Technische Universiteit Delft

Estimation of horizontal deformation rates based on tachymetric measurements

Processing measurements collected yearly in the geothermal area Bjarnarflag in North-East Iceland from 2015 to 2018

Veronica Tollenaar

TU Delft

Estimation of horizontal deformation rates based on tachymetric measurements

Processing measurements collected yearly in the geothermal area Bjarnarflag in North-East Iceland from 2015 to 2018

by

Veronica Tollenaar

as Additional Thesis
at the Delft University of Technology
July 15, 2019

Student number: 4162889
Thesis committee: Dr. ir. F. J. van Leijen, TU Delft
Dr. ir. S. Verhagen, TU Delft
Dr. I. V. Smal, TU Delft

TU Delft

Contents

Abstract

Measurements of a tachymetric network in the geothermal area of Bjarnarflag in North-East Iceland are performed during four consecutive years (2015-2018). For the yearly adjustment of the measurements an alternative iteration scheme to Baarda's 'B-method of testing' is proposed, because the level of significance of the F-test performed in the detection step is too large due to the large redundancy of the measurements. The proposed alternative method detects outliers and verifies the stochastic model simultaneously, resulting in 3 to 4% rejected measurements per year. From the yearly solutions absolute and relative horizontal velocities are estimated. The relative velocities are significant, but have large uncertainties. Adding three years of data reduces the mean standard deviation of the estimated velocities from 2.4 mm in east-direction to 0.9 mm and 1.8 mm in north-direction to 0.8 mm. The improvement of precision can be accelerated by reevaluating the stochastic input parameters, adding more GNSS measurements and reconsidering the network design. Improved horizontal relative velocities can be used to understand the horizontal deformation patterns in the area of study due to extraction of water or steam by the geothermal powerplant, due to the instability of the benchmaks or due to natural processes.

1 Introduction

Complex deformations of the earth's crust in the Krafla volcanic system in North-East Iceland are subject to multiple studies. The deformations are partly due to anthropogenic sources such as extraction of geothermal fluids for two power plants located in the area, Krafla and Bjarnarflag, operated by the national power company of Iceland, Landsvirkjun, and partly due to natural processes such as plate spreading, magma accumulation and magma chamber deflation, post-rifting adjustment, inflation of the Eistareykir-volcano, glacial isostatic adjustment, deflation along the Krafla fissure swarm and seasonal processes [4].

The two geothermal areas in the Krafla volcanic system are surveyed yearly by students of the Delft University of Technology in the educational context of an engineering fieldwork. Various techniques - GNSS, tachymetry, leveling and gravimetry - are used to collect data to measure the deformation for a client, the power company Landsvirkjun. In addition to the in situ measurements, InSAR data is also available for the students. InSAR data has been evaluated thoroughly in the paper of Drouin et al. [4], where the data is also linked to the extensive GPS campaign of the University of Iceland, Reykjavík, overlapping with the small GNSS campaign of the student fieldwork and leveling data provided by Landsvirkjun. Recent and old tachymetric measurements have not been evaluated by Drouin et al. [4] and therefore it is interesting to investigate the value of these data-sets.

The tachymetric measurements are collected in two areas, one related to the Krafla power plant, the other one related to the Bjarnarflag powerplant, resulting in two unattached networks. The measurements at the area of the Krafla power plant are performed using a portable prism held on a stick, measuring along the foundation of a pipeline system of the power plant (see Figure 1.1). This method seems to be very error-prone and therefore my focus lies on the measurements collected in the area of the Bjarnarflag power plant, where tripods are used to hold the prisms.

Figure 1.1: Portable prism used to measure the area of the Krafla power plant.

The Bjarnarflag network contains fourteen benchmarks, but the measured benchmarks are slightly different per year (eleven benchmarks in 2016 and 2018 and twelve benchmarks in 2017 and 2019). In evaluating the quality of the collected data the error sources must be considered, such as instrumental errors, off-center errors, errors caused by unequal atmospheric conditions, idealization error and human errors. The latter ones are relevant considering the retrieval circumstances, namely in the context of an educational fieldwork. There have been about 40 different inexperienced observers and the measurements have been noted down manually, so the assumption of 1 outlier per 100 measurements suggested by the Move3 Software [10, p. 50] is probably an underestimation of the number of outliers. Also the methods and instrument settings have been slightly different over the years and the documentation is fragmentary. A good statistical evaluation of the measurements is essential as well as relating the measurements to the less error-prone GNSS measurements.
The research question is:

How to process yearly repeated measurements of a tachymetric network to estimate a horizontal deformation rate?

In answering this question the aforementioned data collected in the area of Bjarnarflag in North-East Iceland has been used. The details of these measurements are described in Section 2.1. The yearly measurements need to be tested for outliers and also the stochastic model needs to be verified. The commonly used testing procedure has limitations for measurements of large tachymetric networks and therefore an alternative is proposed in Section 2.2. The alternative testing procedure has been applied to the measurements per epoch to come to yearly solutions and from these yearly solutions multiyear trends are derived and presented in Chapter 3, where also the precision and significance of the estimated relative horizontal deformation are evaluated. A conclusion on the results can be found in Chapter 4, where also recommendations for future measurements are given.

