Journal of Global Positioning Systems (*2021*)
Vol. 17, No. 1: 48-64
DOI: 10.5081/jgps.17.1.48

Journal of Global Positioning Systems
2021
© CPGPS

Performance assessment of the generalized DIA estimators applied in GNSS positioning

Ling Yang

College of Surveying and Geo-informatics, Tongji University, Shanghai 200092, China

Corresponding authors: Dr. Ling Yang, lingyang@tongji.edu.cn

Abstract: This paper discusses the performance assessment of the generalized DIA estimators. Based on the unifying framework of the DIA estimators introduced by Teuniseen (2018), quality control indices from three aspects are discussed and formulated, to measure the confidence levels of the testing decisions, to evaluate the reliability of the specified alternative hypothesis models, as well as to compute the biasedness, dispersion and integrity of the estimated parameters under an unconditional case, by taking the uncertainty of the combined estimation-testing procedure into account and performing the propagation of uncertainty accordingly. With a dual-constellation GNSS single-point positioning example, the quality control evaluation of and comparisons among three conventional used DIA estimators are demonstrated for practical applications.

Keywords: Detection, Identification and Adaption (DIA); Hypothesis Model; Stochastical Testing; Reliability; Confidence Level; Integrity Risk (IR); Protection Level (PL)

1 Introduction

Least-squares (LS) estimations, typically applied for geodetic data processing, can only achieve unbiasedness and minimum variance when realistic functional and stochastic models are used [Koch 1999; Teunissen 2006; Li et al. 2011]. Outliers generally result in LS estimators that are biased. The DIA method for the detection, identification and adaptation of functional model misspecifications, together with its associated internal and external reliability measures, finds its origin in the pioneering work of Baarda [1967, 1968], see also, e.g., Alberda [1976], Kok [1984], Teunissen [1985]. A unifying framework that captures the combined estimation and testing scheme of the DIA estimator has recently been introduced by Teunissen [2018].

The DIA estimator firstly makes hypothesis testing between the original/null and a group of alternative hypothesis models. By testing statistics that follow normal distribution, τ distribution, χ^2 distribution, F distribution or others [Pope 1976; Koch 2015; Kok 1984; Xu 1987; Lehmann 2013], the most trustworthy model is selected and is considered to exclude any unmodelled misspecifications. Then, estimation is further conducted under the identified model aiming to remove the potential bias on the unknown parameters. Theoretically, the bias on the unknown parameters can only be completely removed once the hypothesis testing successfully selects the model that can realistically compensate the unmodelled misspecifications. However, missed detection, false alarm and wrong identification, usually cannot be avoided due to the geometry of the observation model [Hekimoglu and Berber 2003; Wieser 2004], the separability among hypothesis models [Yang et al. 2013, 2017], the selected test statistics [Leick and Emmons 1994; Koch 2015] and the predetermined critical values for testing [Lehmann 2012]. These incorrect decisions would still introduce biases on the final parameter estimation. It implies

Editor in charge: Dr. Zhizhao Liu

that the DIA method has to take the contributions of the various decisions and estimators into consideration. The consequence of this practice is that the DIA estimator is neither estimation only, nor testing only, but one where estimation and testing are combined [Teunissen 2018].

Since incorrect testing decisions and biases on estimated parameters cannot be completely avoided via a DIA estimator, the reliability analysis has become a fundamental part of the procedure, which gives a measure for the ability of an estimator to successfully detect the misspecification and to resist the bias on parameter estimation if the misspecification is failed to detect. The internal reliability is usually measured by the Minimal Detectable Bias (MDB). MDB describes the minimal size of model misspecification or bias specified by an alternative hypothesis that can be detected with a certain preset probability, when employing a DIA test with user-defined false alarm rate. Theoretically, the MDBs measure the capability of correct detection rather than correct identification. For a multiple alternative hypotheses case, concepts of Minimal Separable Bias (MSB) or Minimal Identifiable Bias (MIB) have been further investigated and defined to measure the minimal bias sizes that can be successfully separated with another alternative hypothesis or can be correctly identified among multiple alternative hypotheses, with a certain preset probability [Förstner 1983; Yang et al. 2013, 2017]. External reliability is accordingly obtained by substitution of the MDB into the parameter solution [Wang and Chen 1999; Ryan and Lachapelle 2001], aiming to measure the significance of the influence on the parameter estimation caused by the undetected misspecification. Besides, other related measures of reliability, including reliability numbers [Pelzer 1980; Wang and Chen 1994; Chen and Wang 1996; Schaffrin 1997; Ou 1999] and controllability [Pelzer 1980; Förstner 1985], have also been generalized for single and multiple outliers. Generally, these indices mainly work to qualitatively describe the ability of successfully detect a specific misspecification against to the null hypothesis in a binary case, when next to the null hypothesis only a single alternative hypothesis is considered.

In the safety-critical navigation application of aviation, Receiver Autonomous Integrity Monitoring (RAIM) was specifically developed to safeguard the navigation integrity by means of self-contained fault detection at the GNSS navigation receiver [Lee 1986; Parkinson and Axelrad 1988]. Theoretically, RAIM commonly uses a specified DIA estimator relying on statistical hypothesis testing on the positioning domain. In RAIM without fault exclusion, only the null hypothesis is conducted, and therefore the Gaussian distributional properties of the estimator can directly be used to compute the probability of hazardous misleading information (PHMI) and to determine the protect levels (PLs). However, when exclusion is included in real-time navigation application, the DIA estimator for positioning is actually dealing with a combination of multiple estimators, one for each hypothesis model specified. Therefore, the integrity evaluation has to take the uncertainty of the combined estimation-testing procedure into account and perform the propagation of uncertainty accordingly [Teunissen et al. 2017].

Various reliability and integrity indices have been defined and investigated under a conditional case when one single hypothesis is considered. However, aiming to evaluate the overall performance of a DIA estimator with a group of specified hypothesis models for quality control purpose, the uncertainty of the estimated parameters produced from the combined estimation-testing procedure should be stochastically considered. In this paper, the overall performances of conventional used DIA estimators are evaluated via three groups of indices, which are, (1) the confidence levels of the testing decisions; (2) the reliability of the hypothesis models; (3) the biasedness, dispersion, and integrity of the estimated parameters. The rest of this paper is organized as follows. Section 2 briefly reviews the generalized Detection, Identification and Adaption (DIA) methods and principles proposed by Teunissen [2018]. Section 3 presents the three groups of quality control indices of a DIA estimator. Section 4 gives numerical examples of a dual-constellation GNSS Single Point Positioning (SPP) to demonstrate the quality control of and comparison among three

conventional DIA methods. Section 5 presents the conclusions of this study.

2 DIA theory

In this section, the Detection, Identification and Adaption (DIA) methods and principles introduced in Teunissen [2018] are reviewed.

The DIA methods rely on making decision among a null and a set of alternative hypotheses $\mathcal{H}_0$ an $\mathcal{H}_i$, respectively. The null hypothesis that one believes to be valid under normal working conditions is usually formed as

$$\mathcal{H}_0: E(\boldsymbol{y}) = \boldsymbol{A}\boldsymbol{x} \text{ and } D(\boldsymbol{y}) = \boldsymbol{\Sigma}_{yy} \quad (1)$$

Where $E(\cdot)$ and $D(\cdot)$ are the expectation and dispersion operator, respectively, $\boldsymbol{y} \in \mathbb{R}^m$ is the vector of observables with normally distributed random errors, $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ is the given design matrix of rank n, $\boldsymbol{x} \in \mathbb{R}^n$ is the to-be-estimated unknown parameter vector, $\boldsymbol{\Sigma}_{yy}$ is the given positive-definite variance matrix of $\boldsymbol{y}$.

Assuming a misspecification confined to the functional model, the alternative hypothesis can be formed as

$$\mathcal{H}_i: E(\boldsymbol{y}) = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{C}_i\boldsymbol{b}_i \text{ and } D(\boldsymbol{y}) = \boldsymbol{\Sigma}_{yy} \quad (2)$$

where $\boldsymbol{b}_i \in \mathbb{R}^q$ is a vector of nonrandom parameters denoting a specific fault type that impacts the mean of $\boldsymbol{y}$, via a matrix $\boldsymbol{C}_i \in \mathbb{R}^{m\times q}$. Through $\boldsymbol{C}_i\boldsymbol{b}_i$, one may model, for instance, the presence of one or more fault (outliers) in the data, cycle slips in GNSS phase data, satellite failures, antenna-height errors, erroneous neglect of atmospheric delays, or any other systematic effect that one failed to take into account.

Since one cannot specify the design matrix $\boldsymbol{C}_i$ beforehand, multiple alternative models in parallel should be set up with different $\boldsymbol{C}_i$ $(i = 1, \cdots, N)$. Accordingly, the DIA testing procedure needs to be devised for handling multiple alternative hypotheses, and therefore usually consists of the following three steps of detection, identification and adaption [Baarda 1968; Lehmann 2014; Teunissen 2018].

In the detection step, a global test on $\mathcal{H}_0$ is performed to diagnose whether an unspecified model fault has occurred. Conventionally, the global test statistic is formed as

$$T_0 = \|\boldsymbol{v}_0\|^2_{\boldsymbol{\Sigma}_{yy}} \quad (3)$$

Where $\|\cdot\|^2_{\boldsymbol{M}} = (\cdot)^{\mathrm{T}}\boldsymbol{M}^{-1}(\cdot)$ is the notation for a weighted squared norm. Under null hypothesis $\mathcal{H}_0$, one can deduce that T_0 follows a centralized Chi-square distribution $\mathrm{T}_0 \sim \chi^2(r, 0)$, therefore, $\mathcal{H}_0$ can be accepted with

$$\mathcal{P}_0 = \{T_0 \le c_{\chi^2}(1-\alpha_0, r)\} \quad (4)$$

Where c_{χ^2} denote the corresponding critical value of Chi-square distribution relying on the probability of false alarm α_0 and the degree of freedom r. Once $\mathcal{H}_0$ has been accepted, least-squares (LS) estimation under model (1) is provided as the best linear unbiased estimate (BLUE) of $\boldsymbol{x}$.

