



Delft University of Technology

## Minimal Detectable and Identifiable Biases for quality control

Imparato, D.; Teunissen, P. J.G.; Tiberius, C. C.J.M.

### DOI

[10.1080/00396265.2018.1437947](https://doi.org/10.1080/00396265.2018.1437947)

### Publication date

2018

### Document Version

Final published version

### Published in

Survey Review

### Citation (APA)

Imparato, D., Teunissen, P. J. G., & Tiberius, C. C. J. M. (2018). Minimal Detectable and Identifiable Biases for quality control. *Survey Review*, 51(367), 289-299. <https://doi.org/10.1080/00396265.2018.1437947>

### Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

### Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

### Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

***Green Open Access added to TU Delft Institutional Repository***

***'You share, we take care!' – Taverne project***

**<https://www.openaccess.nl/en/you-share-we-take-care>**

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.

# Minimal Detectable and Identifiable Biases for quality control

D. Imparato<sup>\*1</sup>, P. J. G. Teunissen<sup>1,2</sup> and C. C. J. M. Tiberius<sup>2</sup>

The Minimal Detectable Bias (MDB) is an important diagnostic tool in data quality control. The MDB is traditionally computed for the case of testing the null hypothesis against a single alternative hypothesis. In the actual practice of statistical testing and data quality control, however, multiple alternative hypotheses are considered. We show that this has two important consequences for one's interpretation and use of the popular MDB. First, we demonstrate that care should be exercised in using the single-hypothesis-based MDB for the multiple hypotheses case. Second, we show that for identification purposes, not the MDB, but the Minimal Identifiable Bias (MIB) should be used as the proper diagnostic tool. We analyse the circumstances that drive the differences between the MDBs and MIBs, show how they can be computed using Monte Carlo simulation and illustrate by means of examples the significant differences that one can experience between detectability and identifiability.

**Keywords:** Quality control, Detection–identification–adaptation (DIA), Minimal Detectable Bias (MDB), Minimal Identifiable Bias (MIB), Global Navigation Satellite Systems (GNSS)

## Introduction

Statistical testing for modelling errors and biases is an important component of any data quality control procedure (Alberda 1976, Amiri-Simkooei 2004, Yetkin et al. 2014, Klein et al. 2015, 2017). Such a testing procedure then usually consists of a detection, identification and adaptation (DIA) step (Baarda 1968, Teunissen 2000, Imparato 2016). In the detection step, one decides whether they can have enough confidence in the null hypothesis  $\mathcal{H}_0$ , which is, at onset, the assumed working model. Once confidence in  $\mathcal{H}_0$  has been declared, this model is eventually used to estimate the unknown parameters, usually in a least-squares sense. In case confidence in  $\mathcal{H}_0$  is lacking, subsequently identification of the potential source of model error is carried out. This implies a search among the specified alternative hypotheses  $\mathcal{H}_i$ ,  $i = 1, \dots, m$ , for the most likely model misspecification. After successful identification of the suspected model error, adaptation is needed to perform a corrective action on earlier  $\mathcal{H}_0$ -based inferences. Thus, with the null hypothesis rejected, the identified alternative hypothesis becomes the new null hypothesis and this model is then effectively used to provide estimates for the unknown parameters. For the distributional theory of the DIA method, including a discussion on how the first moment of the parameter estimators is impacted by the testing procedure, see Teunissen et al. (2017) and Teunissen (2017).

In this contribution, we analyse what diagnostic quantities are actually needed to adequately describe the performance of the testing procedure. We thereby distinguish between detection and identification, and also between single alternative hypothesis testing and multiple alternative hypotheses testing. For detection, Baarda (1967; 1968) introduced the important and well-known concept of the Minimal Detectable Bias (MDB). It describes, for each alternative hypothesis  $\mathcal{H}_i$ , the size of model error or bias that can be detected with a certain pre-set probability, when employing a detector test with user-defined false alarm rate (Van der Marel and Kesters 1990, Salzmann 1991, Teunissen 1998). Although the MDB is usually computed based on Baarda's normally distributed  $w$ -statistic, it can also be linked to the Chi-square distributed overall model test (OMT). The MDB is a general concept that applies in principle to any detection test for validating  $\mathcal{H}_0$ . Hence, when computing and reporting MDBs, one should always clarify which detection test is actually used for validating the null hypothesis. In this contribution, we consider as the alternative hypotheses outliers in individual observations, and the testing procedure itself will be based on Baarda's  $w$ -test statistic for data snooping (Baarda 1967). Thus with  $m$  observations, there will also be  $m$  alternative hypotheses.

There is no difference between detection and identification in case only a single alternative hypothesis is in play. In that case, rejection of  $\mathcal{H}_0$  automatically leads to the acceptance of the only alternative, and the corresponding MDB also describes the smallest bias that can be *identified*. As this equivalence disappears in multiple alternative hypotheses testing, one has to be aware of the following two consequences. First, as a different test

<sup>1</sup>GNSS Research Centre, Curtin University of Technology, Perth, Australia

<sup>2</sup>Department of Geoscience and Remote Sensing, Delft University of Technology, Delft, Netherlands

\*Corresponding author, email [davide.imparato@curtin.edu.au](mailto:davide.imparato@curtin.edu.au)

statistic will be used for *detection*, also the MDB of a hypothesis  $\mathcal{H}_i$  will then change from the single alternative hypothesis  $\text{MDB}_1(i)$  into the  $m$  alternative hypotheses  $\text{MDB}_m(i)$ . Second, as *identification* differs from detection in the  $m$  alternative hypotheses case, the MDBs will not represent the minimal size of biases that one is able to correctly identify. To obtain such information (Teunissen 2017), one will have to consider – in analogy with the Minimal Detectable Bias (MDB) – the Minimal Identifiable Bias (MIB). Earlier it has been recognised already (Prószyński 2015) that the MDB does not cover identification. In this contribution, we show that for the given testing procedure

$$\text{MDB}_m(i) \leq \text{MDB}_1(i) = \text{MIB}_1(i) \leq \text{MIB}_m(i) \quad (1)$$

This relationship has two implied consequences. First, it shows that the easy-to-compute Minimal Detectable Bias  $\text{MDB}_1(i)$  can be used as a safe upper bound for the MDB when  $m$  alternative hypotheses are considered, i.e.  $\text{MDB}_m(i)$ . We will analyse the circumstances under which such an approximation can be considered satisfactory. Second, the relationship implies that identification is more difficult than detection. This is something one has to be aware of when using identified hypotheses for adaptation purposes. The MDBs cannot be used as safe bounds for the MIBs. We will analyse the circumstances that drive the differences between the MDBs and MIBs, and in support of this analysis we also show how the  $m$ -hypotheses MDBs and MIBs can actually be computed using Monte Carlo simulation.

This contribution is organised as follows. We start in the section ‘Detection and the minimal detectable bias’ with a brief review of the necessary estimation and testing results of linear model theory. Then in the section ‘Detection and identification’, we introduce our testing procedure, give a proof of relation (1) and show how the multivariate minimal biases  $\text{MDB}_m(i)$  and  $\text{MIB}_m(i)$  can be numerically computed by means of Monte Carlo simulation. In the section ‘A qualitative description’, we provide a qualitative description, as a function of the normalised bias size and overall false alarm rate, of the behaviour of the univariate and multivariate probabilities of missed detection and correct identification. Then in the section ‘Terrestrial and satellite positioning’, we illustrate by means of two surveying examples the significant differences that one can indeed experience in practice between detectability and identifiability. The conclusions and findings are summarised in the section ‘Conclusions’.

## Detection and the MDB

In this section, we briefly review necessary estimation and testing results of linear model theory.

### Detection

Before any start can be made with statistical model validation, one needs to have a clear idea of the models one wants to work with. The null hypothesis  $\mathcal{H}_0$ , also referred to as working hypothesis, consists of the model one believes to be valid under nominal or normal working conditions. We assume the null hypothesis to be of the form

$$\mathcal{H}_0 : \mathbf{E}(y) = Ax, \quad \mathbf{D}(y) = Q_{yy} \quad (2)$$

with  $\mathbf{E}(\cdot)$  being the expectation operator,  $y \in \mathbb{R}^m$  the normally distributed random vector of observables,  $A \in \mathbb{R}^{m \times n}$  the known design matrix of rank  $n$  and  $x \in \mathbb{R}^n$  the unknown parameter vector.  $\mathbf{D}(\cdot)$  is the dispersion operator and  $Q_{yy} \in \mathbb{R}^{m \times m}$  the known positive-definite variance matrix of the observables. As  $A$  is assumed to be of full rank, the redundancy of the model is  $r=m-n$ . For example, in case of GNSS satellite positioning, model (2) is formed by the linear (ised) pseudorange code and/or carrier-phase observation equations; see, e.g. Leick (2004), Hofmann-Wellenhof et al. (2008) and Teunissen and Montenbruck (2017).

