF (1981) The theory of linear models and multivariate analysis. Wiley, New York
- Baarda W (1967) Statistical concepts in geodesy. Publications on Geodesy, vol. 2, nr. 4. Netherlands Geodetic Commission, Delft
- Baarda W (1968) A testing procedure for use in geodetic networks. Publications on Geodesy, vol. 2, nr. 5. Netherlands Geodetic Commission, Delft
- Baselga S (2011) Nonexistence of rigorous tests for multiple outlier detection in least-squares adjustment. *J Surv Eng* 137:109–112
- Beckman RJ, Cook RD (1983) Outlier..s. *Technometrics* 25:119–149
- Cox MG (1972) The numerical evaluation of B-splines. *J Inst Math Appl* 10:134–149
- de Boor C (1972) On calculating with B-splines. *J Approx Theory* 6:50–62
- Gaida W, Koch KR (1985) Solving the cumulative distribution function of the noncentral F -distribution for the noncentrality parameter. *Sci Bull Stanislaw Staszic Univ Min Metall Geodesy* b.90 (1024): 35–44
- Golub GH, van Loan CF (1984) Matrix computations. The Johns Hopkins University Press, Baltimore
- Guo J-F, Ou J-K, Yuan Y-B (2011) Reliability analysis for a robust M-estimator. *J Surv Eng* 137:9–13
- Kargoll B (2008) On the theory and application of model misspecification tests in geodesy. Universität Bonn, Institut für Geodäsie und Geoinformation, Schriftenreihe 8, Bonn
- Knight NL, Wang J, Rizos C (2010) Generalised measures of reliability for multiple outliers. *J Geodesy* 84:625–635
- Koch KR (1983) Ausreißertests und Zuverlässigkeitsmaße. *Vermessungswesen und Raumordnung* 45:400–411
- Koch KR (1999) Parameter estimation and hypothesis testing in linear models, 2nd edn. Springer, Berlin
- Koch KR (2007) Introduction to Bayesian statistics, 2nd edn. Springer, Berlin
- Koch KR (2009) Identity of simultaneous estimates of control points and of their estimates by the lofting method for NURBS surface fitting. *Int J Adv Manuf Technol* 44:1175–1180
- Koch KR (2011) Data compression by multi-scale representation of signals. *J Appl Geod* 5:1–12
- Kok JJ (1982) Statistical analysis of deformation problems using Baarda's testing procedures. In: "Forty Years of Thought". Anniversary Volume on the Occasion of Prof. Baarda's 65th Birthday, vol 2. Delft, pp 470–488
- Kok JJ (1984) On data snooping and multiple outlier testing. In NOAA Technical Report NOS NGS 30. US Department of Commerce, National Geodetic Survey, Rockville, Maryland
- Lehmann R (2012) Improved critical values for extreme normalized and studentized residuals in Gauss-Markov models. *J Geodesy* 86:1137–1146
- Lehmann R (2013) On the formulation of the alternative hypothesis for geodetic outlier detection. *J Geodesy* 87:373–386
- Piegl L, Tiller W (1997) The NURBS book, 2nd edn. Springer, Berlin
- Pope AJ (1976) The statistics of residuals and the detection of outliers. In: NOAA Technical Report NOS65 NGS 1. US Department of Commerce, National Geodetic Survey, Rockville, Maryland
- Proszynski W (2010) Another approach to reliability measures for systems with correlated observations. *J Geodesy* 84:547–556
- Schader M, Schmid F (1986) Distribution function and percentage points for the central and noncentral F -distribution. *Stat Papers* 27:67–74
- Staff of the Geodetic Computing Center (1982) The Delft approach for the design and computation of geodetic networks. In: "Forty Years of Thought". Anniversary Volume on the Occasion of Prof. Baarda's 65th Birthday, vol 1. Delft, pp 202–274
- Teunissen PJG (1986) Adjusting and testing with the models of the affine and similarity transformation. *Manusc Geod* 11:214–225
- Teunissen PJG (2000) Testing theory; an introduction. Delft University of Technology, Department of Mathematical Geodesy and Positioning, Delft, MGP
- Teunissen PJG, de Bakker PF (2013) Single-receiver single-channel multi-frequency GNSS integrity: outliers, slips, and ionospheric disturbances. *J Geodesy* 87:161–177