If $\mathcal{H}_0$ has been rejected, an identification procedure should be further executed by searching among the specified alternative hypotheses, $\mathcal{H}_i, (i = 1, \cdots, N)$ with a set of parallel local tests, for the most likely model misspecification. The local test statistic for $\mathcal{H}_i$ is usually formed as

$$T_i = \left\|\widehat{\boldsymbol{b}}_i\right\|_{\boldsymbol{\Sigma}_{\hat{b}_i\hat{b}_i}} \quad (5)$$

Where $\widehat{\boldsymbol{b}}_i$ and $\boldsymbol{\Sigma}_{\hat{b}_i\hat{b}_i}$ are the LS estimation of misspecification parameter $\boldsymbol{b}_i$ under model (2). If the misspecification is significant, it can be deduced that T_i follows a noncentralized Chi-square distribution $T_i \sim \chi^2(q, \lambda_i^2)$, with $\lambda_i^2 = \|\boldsymbol{b}_i\|^2_{\boldsymbol{\Sigma}_{\hat{b}_i\hat{b}_i}}$ [Koch 1999; Teunissen 2018; Yang and Shen 2020]. Therefore, $\mathcal{H}_i$ can be accepted with

$$\mathcal{P}_i = \{T_i > c_{\chi^2}(1-\alpha_0, q)\} \quad (6)$$

However, for practical applications, identification usually involves making decision among multiple alternative hypotheses after conducting detection, therefore a complex region is usually defined for the acceptance of $\mathcal{H}_i$ as

$$\mathcal{P}_{i\neq 0} = \left\{T_i = \max_{j\in\{1,\cdots,N\}} T_j^2\right\}, \qquad i = 1, \cdots, N \quad (7)$$

After identification of the suspected misspecification, the accepted hypothesis $\mathcal{H}_i$ would become the new null hypothesis. Either a direct solution under model (2) or a correction on the $\mathcal{H}_0$-based solution under model (1) is adopted.

3 Discussion on quality control of the generalized DIA estimators

3.1 Three groups of quality control indices for DIA estimators

It is considered that the biases of the LS estimation of $\boldsymbol{x}$ caused by outliers can be controlled by above DIA procedure. However, the BLUE-property intrinsically lose due to the combination of estimation and testing. Therefore, quality control of the DIA estimators is essential. Proposed by Yang et al. [2021], quality control of DIA estimators can generally be conducted by indices from three aspects: 1) Confidence levels of the hypothesis testing decisions; 2) Reliability of the alternative hypothesis models; 3) Biasedness, dispersion, and integrity of the estimated parameters. In this subsection, we will discuss the main factors that impact these indices.

3.1.1 Confidence levels of the hypothesis testing decisions

The confidence levels of the hypothesis testing decisions are measured by the probabilities of different testing decisions, including Correct Acceptance (CA) and False Alarm (FA) under null hypothesis, and Missed Detection (MD), Correct Detection (CD), Correct Identification (CI), and Wrong Identification (WI) under a specific alternative hypothesis $\mathcal{H}_i$. For simplification, the probabilities of these decisions can be put into a probability matrix as

$$P_{ij} = P\left(T_i \in \mathcal{P}_j | \mathcal{H}_i\right), i, j = 0,1, \cdots, N \tag{8}$$

Where indicator i and j denote the unknown reality and the testing decision, respectively. It satisfies $\sum_{j=0}^{N} P_{ij} = 1, \forall i$. Computing such probabilities requires information about probability density function (PDF) of the test statistics T_i, and depends on the complexity of the acceptance regions for each of these events [Teunissen 2018]. Usually, numerical evaluation methods via Monte Carlo simulation are applied [Yang et al. 2013, 2017, 2021; Zaminpardaz and Teunissen 2019].

Among above testing decisions, the wrong identification is the severest incorrect decision, since biases on the estimated parameters would be further enlarged by weaker geometric strength. According to the definitions of Eq. (7), wrong identification will be triggered when

$$\mathcal{P}_{j\neq 0,i} = \left\{ T_{j\neq 0,i}^2 = \max_{s\in\{1,\cdots,N\}} T_s^2 \,|\mathcal{H}_i \right\} \tag{9}$$

Derived from Eq. (5), the test statistic for $\mathcal{H}_{j\neq 0,i}$ when $\mathcal{H}_i$ holds true follows

$$T_{j\neq 0,i}|\mathcal{H}_i = \left\|\widehat{\boldsymbol{b}}_j\right\|_{\boldsymbol{\Sigma}_{\hat{b}_j\hat{b}_j}} \sim \chi^2\left(q, \lambda_{ij}^2\right) \tag{10}$$

with the noncentrality parameter λ_{ij}^2 satisfying

$$\lambda_{ij}^2 = \left(\left\| \boldsymbol{\Sigma}_{\hat{b}_j\hat{b}_i} \boldsymbol{\Sigma}_{\hat{b}_i\hat{b}_i}^{-1} \boldsymbol{b}_i \right\|_{\boldsymbol{\Sigma}_{\hat{b}_j\hat{b}_j}} \right)^2 \tag{11}$$

and the covariance $\boldsymbol{\Sigma}_{\hat{b}_j\hat{b}_i}$ being

$$\boldsymbol{\Sigma}_{\hat{b}_j\hat{b}_i} = \boldsymbol{\Sigma}_{\hat{b}_j\hat{b}_j} \boldsymbol{C}_j^T \boldsymbol{\Sigma}_{yy}^{-1} \boldsymbol{P}_{\boldsymbol{A}_0}^{\perp} \boldsymbol{C}_i \boldsymbol{\Sigma}_{\hat{b}_i\hat{b}_i} \tag{12}$$

with

$$\boldsymbol{P}_{\boldsymbol{A}_0}^{\perp} = \boldsymbol{I}_m - \boldsymbol{P}_{\boldsymbol{A}_0},\ \boldsymbol{P}_{\boldsymbol{A}_0} = \boldsymbol{A}\boldsymbol{A}_0^{+},\ \boldsymbol{A}_0^{+} = \left(\boldsymbol{A}^T \boldsymbol{\Sigma}_{yy}^{-1} \boldsymbol{A}\right)^{-1} \boldsymbol{A}^T \boldsymbol{\Sigma}_{yy}^{-1} \tag{13}$$

Defining the general correlation coefficient ρ_{ij} satisfying

$$\rho_{ij}^2 = \lambda_{ij}^2 / \lambda_i^2 \tag{14}$$

It is easy to prove that

$$0 < \rho_{ij}^2 < 1 \tag{15}$$

Therefore, we can conclude that larger ρ_{ij}^2 indicates a larger $P_{\mathrm{WI}_{ij}}$, and accordingly a larger discrepancy between P_{CD_i} and P_{CI_i}.

3.1.2 Reliability of the alternative hypothesis models

For a specified alternative hypothesis model, larger $\boldsymbol{b}_i$ will be correctly detected and identified with higher probabilities. With an identical probability threshold, the size of $\boldsymbol{b}_i$ will vary among multiple alternative hypothesis models, as well as the corresponding influences on the parameter estimations are also diverse. Therefore, the reliability indices are defined to measure the significance discrepancies between each alternative hypothesis models and the null hypothesis models, or between any two alternative hypotheses, with the internal reliability describing the significance of $\boldsymbol{b}_i$, and external reliability describing the significance of the biases on the estimated parameters caused by $\boldsymbol{b}_i$.

There are several parameters are defined to describe the internal reliability of an alternative hypothesis models, including Minimal Detectable Bias (MDB), Minimal Identifiable Bias (MIB), and Minimal Separable Bias (MSB).

The MDB of an alternative hypothesis $\mathcal{H}_i$ is defined as the (in absolute value) smallest bias that leads to rejection of $\mathcal{H}_0$ for a given CD probability [Teunissen 2018]. For the $\mathcal{P}_0$ given in Eq. (4), the MDB of $\mathcal{H}_i$ can be computed from 'inverting'

$$P\left(T_0^2 > c_{\chi^2}(1-\alpha_0, r)|\mathcal{H}_i\right) = \gamma_{CD} \tag{16}$$

using the fact that under $\mathcal{H}_i$, $T_0^2 \sim \chi^2\left(r, \lambda_i^2\right)$ with $\lambda_i^2 = \|\boldsymbol{b}_i\|^2_{\boldsymbol{\Sigma}_{\hat{b}_i\hat{b}_i}}$. Then the MDB of $\mathcal{H}_i$ can be resolved as

$$MDB(\boldsymbol{b}_i) = \frac{\lambda_0(\alpha_0, \gamma_{CD}, r)}{\sqrt{\boldsymbol{d}^T \boldsymbol{\Sigma}^{-1}_{\hat{b}_i\hat{b}_i}\boldsymbol{d}}}\boldsymbol{d} \tag{17}$$

where $\boldsymbol{d} \in \mathbb{R}^q$ is a unit vector, and $\lambda_0^2(\alpha_0, \gamma_{\mathrm{CD}}, r)$ is the noncentrality parameter that captures the dependency on the α_0, r and γ_{CD} derived from Eq. (16).

Similarly, if only single alternative hypothesis $\mathcal{H}_1$ is concerned, with the acceptance region for $\mathcal{H}_0$ and $\mathcal{H}_1$ given in Eq. (6), the MDB of $\mathcal{H}_1$ can be computed by

$$MDB(\boldsymbol{b}_1) = \frac{\lambda_0(\alpha_0, \gamma_{CD}, q)}{\sqrt{\boldsymbol{d}^T \boldsymbol{\Sigma}^{-1}_{\hat{b}_i\hat{b}_i}\boldsymbol{d}}}\boldsymbol{d} \tag{18}$$

where $\lambda_0(\alpha_0, \gamma_{\mathrm{CD}}, q)$ is derived from Eq. (6) with $\mathrm{P}\left(T_i^2 > c_{\chi^2}(1-\alpha_0, q)|\mathcal{H}_i\right) = \gamma_{\mathrm{CD}}$.

It is noted that Eq. (18) is the one that the conventional MDB concept are defined as, other than Eq. (17) [Baarda 1968, Teunissen 2006, Lehmann 2014]. Apparently, although the identical α_0 and γ_{CD} are defined, the noncentrality parameter in Eq. (17) and (18) are discrepant due to the unequal freedom degrees of the two test statistics. Since it satisfies $q < r$, the MDB under a single alternative hypothesis (Eq. (18)) should be always smaller than its counterpart under multiple alternative hypotheses (Eq. (17)).