The first step in testing usually consists of a check on the overall validity of the null hypothesis  $\mathcal{H}_0$ . Such a detection step provides information on whether one can have confidence in the assumed null hypothesis. By choosing an acceptance region  $\mathcal{A}_\alpha \subset \mathbb{R}^m$ , the test takes the form

$$\text{Accept } \mathcal{H}_0 \text{ if } y \in \mathcal{A}_\alpha, \text{ reject otherwise} \quad (3)$$

The performance of such a test is described by its probabilities of false alarm  $\text{P}_{\text{FA}}$  and missed detection  $\text{P}_{\text{MD}}$ ,

$$\begin{aligned} \text{P}_{\text{FA}} &= P[y \notin \mathcal{A}_\alpha | \mathcal{H}_0 = \text{true}] = \alpha \\ \text{P}_{\text{MD}} &= P[y \in \mathcal{A}_\alpha | \mathcal{H}_0 = \text{false}] = \beta \end{aligned} \quad (4)$$

The probability of false alarm  $\alpha$ , also known as significance level, is the probability of rejecting the null hypothesis while it is true in fact (also known as a type-I error), while the probability of missed detection  $\beta$  is the probability of wrongfully accepting the null hypothesis (also known as a type-II error). The complement to the probability of missed detection,  $\text{P}_{\text{CD}} = 1 - \text{P}_{\text{MD}} = \gamma$ , is the probability of correct detection, also known as the power of the test. The false alarm probability can be computed by means of a multivariate integration of the null-hypothesis’ probability density function of  $y$  over the complement of  $\mathcal{A}_\alpha$ . For the computation of the missed detection, however, we first need to specify the alternative hypothesis  $\mathcal{H}_a$  to which  $\mathcal{H}_0$  is tested against. Here, we consider the following type of alternative hypotheses:

$$\mathcal{H}_a : \mathbf{E}(y) = Ax + Cb, \quad \mathbf{D}(y) = Q_{yy} \quad (5)$$

with  $b_y = Cb$  being the additional observation bias vector, in which  $b \in \mathbb{R}^q$  is unknown and matrix  $C \in \mathbb{R}^{m \times q}$  known. We have  $q \leq m - n = r$ , as the design matrix under  $\mathcal{H}_a$ ,  $[A, C] \in \mathbb{R}^{m \times (n+q)}$ , is assumed to be of full rank. Through the choice of matrix  $C$  one may model, for instance in GNSS applications, the presence of one or more blunders (outliers) in the code (pseudorange) data, cycle-slips in carrier-phase data, satellite failures, antenna-height errors, or incorrect neglectance of atmospheric delays.

For each possible alternative hypothesis  $\mathcal{H}_a$ , i.e. for each choice of matrix  $C$  and set values for the elements of vector  $b$ , one may compute the probability of missed detection  $\text{P}_{\text{MD}}$ . This boils down to a multivariate integration of the alternative hypothesis’ probability density function of  $y$  over the acceptance region  $\mathcal{A}_\alpha \subset \mathbb{R}^m$ . Alternatively, as one is often interested in the ‘size’ of bias vector  $b$  that can be detected with a certain pre-set probability, one may instead follow the inverse route. In that case one assumes a certain probability of missed detection,  $\text{P}_{\text{MD}} = \beta$  say, and then computes for a given  $C$ -matrix, the corresponding  $b$ -vector. This latter

approach then leads to the concept of the MDB as introduced by Baarda (1967).

### Minimal Detectable Bias (MDB)

No particular choice for the acceptance region  $\mathcal{A}_\alpha$  of  $\mathcal{H}_0$  has yet been made. Different choices lead to different probabilities of missed detection for the same  $\mathcal{H}_a$ . A commonly used choice is given by Baarda (1968) and Teunissen (2000)

$$\mathcal{A}_\alpha = \{y \in \mathbb{R}^m \mid \|\hat{e}\|_{Q_{yy}}^2 \leq \chi_\alpha^2(r, 0)\} \quad (6)$$

with the least-squares residual vector  $\hat{e} = P_A^\perp y$ , orthogonal projector  $P_A^\perp = I_m - A(A^T Q_{yy}^{-1} A)^{-1} A^T Q_{yy}^{-1}$ ,  $\|\hat{e}\|_{Q_{yy}}^2 = \hat{e}^T Q_{yy}^{-1} \hat{e}$ , and critical value  $\chi_\alpha^2(r, 0)$ . The critical value is computed from the central Chi-square distribution with  $r=m-n$  degrees of freedom and a chosen false alarm rate, or level of significance,  $\alpha$ . The detection test with the above acceptance region (6) is referred to as the OMT. It is known to be a uniformly most-powerful-invariant (UMPI) detector test, from Arnold (1981), Teunissen (2000) and Lehmann and Voß-Böhme (2017), when testing equation (2) against equation (5).

In order to determine the OMT's probability of missed detection, we need the distribution of the 'weighted-sum-of-squared-residuals' test statistic  $\|\hat{e}\|_{Q_{yy}}^2$  under  $\mathcal{H}_a$ . It is given as

$$\|\hat{e}\|_{Q_{yy}}^2 \sim \chi^2(r, \lambda_q) \quad (7)$$

with non-centrality parameter  $\lambda_q = b^T Q_{\hat{b}\hat{b}}^{-1} b$ , in which  $Q_{\hat{b}\hat{b}} = (\bar{C}^T Q_{yy}^{-1} \bar{C})^{-1}$  is the variance matrix of the best linear unbiased estimator (BLUE) of  $b$  under  $\mathcal{H}_a$ , and  $\bar{C} = P_A^\perp C$ . From combining equations (6) and (7), the probability of missed detection for hypothesis  $\mathcal{H}_a$  follows as  $P_{MD} = P[\chi^2(r, \lambda_q) \leq \chi_\alpha^2(r, 0)] = \beta$ .

This relation can also be used as a starting point for computing the OMT's Minimal Detectable Biases. First, the non-centrality parameter  $\lambda_q$  is computed from this relation as a function of  $\alpha, \beta$  and  $r$ , denoted as  $\lambda_q(\alpha, \beta, r)$ , the value of which provides then a yardstick for the detectable bias,

$$\lambda_q(\alpha, \beta, r) = b^T Q_{\hat{b}\hat{b}}^{-1} b = \|\bar{C}b\|_{Q_{yy}}^2 \quad (8)$$

This quadratic equation describes a hyper-ellipsoid in  $\mathbb{R}^q$  for the bias  $b$ . Biases outside this region will have a larger than  $\gamma = 1 - \beta$  probability of being detected under  $\mathcal{H}_a$  using the OMT, while biases inside it will have a smaller probability. A further inversion is possible to find the corresponding biases themselves. Let  $u$  be a unit vector ( $\|u\| = 1$ ) and parametrise the bias vector as  $b = \|b\|u$ . Substitution into equation (8) followed by inversion gives then,

$$b = \sqrt{\frac{\lambda_q(\alpha, \beta, r)}{\|\bar{C}u\|_{Q_{yy}}^2}} u \quad \text{with } \|u\| = 1 \quad (9)$$

This is the vectorial form of Baarda's Minimal Detectable Bias (MDB) (Baarda 1967, 1968). Baarda referred to his MDBs as 'boundary values' (in Dutch: grenswaarden); nowadays, more customary term MDB was introduced

in Teunissen (1989). The length  $\|b\|$  of the MDB vector is the smallest magnitude of the bias vector  $b$  that can be found with at least probability  $\gamma = 1 - \beta$  in the direction  $u$  using the detection test (6). By letting  $u$  vary over the unit sphere in  $\mathbb{R}^q$ , one obtains the whole range of MDBs that can be detected with probability  $\gamma$ . Applications can be found, for example, in Van Mierlo (1980) and Kok (1982) for deformation analysis, in Salzmann (1991) and Tiberius (1998) for GNSS positioning and navigation and in Teunissen (1986) for trend testing.

Baarda, in his work on the strength analysis of general purpose geodetic networks, applied his general MDB form to data snooping for outlier detection (Staff of DGCC 1982). In that case,  $q=1$ , and the  $m$  alternative hypotheses take the form

$$\mathcal{H}_i : E(y) = Ax + c_i b_i, \quad i = 1, \dots, m, \quad (10)$$

$$D(y) = Q_{yy}$$

with  $c_i$  being the canonical unit vector having 1 as its  $i$ th entry and zeroes otherwise. Since  $\sigma_{b_i}^2 = \|\bar{c}_i\|_{Q_{yy}}^2$  when  $q=1$ , the corresponding scalar MDBs,  $MDB(i) = |b_i|$ , for each of these  $m$  hypotheses can then be computed as

$$MDB(i) = \sigma_{b_i} \sqrt{\lambda_{q=1}(\alpha, \beta, r)}, \quad i = 1, \dots, m \quad (11)$$

The MDB gets smaller for larger  $\alpha$ , larger  $\beta$  (smaller power  $\gamma$ ) and smaller  $\sigma_{b_i}$  (more precisely estimated  $\hat{b}_i$ ).

