The probability of type I and type II errors in imprecise hypothesis testing

Ingo Neumann and Hansjörg Kutterer
Geodetic Institute, Leibniz University of Hannover, Nienburger Straße 1, D-30419 Hannover
e-mail: [neumann, kutterer]@gih.uni-hannover.de

Abstract: In many engineering disciplines the interesting model parameters are estimated from a large number of heterogeneous and redundant observations by a least-squares adjustment. The significance of the model parameters and the model selection itself are checked with statistical hypothesis tests. After formulating a null hypothesis, the test decision is based on the comparison of a test value with a quantile value. The acceptance and the rejection of the null hypothesis are strongly related with two types of errors. A type I error occurs if the null hypothesis is rejected, although it is true. A type II error occurs if the null hypothesis is accepted, although it is false. This procedure is well known in case of only random errors for the observations.

If the uncertainty budget of the observations is assumed to comprise both random variability (probabilistic errors) and imprecision (interval errors), the classical test strategies have to be extended accordingly. In this study we focus on the relation of imprecision and the probability of type I and type II errors. These steps are based on newly developed one- and multidimensional hypothesis tests in case of imprecise data. The applied procedure is outlined in detail showing both theory and one numerical example for the parameterization of a geodetic monitoring network. Its main benefit is an improved interpretation of the influence of imprecision in model selection and significance tests. In addition the well known sensitivity analysis in parameter estimation can now generally be treated in terms of imprecise data.

Keywords: hypothesis testing, imprecision, probability, type I/II error

1. Introduction

Hypothesis tests are of wide interest for many applications in engineering and mathematical science. Different approaches to hypothesis testing exist, which are due to different methods for the description of the occurring uncertainties, e. g., in the performed measurements and the prior knowledge about the model formulation (for further data processing) and in model selection. The probably most popular approaches are statistical tests in parameter estimation, where interesting model parameters are estimated from a large number of heterogeneous and redundant observations by a least-squares adjustment. The uncertainties are assessed in a stochastic framework: measurement and system errors are modeled using random variables and probability distributions. However, the quantification of the uncertainty budget of empirical measurements is often too optimistic due to, e.g., the ignorance of non-stochastic errors in the analysis process

(Ferson et al., 2007). For this reason in this paper a more general formulation is presented which may be closer to the situation in real-world applications.

The paper is organized as follows: first, the main steps in uncertainty modeling with respect to non-stochastic measurement errors are briefly reviewed, see, e. g. (Kutterer, 2004; Neumann et al., 2006). Second, two linear hypotheses are introduced as a general approach to imprecise hypothesis testing. The main part of the paper deals with the relation of imprecision and the probability of type I and type II errors in imprecise hypothesis testing. The applied procedure is outlined in detail showing both theory and numerical examples for the parameterization of a geodetic monitoring network.

2. Hypothesis testing in parameter estimation under interval-/fuzzy-uncertainty

2.1. Modeling of Uncertainty

In this paper *uncertainty* is treated in terms of fuzzy-intervals (e. g., Bandemer and Näther 1992), see Fig. 1. With a fuzzy-interval $\tilde{A}$ it is possible to describe uncertain quantities by their membership function $m_{\tilde{A}}(x)$ over the set $\mathbb{R}$ of real numbers with a membership degree between 0 and 1:

$$\tilde{A} := \left\{ (x, m_{\tilde{A}}(x)) \middle| x \in \mathbb{R} \right\} \quad \text{with} \quad m_{\tilde{A}} : \mathbb{R} \to [0,1]. \tag{1}$$

The membership function of a fuzzy interval can be described by its left *(L)* and right (*R*) reference functions (see also Fig. 1)

$$m_{\tilde{A}}(x) = \begin{cases} L\left(\dfrac{x_m - x - r}{c_l}\right), & x < x_m - r \\ 1, & x_m - r \le x \le x_m + r \\ R\left(\dfrac{x - x_m - r}{c_r}\right), & x > x_m + r \end{cases} \tag{2}$$

with x_m denoting the midpoint, r the radius, and c_l, c_r the spread parameters of the monotonously decreasing reference functions (convex fuzzy intervals).

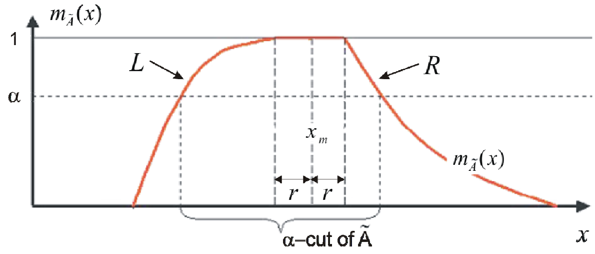


Figure 1. Fuzzy interval and its α -cut

The α - *cut* of a fuzzy-interval $\tilde{A}$ is defined by:

$$\tilde{A}_\alpha := \left\{ x \in X \middle| m_{\tilde{A}}(x) \geq \alpha \right\}, \tag{3}$$

with $\alpha \in [0,1]$. Each α - *cut* represents in case of monotonously decreasing reference functions a classical interval. The lower bound $\tilde{A}_{\alpha,\min}$ and upper bound $\tilde{A}_{\alpha,\max}$ of an α - *cut* are obtained as:

$$\tilde{A}_{\alpha,\min} = min\left(\tilde{A}_\alpha\right), \tag{4}$$

$$\tilde{A}_{\alpha,\max} = max\left(\tilde{A}_\alpha\right). \tag{5}$$

Throughout the paper we assume symmetric fuzzy intervals. Hence, an equivalent representation of symmetric α - *cuts* can be found by the midpoint A_m and radius $\tilde{A}_{\alpha,r}$ representation:

$$\tilde{A}_{\alpha,\min} = A_m - \tilde{A}_{\alpha,r}, \tag{6}$$

$$\tilde{A}_{\alpha,\max} = A_m + \tilde{A}_{\alpha,r}. \tag{7}$$

The integral over all α - *cuts* equals the membership function:

$$m_{\tilde{A}}(x) = \int_0^1 m_{\tilde{A}_\alpha}(x) d\alpha. \tag{8}$$

Furthermore, basic operations on fuzzy intervals are the *intersection* and the *complement*; they are defined through the following membership functions:

Intersection: $\tilde{C} = \tilde{A} \cap \tilde{B} \Leftrightarrow m_{\tilde{A}\cap\tilde{B}}(x) = \min\left(m_{\tilde{A}}(x), m_{\tilde{B}}(x)\right) \quad \forall\, x \in \mathbb{R}$ (9_a)

Complement: $\tilde{C} = \tilde{A}^C \Leftrightarrow m_{\tilde{A}^C}(x) = 1 - m_{\tilde{A}}(x) \quad \forall\, x \in \mathbb{R}$ (9_b)

Fuzzy intervals serve as basic quantities: *Random variability* is introduced through the fuzzy-interval midpoint which is modeled as a random variable and hence treated by methods of stochastics. Here random variability is superposed by *imprecision* which is due to non-stochastic errors of the measurements and the physical model with respect to reality. The standard deviation σ_x is the carrier of the stochastic uncertainty, and the spread of the fuzzy-intervals describes the range of *imprecision*.