Generally, it is important to note that the MDB concept describes the sensitivity of rejecting the null hypothesis when the specified alternative hypothesis holds true. It means the existence of a model misspecification larger than its MDB will be successfully detected with a probability higher than γ_{CD}, but the identification cannot be always ensured with the same probability for multiple alternative hypotheses case.

The MIB of an alternative hypothesis $\mathcal{H}_i$ is defined as the smallest bias that leads to acceptance of $\mathcal{H}_i$ for a given CI probability. For multiple alternative hypotheses case, the correct identification will be triggered with the acceptance region defined in Eq. (7). Therefore, for a certain CI probability, say $P_{\mathrm{CI}_i} = \gamma_{\mathrm{CI}}$, the MIB($\boldsymbol{b}_i$), as a function of α_0, γ_{CI}, q and r, is calculated relying on the noncentrality parameter from resolving

$$\begin{aligned} P\Big(T_i^2 = \max_{j\in\{1,\cdots,N\}} T_j^2 \cap T_0^2 \\ > c_{\chi^2}(1-\alpha_0, r)|\mathcal{H}_i\Big) \\ = \gamma_{CI} \end{aligned} \tag{19}$$

As defined in Section 3.1.1, the discrepancy between detection and identification is caused by the unavoidability of the wrong identification for multiple alternative hypotheses. The MSB of an alternative hypothesis $\mathcal{H}_i$ corresponding to another $\mathcal{H}_{j\neq 0,i}$ is defined as the smallest bias that leads to incorrect acceptance of $\mathcal{H}_j$ for a given upper boundary of the WI_{ij} probability. For multiple alternative hypotheses case, the wrong identification of $\mathcal{H}_{j\neq 0,i}$ when $\mathcal{H}_i$ holds true will be triggered with an acceptance region defined by Eq. (9). If an upper threshold of WI probability is defined, say $\mathrm{P}_{\mathrm{WI}_{ij}} \leq \gamma_{\mathrm{WI}}$, the MSB($\boldsymbol{b}_{ij}$), as a function of α_0, γ_{WI}, q and r, is calculated relying the noncentrality parameter from resolving

$$\begin{aligned} P\Big(T_{j\neq 0,i}^2 = \max_{s\in\{1,\cdots,N\}} T_s^2 \cap T_0^2 \\ > c_{\chi^2}(1-\alpha_0, r)|\mathcal{H}_i\Big) \\ = \gamma_{WI} \end{aligned} \tag{20}$$

With above definition of MDB, MIB and MSB, a discrepant smallest bias for the alternative hypothesis model $\mathcal{H}_i$ is determined to satisfy specific requirements on the probabilities of CD_i, CI_i and WI_{ij}. Intrinsically, once $\boldsymbol{b}_i$ is fixed, the corresponding

probabilities of CD_i, CI_i and WI_{ij} are all determined. Therefore, once can select one of the three to control the confidence levels of the testing decisions. Or else, by comparing the sizes of the three, the discrepancies among different decisions can be analyzed. Usually, due to the unavoidability of the WI, the $\mathrm{MIB}(\boldsymbol{b}_i)$ is always larger than $\mathrm{MDB}(\boldsymbol{b}_i)$ for an identical $P_{\mathrm{CD}_i} = P_{\mathrm{CI}_i} = \gamma$. For the k alternative hypothesis models, there are k-1 $\mathrm{MSB}(\boldsymbol{b}_{ij})$ with $j = 1, \cdots, k$ and $j \neq i$, which indicate the separability between the alternative hypothesis $\mathcal{H}_i$ and another alternative hypothesis $\mathcal{H}_j$ [Yang et al. 2013, 2017]. Among these k-1 $\mathrm{MSB}(\boldsymbol{b}_{ij})$, the largest one corresponds to the hypothesis with the largest ρ_{ij} as given in Eq. (14). This means larger ρ_{ij} implying a weaker separability between two alternative hypotheses. When $\rho_{ij} = 1$ the two hypotheses are not separable, and identification between $\mathcal{H}_i$ and $\mathcal{H}_j$ becomes invalid. In this case, either the alternative hypotheses set should be reconstructed or decisions of accepting $\mathcal{H}_i$ or $\mathcal{H}_j$ should be considered with a low confidence level.

The conventional external reliability is defined as the maximum effect of any nondetectable outlier on the estimated parameters [Schaffrin 1997], which can be quantified by

$$\widehat{\boldsymbol{b}}_{x_0}(MDB) = \boldsymbol{A}_0^+ \boldsymbol{C}_i \cdot MDB(\boldsymbol{b}_i) \tag{21}$$

Accordingly, the bias-to-noise ratios (BNRs) introduced by Baarda in his reliability theory to measure the external reliability for single bias is defined as [Baarda 1967, 1976; Teunissen 2006]

$$BNR_i = \boldsymbol{A}_0^+ \boldsymbol{c}_i \tag{22}$$

Which describes the propagation ratio from bias on measurement to the bias on parameter.

Besides, the quadratic form of $\widehat{\boldsymbol{b}}_{x_0}$(MDB)

$$\begin{aligned} \lambda_{x_0} &= \boldsymbol{A}_0^+ \boldsymbol{C}_i \cdot MDB(\boldsymbol{b}_i) \\ \lambda_{x_0}^2 &= \left\| \widehat{\boldsymbol{b}}_{x_0}(MDB) \right\|^2_{\boldsymbol{\Sigma}_{\hat{x}_0 \hat{x}_0}} \end{aligned} \tag{23}$$

can also be used as a measure of the external reliability [Koch 1999, Teunissen 2006].

Another measure related to the external reliability for single bias is defined by

$$slope_i = \boldsymbol{A}_0^+ \boldsymbol{c}_i \Big/ \sqrt{\boldsymbol{\Sigma}^{-1}_{\hat{b}_i \hat{b}_i}} \tag{24}$$

Where $slope_i$ describes the propagation ratio from the test statistic to the unknown parameter. Generally, above external reliability indices are aiming to measure the effect on the estimated parameters under missed detection.

3.1.3 Biasedness, dispersion, and integrity of the estimated parameters

With the DIA method, the outcome of testing determines how the parameter vector $\boldsymbol{x}$ gets estimated. By making use of the indicator function $p_i(\boldsymbol{t})$ of the regions $\mathcal{P}_i$ (i.e. $p_i(\boldsymbol{t}) = 1$ for $\boldsymbol{t} \in \mathcal{P}_i$ and $p_i(\boldsymbol{t}) = 0$ elsewhere), the DIA estimator $\overline{\boldsymbol{x}}$ can be written in the compact form

$$\overline{\boldsymbol{x}} = \sum_{i=0}^{N} \widehat{\boldsymbol{x}}_i p_i(\boldsymbol{t}) \tag{25}$$

Due to the inevitability of committing false alarm, missed detection, and wrong identification, the DIA estimator $\overline{\boldsymbol{x}}$ will be a combination of multiple individual estimators $\widehat{\boldsymbol{x}}_i, i = 0, \cdots, k$. Although the individual estimators $\widehat{\boldsymbol{x}}_i$ are normally distributed with different expectations and dispersions, the combination, $\overline{\boldsymbol{x}}$, are not normally distributed anymore. Based on Monte Carlo simulation with N_t independent experiments, the biasedness and dispersion of $\overline{\boldsymbol{x}}$ are evaluated as

$$\begin{cases} E(\overline{\boldsymbol{x}}|\mathcal{H}_i) = \dfrac{1}{N_t} \displaystyle\sum_{k=1}^{N_t} (\overline{\boldsymbol{x}}|\mathcal{H}_i)_k \\ \boldsymbol{b}_{\overline{\boldsymbol{x}}|\mathcal{H}_i} = \boldsymbol{x} - E(\overline{\boldsymbol{x}}|\mathcal{H}_i) \\ D(\overline{\boldsymbol{x}}|\mathcal{H}_i) = \dfrac{1}{N_t} \displaystyle\sum_{k=1}^{N_t} [(\overline{\boldsymbol{x}}|\mathcal{H}_i)_k - E(\overline{\boldsymbol{x}}|\mathcal{H}_i)]^2 \end{cases} \tag{26}$$

Apart from biasedness and dispersion, integrity is also important for the quality control of the parameter estimation. The integrity risk defined as the Probability of Hazardously Misleading Information (PHMI) is the probability that the estimator of unknown parameters is outside the error bound (or so called the Protection Level (PL)) around its true value, whatever hypothesis is true in reality [Blanch et al. 2015, Zaminpardaz et al. 2019]. As integrity plays a crucial role in critical and safe-of-life applications, for

instance in aviation, when GNSS positioning is used to approach to an airport, stringent requirements and monitoring on integrity obviously apply.

Being a good estimator of $\boldsymbol{x}$, the DIA estimator $\bar{\boldsymbol{x}}$ is likely close to $\boldsymbol{x}$ with a sufficiently large probability, or with a sufficiently small PHMI. Defining an $\boldsymbol{x}$-centered region $\mathcal{B}_{\boldsymbol{x}}$, the PHMI can be expressed as [Blanch et al. 2010]

$$\begin{aligned} PHMI = P(\bar{\boldsymbol{x}} \in \mathcal{B}_{\boldsymbol{x}}^{C}) &= \sum_{i=0}^{N} P(HMI|\mathcal{H}_i) P_{\mathcal{H}_i} \\ &= \sum_{i=0}^{N} P(\bar{\boldsymbol{x}} \in \mathcal{B}_{\boldsymbol{x}}^{C}|\mathcal{H}_i) P_{\mathcal{H}_i} \end{aligned} \tag{27}$$

Where $\mathcal{B}_{\boldsymbol{x}}^{C} = \mathbb{R}^n / \mathcal{B}_{\boldsymbol{x}}$ and $P(HMI|\mathcal{H}_i)$ is the conditional PHMI when $\mathcal{H}_i$ holds true. With a predetermined region $\mathcal{B}_{\boldsymbol{x}}$, the PHMI is determined by the priori probability $\mathrm{P}_{\mathcal{H}_i}$ and the PDF of $\bar{\boldsymbol{x}}$. Inversely, with a predetermined threshold on $PHMI$ and a specific definition on the shape, the region $\mathcal{B}_{\boldsymbol{x}}$ can also be determined. The latter is usually applied in RAIM/ARAIM algorithms to calculate the Horizontal PL (HPL) and Vertical PL (VPL) for precision approaches in aviation applications [Blanch et al. 2010, 2012, 2015], where the region $\mathcal{B}_{\boldsymbol{x}}$ is defined as

$$\begin{aligned} \mathcal{B}_{\boldsymbol{x}} = \{ &|\bar{\boldsymbol{x}}_1 - \boldsymbol{x}_1| < HPL_1 \cap |\bar{\boldsymbol{x}}_2 - \boldsymbol{x}_2| \\ &< HPL_2 \cap |\bar{\boldsymbol{x}}_3 - \boldsymbol{x}_3| \\ &< VPL \} \end{aligned} \tag{28}$$

with $\bar{\boldsymbol{x}}_*, '*' = 1,2,3$ denotes the horizontal and vertical position solution of a DIA estimator for GNSS positioning applications, respectively. Rigorously, the PDF of $\bar{\boldsymbol{x}}$ is not normal and this further complicates the calculation of the PHMI or the PL parameters.