### Detection and identification

In the remainder of this contribution, we restrict our attention to alternative hypotheses of the form (10), i.e. we restrict our attention to data snooping. And instead of detection only (accept or reject  $\mathcal{H}_0$ ), we are now also interested in the identification of the correct hypothesis.

### Test procedure

There would be no difference between detection and identification in case there would be only a single alternative hypothesis, i.e. if one would assume that only one of two hypotheses could be true,  $\mathcal{H}_0$  or  $\mathcal{H}_a$ . In that case, rejection of  $\mathcal{H}_0$  automatically implies acceptance of  $\mathcal{H}_a$ . In our case, however, we will be working with multiple alternatives. This implies that rejection of  $\mathcal{H}_0$  still requires a further identification as to which of the  $m$  alternatives to accept.

For data snooping, we make use of Baarda's  $w$ -test statistic, see Baarda (1968),

$$w_i = \frac{c_i^T Q_{yy}^{-1} \hat{e}}{\sqrt{c_i^T Q_{yy}^{-1} Q_{\hat{e}\hat{e}} Q_{yy}^{-1} c_i}} \sim N\left(\frac{b_i}{\sigma_{b_i}}, 1\right) \quad (12)$$

The UMPI test for testing  $\mathcal{H}_0$  against a single alternative  $\mathcal{H}_i$  of equation (10) would then be to reject  $\mathcal{H}_0$  in favour of  $\mathcal{H}_i$  if  $y \notin \mathcal{A}_{\alpha_1}^i$ , with the acceptance region given as  $\mathcal{A}_{\alpha_1}^i = \{y \in \mathbb{R}^m \mid |w_i| \leq \chi_{\alpha_1}(1, 0)\}$ , with  $\chi_{\alpha_1}(1, 0) = \sqrt{\chi_{\alpha_1}^2(1, 0)}$ . In our case, however, we do not have a single alternative, but multiple alternative hypotheses. We therefore take as the acceptance region for  $\mathcal{H}_0$  the intersection of the  $m$  individual acceptance regions  $\mathcal{A}_{\alpha_1}^i$ ,  $i = 1, \dots, m$ . We thus accept the null hypothesis if  $y \in \mathcal{A}_{\alpha_m} = \bigcap_{i=1}^m \mathcal{A}_{\alpha_1}^i$  (i.e. if none of the  $m$

$w$ -tests gets rejected), and reject otherwise. But as this is only the detection-part of the testing procedure, we still need to decide which  $\mathcal{H}_i$  to choose once  $\mathcal{H}_0$  is rejected. Here, we choose for  $\mathcal{H}_i$  if its  $w_i$  is largest in absolute value. The testing procedure can be summarised as

$$\text{Accept } \mathcal{H}_0 \text{ if } \max_{j \in \{1, \dots, m\}} |w_j| \leq k = \chi_{\alpha_1}(1, 0) \quad (13)$$

otherwise

$$\begin{aligned} \text{Accept } \mathcal{H}_i \text{ if } |w_i| &= \max_{j \in \{1, \dots, m\}} |w_j| > k \\ &= \chi_{\alpha_1}(1, 0) \end{aligned} \quad (14)$$

Hence, equation (13) is the detection step and equation (14) the identification step. Note, since we are here solely working with the  $m$   $w_i$ -statistics and alternative hypotheses of the form (10), that detector (13) is different from the one based on equation (6). The detector based on equation (6) is ellipsoidal shaped, while detector (13) is constructed from the intersections of  $m$  banded regions  $\mathcal{A}_{\alpha_1}^i$ ,  $i = 1, \dots, m$ . The two regions are shown in Fig. 1, with equation (6) on the left, and equation (13) on the right (note: the region  $\mathcal{A}_{\alpha_1}^i$  is shown to be spherical (circular) in the figure as the used metric is with respect to  $Q_{yy}$ ).

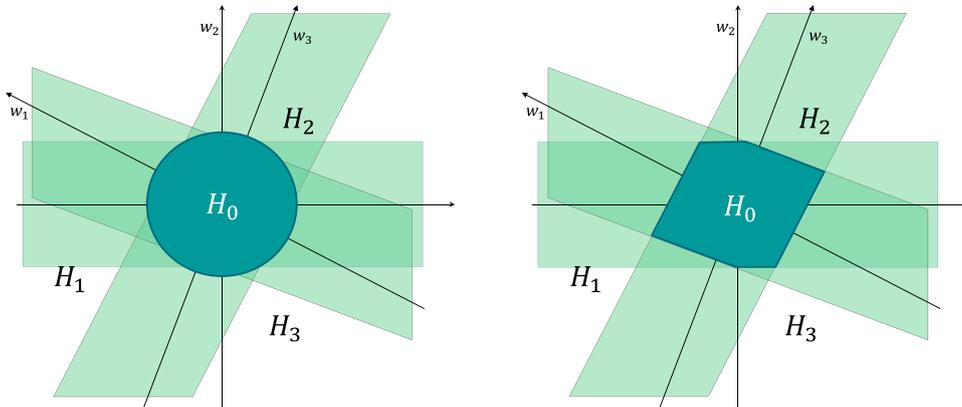
As the two acceptance regions are different, their MDBs for the same alternative hypothesis  $\mathcal{H}_i$  will be different as well, even when the same false alarm  $P_{FA}$  and the same missed detection probability  $P_{MD}$  are chosen. It is therefore important to explicitly describe the underlying testing procedure when presenting MDBs.

### Minimal Detectable Bias revisited

As the MDB depends on the chosen detector, the MDBs for equation (13) will differ from the OMT-based ones of equation (11). For the probability of missed detection through equation (13)  $P_{MD_m}$  we have, since  $\mathcal{A}_{\alpha_m} = \cap_{i=1}^m \mathcal{A}_{\alpha_1}^i$ ,

$$\begin{aligned} P_{MD_m}(i) &= P[y \in \mathcal{A}_{\alpha_m} | \mathcal{H}_i] \leq P[y \in \mathcal{A}_{\alpha_1} | \mathcal{H}_i] \\ &= P_{MD_1}(i) \end{aligned} \quad (15)$$

in which  $P_{MD_1}(i)$  is the missed detection probability in case one would have only one alternative hypothesis  $\mathcal{H}_i$



1 Two examples of acceptance regions for  $\mathcal{H}_0$ : ellipsoidal acceptance region  $\mathcal{A}_{\alpha}$  (cf. 6) of the OMT (left; dark circle) and polygonal acceptance region  $\mathcal{A}_{\alpha_m}$  (cf. 13) of the multiple  $w_i$ -test (right; dark polygon). The banded regions are those of  $\mathcal{A}_{\alpha_1}^i$ ,  $i=1,2,3$ , for respectively  $\mathcal{H}_1$ ,  $\mathcal{H}_2$  and  $\mathcal{H}_3$  ( $m=3$ )

( $P_{MD}$  with index 1 denotes the univariate probability of missed detection, and  $P_{MD}$  with index  $m$  the multivariate probability of missed detection). To contribute to the  $P_{MD_m}(i)$  the observation  $y$  needs to lie in the intersection of all  $m$  individual acceptance regions, so that none of them gets rejected, while, in fact,  $\mathcal{H}_i$  is true. The intersection is a smaller area than each of the individual ones; hence, the probability gets smaller,  $P_{MD_m}(i) \leq P_{MD_1}(i)$ . Thus, the missed detection probability gets smaller in case of multiple testing. The false alarm rate, however, will generally get larger,  $\alpha_m \geq \alpha_1$ . Would the  $w_i$ -test statistics all be independent, then  $1 - \alpha_m = (1 - \alpha_1)^m$ , which for small  $\alpha_1$  can be approximated as  $1 - \alpha_m \approx 1 - m\alpha_1$  or  $\alpha_m \approx m\alpha_1$ . This approximation generally works quite well in practice, especially for small false alarm probabilities, unless the underlying model is too weak, having low redundancy and/or very correlated  $w_i$ -test statistics. For the correlated case, we have the upper bound,  $\alpha_m \leq m\alpha_1$ , which is known as the Bonferroni (1936) inequality.

Although the right-hand side of the inequality (15) is usually easy to compute (with a one-dimensional normally distributed random variable,  $P_{MD_1}(i) = P[|w_i| < k | \mathcal{H}_i]$ ), the computation of the multivariate missed detection probability  $P_{MD_m}(i)$  itself generally needs a Monte Carlo integration. This is due to the geometric complexity of the acceptance region  $\mathcal{A}_{\alpha_m}$  (see Fig. 1 at right). In the next section, we will provide a qualitative analysis of the upper bound (15) together with a quantitative investigation of its closeness.