For the modeling of imprecision it is important to know that the original measurement results are typically preprocessed before they are used in the further calculations. These preprocessing steps comprise several factors $\mathbf{p}$ influencing the observations (see also Fig. 2):

- Physical parameters (model constants) for the reduction and correction steps from the original to the reduced measurements
- Sensor parameters (e. g., remaining error sources that cannot be modeled)
- Additional information (e. g., temperature and pressure measurements for the reduction steps of a distance measurement)

Most of these influence factors are uncertain realisations of random variables; their imprecision is meaningful by many reasons:

- The model constants are only partially representative for the given situation (e. g., the model constants for the refraction index for distance measurements).
- The number of additional information (measurements) may be too small to estimate reliable distributions.
- Displayed measurement results are affected by rounding errors.
- Other non-stochastic errors of the reduced observations occur due to neglected correction and reduction steps and for effects that cannot be modeled.

Figure 2 shows the interaction between the observation and analysis model and their influence factors. While correction and reduction steps are systematic, the imprecision of the influence parameters is directly transferred to the measurements, which are now carrier of random variability and imprecision.

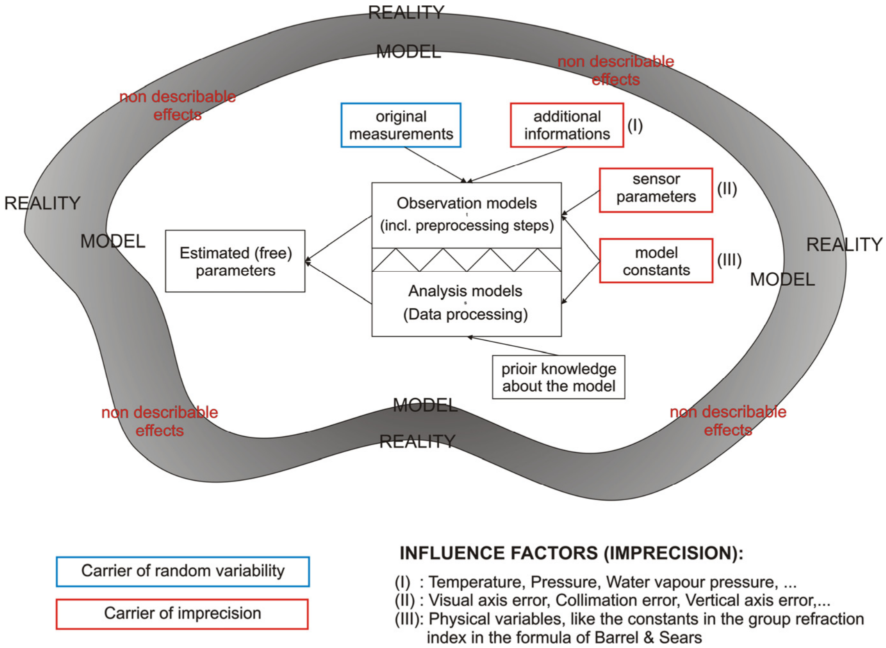


Figure 2. Interaction between the observation/analysis model and their influence factors

The non-stochastic part of the influence factors is described by fuzzy-intervals. This step is based on expert knowledge and on error models concerning the deterministic behavior of these parameters. The propagation of uncertainty is then separated into two parts. The stochastic part is treated with the law of variance-covariance propagation. Based on the assumption that imprecision is small in comparison with the measured values, we derive the data imprecision by means of a sensitivity analysis of the mostly sophisticated observation models (Neumann et al., 2006).

2.2. GENERAL FORM OF A LINEAR HYPOTHESIS IN IMPRECISE HYPOTHESIS TESTING

2.2.1. *The pure stochastic case*

In this subsection a general approach to imprecise hypothesis testing in parameter estimation is presented. We focus on the standard case where the vector $\mathbf{y}$ is assumed to be normal distributed with expectation vector

$$\mathrm{E}(\mathbf{y}) = \boldsymbol{\mu}_{\mathrm{y}}, \tag{10\_a}$$

and (positive definite) variance-covariance matrix $\boldsymbol{\Sigma}_{\mathrm{yy}}$ (vcm)

$$\mathrm{D}(\mathbf{y}) = \boldsymbol{\Sigma}_{\mathrm{yy}} = \sigma_0^2 \mathbf{Q}_{\mathrm{yy}}, \tag{10\_b}$$

where σ_0^2 denotes the variance of the unit weight and $\mathbf{Q}_{\mathrm{yy}}$ the associated cofactor matrix. Such a random vector may either be an observable quantity or a derivable quantity such as the parameters estimated by means of a least-squares (LS) adjustment. The next steps of these well-known test procedure leads to a quadratic form, which may be given by

$$\mathbf{y}^T \boldsymbol{\Sigma}_{\mathrm{yy}}^{-1} \mathbf{y} \sim \chi^2(f, \lambda). \tag{11}$$

In general, the quadratic form follows a non-central chi-square distribution with $\mathrm{f} = \mathrm{rank}(\boldsymbol{\Sigma}_{\mathbf{yy}})$ degrees of freedom and the non-centrality parameter λ. In the following, the vector $\mathbf{y}$ is assumed as the vector of reduced observations $\mathbf{y} = \mathbf{l} - \mathbf{a_0}$ within a least-squares adjustment, with the random vector of observations $\mathbf{l}$ and the deterministic vector of approximate observations $\mathbf{a_0}$. Then the estimated parameters $\hat{\mathbf{x}}$ of a least-squares adjustment (Gauß-Markov model) are given by the following equation:

$$\hat{\mathbf{x}} = \boldsymbol{f}(\mathbf{l}, \mathbf{x}_0) = \mathbf{x}_0 + (\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A})^{+} \mathbf{A}^{\mathrm{T}}\mathbf{P}(\mathbf{l} - \mathbf{a}_0), \tag{12}$$

with the $n \times u$ column regular design matrix $\mathbf{A}$, the $n \times 1$ vector of approximate values $\mathbf{x}_0$ of the parameters $\mathbf{x}$, the $n \times n$ regular weight matrix $\mathbf{P} = \mathbf{Q}_{yy}^{-1}$. The number of observations is n and the number of parameters is u. In geodetic networks the normal equations matrix $\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}$ can be rank-deficient due to an incomplete definition of the coordinate frame through the configuration. If for example such a network is composed of distance observations only, it is not possible to estimate coordinates which are required in practice. This problem can be overcome if the pseudo-inverse matrix $(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A})^{+}$ is used; see, e. g., (Koch, 1999) which is a standard reference in geodetic literature on parameter estimation (and hypotheses tests). Finally, the imprecise vector of estimated parameters $\tilde{\hat{\mathbf{x}}}$ is constructed, based on a sufficient number of α *- cuts* :

$$\tilde{\hat{\mathbf{x}}}_{\alpha,\min} = \mathbf{x}_0 + (\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A})^{+} \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{y} - |\mathbf{F}|\left(\tilde{\mathbf{p}}_{\alpha,r}\right), \tag{13\_a}$$

$$\tilde{\hat{\mathbf{x}}}_{\alpha,\max} = \mathbf{x}_0 + (\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A})^{+} \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{y} + |\mathbf{F}|\left(\tilde{\mathbf{p}}_{\alpha,r}\right), \tag{13\_b}$$

$$m_{\tilde{\mathbf{x}}}(x) = \int_0^1 m_{\tilde{\mathbf{x}}_\alpha}(x)d\alpha \qquad \text{and} \qquad m_{\tilde{\mathbf{x}}_\alpha} = \left[\tilde{\hat{\mathbf{x}}}_{\alpha,\min}, \tilde{\hat{\mathbf{x}}}_{\alpha,\max}\right], \tag{13\_c}$$

with the matrix of partial derivatives $\mathbf{F} = \frac{\partial \mathbf{x}}{\partial \mathbf{p}}$ and $|\bullet|$ denoting the element-by-element absolute value of the matrix.