3.2 Interdependence of quality control indices when applying a DIA estimator

Above discussion shows that the confidence levels of the testing decisions, the reliability of the hypothesis models, as well as the biasedness, dispersion, and integrity of the estimated unknown parameters are interdependent. The straightforward flowchart for determining these parameters is demonstrated in Figure 1.

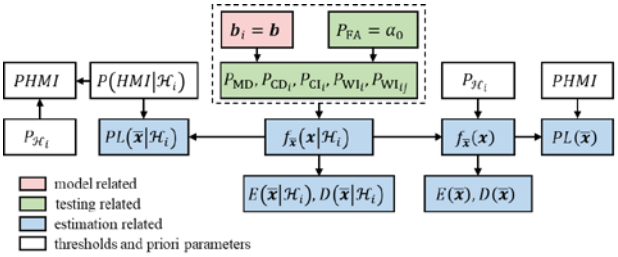

Figure 1 Demonstration of the parameters' relationship in a DIA estimator

Figure 1 shows that by directly specifying the bias of the alternative hypothesis model $\mathcal{H}_i$ and the probability of false alarm P_{FA}, the probabilities of each type of testing decisions, and the conditional PDF, expectation, dispersion of the DIA estimator $\bar{\boldsymbol{x}}$ under $\mathcal{H}_i$ are determined accordingly. With a requirement on $P(HMI|\mathcal{H}_i)$, the conditional PL of $\bar{\boldsymbol{x}}$ under $\mathcal{H}_i$ can further be calculated, which together with the priori probability of the event $\mathcal{H}_i$ ($P_{\mathcal{H}_i}$) are used to determine the unconditional PDF, PL and total PHMI of the DIA estimator $\bar{\boldsymbol{x}}$ ultimately. With these parameters the overall performance of the DIA estimator applied on the specific null and alternative hypothesis models are evaluated. For the same size of bias $\boldsymbol{b}_i$, the values of P_{CD}, P_{CI}, P_{WI} of different alternative hypotheses, as well as the corresponding biases projected onto the unknown parameters are compared. It is noted that to implement the DIA procedure, the preset value of P_{FA} are usually required to determine the critical value for the detection and identification testing.

Intrinsically, the value of P_{CD_i}, P_{CI_i}, P_{WI_i}, and $P_{\mathrm{WI}_{ij}}$ are mutual dependent. Once the mathematical models of the null and alternative hypotheses are specified, the ratios of these probabilities are uniquely dependent on the size of $\boldsymbol{b}_i$. Alternatively, if the value of P_{FA} and one of another testing decision probabilities under $\mathcal{H}_i$ are specified, the size of $\boldsymbol{b}_i$ can be inversely determined, so do the probabilities of other testing decisions under $\mathcal{H}_i$. Since this $\boldsymbol{b}_i$ can produce the same value of either P_{CD}, P_{CI}, or P_{WI} for each alternative hypothesis, it again makes comparison among the hypotheses possible, by ranking the alternative hypotheses in terms of an equal capability of detection or identification. It is also noted that the biases projected onto the unknown

parameters would be significant discrepant when the same size of $\boldsymbol{b}_i$, or the same size of either P_{CD}, or P_{CI}, or P_{WI} is predefined under different alternative hypotheses. Generally, achieving smaller estimation bias, higher precision and lower protection level of the unknown parameters are the goal for a good DIA estimator. Thus, determining the statistical distribution characteristics of $\bar{\boldsymbol{x}}$ is essential, which is still less considered when applied a DIA procedure.

4 Examples of integrity monitoring for three conventional DIA procedures

Evaluating the confidence levels of the testing decisions, the reliability of the hypotheses, as well as the biasedness, precision and integrity of the estimated parameters can describe the overall performance of a DIA estimator, and therefore can be used as a quality control procedure. In order to study the actual performances of different DIA methods, a Monte Carlo simulation was conducted using some numerical examples. Such a simulation allows for a large number of independent realizations of the same experiment and a controlled environment, which can hardly be achieved using real data. Three practical DIA procedures are considered in the study:

1. DIA_1: Detection and identification are restricted to test between the null hypothesis $\mathcal{H}_0$ and a single alternative hypothesis $\mathcal{H}_1$, as defined in Eq. (6). In this procedure, the detection and identification become one step. However, identifying the unique alternative hypothesis model beforehand is a prerequisite.
2. DIA_2: Detection is performed using a global test of Eq. (4). Once the null hypothesis is rejected, the identification tries to accept the alternative hypothesis with the largest value of T_j^2, as defined in Eq. (7).
3. DIA_3: Detection is performed using another global test, as defined as

$$\mathcal{P}_0 = \left\{ \boldsymbol{t} \in \mathbb{R}^r \,\middle|\, \max_{j\in\{1,\cdots,N\}} T_j^2 \le c(1-\alpha_0) \right\} \tag{29}$$

Once the detection is triggered, the identification is further performed with the same local test as DIA_2.

4.1 GNSS single-point positioning data and preset thresholds

The performances of above three DIA estimators are evaluated via a dual-constellation GNSS single-point positioning (SPP) example with the geometry defined by

$$\boldsymbol{A} = \begin{bmatrix} 0.0225 & 0.9951 & -0.0966 & 1 & 0 \\ 0.6750 & -0.6900 & -0.2612 & 1 & 0 \\ 0.0723 & -0.6601 & -0.7477 & 1 & 0 \\ -0.9398 & 0.2553 & -0.2269 & 1 & 0 \\ -0.5907 & -0.7539 & -0.2877 & 1 & 0 \\ -0.3236 & -0.0354 & -0.9455 & 0 & 1 \\ -0.6748 & 0.4356 & -0.5957 & 0 & 1 \\ 0.0938 & -0.7004 & -0.7075 & 0 & 1 \\ 0.5571 & 0.3088 & -0.7709 & 0 & 1 \\ 0.6622 & 0.6958 & -0.2780 & 0 & 1 \end{bmatrix} \tag{30}$$

with the variance matrix of observation $\boldsymbol{y}$ is given by

$$\boldsymbol{\Sigma}_{yy} = diag\left(\begin{bmatrix} 3.8865 & 1.4377 & 0.8604 & 1.6383 & 1.3229 \\ 0.8434 & 0.8963 & 0.8669 & 0.8573 & 1.3616 \end{bmatrix}\right) \tag{31}$$

To compute those performance evaluation parameters mentioned above, a single bias of its MDB size is specifically introduced on one of the original observations. The MDB is computed from Eq. (18) with a specific value of noncentrality parameter λ_0. The study yields estimated probabilities of the testing decisions defined in Section 3.1, as well as the bias, dispersion and PL of $\bar{\boldsymbol{x}}$ defined in Section 3.3. The critical value of the testing is determined by presetting $P_{\mathrm{FA}} = \alpha_0$. The conditional PHMI is preset as $P(HMI|\mathcal{H}_i) = IR_0 = 1 \times 10^{-4}$.

By setting $\alpha_0 = 0.1\%$, the critical value of DIA_1, and DIA_2, are calculated as $c_1 = c_{\chi^2}(1-\alpha_0, 1) = 10.83$ and $c_2 = c_{\chi^2}(1-\alpha_0, 5) = 20.52$. While it is complicate to compute the critical value of DIA_3 algebraically, therefore we get the value of c_3 by the Monte Carlo simulation. With $N_t = 1 \times 10^6$ times independent experiments, the PDF and Cumulative Distribution Function (CDF) of test statistics T_i^2 defined as

$$P\left(T_i^2 = \max_{j\in\{1,\cdots,N\}} T_j^2 \le c(1-\alpha_0)|\mathcal{H}_0\right) = \alpha_0 \tag{32}$$

can be drawn and used to determine the value of c_3 with specific value of α_0, as illustrated in Figure 2. It shows that $c_3 = c(1-\alpha_0) = 15.14$.

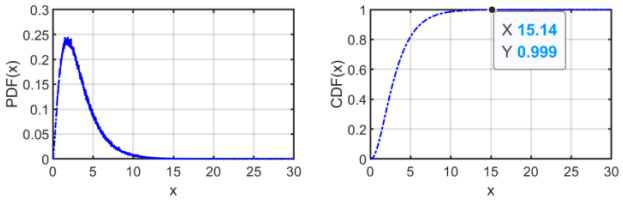

Figure 2 PDF and CDF of the global test statistics for DIA_3

4.2 Analyses under two bias models with a unique internal reliability

By setting $\alpha_0 = 0.1\%$, and $P_{\mathrm{CD}} = 80\%$ for DIA_1, the noncentrality parameter $\lambda_0 = 4.13$ is predetermined, and a bias with a size of its MDB at y_1 and y_9 is simulated respectively. With $N_t = 1 \times 10^6$ times independent experiments, the performances of three DIA methods are analyzed in terms of the probabilities of different testing decisions, the PDF and absolute CDF of the parameter estimation errors $(\bar{\boldsymbol{x}} - \boldsymbol{x})$, as well as the conditional PLs, biases and dispersions of the estimated unknown parameters under $\mathcal{H}_1$ and $\mathcal{H}_9$. Since we care more about the positioning parameters than the receiver clock error parameter, only the results of horizontal positioning parameters $(\boldsymbol{x}_1, \boldsymbol{x}_2)$ and vertical positioning parameters$(\boldsymbol{x}_3)$ are presented below.