Similar to the  $P_{MD}$  inequality (15), we have the MDB inequality (with 11)

$$MDB_m(i) \leq MDB_1(i) = \sigma_{b_i} \sqrt{\lambda_{q=1}(\alpha_1, \beta, 1)} \quad (16)$$

The right-hand side is the easy-to-use MDB formula one often uses for computing the MDBs of data snooping. Hence, this result shows that these MDBs can be used as safe bounds for the actual values  $MDB_m(i)$ . In the next section, the two will be compared for various examples.

### The Minimal Identifiable Bias (MIB)

Next to detection, we have identification (cf. 14). The  $MDB_m$ 's of equation (16) are minimal detectable biases,

but not minimal identifiable biases (Teunissen 2017). To determine the MIBs, we start from the complement of  $P_{MD}$  being the probability of correct detection,  $P_{CD} = 1 - P_{MD}$ . As correct detection does not necessarily imply correct identification (this is only the case with  $m=1$ ), the probability of correct detection is the sum of the probability of correct identification (selecting the single correct alternative) and the probability of wrong identification (selecting one of the  $m-1$  other alternatives):  $P_{CD} = P_{CI} + P_{WI}$ . Thus, the probability of correct identification is smaller than that of correct detection. As a consequence, the MIBs will be larger than their MDB counterparts

$$P_{CI} \leq P_{CD} \implies \text{MIB} \geq \text{MDB} \quad (17)$$

This shows that one has to be careful in using MDBs as MIBs. It is one thing to be able to *detect* a bias, but an all together other thing is to also be able to *identify* it.

For the probability of correct identification, we have similar to equation (15), the inequality

$$\begin{aligned} P_{CI_m}(i) &= P[|w_i| = \max_j |w_j| > k | \mathcal{H}_i] \leq P[|w_i| \\ &> k | \mathcal{H}_i] = P_{CI_1}(i) = 1 - P_{MD_1}(i) \end{aligned} \quad (18)$$

On the right-hand side, test-statistic  $w_i$  shall exceed, in absolute sense, the critical value  $k$ , and on the left-hand side, in addition,  $w_i$  needs to be the largest of all  $m$   $w$ -test statistics, resulting in a smaller probability. The last equality follows from the fact that correct identification equals correct detection in case one would have only the one alternative hypothesis  $\mathcal{H}_i$ . From the probabilistic inequality (18) follows for the Minimal Identifiable Biases

$$\text{MIB}_m(i) \geq \text{MIB}_1(i) = \text{MDB}_1(i) \quad (19)$$

### Computing the MDBs and MIBs

The univariate Minimal Detectable Bias  $\text{MDB}_1(i)$  is relatively easy to compute using the expression (16). The computation of  $\text{MDB}_m(i)$  and  $\text{MIB}_m(i)$ , however, is more involved and needs to be based on simulation. First, we need to be able to determine the acceptance region  $\mathcal{A}_{\alpha_m}$  for a user-given value of the overall false alarm rate  $\alpha_m$ . Thus given  $\alpha_m$ , we need to be able to compute  $\alpha_1$  or  $k = \chi_{\alpha_1}(1, 0)$ , thereby giving  $\mathcal{A}_{\alpha_m} = \cap_{i=1}^m \mathcal{A}_{\alpha_1}^i$ . The procedure of computing  $\alpha_1$  from  $\alpha_m$  is a problem already addressed in Lehmann (2010, 2011) and can briefly be described as follows: simulate  $N$  samples from the  $\mathcal{H}_0$ -distribution of the  $m$ -vector of  $w$ -test statistics  $w = [w_1, w_2, \dots, w_m]^T \sim \mathcal{N}(0, Q_{ww})$ , with  $(Q_{ww})_{ii} = 1$  and  $(Q_{ww})_{ij} = \rho(w_i, w_j)$  being the correlation between  $w_i$  and  $w_j$ . The  $i$ th entry of the  $v$ th such sample is denoted as  $w_i^v$ . Then sort out in ascending order the maximum components  $\tilde{w}^v = \max_i w_i^v$ , with  $v \in \{1, 2, \dots, N\}$ , and compute (Lemeshko and Lemeshko 2005, Lehmann 2012):

$$k = 1/2(\tilde{w}_{\lfloor(1-\alpha_m)N\rfloor} + \tilde{w}_{\lfloor(1-\alpha_m)N+1\rfloor}) \quad (20)$$

where the subscript of  $\tilde{w}$  ( $\lfloor(1-\alpha_m)N\rfloor$  and  $\lfloor(1-\alpha_m)N+1\rfloor$ ) is the rounded ordinal number that indicates the position of the selected elements in the ascendingly sorted list of  $\tilde{w}^v$ . From the above computed  $k$ , the corresponding univariate false alarm rate is computed as  $\alpha_1 = \chi_k^{-1}(1, 0)$ .

With  $k$  (or  $\alpha_1$ ) known,  $\mathcal{A}_{\alpha_m} = \cap_{i=1}^m \mathcal{A}_{\alpha_1}^i$  is known, thus making it possible to compute the multivariate missed detection probability  $P_{MD}$  through simulation. Again  $N$  samples are drawn, but now from the  $\mathcal{H}_j$ -distribution of  $w = [w_1, w_2, \dots, w_m]^T \sim \mathcal{N}(b_w, Q_{ww})$ , in which the  $i$ th entry of  $b_w$  is given as

$$\mathbb{E}(w_i | \mathcal{H}_j) = \rho(w_i, w_j) \frac{b_j}{\sigma_{b_j}} \quad (21)$$

By counting the number of missed detections in the generated samples, one can determine the corresponding probability of missed detection. Thus if  $N$  is the total number of samples simulated, we first count the number of times  $n_{MD}(b_j)$  that  $\cap_{i=1}^m |w_i^v| < k$  for  $v = 1, \dots, N$ , and then approximate the probability of missed detection as

$$P_{MD_m}(b_j) \approx \frac{n_{MD}(b_j)}{N} \quad (22)$$

The probability of correct identification  $P_{CI_m}(b_j)$  is computed in a similar manner, although the counting is now based on the number of times that  $\mathcal{H}_j$  is correctly identified, i.e. how often  $|w_i^v| = \max_i |w_i^v| > k$  for  $v = 1, \dots, N$ . As the approximation (22) gets better for larger  $N$ , the question is how large a sample should be chosen. Using the results of Tanizaki (2004), we found that for our purposes  $N = 10^5$  is sufficient to compute probabilities of order  $10^{-3}$  or larger, with high accuracy.

The computed probabilities  $P_{MD_m}(b_j)$  and  $P_{CI_m}(b_j)$  depend on  $b_j$ , i.e. on the assumed size of the bias under  $\mathcal{H}_j$ . Thus in order to extract the MDB and MIB, the above simulations need to be repeated for a range of  $b_j$  values, so that finally the minimal detectable and identifiable biases can be computed for a certain chosen fixed reference probability  $\gamma$  (e.g.  $\gamma = 0.8$ ) as

$$\begin{aligned} \text{MDB}_m &= \arg \min_{b_j} P_{MD_m}(b_j) < 1 - \gamma \\ \text{MIB}_m &= \arg \min_{b_j} P_{CI_m}(b_j) > \gamma \end{aligned} \quad (23)$$

### A qualitative description

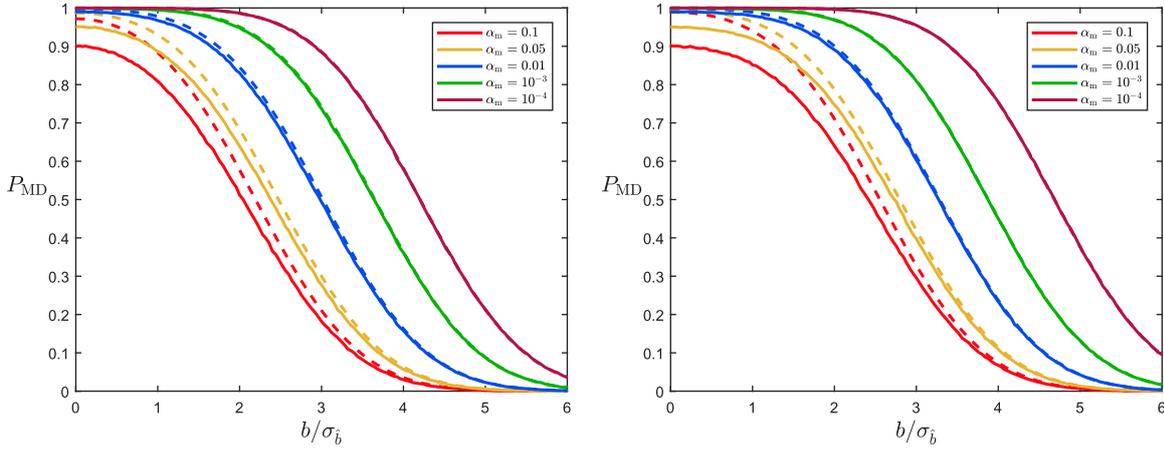
In this section, we provide a qualitative description of the behaviour of the univariate and multivariate probabilities of missed detection and correct identification.