2.2.2. *A linear hypothesis for the standard model in parameter estimation*

The standard model in parameter estimation is given by

$$\mathrm{E}(\mathbf{y}) = \mathbf{Ax}, \tag{14}$$

where the expected value of the reduced observations $\mathrm{E}(\mathbf{y})$ equals $\mathbf{Ax}$. The null hypothesis of a linear hypothesis is then introduced as:

$$H_0: \quad \mathbf{Cx} = \mathbf{w}, \tag{15\_a}$$

provided that $\mathbf{Cx}$ must be a testable hypothesis, cf. (Koch, 1999) for details concerning the matrix $\mathbf{C}$ and the vector $\mathbf{w}$. The null hypothesis must be compared with the alternative hypothesis

$$H_A: \quad \mathbf{Cx} = \overline{\mathbf{w}} \neq \mathbf{w}. \tag{15\_b}$$

This leads after a few calculation steps to a quadratic form:

$$T = (\mathbf{C}\hat{\mathbf{x}} - \mathbf{w})^T \left[\mathbf{C}\left(\mathbf{A}^{\mathrm{T}}\mathbf{PA}\right)^+ \mathbf{C}^{\mathrm{T}}\right]^+ (\mathbf{C}\hat{\mathbf{x}} - \mathbf{w}) \sim \chi^2(h,0) \text{ under } \mathrm{H}_0, \tag{16}$$

that follows under the null hypothesis a central chi-square distribution ($\lambda = 0$) with $h = \mathrm{rank}\left[\mathbf{C}\left(\mathbf{A}^{\mathrm{T}}\mathbf{PA}\right)^+ \mathbf{C}^{\mathrm{T}}\right]$ degrees of freedom. In order to avoid overestimation in imprecise hypothesis testing, the general form of a linear hypothesis has to be converted to a quadratic form of imprecise influence parameters $\tilde{\mathbf{p}}$; it is obtained as:

$$\tilde{T}_{\alpha,\min} = \min\left(\begin{bmatrix}\Delta\mathbf{p}\\ \mathbf{y}_m\\ \mathbf{w}\end{bmatrix}^T \begin{bmatrix}\mathbf{F}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\mathbf{DKF} & \mathbf{F}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\mathbf{DK} & -\mathbf{F}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\mathbf{D}\\ \mathbf{K}^{\mathrm{T}}\mathbf{DKF} & \mathbf{K}^{\mathrm{T}}\mathbf{DK} & -\mathbf{K}^{\mathrm{T}}\mathbf{D}\\ -\mathbf{DKF} & -\mathbf{DK} & \mathbf{D}\end{bmatrix} \begin{bmatrix}\Delta\mathbf{p}\\ \mathbf{y}_m\\ \mathbf{w}\end{bmatrix}\right), \tag{17\_a}$$

$$\tilde{T}_{\alpha,\max} = \max\left(\begin{bmatrix}\Delta\mathbf{p}\\ \mathbf{y}_m\\ \mathbf{w}\end{bmatrix}^T \begin{bmatrix}\mathbf{F}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\mathbf{DKF} & \mathbf{F}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\mathbf{DK} & -\mathbf{F}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\mathbf{D}\\ \mathbf{K}^{\mathrm{T}}\mathbf{DKF} & \mathbf{K}^{\mathrm{T}}\mathbf{DK} & -\mathbf{K}^{\mathrm{T}}\mathbf{D}\\ -\mathbf{DKF} & -\mathbf{DK} & \mathbf{D}\end{bmatrix} \begin{bmatrix}\Delta\mathbf{p}\\ \mathbf{y}_m\\ \mathbf{w}\end{bmatrix}\right), \tag{17\_b}$$

$$m_{\tilde{T}}(x) = \int_0^1 m_{\tilde{T}_\alpha}(x)d\alpha \qquad \text{and} \qquad m_{\tilde{T}_\alpha} = \left[\tilde{T}_{\alpha,\min}, \tilde{T}_{\alpha,\max}\right]. \tag{17\_c}$$

with $\mathbf{\Delta p} \in \tilde{\mathbf{p}}_{\alpha} = \left[\tilde{\mathbf{p}}_{\alpha,\min} - \mathbf{p}_{\mathrm{m}} \;,\; \tilde{\mathbf{p}}_{\alpha,\max} - \mathbf{p}_{\mathrm{m}}\right]$, $\mathbf{K} = \mathbf{C}^{\mathrm{T}}\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\right)^{+}\mathbf{A}^{\mathrm{T}}\mathbf{P}$, $\mathbf{D} = \left[\mathbf{C}\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\right)^{+}\mathbf{C}^{\mathrm{T}}\right]^{+}$ and $\mathbf{y}_{m}$ the midpoint of the reduced observations.

2.2.3. *A linear hypothesis for an extended model in parameter estimation*

The presented strategy from Section 2.2.2 has some shortcomings concerning the computational complexity. If additional parameters $\mathbf{z}$, e. g., in model selection and outlier detection shall be tested in the given enviroment, the model from Equation (15) has to be reformulated and must be fully analyzed (including the inversion of the normal equations). This problem can be overcome by an extended model in parameter estimation:

$$\mathrm{E}(\mathbf{y}) = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{z}\,. \tag{18}$$

The linear hypothesis may then be given by:

$$H_0:\quad \mathbf{C}\begin{bmatrix}\mathbf{x}\\ \mathbf{z}\end{bmatrix} = \mathbf{w} \qquad \text{versus} \qquad H_A:\quad \mathbf{C}\begin{bmatrix}\mathbf{x}\\ \mathbf{z}\end{bmatrix} = \overline{\mathbf{w}} \neq \mathbf{w}\,. \tag{19}$$

Starting with the extended normal equations (Koch, 1999)

$$\begin{bmatrix}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A} & \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{B}\\ \mathbf{B}^{\mathrm{T}}\mathbf{P}\mathbf{A} & \mathbf{B}^{\mathrm{T}}\mathbf{P}\mathbf{B}\end{bmatrix}\begin{bmatrix}\hat{\mathbf{x}}\\ \hat{\mathbf{z}}\end{bmatrix} = \begin{bmatrix}\mathbf{A}^{\mathrm{T}}\mathbf{P}(\mathbf{l}-\mathbf{a}_0)\\ \mathbf{B}^{\mathrm{T}}\mathbf{P}(\mathbf{l}-\mathbf{a}_0)\end{bmatrix}, \tag{20}$$

this procedure leads after a few calculation steps to a modified quadratic form:

$$T = (\mathbf{C}\begin{bmatrix}\hat{\mathbf{x}}\\ \hat{\mathbf{z}}\end{bmatrix} - \mathbf{w})^{T}\left[\mathbf{C}\begin{bmatrix}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A} & \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{B}\\ \mathbf{B}^{\mathrm{T}}\mathbf{P}\mathbf{A} & \mathbf{B}^{\mathrm{T}}\mathbf{P}\mathbf{B}\end{bmatrix}^{+}\mathbf{C}^{\mathrm{T}}\right]^{+}(\mathbf{C}\begin{bmatrix}\hat{\mathbf{x}}\\ \hat{\mathbf{z}}\end{bmatrix} - \mathbf{w}) \sim \chi^2(j,0) \;\text{ under } \mathrm{H}_0\,. \tag{21}$$