The probabilities P_{MD}, P_{CI}, P_{WI} of the three DIA methods are listed in Table 1. It shows obvious discrepancies among the three methods, and probabilities under $\mathcal{H}_1$ and $\mathcal{H}_9$ are quite close. Among the three methods, the P_{MD}s of DIA_1 are the smallest, and are coincident with the preset theoretical value, $1 - \mathrm{P_{CD}} = 20\%$. The P_{MD}s of DIA_2 are the largest, up to 47.3%, which are also coincident with the theoretical value, computed as $\int_0^{c_2} \chi^2(r=5, \lambda_0^2 = 4.13^2)$. The P_{MD}s of DIA_3 are much smaller than those of DIA_2, but still as large as 39.9%. Both the DIA_2 and DIA_3 introduce wrong identifications. Generally, the P_{WI}s under $\mathcal{H}_1$ are slightly smaller than that under $\mathcal{H}_9$ when the estimator DIA_2 or DIA_3 is implemented. Accordingly, the P_{CI} under $\mathcal{H}_1$ is somewhat higher than that under $\mathcal{H}_9$. Probabilities listed in Table 1 shows the discrepancies of the three different DIA methods on the confidence levels of the testing decisions.

The PDFs and absolute CDFs of the estimation errors $(\bar{\boldsymbol{x}} - \boldsymbol{x})$ on the two unknown parameters are plotted in Figure 3. It shows significant differences between $\mathcal{H}_1$ and $\mathcal{H}_9$ no matter which DIA method is implemented. Apparently, the distributions of estimation errors $\bar{\boldsymbol{x}} - \boldsymbol{x}$ under $\mathcal{H}_1$ still keep normal and show little discrepancies among different DIA methods, although there are higher probabilities of committing missed detection when DIA_2 or DIA_3 is conducted. Instead, the results under $\mathcal{H}_9$ are apparently not normally distributed, and show significant discrepancies when different DIA method is conducted. The bimodal distribution characteristics of the estimation errors under $\mathcal{H}_9$ become much more conspicuous for DIA_2 and DIA_3. As the discussion in Section 3, the bimodal distributions are mainly composed by the two different normal distributions resulting from correct identification and missed detection, since the probabilities of wrong identification are still trivial under the simulated bias models.

Table 1 Probabilities of testing decisions under $\mathcal{H}_1$ and $\mathcal{H}_9$

	P_{MD} (%)		P_{CI} (%)		P_{WI} (%)	
$\mathcal{H}_i$	$\mathcal{H}_1$	$\mathcal{H}_9$	$\mathcal{H}_1$	$\mathcal{H}_9$	$\mathcal{H}_1$	$\mathcal{H}_9$
DIA_1	20.0	20.0	80.0	80.0	0.0	0.0
DIA_2	47.3	47.3	51.2	50.6	1.6	2.1
DIA_3	39.9	39.7	58.9	58.6	1.2	1.7

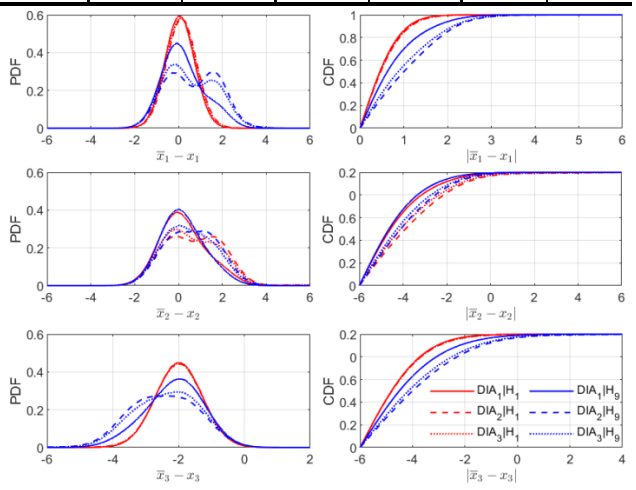

Figure 3 PDFs and CDFs of the parameter estimation errors of three DIA methods

Table 2 PLs, biases and dispersions of the estimated unknown parameters under $\mathcal{H}_1$ and $\mathcal{H}_9$

	$\bar{\boldsymbol{x}}_1$					
	PL($\bar{\boldsymbol{x}}\|\mathcal{H}_i$)		$\boldsymbol{b}_{\bar{\boldsymbol{x}}\|\mathcal{H}_i}$		std($\bar{\boldsymbol{x}}\|\mathcal{H}_i$)	
$\mathcal{H}_i$	$\mathcal{H}_1$	$\mathcal{H}_9$	$\mathcal{H}_1$	$\mathcal{H}_9$	$\mathcal{H}_1$	$\mathcal{H}_9$
DIA_1	2.61	3.75	0.04	0.21	0.67	0.99
DIA_2	3.07	4.68	0.13	0.66	0.69	1.19
DIA_3	3.00	4.67	0.11	0.53	0.69	1.17
	$\bar{\boldsymbol{x}}_2$					
	PL($\bar{\boldsymbol{x}}\|\mathcal{H}_i$)		$\boldsymbol{b}_{\bar{\boldsymbol{x}}\|\mathcal{H}_i}$		std($\bar{\boldsymbol{x}}\|\mathcal{H}_i$)	
$\mathcal{H}_i$	$\mathcal{H}_1$	$\mathcal{H}_9$	$\mathcal{H}_1$	$\mathcal{H}_9$	$\mathcal{H}_1$	$\mathcal{H}_9$
DIA_1	4.24	4.03	0.22	0.19	1.10	1.03
DIA_2	7.46	4.82	0.69	0.59	1.34	1.18
DIA_3	7.46	4.75	0.56	0.47	1.32	1.16
	$\bar{\boldsymbol{x}}_3$					
	PL($\bar{\boldsymbol{x}}\|\mathcal{H}_i$)		$\boldsymbol{b}_{\bar{\boldsymbol{x}}\|\mathcal{H}_i}$		std($\bar{\boldsymbol{x}}\|\mathcal{H}_i$)	
$\mathcal{H}_i$	$\mathcal{H}_1$	$\mathcal{H}_9$	$\mathcal{H}_1$	$\mathcal{H}_9$	$\mathcal{H}_1$	$\mathcal{H}_9$
DIA_1	6.91	8.83	0.00	-0.39	1.78	2.25
DIA_2	8.96	14.33	-0.03	-1.20	1.84	2.62
DIA_3	8.77	14.32	-0.03	-0.96	1.83	2.57

To further analyze the performances of the estimated parameters, the PLs, biases on expectation, and standard deviations (stds) of $\bar{\boldsymbol{x}}|\mathcal{H}_i$ are drawn from the PDFs and CDFs given in Figure 3 and listed in Table 2, with the conditional PHMI being given as $P(HMI|\mathcal{H}_i) = IR_0 = 1 \times 10^{-4}$. It shows the best performances on parameter estimation are achieved by DIA_1, with smallest PLs, biases and stds under both $\mathcal{H}_1$ and $\mathcal{H}_9$. The performances of DIA_2 and DIA_3 are much similar, except that the biases on expectation of DIA_2 are slightly larger. Comparisons between $\mathcal{H}_1$ and $\mathcal{H}_9$ show that all the three DIA estimators display better performances on $\bar{\boldsymbol{x}}_1$ and $\bar{\boldsymbol{x}}_3$ and worse performances on $\bar{\boldsymbol{x}}_2$ under the former. Generally, although the two bias models are simulated with the equivalent internal reliability and produce similar confidence levels on testing decisions, the performances on the parameter estimations are quite discrepant. As discussions in Section 3.1.3, the reason is the conditional normal-distributed PDF of the estimated parameters would be characterized with discrepant expectations and dispersions under different bias models.

4.3 Analyses under two bias models with the influence of changeable internal reliability

Apparently, the DIA estimator would account for a worse performance on the parameter estimation when the bias $\boldsymbol{b}_i$ is neither too large nor too small. Since a small bias would impact the parameter insignificantly, and a significantly large bias can be correctly detected and identified with high confidence levels. To demonstrate the performance discrepancies on the parameter estimations with a different size of $\boldsymbol{b}_i$, the noncentrality parameter, λ_0, used to calculate the MDB of $\boldsymbol{b}_i$ under $\mathcal{H}_1$ and $\mathcal{H}_9$ are simulated from 0 to 10 with a step length of 0.2. By setting $\alpha_0 = 0.1\%$ and $P(HMI|\mathcal{H}_i) = 1 \times 10^{-4}$, the Probabilities of CI, MD and WI of three DIA methods are shown in Figure 4, and the PLs, biases and dispersions of the estimated unknown parameters under $\mathcal{H}_1$ and $\mathcal{H}_9$ are shown in Figure 6. Table 3 lists the estimated squared variances of the unknown parameters when applying different hypothesis models. Table 4 presents the values of BNRs and slopes for the unknown parameters, as well as $\left(\boldsymbol{P}_{\boldsymbol{A}_0}^{\perp}\right)_{ii}$-the diagonal value of matrix $\boldsymbol{P}_{\boldsymbol{A}_0}^{\perp}$, which are conventionally used as reliability indices in previous studies.

Figure 4 shows that the corresponding probabilities of CI and MD for $\mathcal{H}_1$ and $\mathcal{H}_9$ are similar. As λ_0 increases, P_{CI} gradually increase from 0 to 1 and P_{MD} encounter adverse tendencies. For a specific λ_0, P_{CI} of DIA_1 is the largest, and P_{CI} of DIA_2 is the smallest. This indicates that the overall confidence level of testing decisions for DIA_1 is the highest, and that for DIA_2 is the lowest. Theoretically, P_{MD} for each DIA estimator under $\mathcal{H}_1$ and $\mathcal{H}_9$ should be equivalent for the same λ_0. Generally, the probabilities of WI show parabolic change tendencies, increasing gradually at first and then decrease to 0 as the bias is large enough. Also, P_{WI} for DIA_2 are much higher than those for DIA_3 under both $\mathcal{H}_1$ and $\mathcal{H}_9$. When the same DIA estimator is applied, the P_{WI} for $\mathcal{H}_9$ are always much large than those for $\mathcal{H}_1$. It is noted that WI are always not considered for DIA_1.

The discrepancies on P_{WI} under $\mathcal{H}_1$ and $\mathcal{H}_9$ can be further explained by the discrepant correlation

coefficients between the implemented model ($\mathcal{H}_1$ or $\mathcal{H}_9$) and another one. As shown in Figure 5, the correlation coefficients between $\mathcal{H}_1$ and another alternative hypothesis are generally much smaller than those between $\mathcal{H}_9$ and another one. Since larger correlation coefficient will result in a higher probability of wrong identification, the total P_{WI} under $\mathcal{H}_9$ will be larger than that under $\mathcal{H}_1$. Also, since the correlation coefficients between any two alternative hypotheses are not significantly large (i.e. >0.8), the total probabilities of WI for this example are relatively unremarkable, always under 2%.