#### Detection

We consider a simple model with  $A = [1, \dots, 1]^T$  and  $D(y) = I_m$ . The data have been simulated with a bias in the first observable, i.e. the true hypothesis is  $\mathcal{H}_{i=1}$ . The bias-size  $b$  was varied and the simulations were done for  $m=4$  and  $m=10$ . Figure 2 (at left for  $m=4$ , and at right for  $m=10$ ) shows the graphs of  $P_{MD_m}$  (full line) and  $P_{MD_1}$  (dashed line) as a function of the normalised bias-size  $b/\sigma_b$  for five different values of  $\alpha_m$ . The general characteristic of the probability of missed detection as a function of bias-size  $b$ , precision  $\sigma_b$ , false alarm rate  $\alpha_m$  or critical value  $k$  is as follows:

$$P_{MD} \downarrow \text{ if } b/\sigma_b \uparrow \text{ and/or } \alpha_m \uparrow \text{ (} k \downarrow \text{)} \quad (24)$$

Thus, the  $P_{MD}$  gets smaller for larger biases, more precisely estimated biases, larger false alarm rates and



2  $P_{MD_m}$  (full line) as a function of the normalised bias-size  $b/\sigma_b$ , for different values of  $\alpha_m$ , in comparison with the upper bound  $P_{MD_1}$  (dashed lines). For the model  $A = [1, \dots, 1]^T \in \mathbb{R}^m$  and  $D(y) = I_m$  with on the left the case  $m=4$  ( $\sigma_b = 1.155$  m) and on the right the case  $m=10$  ( $\sigma_b = 1.054$  m)

smaller critical values. This is explained by the fact that in each of these cases the probability mass of the  $w$ -test statistic over the acceptance region gets smaller. Either the PDF of the  $w$ -statistic moves away from the acceptance region when  $b/\sigma_b$  gets larger, or the acceptance region itself gets smaller when either  $\alpha$  gets larger, or  $k$  gets smaller.

Also note that  $P_{MD_1}$  (dashed line) gets larger when  $m$  gets larger (compare left and right graphs of Fig. 2). This is due to the decrease in  $\alpha_1$ . When the overall false alarm rate  $\alpha_m$  is kept fixed, a smaller  $\alpha_1$  (and thus larger critical value  $k$ ) needs to be chosen when the number of hypotheses  $m$  increases. Increasing  $k$  increases the acceptance region and thus  $P_{MD}$ .

From Fig. 2 one can also observe the difference  $P_{MD_1} - P_{MD_m}$ . It follows that

$$P_{MD_1} - P_{MD_m} \downarrow \text{ if } b/\sigma_b \uparrow \text{ and/or } \alpha_m \downarrow \text{ (} k \uparrow \text{)} \quad (25)$$

This behaviour can be explained by means of the illustration given in Fig. 3 at left. It shows for  $m=2$ , the regions  $\mathcal{A}_{\alpha_m}$  (dark green),  $\mathcal{A}_{\alpha_1}^1$  (light green) and  $\mathcal{A}_{\alpha_1}^2$  (light green) in case the two  $w$ -test statistics are independent. When  $b/\sigma_b$  gets larger, the PDF of  $w_1$  under  $\mathcal{H}_1$  (illustrated by a circle) translates to the right and the probability mass

over the region  $\mathcal{A}_{\alpha_1}^1/\mathcal{A}_{\alpha_m}$  (indicated by the yellow area) gets smaller, therefore making the difference  $P_{MD_1} - P_{MD_m}$  smaller too. A similar explanation can be given for when  $k$  gets larger, or when  $\alpha_m$  gets smaller.

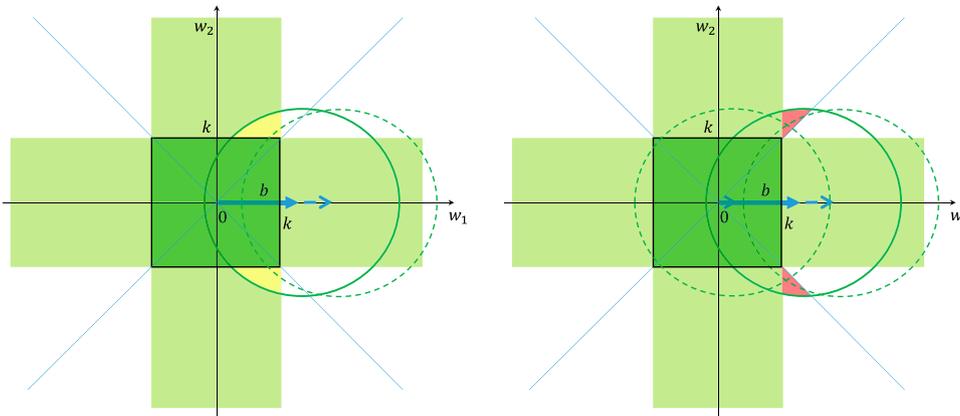
### Identification

We now consider identification. Figure 4 shows for the same model the graphs of  $P_{CI_m}$  (full line) and  $P_{CI_1}$  (dashed line) as a function of the normalised bias-size  $b/\sigma_b$  for five different values of  $\alpha_m$  (at left for  $m=4$  and at right for  $m=10$ ). The general characteristic of the probability of correct identification as a function of bias-size  $b$ , precision  $\sigma_b$ , false alarm rate  $\alpha_m$  or critical value  $k$  is as follows:

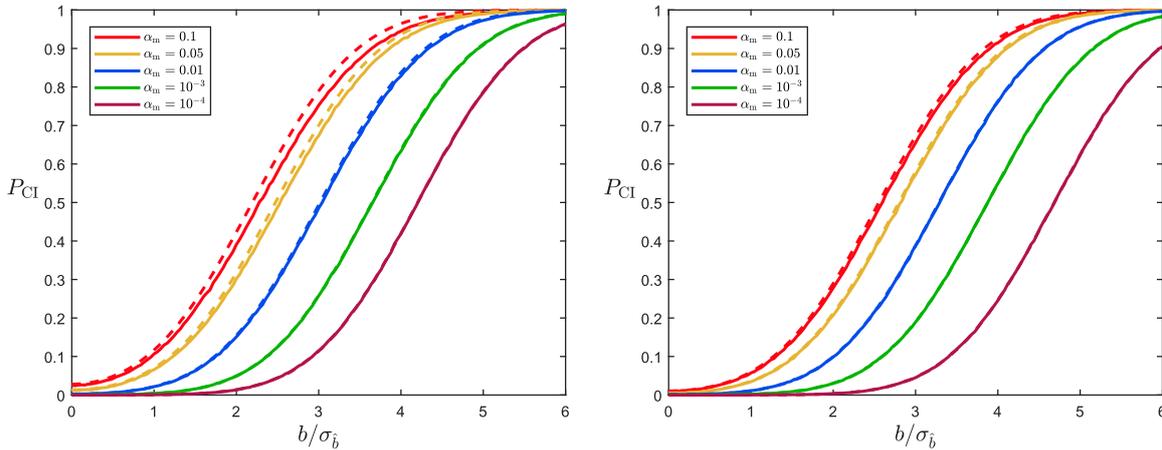
$$P_{CI} \uparrow \text{ if } b/\sigma_b \uparrow \text{ and/or } \alpha_m \uparrow \text{ (} k \downarrow \text{)} \quad (26)$$

Larger biases and/or more precisely estimated biases have a larger chance of being correctly identified. Also increasing the false alarm rate, thus reducing the acceptance region, obviously leads to larger probabilities of correct identification.

Also note that  $P_{CI_1}$  (dashed line) gets smaller when  $m$  gets larger (compare left and right graphs of Fig. 4). This is due to the decrease in  $\alpha_1$ . When the overall false alarm rate  $\alpha_m$  is kept fixed, a smaller  $\alpha_1$  (and thus larger



3 Left:  $P_{MD_1} - P_{MD_2}$  gets smaller if  $b/\sigma_b$  gets larger. Right:  $P_{CI_1} - P_{CI_m}$  gets smaller if a small  $b/\sigma_b$  gets even smaller and if a large  $b/\sigma_b$  gets even larger. The joint PDF of  $w_1$  and  $w_2$  is represented by a circle. The two diagonal lines in each figure denote the boundaries of identification between the two hypotheses, cf. (14)



4  $P_{CI_m}$  (full line) as a function of the normalised bias-size  $b/\sigma_b$ , for different values of  $\alpha_m$ , in comparison with the upper bound  $P_{CI_1}$  (dashed lines). For the model  $A = [1, \dots, 1]^T \in \mathbb{R}^m$  and  $D(y) = I_m$ , with on the left the case  $m=4$  ( $\sigma_b = 1.155$  m) and on the right the case  $m=10$  ( $\sigma_b = 1.054$  m)

critical value  $k$ ) needs to be chosen when the number of hypotheses  $m$  increases. Increasing  $k$  increases the acceptance region and thus decreases  $P_{CI}$ .