This quadratic form follows under the null hypothesis a central chi-square distribution ($\lambda = 0$) with $j = \operatorname{rank}\left[\mathbf{C}\begin{bmatrix}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A} & \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{B}\\ \mathbf{B}^{\mathrm{T}}\mathbf{P}\mathbf{A} & \mathbf{B}^{\mathrm{T}}\mathbf{P}\mathbf{B}\end{bmatrix}^{+}\mathbf{C}^{\mathrm{T}}\right]$ degrees of freedom. If only the additional parameters $\mathbf{z}$ have to be tested, the null hypothesis H_0 can be reformulated as follows:

$$H_0:\quad \mathbf{C}\begin{bmatrix}\mathbf{x}\\ \mathbf{z}\end{bmatrix} = \begin{bmatrix}\mathbf{C}_1 & \mathbf{C}_2\end{bmatrix}\begin{bmatrix}\mathbf{x}\\ \mathbf{z}\end{bmatrix} = \begin{bmatrix}\mathbf{0} & \mathbf{C}_2\end{bmatrix}\begin{bmatrix}\mathbf{x}\\ \mathbf{z}\end{bmatrix} = \mathbf{w} = \begin{bmatrix}\mathbf{0}\\ \mathbf{w}_2\end{bmatrix}, \tag{22}$$

and the quadratic form is now easy to handle (Koch, 1999):

$$T = (\mathbf{C}_2\hat{\mathbf{z}} - \mathbf{w}_2)^{\mathrm{T}}\left[\mathbf{C}_2\left(\mathbf{B}^{\mathrm{T}}\left(\mathbf{P} - \mathbf{P}\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A})^{+}\mathbf{A}^{\mathrm{T}}\mathbf{P}\right)\mathbf{B}\right)^{-1}\mathbf{C}_2^{\mathrm{T}}\right]^{+}(\mathbf{C}_2\hat{\mathbf{z}} - \mathbf{w}_2) \sim \chi^2(j,0) \;\text{ under } \mathrm{H}_0\,. \tag{23}$$

According to Section 2.2.2 this quadratic form from Equation (23) has to be converted to a quadratic form of imprecise influence parameters $\tilde{\mathbf{p}}$. With $\hat{\mathbf{z}} = \left(\mathbf{B}^{\mathrm{T}}\mathbf{P}\mathbf{Q}_{\hat{\mathbf{v}}\hat{\mathbf{v}}}\mathbf{P}\mathbf{B}\right)^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{P}\mathbf{Q}_{\hat{\mathbf{v}}\hat{\mathbf{v}}}\mathbf{P}(\mathbf{l}-\mathbf{a}_0)$

and $\mathbf{Q}_{\hat{v}\hat{v}} = \left(\mathbf{P}^{-1} - \mathbf{A}(\mathbf{A}^{T}\mathbf{P}\mathbf{A})^{+}\mathbf{A}^{T}\right)$, $\mathbf{J} = \left(\mathbf{B}^{T}\mathbf{P}\mathbf{Q}_{\hat{v}\hat{v}}\mathbf{P}\mathbf{B}\right)^{-1}\mathbf{C}_{2}\mathbf{B}^{T}\mathbf{P}\mathbf{Q}_{\hat{v}\hat{v}}\mathbf{P}$ and $\mathbf{M} = \left[\mathbf{C}_{2}\left(\mathbf{B}^{T}\mathbf{P}\mathbf{Q}_{\hat{v}\hat{v}}\mathbf{P}\mathbf{B}\right)^{-1}\mathbf{C}_{2}^{T}\right]^{+}$ we obtain:

$$\tilde{T}_{\alpha,\min} = \min\left(\begin{bmatrix}\Delta\mathbf{p}\\ \mathbf{y}_{m}\\ \mathbf{w}_{2}\end{bmatrix}^{T}\begin{bmatrix}\mathbf{F}^{T}\mathbf{J}^{T}\mathbf{MJF} & \mathbf{F}^{T}\mathbf{J}^{T}\mathbf{MJ} & -\mathbf{F}^{T}\mathbf{J}^{T}\mathbf{M}\\ \mathbf{J}^{T}\mathbf{MJF} & \mathbf{J}^{T}\mathbf{MJ} & -\mathbf{J}^{T}\mathbf{M}\\ -\mathbf{MJF} & -\mathbf{MJ} & \mathbf{M}\end{bmatrix}\begin{bmatrix}\Delta\mathbf{p}\\ \mathbf{y}_{m}\\ \mathbf{w}_{2}\end{bmatrix}\right), \tag{24_a}$$

$$\tilde{T}_{\alpha,\max} = \max\left(\begin{bmatrix}\Delta\mathbf{p}\\ \mathbf{y}_{m}\\ \mathbf{w}_{2}\end{bmatrix}^{T}\begin{bmatrix}\mathbf{F}^{T}\mathbf{J}^{T}\mathbf{MJF} & \mathbf{F}^{T}\mathbf{J}^{T}\mathbf{MJ} & -\mathbf{F}^{T}\mathbf{J}^{T}\mathbf{M}\\ \mathbf{J}^{T}\mathbf{MJF} & \mathbf{J}^{T}\mathbf{MJ} & -\mathbf{J}^{T}\mathbf{M}\\ -\mathbf{MJF} & -\mathbf{MJ} & \mathbf{M}\end{bmatrix}\begin{bmatrix}\Delta\mathbf{p}\\ \mathbf{y}_{m}\\ \mathbf{w}_{2}\end{bmatrix}\right), \tag{24_b}$$

$$m_{\tilde{T}}(x) = \int_{0}^{1} m_{\tilde{T}_{\alpha}}(x)\,d\alpha \quad \text{and} \quad m_{\tilde{T}_{\alpha}} = \left[\tilde{T}_{\alpha,\min}, \tilde{T}_{\alpha,\max}\right]. \tag{24_c}$$

The quadratic form from the Equations (23) and (24) is computable from the residuals without a new parameter estimation. Therefore the computational complexity is significant reduced.

2.2.4. *Final Test decision based on the card criterion*

The fuzzy evaluation of the quadratic forms from the Equations (17) and (24) is based on Zadeh's extension principle (Zadeh 1965), which can be equivalently replaced by the min-max operator of an optimization algorithm, cf. (Dubois and Prade, 1980, p. 37) for the theoretical concept and (Möller and Beer, 2004) for applications in civil engineering. The optimization problem can be solved, e. g., with a standard Newton algorithm, cf. (Coleman and Li, 1996). Figure 3 shows a constructed test value $\tilde{T}$ and the comparison of the imprecise test value with the imprecise regions of acceptance $\tilde{A}$ and rejection $\tilde{R}$ (Neumann et al., 2006).