Figure 6 shows distinguished differences on the parameter estimation performances when the three DIA estimators are implemented under either $\mathcal{H}_1$ or $\mathcal{H}_9$. Generally, PLs, biases and stds of $\bar{\boldsymbol{x}}_1$ and $\bar{\boldsymbol{x}}_3$ under $\mathcal{H}_9$ are apparently larger than their counterparts under $\mathcal{H}_1$. Also, the influences of different $\boldsymbol{b}_i$ on $\bar{\boldsymbol{x}}_1$ and $\bar{\boldsymbol{x}}_3$ under $\mathcal{H}_9$ are relatively more significant and discrepancies among the three DIA estimators are much larger than under $\mathcal{H}_1$. However, the performances on $\bar{\boldsymbol{x}}_2$ are reversed. Comparisons among the three DIA estimators show that DIA_1 always performs much better than DIA_2 and DIA_3, and performances of DIA_2 are slightly worse than DIA_3. Specifically, PLs of $\bar{\boldsymbol{x}}_1|\mathcal{H}_9$ increase gradually from around 2.6 to 3.8 for DIA_1 and to 5.4 for DIA_2 and DIA_3, and then all decrease more rapidly to around 3 as the size of λ_0 increases to 10. Similarly, the biases of $\bar{\boldsymbol{x}}_1|\mathcal{H}_9$ increase from 0 to 0.65, 0.93, and 0.87 respectively and then gradually fall back to 0 for the three methods. The stds of $\bar{\boldsymbol{x}}_1|\mathcal{H}_9$ increase from 0.65 to 1.05, 1.20 and 1.17 respectively and then all fall to 0.76 for the three methods. Performances on $\bar{\boldsymbol{x}}_2|\mathcal{H}_9$ and $\bar{\boldsymbol{x}}_3|\mathcal{H}_9$ are similar with $\bar{\boldsymbol{x}}_1|\mathcal{H}_9$, with DIA_2 being the worst and DIA_1 being the best. It is noted that the sizes of λ_0 for the PLs, biases and stds reaching their largest values are different for each DIA estimator. Generally, the sizes of λ_0 corresponding to the largest PLs, biases and stds for DIA_1 are the smallest, and those for of DIA_2 are slightly larger than DIA_3. The largest PLs of $\bar{\boldsymbol{x}}_1|\mathcal{H}_9$ appear at $\lambda_0 = 5.0$ for DIA_1 and $\lambda_0 = 6.2$ for DIA_2 and DIA_3; The largest biases of $\bar{\boldsymbol{x}}_1|\mathcal{H}_9$ appear at $\lambda_0 = 2.4$, 3.0 and 2.8, and the largest stds appear at $\lambda_0 = 3.4$, 4.4 and 4.2 for the three DIA estimators respectively. Comparably, PLs of $\bar{\boldsymbol{x}}_1|\mathcal{H}_1$ are always smaller than 3, the largest bias on $\bar{\boldsymbol{x}}_1|\mathcal{H}_1$ is just around 0.20, and stds change rarely between 0.66 and 0.7 for all the three DIA estimators.

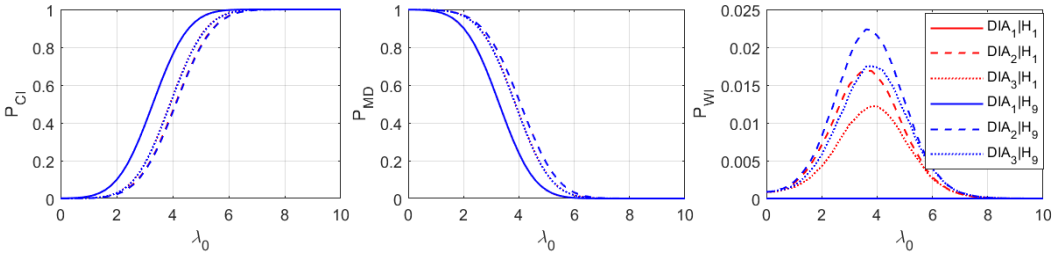

Figure 4 Probabilities of CI, MD and WI of three DIA methods under $\mathcal{H}_1$ and $\mathcal{H}_9$

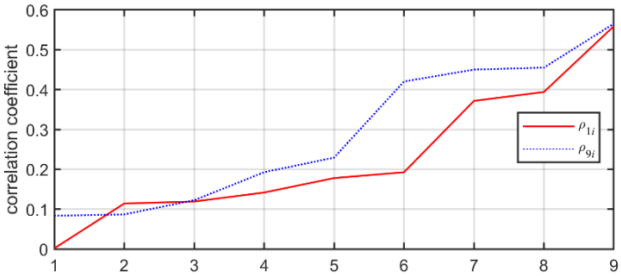

Figure 5 correlation coefficients between the implemented model ($\mathcal{H}_1$ or $\mathcal{H}_9$) and another one

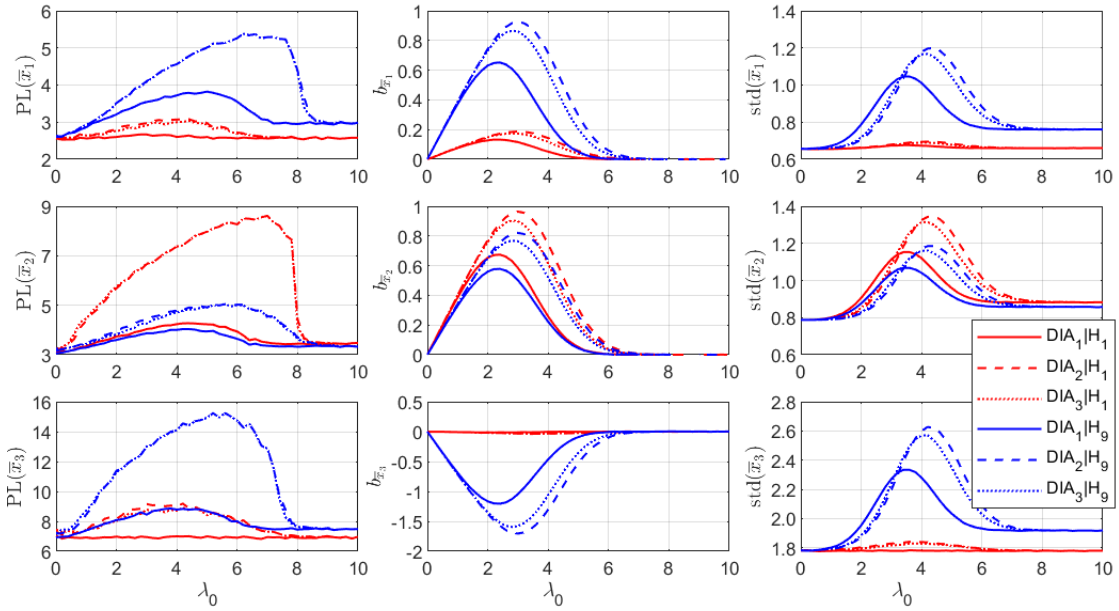

Figure 6 PLs, biases and dispersions of the parameter estimation errors of three DIA methods

Table 3 Estimated squared variances of the parameters under each alternative hypothesis model

$\mathcal{H}_i$		**0**	**1**	2	3	4	5	6	7	8	**9**	10
$\sqrt{\boldsymbol{\Sigma}_{\hat{x}_i\hat{x}_i}}$	$\boldsymbol{x}_1$	**0.66**	**0.66**	0.74	0.69	0.74	0.71	0.68	0.78	0.66	**0.76**	0.74
	$\boldsymbol{x}_2$	**0.79**	**0.88**	0.83	0.81	0.82	0.88	0.79	0.80	1.16	**0.86**	0.80
	$\boldsymbol{x}_3$	**1.78**	**1.78**	1.89	2.51	1.80	1.96	2.03	1.80	1.99	**1.92**	2.05

It is interesting to notice that as the size of λ_0 enlarges, PLs, biases and stds of both $\bar{\boldsymbol{x}}_1|\mathcal{H}_1$ and $\bar{\boldsymbol{x}}_1|\mathcal{H}_9$ tend to an equivalent value, no matter which DIA estimator is applied. The reason is $\bar{\boldsymbol{x}}|\mathcal{H}_i$ for $\lambda_0 = 0$ are dominant by $\hat{\boldsymbol{x}}_0$ (solution of Eq. (1)), and $\bar{\boldsymbol{x}}|\mathcal{H}_i$ for a significantly large λ_0 are dominant by $\hat{\boldsymbol{x}}_i$ (solution of Eq. (2)). Therefore, in both cases, $\bar{\boldsymbol{x}}|\mathcal{H}_i$ will be unbiased, stds of $\bar{\boldsymbol{x}}|\mathcal{H}_i$ are the squared value of $\boldsymbol{\Sigma}_{\hat{x}_0\hat{x}_0}$ or $\boldsymbol{\Sigma}_{\hat{x}_i\hat{x}_i}$ as listed in Table 3, and PLs of $\bar{\boldsymbol{x}}|\mathcal{H}_i$ are determined by the std values and the preset value of $P(HMI|\mathcal{H}_i)$.

The performance discrepancies between $\mathcal{H}_1$ and $\mathcal{H}_9$ can be explained qualitatively by those reliability indices given in Table 4, which are conventional investigated in previous studies. The discrepancies on the PLs and biases under $\mathcal{H}_1$ and $\mathcal{H}_9$ are explained by the values of BNR and slope listed in the 5nd to 10th row of Table 4. As defined in Eq.(22) and (24), BNR is the ratio between $\boldsymbol{b}_i$ and $\boldsymbol{b}_{x_{i0}}$, and slope is the ratio between λ_0 and the $\boldsymbol{b}_{x_{i0}}$. Therefore, larger BNR or slope indicates that larger bias would be introduced on the estimated parameter for the same size of $\boldsymbol{b}_i$ or λ_0. Also, $\left(\boldsymbol{P}_{\boldsymbol{A}_0}^{\perp}\right)_{ii}$ known as the redundancy of the measurement indicates the impact of the measurement on the estimated parameters. Measurement with redundancy closer to 1 would less impact the biasedness and dispersion of the parameter estimation. The consistent results for $\mathcal{H}_1$ and $\mathcal{H}_9$ are observed in Figure 6. Although the confidence level of testing decisions under $\mathcal{H}_1$ and $\mathcal{H}_9$ are similar, the performances on the parameter estimation are significantly discrepant, where bias model $\mathcal{H}_1$ impacts the PLs, biases and stds of $\bar{\boldsymbol{x}}_1$ and $\bar{\boldsymbol{x}}_3$ much less, and impacts $\bar{\boldsymbol{x}}_2$ much more.