We also see that  $P_{CI} \neq 0$  for  $b=0$ . In that case,  $P_{CI}$  reduces to the false alarm rate for single alternative hypothesis testing, while for multiple alternative hypothesis testing it then reduces to a fraction of the false alarm rate, namely the probability by which the particular hypothesis will lead to wrongful detection. For  $m$  hypotheses, an approximation of this fraction is  $\alpha_1 \approx (1/m)\alpha_m$ .

For the difference  $P_{CI_1} - P_{CI_m}$ , we have

$$P_{CI_1} - P_{CI_m} \downarrow \text{ if } b/\sigma_b \downarrow 0 \text{ or } b/\sigma_b \uparrow \infty \text{ and/or } \alpha_m \downarrow (k \uparrow) \quad (27)$$

This behaviour can be explained by means of the illustration given in Fig. 3 at right. It shows that the difference  $P_{CI_1} - P_{CI_2}$  gets smaller if an already small  $b/\sigma_b$  gets smaller (purple area) or when an already large  $b/\sigma_b$  gets larger (red area). But note that the purple area and red area both get larger when  $k$  gets smaller, thus explaining why the difference (27) gets smaller when  $k$  gets larger or  $\alpha_m$  gets smaller.

### MDBs and MIBs compared

It follows from equations (16) and (19) that

$$MDB_m(i) \leq MDB_1(i) = MIB_1(i) \leq MIB_m(i) \quad (28)$$

This clearly shows the impact of testing multiple alternative hypotheses. There is no difference between detection and identification in case of a single alternative hypothesis. With an increase in the number of hypotheses, however, the MDBs get smaller, while the MIBs get larger. It may seem strange that the MDBs get smaller by merely increasing the number of hypotheses. This, however, is a dimensionality effect and due to the increase in intersections of  $\mathcal{A}_{\alpha_m} = \bigcap_{i=1}^m \mathcal{A}_{\alpha_i}^i$  when  $m$  gets larger. With this increase in  $m$ , also the overall false alarm rate  $\alpha_m$  gets larger.

The relation (28) shows that the easy-to-compute  $MDB_1(i)$  can be used as a safe upper bound for its multivariate counterpart. This is not the case, however, for the MIBs. Figure 5 shows for the same earlier example, a comparison between  $MDB_m(1)$ ,  $MDB_1(1)$  and  $MIB_m(1)$

as a function of  $\alpha_m$  for  $P_{MD} = 0.2$ . First, note that all get smaller if either  $\alpha_m$  or  $m$  gets larger. This is explained as follows. If the false alarm rate gets larger, then  $k$  gets smaller and thus  $\mathcal{A}_{\alpha_m}$  and  $\mathcal{A}_{\alpha_1}$  get smaller. This implies, when  $P_{MD}$  is kept unchanged, that the MDBs and MIBs get smaller. The impact of  $m$  is due to the improved precision with which the bias can be estimated.

Figure 5 also shows that the difference between  $MDB_m(i)$  and  $MIB_m(i)$  is small for small  $\alpha_m$ , but that it gets more pronounced when  $\alpha_m$  gets larger. This is due to the difference between the probabilities of correct detection and correct identification,  $P_{CD} - P_{CI}$ . This difference gets more pronounced if  $\alpha_m$  gets larger. If  $\alpha_m$  gets larger, then the acceptance region gets smaller as a result of which  $P_{CD}$  increases more than an individual  $P_{CI}$ .

## Terrestrial and satellite positioning

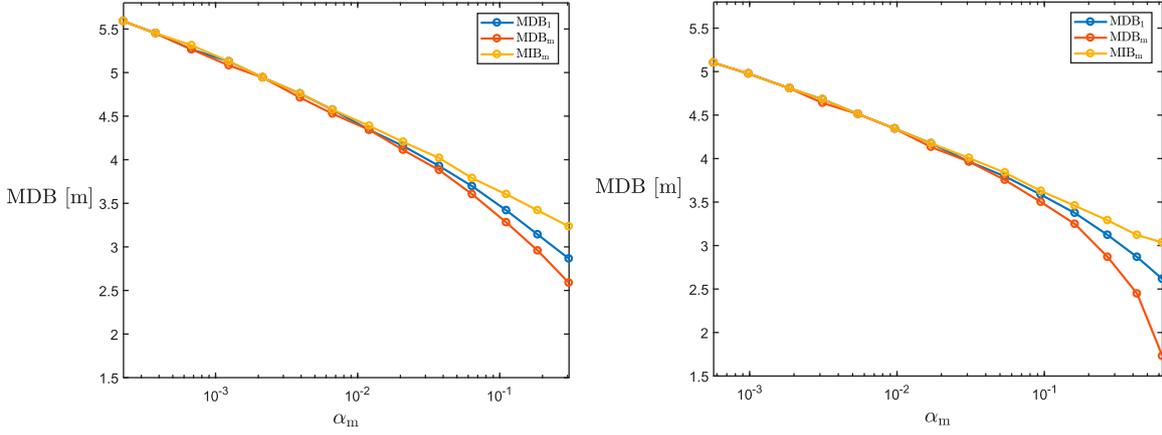
In this section, we illustrate by means of two practical surveying examples the significant differences one can experience between detectability and identifiability.

### A planar surveying triangle

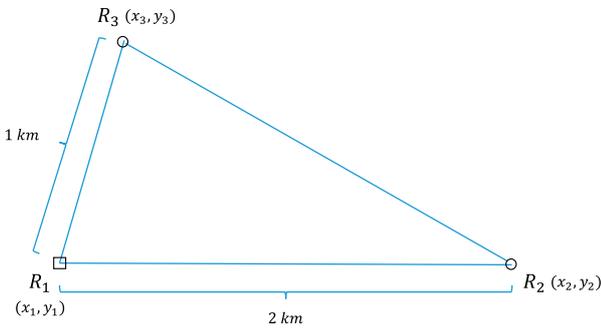
Consider the two-dimensional surveying triangle of Fig. 6. We assume distance and azimuth measurements. The linearised observation equations for distance  $l_{ij}$  and azimuth  $a_{ij}$  between points  $i$  and  $j$  read then

$$\begin{aligned} \Delta l_{ij} &= \frac{x_{ij}^0}{l_{ij}} \Delta x_{ij} + \frac{y_{ij}^0}{l_{ij}} \Delta y_{ij} \\ \Delta a_{ij} &= \frac{y_{ij}^0}{l_{ij}^2} \Delta x_{ij} - \frac{x_{ij}^0}{l_{ij}^2} \Delta y_{ij} \end{aligned} \quad (29)$$

in which  $\Delta x_{ij}$  and  $\Delta y_{ij}$  are the increments of the unknown coordinate differences  $x_{ij} = x_j - x_i$  and  $y_{ij} = y_j - y_i$  respectively. The coordinates of point 1 are assumed to be known. In total, nine observations are available (three distances and six azimuth measurements). Hence, the redundancy equals 5 as there are 4 unknowns (the coordinates of two stations). The observations are assumed normally distributed, uncorrelated and with



5  $MDB_m(1)$ ,  $MDB_1(1)$  and  $MIB_m(1)$  as functions of  $\alpha_m$  for  $P_{CD}$  and  $P_{CI}$  set at  $\gamma = 0.8$ . For the model  $A = [1, \dots, 1]^T \in \mathbb{R}^m$  and  $D(y) = I_m$ , with on the left the case  $m=4$  ( $\sigma_b = 1.155$  m) and on the right the case  $m=10$  ( $\sigma_b = 1.054$  m)



6 A two-dimensional planar surveying triangle

standard deviations given as

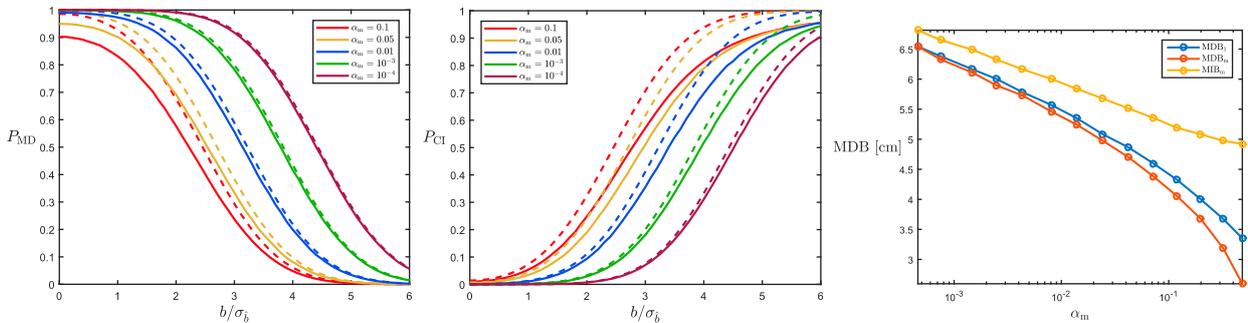
$$\begin{aligned} \sigma_{l_{ij}} &= 1 \text{ cm} \\ \frac{\sigma_{d_{ij}}}{l_{ij}} &= 1 \text{ cm} \end{aligned} \quad (30)$$