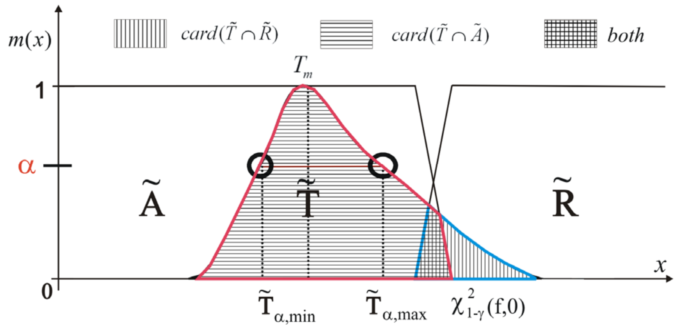


Figure 3. Comparison of the constructed test value $\tilde{T}$ with the regions of acceptance $\tilde{A}$ and rejection $\tilde{R}$

Whereas the influence of imprecision on the test decision for a smaller number of observations is unimportant, it gets more important for a larger number of observations. This is in full accordance to the theoretical concept, because the goodness of fit for the stochastic uncertainty of the parameters increases with the number of observations.
The final test decision is based on the set-theoretical comparison of the imprecise test value (constructed using an α-*cut* optimization algorithm) with the region of acceptance $\tilde{A}$ and the region of rejection $\tilde{R}$ (see Fig. 3), cf. (Kutterer 2004) and (Neumann et al. 2006) for detailed explanations. The hypotheses are defined by

$$T_m \sim \chi^2(k,\lambda)\ ;\quad \lambda\begin{cases} =0 & \mid \mathrm{H}_0 \text{ the null hypothesis,} \\ \neq 0 & \mid \mathrm{H}_\mathrm{A} \text{ the alternative hypothesis,} \end{cases} \tag{25}$$

with the non-centrality parameter λ. The midpoint of the test value follows under the null hypothesis a central chi-square distribution with $k \in \{h\ ,j\}$ degrees of freedom. The regions of acceptance $\tilde{A}$ and rejection $\tilde{R} = \tilde{A}^C$ are defined as fuzzy intervals. The degree of the rejectability $\rho_R(\tilde{T})$ of the null hypothesis H_0 under the condition of $\tilde{T}$ is computed based on the degree of agreement of the test value with the region of rejection $\gamma_{\tilde{R}}(\tilde{T})$ and the degree of disagreement of the test value with the region of acceptance $\delta_{\tilde{A}}(\tilde{T})$. We use the card criterion, because it allows a more suitable description of the degree of agreement between two fuzzy intervals. This leads to the equations given below (see also Fig. 3):

$$\gamma_{\tilde{R}}(\tilde{T}) = \frac{card\left(\tilde{T}\cap\tilde{R}\right)}{card\left(\tilde{T}\right)} \quad \text{and} \quad \delta_{\tilde{A}}(\tilde{T}) = 1 - \frac{card\left(\tilde{T}\cap\tilde{A}\right)}{card\left(\tilde{T}\right)} \tag{26\_a}$$

$$\rho_{\tilde{R}}(\tilde{T}) = \min\left(\gamma_{\tilde{R}}(\tilde{T}), \delta_{\tilde{A}}(\tilde{T})\right) \tag{26\_b}$$

For the final test decision, the degree of rejectability $\rho_{\tilde{R}}(\tilde{T})$ of the null hypothesis has to be compared with a suitable critical value $\rho_{crit} \in [0,1]$:

$$\rho_{\tilde{R}}\left(\tilde{T}\right)\begin{Bmatrix}\leq \\ >\end{Bmatrix}\rho_{\mathrm{crit}} \in \left[0,1\right] \quad \Rightarrow \quad \begin{cases}\text{Do not reject } \mathrm{H}_0 \\ \text{Reject } \mathrm{H}_0\end{cases} \tag{27}$$

The test is only rejected, if the test value agrees with the region of rejection and disagrees with the region of acceptance. This is in full accordance with the theoretical expectations, where observation imprecision is an additive term of uncertainty during the measurement process. The choice of ρ_{crit} depends on the particular application and must be based on expert knowledge. For outlier detection we propose to choose $\rho_{crit} \to 1$ and for safety-relevant measures $\rho_{crit} \to 0$.

3. Probability of type I errors in imprecise hypothesis testing

In this subsection we focus on the relation of imprecision and the probability of a type I error. The probability γ_{impr} of a type I error in the imprecise case is defined by:

$$\gamma_{impr} = P\left(\rho_{\tilde{R}}\left(\tilde{T}\right) > \rho_{crit} \left| H_0 \right.\right). \tag{28}$$

The index „impr" denotes the case of imprecision. Equation (28) can be reformulated as follows

$$\gamma_{impr} = P\left(\mathrm{f}\,(T_m) > \rho_{crit} \left| H_0 \right.\right), \tag{29}$$

with the degree of rejectability $\rho_{\tilde{R}}(\tilde{T})$ of the null hypothesis under the condition of $\tilde{T}$ as a function f of the midpoint T_m of the imprecise test value $\tilde{T}$.

$$\gamma_{\text{impr}} = \mathrm{P}\left(\mathrm{f}^{-1}(\rho_{\text{crit}}) > \mathrm{T}_{\text{m}} \left| \mathrm{H}_0 \right.\right). \tag{30}$$

This leads with respect to Equation (30) after a few calculation steps to the quantile value $\chi^2_{1-\gamma_{\text{impr}}}$ of the chi-square distribution (k degrees of freedom) in the imprecise case

$$\chi^2_{1-\gamma_{impr}}(k,0) = \mathrm{f}^{-1}(\rho_{crit}), \tag{31}$$

with f^{-1} denoting the inverse function of f . In order to illustrate the theoretical concept, an example will be shown in Section 5. See (Kutterer, 2004) for a close mathematical formulation in case of classical regions of acceptance and rejection in the one-dimensional case. Based on the quadratic form from Equation (24), the influence of imprecision on the tests decision is analyzed for different positions for the midpoint T_m of the test value (see figure 4).

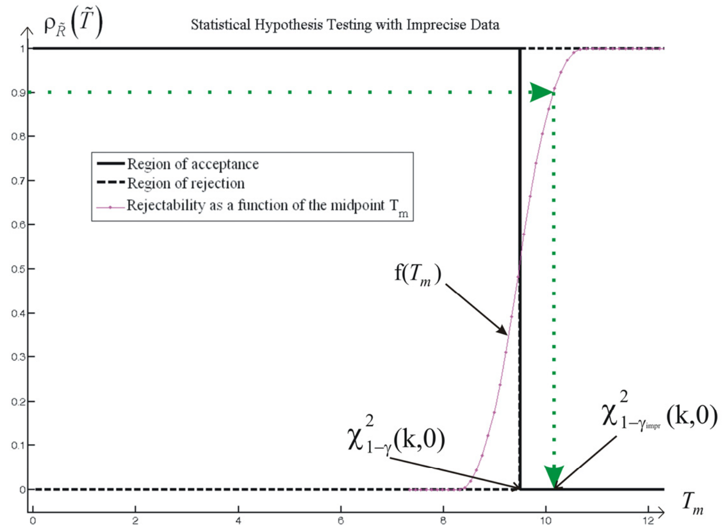


Figure 4. Calculation of the probability of a type I error in the imprecise case (for $\rho_{crit} = 0.9$)

The calculation of the probability of a type I error is then easy to handle and can be solved by the followings steps:

Step 1: Choose an adequate value for ρ_{crit} (see Section 2.2.4).