Table 4 BNRs, slopes and redundancies of each measurement

Measurement No.		**1**	2	3	4	5	6	7	8	**9**	10
BNR	x_1	**0.03**	0.21	0.14	-0.20	-0.18	-0.14	-0.32	-0.03	**0.30**	0.19
	x_2	**0.17**	-0.15	0.11	0.13	-0.26	0.09	0.10	-0.55	**0.27**	0.09
	x_3	**0.00**	0.39	-1.06	0.14	0.53	-0.79	0.18	0.58	**-0.56**	0.58
slope	x_1	**0.08**	0.35	0.23	-0.34	-0.28	-0.18	-0.42	-0.05	**0.39**	0.34
	x_2	**0.40**	-0.25	0.18	0.23	-0.40	0.11	0.13	-0.85	**0.34**	0.17
	x_3	**-0.01**	0.65	-1.77	0.25	0.83	-0.98	0.24	0.90	**-0.72**	1.03
$\left(\boldsymbol{P}_{\boldsymbol{A}_0}^{\perp}\right)_{ii}$		**0.69**	0.53	0.31	0.55	0.54	0.54	0.52	0.36	**0.52**	0.44

4.4 Analyses under two bias models with the influence of changeable P_{FA}

In this subsection, we discuss the impacts of the preset value of P_{FA}, which are used to determine the critical value for detection. With the preset value of P_{FA}=α_0=0.5%, 1%, 5%, 10%, probabilities of CI, MD and WI of the three DIA methods under $\mathcal{H}_1$ and $\mathcal{H}_9$ respected to $\lambda_0 = 0$ to 10 are plotted in Figure 7, to demonstrate the confidence level of testing decisions. The corresponding values of PLs, biases and dispersions of parameter $\bar{\boldsymbol{x}}_1$ are plotted in Figure 8, to demonstrate the performances on parameter estimation of the three DIA estimators. The results of parameter $\bar{\boldsymbol{x}}_2$ and $\bar{\boldsymbol{x}}_3$ display similar tendencies, therefore are not presented.

Figure 7 shows that larger P_{FA} will produce higher P_{CI} and lower P_{MD} no matter which DIA estimator is applied. Specifically, as P_{FA} changes from 0.5% to 10%, P_{CI} of DIA_1 will increase from 21% to 64%, and P_{CI} of DIA_2 and DIA_3 will increase from 7% to 26% for $\lambda_0 = 2$. P_{CI} of all the three estimators will go up to close 1 when λ_0 increase to 6. P_{MD} of the three methods generally show adverse tendencies as P_{CI}. P_{WI} of DIA_2 and DIA_3 show nonmonotonic tendencies, firstly increasing and then falling to 0 with λ_0 becoming large enough. Also, P_{WI} become much higher for larger P_{FA}. This means the reduction on P_{MD} caused by a larger P_{FA} will be allocated to both P_{CI} and P_{WI}. Note that as P_{FA} enlarges P_{CI} of DIA_1 will increase gradually from 0 to 10%, while P_{CI} of DIA_2 and DIA_3 always keep around 0% for $\lambda_0 = 0$. Similarly, P_{MD} of the three estimators all decrease from 99.5% to 90% for $\lambda_0 = 0$, and discrepancies among estimators are insignificant. While, P_{WI} of DIA_2 and DIA_3 increase from 0.5% to 9% for $\lambda_0 = 0$. These discrepancies can be explained by the definitions given in Eq.(8). Theoretically, P_{MD} presents the probability of correct acceptance of $\mathcal{H}_0$ when $\lambda_0 = 0$, so it should be equal to the value of 1-P_{FA} for all the three estimators. Also, the sums of P_{CI} and P_{WI} for $\lambda_0 = 0$ should be the probability of false alarm. Since P_{WI} of DIA_1 are 0, P_{CI} of DIA_1 are always higher than the other estimators. It is also noted that the minimal values of λ_0 for P_{MD} declining to 0, and for P_{WI} reaching the maximum, become smaller for larger P_{FA}. This indicates that the confidence level of the testing decisions for smaller biases will increase when P_{FA} getting larger.

Figure 8 shows the change tendencies of PLs, biases and dispersions of parameter $\bar{\boldsymbol{x}}_1$ with different values of P_{FA}. Generally, performances of PLs for different P_{FA} are similar, with the maximum values of PLs unchanged. Comparably, as P_{FA} enlarges the biases and stds of $\bar{\boldsymbol{x}}_1$ will all reduce for the three estimators. This means under a specified bias model, a larger P_{FA} will benefits the DIA estimator in terms of smaller biases and dispersions on the parameter estimation. When the bias is larger enough, PLs, biases and dispersions of the parameter will tend to a fixed value which is independent to the value of P_{FA} and by the DIA estimator. This indicates that different DIA estimators and different critical values for hypothesis testing generally show discrepant effects when the bias is of medium sizes. Inherently, how a DIA estimator would perform for a bias with medium size could be an effective criterion to measure its general robustness, since an unremarkable bias will

not influence the parameter estimation severely, and a much easier.
significantly large bias will be correctly identified

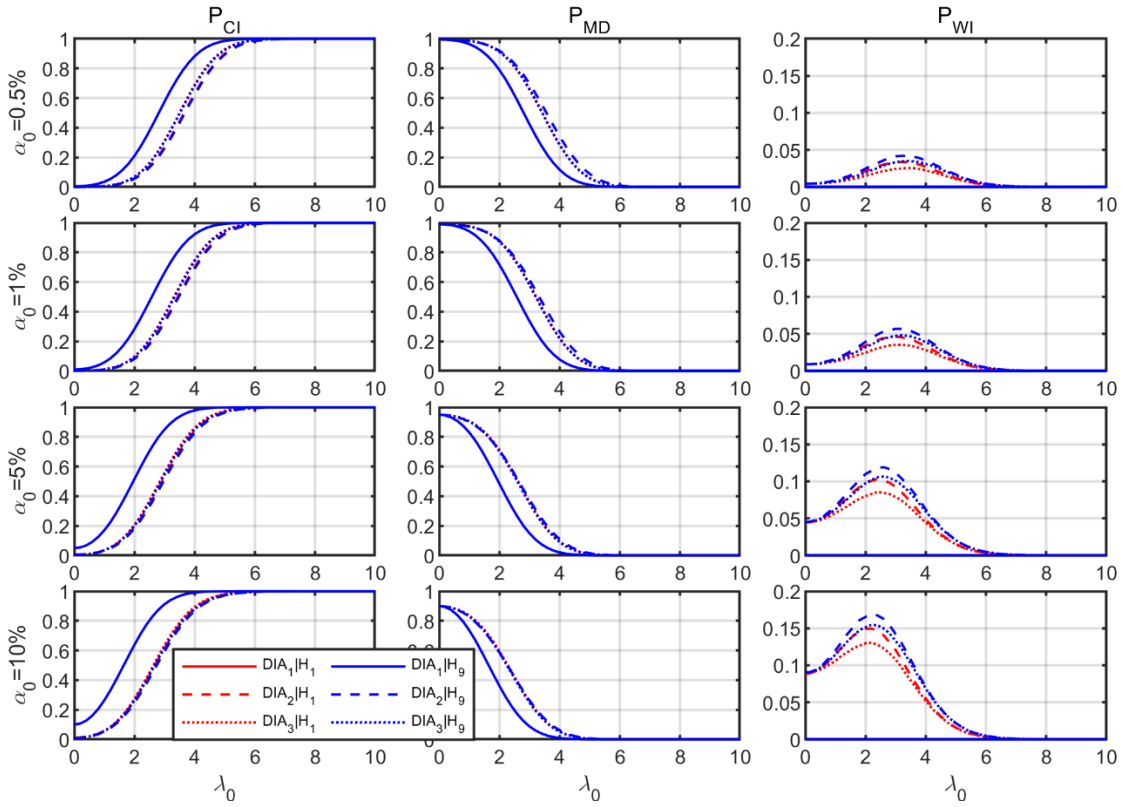

Figure 7 Probabilities of CI, MD and WI of three DIA methods under $\mathcal{H}_1$ and $\mathcal{H}_9$ with different P_{FA} (0.5%, 1%, 5%, 10%)