Figure 7 shows the univariate and multivariate probabilities of missed detection ( $P_{MD_m}$ ,  $P_{MD_1}$ ) and correct identification ( $P_{CI_m}$ ,  $P_{CI_1}$ ), together with their minimal biases ( $MDB_m$ ,  $MDB_1$ ,  $MIB_m$ ), as functions of the normalised bias-size  $b/\sigma_b$  for five different values of  $\alpha_m$ . We observe a very similar behaviour as discussed for the averaging model of the previous section. The difference

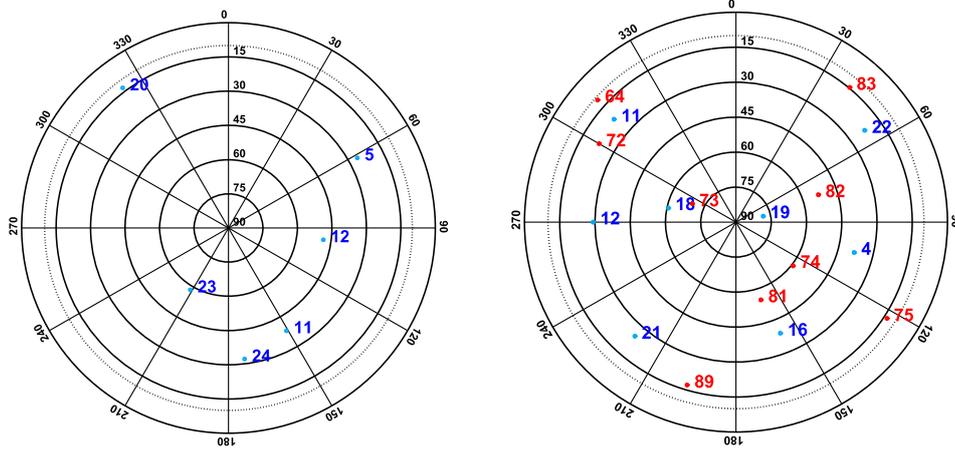
between the actual  $P_{MD_m}$  and its upper bound  $P_{MD_1}$  increases with decreasing bias size and increasing false alarm rate  $\alpha_m$ . Also for identification, the behaviour of the difference between  $P_{CI_m}$  and its upper bound  $P_{CI_1}$  is very similar to what we saw earlier. However, note that this difference is now much larger. This can also be seen in the difference between  $MDB_m$  and  $MIB_m$ . For instance for  $\alpha_m = 0.1$ ,  $MDB_m = 4.2$ m and  $MIB_m = 5.3$ m, an over 25% difference. The reason for these larger differences lies in the strength of the underlying model and the available redundancy. In the averaging model of the previous section, redundancy took a larger fraction of the number of observations, namely  $\frac{3}{4} = 0.75$  and  $\frac{9}{10} = 0.9$  respectively than in the current surveying example, for which this fraction is  $\frac{5}{9} = 0.56$ . The lower this fraction is, the more difficult one can expect correct identification to be. In the following section, we show a more extreme example of this effect.

### GNSS single point positioning

We consider GNSS single point positioning for a GPS-only scenario and compare it to a GPS+Galileo scenario. The two different receiver-satellite geometries are depicted in Fig. 8. In the GPS-only case, the  $m \times 4$  design matrix  $A$  and pseudorange variance matrix  $Q_{yy}$  are



7 Surveying triangle. On the left,  $P_{MD_m}$  (full) and  $P_{MD_1}$  (dashed), and in the middle,  $P_{CI_m}$  (full) and  $P_{CI_1}$  (dashed), as functions of the normalised bias  $b/\sigma_b$  (assuming outlier in distance observation  $l_{23}$ ), for different values of  $\alpha_m$  ( $\sigma_b = 1.4$  cm). At the right, MDBs and MIB as functions of  $\alpha_m$  for  $\gamma = 0.8$



8 Two GNSS skyplots for Delft, the Netherlands. Left: GPS-only skyplot with  $m=6$  and redundancy  $r=2$ ; Right: GPS+Galileo skyplot with  $m=17$  and redundancy  $r=12$ . PRNs up to 32 are by GPS, others by Galileo

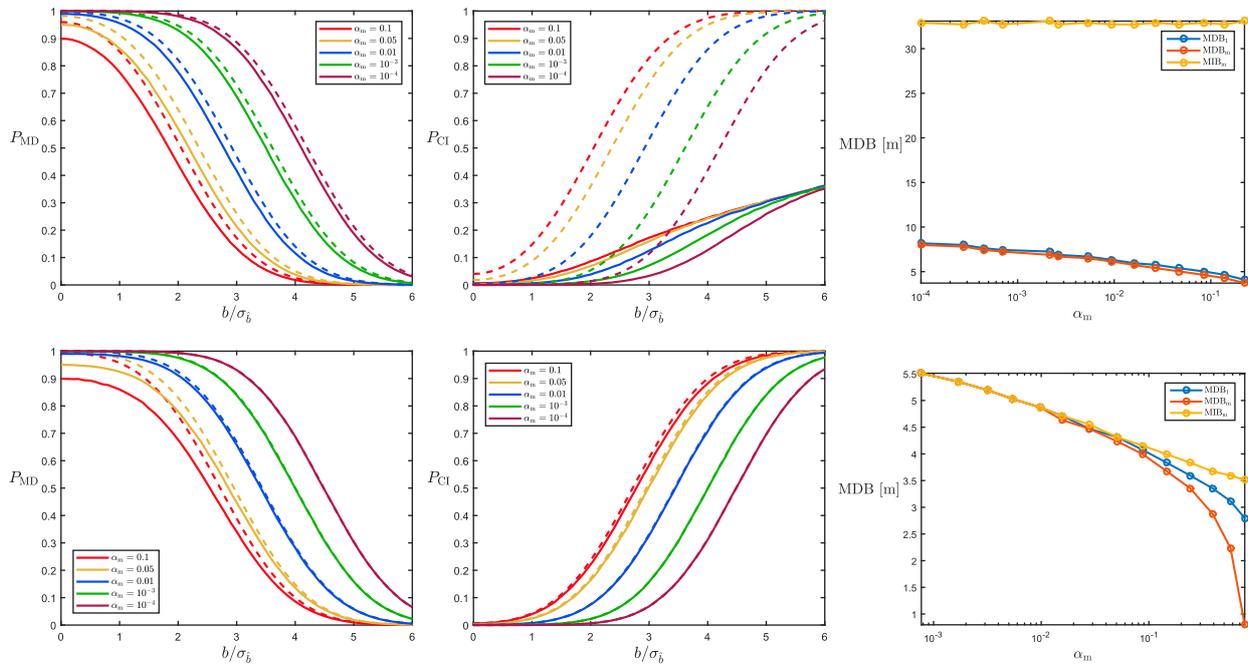
structured as

$$A = \begin{bmatrix} 1 & -u_1^T \\ \vdots & \vdots \\ 1 & -u_m^T \end{bmatrix}, \quad Q_{yy} = \text{diag}(\sigma_1^2, \dots, \sigma_m^2) \quad (31)$$

with  $u_i$  being the  $i$ th receiver-satellite unit direction vector. The unknown parameter vector consists of the receiver clock offset and the increments to the three receiver position coordinates. The stochastic model is based on ionosphere-free observations (from dual frequencies L1 and L5), with the entries of the diagonal variance matrix constructed according to Choi et al. (2011) and Rippl et al. (2014), see also chapter 7 of Imparato (2016). For the design matrix of the dual-

system GPS+Galileo, an additional column is added to the above design matrix  $A$  so as to take care of the inter-system bias or system-specific receiver clock offset.