Step 2: Compute $f^{-1}(\rho_{crit}) \rightarrow \chi^2_{1-\gamma_{impr}}(k,0)$.

Step 3: Find γ_{impr} in such a way that Equation (31) is fulfilled within a negligible threshold.

4. Probability of type II errors in imprecise hypothesis testing

The probability β_{impr} of a type II error in the imprecise case can be derived by

$$\beta_{impr} = \mathrm{P}\left(\rho_{\tilde{\mathrm{R}}}\left(\tilde{T}\right) \leq \rho_{\mathrm{crit}} \left| \mathrm{H_A} \right.\right). \tag{32}$$

According to the probability of a type I error, Equation (32) can be reformulated as follows

$$\beta_{impr} = P\left(\mathrm{f}(T_m) \leq \rho_{crit} \left| H_A \right.\right), \tag{33}$$

with the degree of rejectability $\rho_{\tilde{R}}(\tilde{T})$ of the null hypothesis under the condition of $\tilde{T}$ as a function f of the midpoint T_m of the imprecise test value $\tilde{T}$. In order to analyze Equation (33), either the non-centrality parameter λ_{impr} in the imprecise case or the probability β_{impr} of a type II error in the imprecise case must be set in advance. This leads after a few calculation steps to the comparison of two chi-square distributions (with k degrees of freedom). The first central chi-square distribution is related to the probability of a type I error in the imprecise case and the second one (with the non-centrality parameter λ_{impr} in the imprecise case) is related to the probability of a type II error.

$$\chi^2_{1-\gamma_{impr}}(k,0) = \chi^2_{\beta_{impr}}(k,\lambda_{impr}). \tag{34}$$

The calculation of the probability of a type II error and of the non-centrality parameter in the imprecise case can be seen as the following search problem (see figure 5a and 5b):

1. Calculation of the type II error in imprecise hypothesis testing:

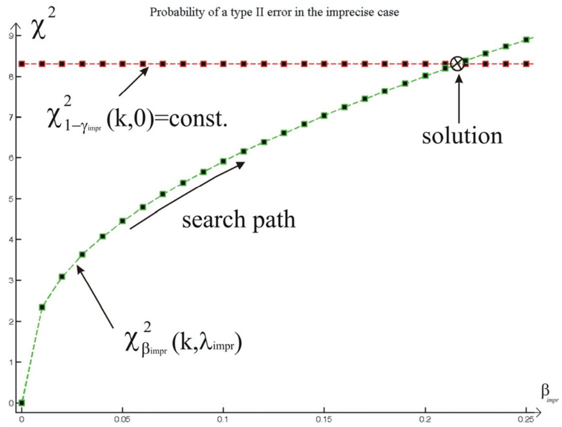
   Step 1: Compute the probability of a type I error in the imprecise case (see Section 3).
   Step 2: Choose an adequate value for λ_{impr}.

   Step 3: Find β_{impr} in such a way that Equation (34) is fulfilled within a negligible threshold.

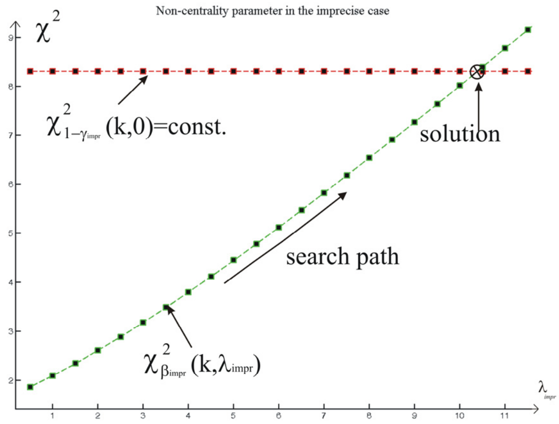
2. Calculation of the non-centrality parameter in imprecise hypothesis testing:

   Step 1: Compute the probability of a type I error in the imprecise case (see Section 3).
   Step 2: Choose an adequate value for β_{impr}.

   Step 3: Find λ_{impr} in such a way that Equation (34) is fulfilled within a negligible threshold.



a. Calculation of a type II error ($\lambda_{impr} = const$)



b. Calculation of the non-centrality parameter ($\beta_{impr} = const$)

Figure 5. Calculation of the probability of a type II error (a) and the non-centrality parameter (b) in the imprecise case

5. Example for the parameterization of a geodetic monitoring network

In order to illustrate the theoretical concept, three exemplary applications in the parameterization of a geodetic monitoring network are presented. The aim of the geodetic monitoring network is to detect significant changes of a lock due to changing water levels inside the lock. Figure 6 shows the lock and the geodetic monitoring network, which consist of four object points on top of the lock (101-104) and eight control points around the lock; see (Neumann et al., 2006) for a detailed description about the geodetic monitoring network.

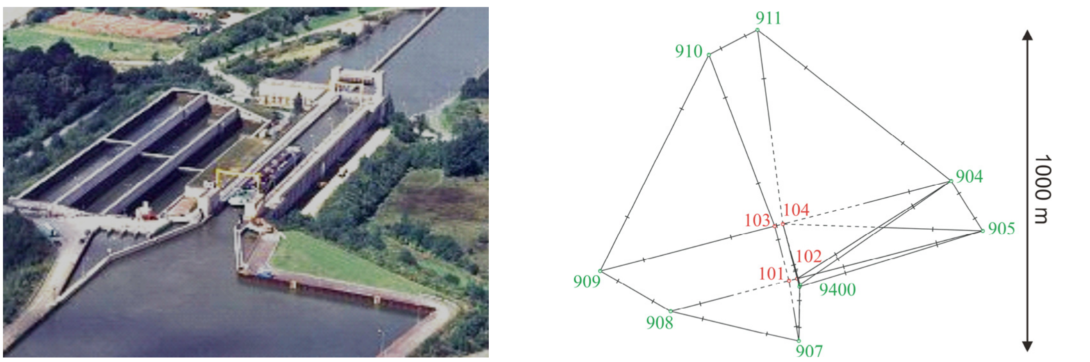


Figure 6 – The lock and the geodetic monitoring network

The coordinates of the object points are estimated within a least-squares adjustment. Therefore special geodetic measurements like horizontal directions (a), zenith angles (b) and distances (c) were carried out between the object and control points. The measurements are affected by different types of uncertainty (see Table 1 and Section 2.1). The non-stochastic uncertainties are analyzed within a sensitivity analysis (see Table 1).

Influence factors **p**	*Interval radii* ($\alpha = 0$) *(imprecision)*	*Affected measurements*
Temperature	1.0 °C	(c)
Pressure	1.0 hPa	(c)
Visual axis error	0.1 mgon	(a)
Collimation error	0.1 mgon	(a)
Vertical axis error	0.2 mgon	(a) and (b)

a. Main influence factors for the observations

Observations	*Interval radii* ($\alpha = 0$) *(imprecision)*	*Standard deviation*
Horizontal direction	0.1 mgon	0.5 mgon
Zenith angle	0.5 mgon	1.5 mgon
Distance	0.5 mm	3 mm

b. Uncertainties of the observations

Table 1. Influence factors and uncertainties of the observations

First we focus on a single and multiple outlier test. Then a congruence test is evaluated in terms of imprecision. For a straightforward comparison to the pure stochastic case, the region of acceptance is given by a classical interval with a significance level of $\gamma = 5\%$. All computations are based on 11 different $\alpha - cuts$.