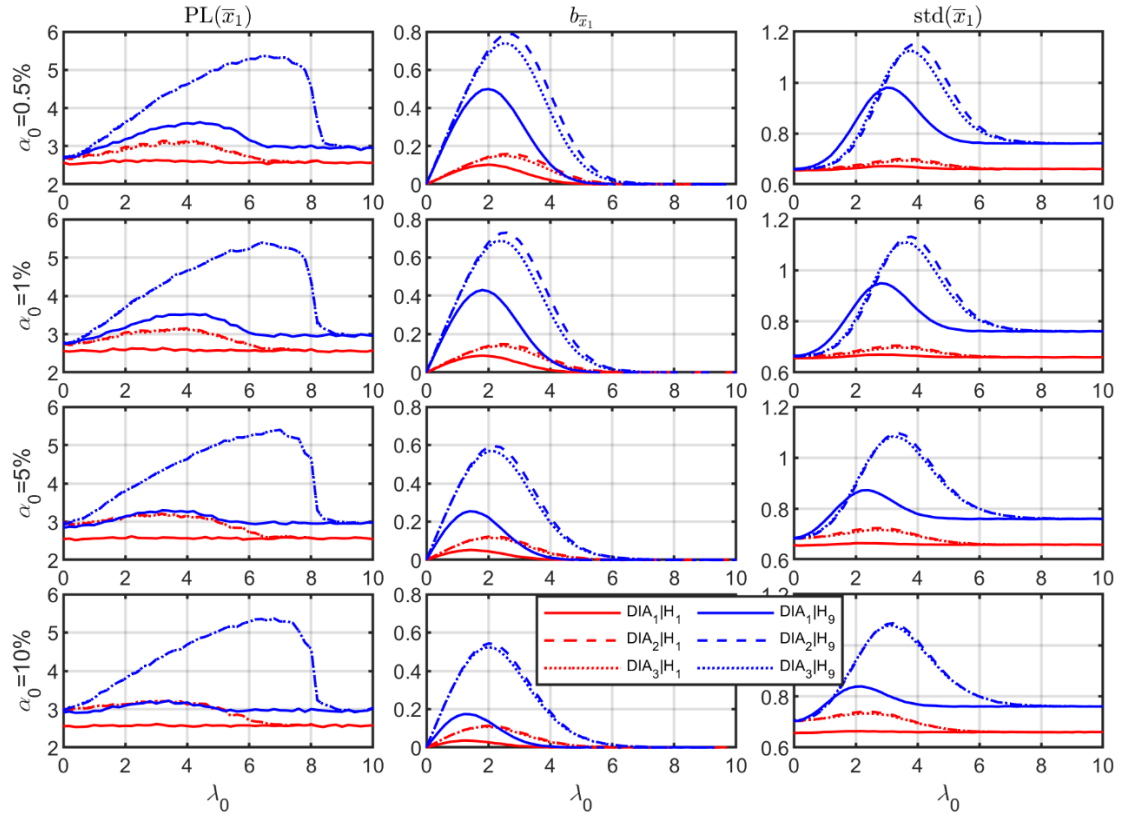

Figure 8 PLs, biases and dispersions of $\bar{\boldsymbol{x}}_1$ of three DIA methods under $\mathcal{H}_1$ and $\mathcal{H}_9$ with different P_{FA} (0.5%, 1%, 5%, 10%)

5 Conclusions

In this paper, we discuss the integrity assessment for generalized DIA estimators. Performance indices of a DIA estimator are defined from three aspects, those are, the confidence levels of the hypothesis testing decisions, the reliability of the alternative hypothesis models, as well as the biases, dispersions, and PLs of the estimated parameters. The interdependences of these indices are detailed explained and discussed.

By these indices, the quality of three conventional used DIA estimators is evaluated and compared via a GNSS SPP numerical example. To ensure a scientifically fair comparison, all computations are based on the same noncentrality parameter, λ_0, and P_{FA}. Firstly, the confidence levels of the testing decisions by using the three DIA estimators are evaluated and compared, which confirms that different methods provide discrepant probabilities of correct detection, correct identification and wrong identification. Secondly, the PDFs and CDFs of the estimated parameters under a specified hypothesis model are displayed, which probably show significant non-Guassian properties, and exhibit remarkable discrepancies among different hypothesis models. Finally, based on the PDFs, the PLs, biases and dispersions of the estimated parameters are further computed to demonstrate the performances of the DIA methods on the parameter estimation.

By shifting the value of λ_0, it shows that for all DIA estimators, controlling the performance for bias with medium size would be the most challenging task. Also, different DIA estimators could exhibit much more remarkable discrepancies in this case. By shifting the value of P_{FA}, it shows that when a bias model occurs larger P_{FA} could certainly improve the quality of the DIA estimator, by increasing the probability of correct identification, as well as reducing biases and dispersions of the estimated parameters. However, since biases usually occur with a relatively low possibility in a practical application, larger P_{FA}, on the other hand, would wrongly exclude more normal observations, and therefore may reduce the quality of the DIA estimator. Generally, quality of a DIA estimator is ultimate measured based on the PDF of the estimated parameters $\bar{\boldsymbol{x}}$, which would be uncertainly influenced by the priori probability of occurrence on a specific bias model $\mathcal{H}_i$.

Data availability: Necessary data is accessible in the Numerical Example Section of the manuscript.

References

Alberda, J. E. (1976): Quality control in surveying, Chart Surv 4(2):23–28.

Baarda, W. (1967): Statistical concepts in geodesy, Netherlands Geodetic Commission, Publication on geodesy, New series 2(4).

Baarda, W. (1968): A testing procedure for use in geodetic networks, Netherlands Geodetic Commission, Publication on geodesy, New Series 2(5).

Baarda,W. (1976): Reliability and precision of geodetic networks, Tech. rep., VII Int Kurs fur Ingenieursmessungen hoher Prazision, Band I, T.H. Darmstadt, Inst. fur Geodaesie

Blanch, J.; Walter, T. and Enge, P. (2010): RAIM with optimal integrity and continuity allocations under multiple failures, IEEE Trans Aerosp Electron Syst 46(3):1235–1247.

Blanch, J.; Walter, T.; Enge, P.; Lee, Y.; Pervan, B.; Ripplb, M. and Spletter, A. (2012): Advanced RAIM user algorithm description: integrity support message processing, fault detection, exclusion, and protection level calculation, In: Proc. of ION ITM-2012, ION, pp 2828–2849.

Blanch, J.; Walter, T.; Enge, P.; Lee, Y. C.; Pervan, B.; Rippl, M.; Spletter, A.; and Kropp, V. (2015): Baseline advanced RAIM user algorithm and possible improvements, IEEE Trans Aerosp Electron Syst 51(1): 713–732.

Chen, Y.; and Wang, J. (1996): Reliability measure for correlated observations, Z. Vermess 121(5): 211–219.

Förstner, W. (1983): Reliability and discernability of extended Gauss-Markov models, Seminar on Mathematical Models to Outliers and Systematic Errors, Series A, No. 98, Deutsche Geodätische Kommision, Munich, Germany, 79–103.

Förstner, W. (1985): The reliability of block

triangulation, Photogramm Eng Remote Sens 51(6): 1137–1149.

Hekimoglu, S.; Berber, M. (2003): Effectiveness of robust methods in heterogeneous linear models, J Geodesy 76(11–12): 706–713.

Lee, Y. C. (1986): Analysis of range and position comparison methods as a means to provide GPS integrity in the user receiver, In Proceedings of the Annual Meeting of the Institute of Navigation, Seattle, WA, USA, 24–26 June 1986; 1–4.

Leick, A.; and Emmons, M. B. (1994): Quality control with reliability for large GPS networks, J Surv Eng 120(1): 25–41.

Lehmann, R. (2012): Improved critical values for extreme normalized and studentized residuals in Gauss–Markov models, Journal of Geodesy, 86(12): 1137–1146.

Lehmann, R. (2013): On the formulation of the alternative hypothesis for geodetic outlier detection, Journal of Geodesy, 87(4): 373–386.

Lehmann, R. (2014): Transformation model selection by multiple hypotheses testing, Journal of Geodesy, 88: 1117–1130.

Koch, K. R. (1999): Parameter estimation and hypothesis testing in linear models, 2nd Ed., Springer, Berlin.

Koch, K. R. (2015): Minimal detectable outliers as measures of reliability, Journal of Geodesy, 89:483–490.

Kok, J. J. (1984): On data snooping and multiple outlier testing, US Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, Charting and Geodetic Services.

Ou, J. (1999): On the reliability for the situation of correlated observations, Acta Geodaetica et Cartographica Sinica, English Edition, pp 9–17.

Parkinson, B. W.; and Axelrad, P. (1988): Autonomous GPS integrity monitoring using the pseudorange residual, Navigation, 35, 255–274.

Pelzer, H. (1980): Some criteria for the reliability of networks, Deutsche Geodätische Kommission, Reihe B, No. 252, Munchen, Germany.

Pope, A.J. (1976): The statistics of residuals and the detection of outliers, NOAA Technical Report NOS 65 NGS 1, U.S. National Geodetic Survey, MD, USA.

Ryan, S.; and Lachapelle, G. (2001): Marine positioning multiple multipath error detection, Hydrogr. J., 100, 3–11.

Schaffrin, B. (1997): Reliability measures for correlated observations, Journal of Surveying Engineering, 123(3): 126-137.

Teunissen, P. J. G. (1985): Quality control in geodetic networks, In: Grafarend EW, Sanso F (eds) Optimization and design of geodetic networks, pp 526–547.

Teunissen, P. J. G. (2006): Testing theory; an introduction, Series on mathematical geodesy and positioning, 2nd edn. Delft University of Technology, Delft

Teunissen, P. J. G.; Imparato, D.; and Tiberius, C. (2017): Does RAIM with Correct Exclusion Produce Unbiased Positions? Sensors, 17: 1508.

Teunissen, P. J. G.; (2018): Distributional theory for the DIA method, Journal of Geodesy, 92:59–80.

Wang, J.; and Chen, Y. (1999): Outlier detection and reliability measures for singular adjustment models, Geomatics Res. Aust., 71, 57–72.

Wieser, A. (2004): Reliability checking for GNSS baseline and network processing, GPS Solut 8:55–66.

Xu, P. (1987): A test method for many outliers, ITC J 4:324–317.

Yang, L.; Li, B.; Shen, Y.; and Rizos, C. (2017): An extension of internal reliability analysis regarding to separability analysis, Journal of Surveying Engineering, 143(3): 04017002.

Yang, L.; and Shen, Y. (2020): Robust M estimation for 3D correlated vector observations based on modified bifactor weight reduction model, Journal of Geodesy, 94: 31.

Yang, L.; Wang, J.; Knight, N. L.; and Shen, Y. (2013): Outlier separability analysis with a multiple alternative hypotheses test, Journal of Geodesy, 87(6): 591-604.

Yang, L.; Shen, Y.; Li, B.; and Rizos, C. (2021): Simplified algebraic estimation for the quality control of DIA estimator, Journal of Geodesy, 95:14.

Zaminpardaz, S.; and Teunissen, P. J. G. (2019): DIA-datasnooping and identifiability, Journal of Geodesy, 93:85–101.

Zaminpardaz, S.; Teunissen, P. J. G.; and Tiberius, C. C. J. M. (2019): Risking to underestimate the integrity risk, GPS Solutions, 2019, 23:29.

Author

Ling Yang received her B.S. (2008) and M.S. (2010) degrees in Surveying Engineering from Tongji University, Shanghai, China, and the Ph.D. degree in Surveying and Spatial Information from the University of New South Wales (UNSW), Sydney, Australia, in 2014. She is currently an Associate Professor at College of Surveying and Geo-informatics of Tongji University. Her research interests include quality control and integrity monitoring for GNSS+ Positioning and Navigation.