Figure 9 shows the univariate and multivariate probabilities of missed detection and correct identification, together with their minimal biases. Although again we observe globally a similar behaviour as discussed before, we now clearly see what impact the model strength has on the detection and identification capabilities of the testing procedure. In the weak GPS-only case, the  $MIB_m$  exceeds its lower bound  $MDB_1$  by more than a factor of 4. The differences between the GPS-only case and the GPS+Galileo case are very significant. The same is true for the fraction their redundancy takes of the number of



9 GNSS single point positioning. On the left,  $P_{MDm}$  (full) and  $P_{MD1}$  (dashed), and in the middle,  $P_{CI_m}$  (full) and  $P_{CI1}$  (dashed), as functions of the normalised bias  $b/\sigma_b$ , for different values of  $\alpha_m$ . At the right, MDBs and MIB as functions of  $\alpha_m$  for  $\gamma = 0.8$ . Top row, weak GPS-only case of Fig. 8 at left, with outlier in PRN 5 ( $\sigma_b = 1.69$  m); bottom row, strong GPS+Galileo case of Fig. 8 at right, with outlier in PRN 12 ( $\sigma_b = 1.14$  m)

observations, namely  $\frac{2}{6} = 0.33$  versus  $\frac{12}{17} = 0.71$ . These results show how detection and identification fundamentally differ and how the capability of the latter can quickly be compromised when the strength of the model becomes too weak.

## Conclusions

In this contribution, we have discussed and analysed the performance of detection and identification of a DIA-testing procedure. We thereby discriminated between single alternative hypothesis testing and multiple alternative hypotheses testing. For detection, we have Baarda's well-known concept of the MDB. It describes a specific alternative hypothesis  $\mathcal{H}_i$ , the smallest size of model error or bias that can be detected with a certain pre-set probability, when employing a detector test with user-defined false alarm rate. As the MDB depends on the chosen detection test, we pointed out that the MDB for single alternative hypothesis testing may differ from that of multiple alternative hypotheses testing. We analysed the difference and indicated the circumstances under which the easy-to-compute  $\text{MDB}_1(i)$  can be considered to be a good upper bound for  $\text{MDB}_m(i)$ .

Next to detection, we have identification. The MDBs are Minimal Detectable Biases, but not Minimal Identifiable Biases. As the probability of correct identification is smaller than that of correct detection, the MIBs are larger than their MDB counterparts

$$\text{MIB} \geq \text{MDB} \quad (32)$$

This shows that one has to be careful in using MDBs as MIBs. It is one thing to be able to detect a bias, but an all together other thing is to also be able to identify it. We have made this distinction clear by means of a detailed qualitative analysis and description of the probabilities of missed detection and correct identification as a function of the normalised bias size and overall false alarm rate. This was then further illustrated by means of two practical surveying examples with varying model strength.

As there are, in contrast to the single alternative hypothesis case, no easy-to-compute expressions available for the  $m$ -hypotheses MIBs and MDBs, we have also shown how they can be computed by means of Monte Carlo simulation. Such computation first determines the acceptance region from the user-defined overall false alarm rate  $\alpha_m$ . Then, for any given bias range of the hypothesis, the probabilities of correct detection and correct identification are simulated. The so obtained functions are 'inverted' to obtain the required MDB or MIB respectively for the given user-defined reference probability.

Future studies on the subject may want to address the numerical aspects of the MDB and MIB computation, in order to improve the efficiency of the Monte Carlo simulations, and explore ways to obtain easy-to-compute tight bounds for the MDBs/MIBs.

## References

Alberda, J.E. 1976. Quality control in surveying. *Chartered surveyor*, 4 (2), 23–28.

- Amiri-Simkooei, A. 2004. A new method for second order design of geodetic networks: aiming at high reliability. *Survey review*, 37 (293), 552–560.
- Arnold, S.F. 1981. *The theory of linear models and multivariate analysis*. Vol. 2. New York: Wiley.
- Baarda, W. 1967. *Statistical concepts in geodesy*. Technical report Netherlands Geodetic Commission, Publ. on Geodesy, New Series.
- Baarda, W. 1968. *A testing procedure for use in geodetic networks*. Technical report Netherlands Geodetic Commission, Publ. on Geodesy, New Series, Vol. 2(5), Delft.
- Bonferroni, C.E. 1936. Teoria statistica delle classi e calcolo delle probabilità. *Pubblicazioni del R. Istituto Superiore de Scienze Economiche e Commerciali de Firenze*, 8, 3–62.
- Choi, M., Blanch, J., Akos, D., Heng, L., Gao, G., Walter, T., and Enge, P. 2011. Demonstrations of multi-constellation advanced RAIM for vertical guidance using GPS and GLONASS signals. *In: Proceedings ION GNSS 2011*, Portland, OR, 3227–3234.
- Hofmann-Wellenhof, B., Lichtenegger, H., and Wasle, E. 2008. *GNSS: global navigation satellite systems: GPS, Glonass, Galileo, and more*. New York: Springer.
- Imparato, D. 2016. *GNSS based receiver autonomous integrity monitoring for aircraft navigation*. Thesis (PhD). TU Delft.
- Klein, I., Matsuoka, M.T., Guzzato, M.P., and Nievinski, F.G. 2017. An approach to identify multiple outliers based on sequential likelihood ratio tests. *Survey review*, 49 (357), 449–457.
- Klein, I., Matsuoka, M.T., Guzzato, M.P., de Souza, S.F., and Veronez, M.R. 2015. On evaluation of different methods for quality control of correlated observations. *Survey review*, 47 (340), 28–35.
- Kok, J.J. 1982. *Statistical analysis of deformation problems using Baarda's testing procedures*. 'Forty Years of Thought'. Anniversary Volume on the Occasion of Prof. Baarda's 65th Birthday, Delft, Vol. 2, 470–488.
- Lehmann, R. 2010. Normalized residuals — how large is too large? *Allgemeine Vermessungsnachrichten*, 2, 53–61.
- Lehmann, R. 2011. Monte Carlo based data snooping with application to a geodetic network. *Journal of applied geodesy*, 5 (3–4), 123–134.
- Lehmann, R. 2012. Improved critical values for extreme normalized and studentized residuals in Gauss-Markov models. *Journal of geodesy*, 86, 1137–1146.
- Lehmann, R., and Voß-Böhme, A. 2017. On the statistical power of Baarda's outlier test and some alternative. *Journal of geodetic science*, 7 (1), 68–78.
- Leick, A. 2004. *GPS satellite surveying*. 3rd ed. John Wiley & Sons.
- Lemeshko, B.Y., and Lemeshko, S.B. 2005. Extending the application of Grubbs-type tests in rejecting anomalous measurements. *Measurement techniques*, 48 (6), 536–547.
- Prószyński, W. 2015. Revisiting Baarda's concept of minimal detectable bias with regard to outlier identifiability. *Journal of geodesy*, 89 (10), 993–1003.
- Rippl, M., Martini, I., Belabbas, B., and Meurer, M. 2014. ARAIM operational performance tested in flight. *In: Proceedings of the 2014 international technical meeting of the institute of navigation (ITM 2014)*, 27–29 January San Diego, CA, 601–615.
- Salzmann, M. 1991. MDB: a design tool for integrated navigation systems. *Bulletin geodesique*, 65 (2), 109–115.
- Staff of DGCC 1982. *The delft approach for the design and computation of geodetic networks*. Forty years of thought... Anniversary edition on the occasion of the 65th birthday of Professor W. Baarda, Vol. 1, 202–274.
- Tanizaki, H. 2004. *Computational methods in statistics and econometrics*. New York: Marcel Dekker.
- Teunissen, P.J.G. 1986. Adjusting and testing with the models of the affine and similarity transformation. *Manuscripta geodaetica*, 11, 214–225.
- Teunissen, P.J.G. 1989. Quality control in integrated navigation systems. *IEEE aerospace and electronic systems magazine*, 5 (7), 35–41.
- Teunissen, P.J.G. 1998. Minimal detectable biases of GPS data. *Journal of geodesy*, 72 (4), 236–244.
- Teunissen, P.J.G. 2000. *Testing theory: an introduction*. Series on mathematical geodesy and positioning, Delft University Press.
- Teunissen, P.J.G. 2017. Distributional theory for the DIA method. *Journal of geodesy*, 91, 1–22. DOI 10.1007/s00190-017-1045-7
- Teunissen, P.J.G., and Montenbruck, O., eds., 2017. *Handbook of global navigation satellite systems*. Springer.
- Teunissen, P.J.G., Imparato, D., and Tiberius, C.C.J.M. 2017. Does RAIM with correct exclusion produce unbiased positions? *Sensors*, 17 (7), 1508.

- Tiberius, C.C.J.M. 1998. *Recursive data processing for kinematic GPS surveying*. Netherlands Geodetic Commission, Publ. on Geodesy, New Series, 45.
- Van der Marel, H., and Kusters, A.J.M. 1990. *Statistical testing and quality analysis in 3-D networks: (part II) application to GPS. Global positioning system: an overview*. International Association of Geodesy Symposia, 102, 290–297.
- Van Mierlo, J. 1980. A testing procedure for analytic deformation measurements. In: *Proceedings of internationales symposium ueber Deformationsmessungen mit Geodaetischen Methoden*. Stuttgart: Verlag Konrad Wittwer.
- Yetkin, M., Berber, M., and Inal, C. 2014. Importance of probability levels for robustness analysis of geodetic networks. *Survey review*, 46 (335), 132–141.