5.1. Examples in outlier detection

5.1.1. *Testing procedure for a single measurement*

The first example shows an outlier test for a distance measurement. The construction of the test value $\tilde{T}$ is based on the imprecise evaluation of the quadratic form in Equation (24) with an $\alpha - cut$ optimization method.

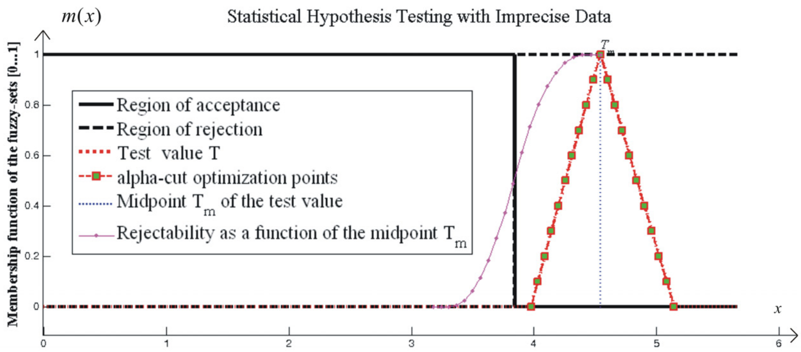


Figure 7 – The degree of rejectability $\rho_{\tilde{R}}(\tilde{T})$ of a single outlier test as a function of the midpoint T_m of the test value $\tilde{T}$

Figure 7 shows the degree of rejectability $\rho_{\tilde{R}}(\tilde{T})$ of the null hypothesis H_0 under the condition of $\tilde{T}$ as a function f of the midpoint T_m of the imprecise test value $\tilde{T}$. Obviously, in this example the observation imprecision is small in comparison to the stochastic uncertainty. For this reason, the test value is tight and close to symmetric.

The probability of a type I error in the imprecise case γ_{impr} is strongly related to the choice of the critical value ρ_{crit} for the test decision, see Equation (31). Figure 8 shows the probability of a type I error in relationship to the choice of ρ_{crit}.

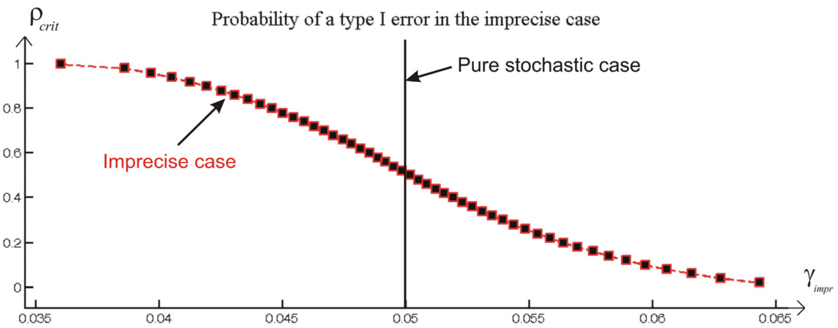


Figure 8 – Probability of a type I error in the imprecise case for a single outlier test (depending on the choice of ρ_{crit})

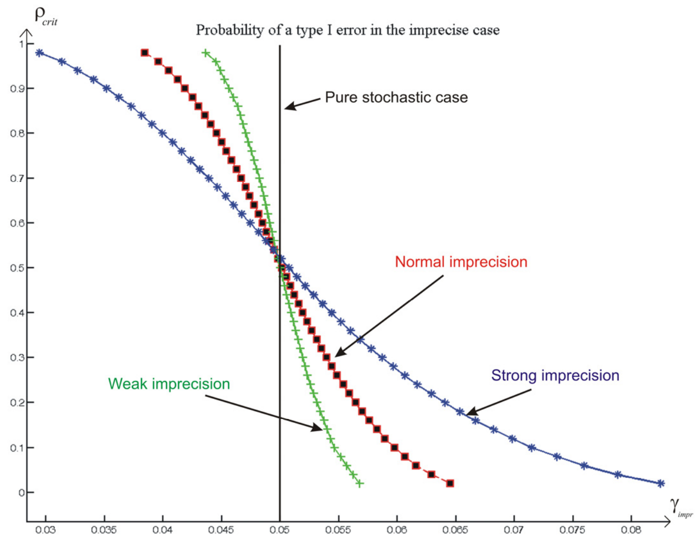


Figure 9 – Variation of the probability of a type I error in the imprecise case (depending on the choice of ρ_{crit} and the order of magnitude of imprecision)

The choice of ρ_{crit} depends on the particular application and must be based on expert knowledge. For outlier detection we propose to choose $\rho_{crit} \to 1$ and for safety-relevant measures $\rho_{crit} \to 0$. The variation of a type I error in the imprecise case γ_{impr} depends also on the order of magnitude of imprecision. If imprecision is more important in comparison to the stochastic uncertainty, the variation of a type I error in the imprecise case increases. Figure 9 shows an example with strong imprecision (twice of the imprecision of Table 1), normal imprecision and small imprecision (half of the imprecision of Table 1):

5.1.2. *Testing procedure for multiple measurements*

The second example shows a multiple outlier test due to an assumed centering error of the instrument, while measuring a set of distances at station 102. The construction of the test value $\tilde{T}$ is based on the imprecise evaluation of the quadratic form in Equation (24) with the $\alpha-cut$ optimization method. In this example, the number of tested observations is four ($j = 4$). Figure 10 shows the probability of a type I error in the imprecise case γ_{impr}.

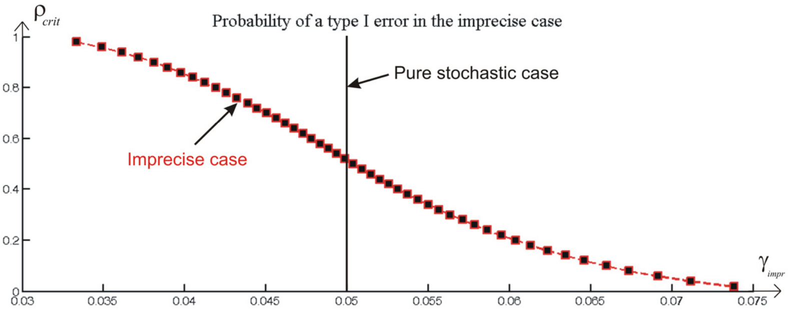


Figure 10 – Probability of a type I error in the imprecise case for a multiple outlier test (depending on the choice of ρ_{crit})

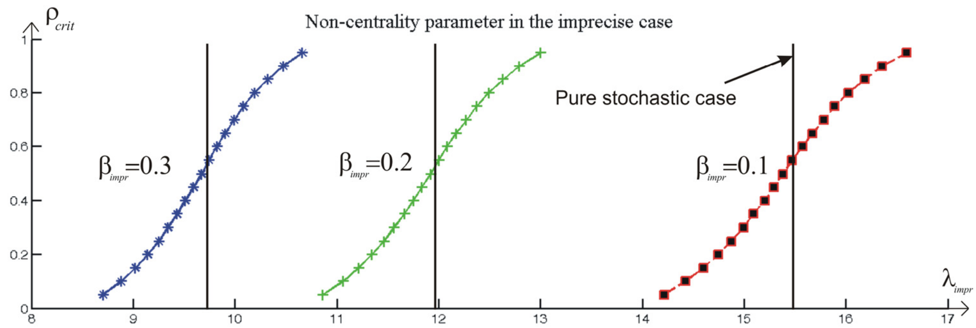


Figure 11 – The non-centrality parameter in the imprecise case (depending on the choice of ρ_{crit} and β_{impr})

To study the influence of imprecision on the probability β_{impr} of a type II error and the non-centrality parameter λ_{impr} in the imprecise case, it is more meaningful to hold the probability of a type II error constant. We focus on three standard cases with $\beta_{impr} = 0.1$, $\beta_{impr} = 0.2$ and $\beta_{impr} = 0.3$. The non-centrality parameter is obtained by the search problem described in Section 4 (see Figure 11). For $\rho_{crit} > 0.5$ the non-centrality parameter in the imprecise case is greater than in the precise case. This leads to a reduced sensitivity regarding the rejection of the null hypothesis.

5.2. Example for a Congruence Test (Epoch Comparison)

The third example demonstrates an epoch comparison between the years 1999 and 2004. Both epochs are estimated within a partially constrained trace minimization with respect to the same six network points. The construction of the test value $\tilde{T}$ is based on the imprecise evaluation of the quadratic form from Equation (17). Figure 12 shows the numerical test situation with the probability of a type I error in the imprecise case and Table 2 some specifications about the two epochs and the geodetic monitoring network. Please note that the configurations in both epochs are different from each other.

Specification	Epoch 1999	Epoch 2004
Observations n	317	144
Parameters u	60	39

Table 2. Specifications about the geodetic monitoring network in the epochs 1999 and 2004

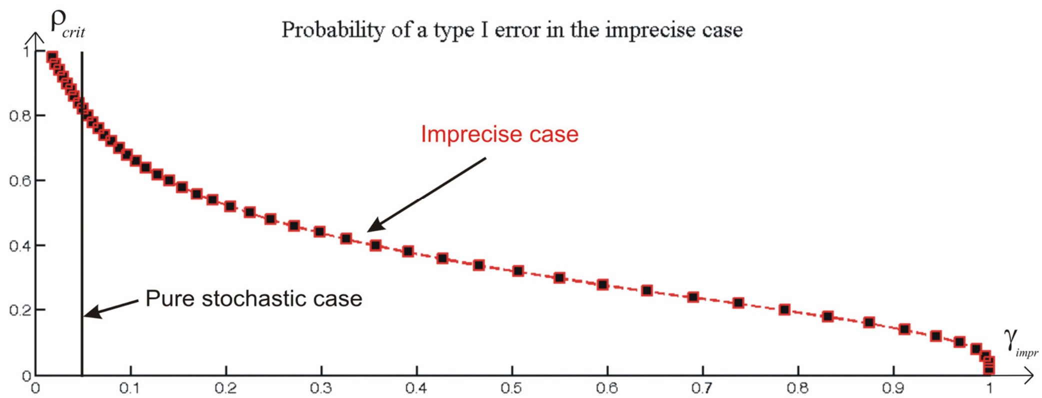


Figure 12 – Probability of a type I error in the imprecise case for a congruence test (depending on the choice of ρ_{crit})

The significant imprecision of the test value in this example is caused by the strong effects of remaining systematics in epoch comparison of a geodetic monitoring network. The influence of imprecision on the given test situation depends also on the geometric configuration of the geodetic monitoring network. Whereas a weak configuration leads to a wider expansion of the test value, a strong configuration decreases the influence of imprecision in the test situation. The strong imprecision leads to a wide variation of the probability of a type I error in the imprecise case. In case of $\rho_{crit} \to 0$ the null hypothesis will be rejected in any rate. For this reason, the probability of a type I error in the imprecise case (for $\rho_{crit} \to 0$) is equal to one.

6. Conclusions

In this paper, we show a joint treatment of stochastic and interval/fuzzy uncertainty (*imprecision*) in hypothesis testing in parameter estimation. Imprecision is considered as an additive term of uncertainty what leads to a more reluctant rejection of the null hypothesis in case of outlier detection and to an earlier rejection of the null hypothesis in case of safety-relevant applications. If imprecision is absent, the results of the pure stochastic case are obtained. We focus on the probability of a type I and type II error and the non-centrality parameter in the imprecise case. In case of outlier detection the probability of a type I error in the imprecise case is lower than in the pure stochastic case and the non-centrality parameter in the imprecise case is greater than in the pure stochastic case. In order to detect the same changes than in the pure stochastic case, e. g., in a risk analysis, more precise measurements have to be carried out.

However, the quantification of the uncertainty budget of empirical measurements is often too optimistic due to, e.g., the ignorance of non-stochastic errors in the analysis process (Ferson et al., 2007). For this reason the above mentioned results in this paper are in our opinion closer to the situation in real-world applications. In addition, the well known sensitivity analysis in parameter estimation can now generally be treated in terms of imprecise data to decide about a suitable model for the collected data.

Further work has to deal with a significant reduction of the computational complexity of the numerical solutions. In addition, it seems to be very promising, that for special types of reference functions analytic solutions for type I and type II errors in the imprecise case can be found.

Acknowledgements

The presented paper shows results and new ideas developed during the research project KU 1250/4-1 "Geodätische Deformationsanalysen unter Berücksichtigung von Beobachtungsimpräzision und Objektunschärfe", which is funded by the German Research Foundation (DFG). This is gratefully acknowledged by the authors.

References

Bandemer, H. and Näther, W.: *Fuzzy Data Analysis*. Kluwer, 1992.

Dubois, D.J. and Prade, H.M.: *Fuzzy Sets and Systems: theory and applications*, Academic Press, New York, 1980.

Ferson, S.; Kreinovich, V.; Hajagos, J.; Oberkampf, W. and Ginzburg, L.: *Experimental Uncertainty Estimation and Statistics for Data Having Interval Uncertainty*. Sandia National Laboratories, SAND2007-0939, 2007.

Koch, K. R.: *Parameter Estimation and Hypothesis Testing in Linear Models*. Springer, Berlin New York, 1999.

Kutterer, H.: *Statistical hypothesis tests in case of imprecise data*. Proc. of the 5. Hotine-Marussi-Symposium, International Association of Geodesy Symposia, Springer, Berlin and New York, 2004, pp. 49-56.

Möller, B. and Beer, M.: *Fuzzy Randomness -Uncertainty in Civil Engineering and Computational Mechanics-*. Springer, Berlin and New York, 2004.

Neumann, I.; Kutterer, H. and Schön, S.: *Outlier detection in geodetic applications with respect to observation imprecision*. Proceedings of the NSF Workshop on Reliable Engineering Computing - Modeling Errors and Uncertainty in Engineering Computations -, Savannah, Georgia, USA, 2006, pp. 75-90.

Zadeh, L.A.: *Fuzzy sets.* In: Information Control, Vol. 8, 1965, pp. 338-353.