2

Method

The details of the measurements of the tachymetric network are described in Section 2.1. To come from the measurements to yearly solutions of the network, the observations need to be adjusted and simultaneously tested to detect outliers and verify the stochastic model. Because the redundancy of the measurements is high, the commonly used Detection, Identification and Adaptation-method for observation testing does not suffice and an alternative procedure is described (Section 2.2).

2.1. Description of measurements

The tachymetry network is located in the area of the Bjarnarflag geothermal powerplant. The locations of the networks' benchmarks are shown in Figure 2.1.

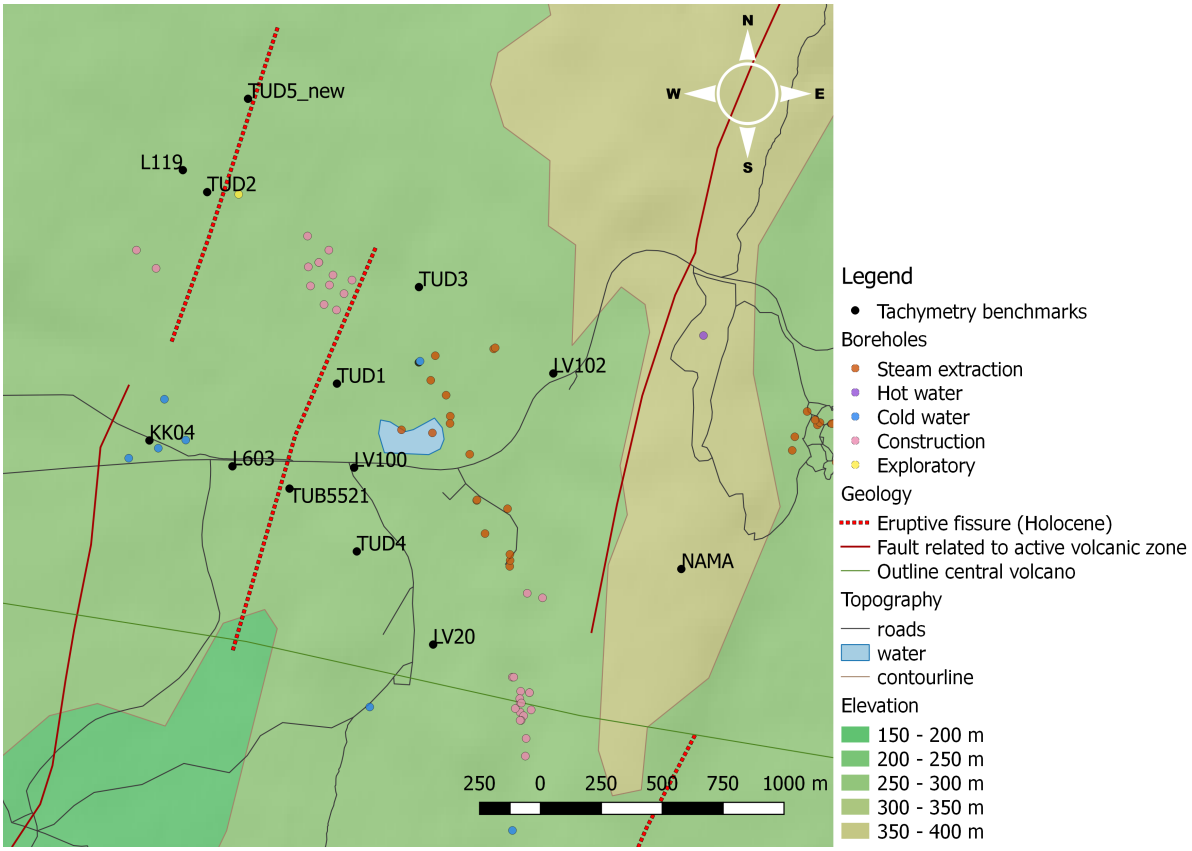


Figure 2.1: Location of the benchmarks with respect to the various boreholes and the geolocial features of the area. Data on the boreholes is extracted from the Orkustofnun Geoportal of the Icelandic National Energy Authority (`http://map.is/os/`) .

The tachymetry network is slightly different every year (see Figure 2.2) and benchmark names have not been consistent over the years (see Table 2.1). In 2015 the wrong benchmark for TUD4 has been used. This data point is used for the network adjustment, but not for estimating the multiyear trends. GNSS measurements are available for the stations L119, L603, LV102 (L102), LV20, LV100 (BF19) and NAMA (see appendix C).

name	used in	corrected to
'L102'	2017	'LV102'
'L965100'	2017	'LV100'
'L100'	2018	'LV100'
'L102'	2018	'LV102'
'TUD5'	2018	'TUD5_new'

Table 2.1: Corrected inconsistencies in benchmark naming.

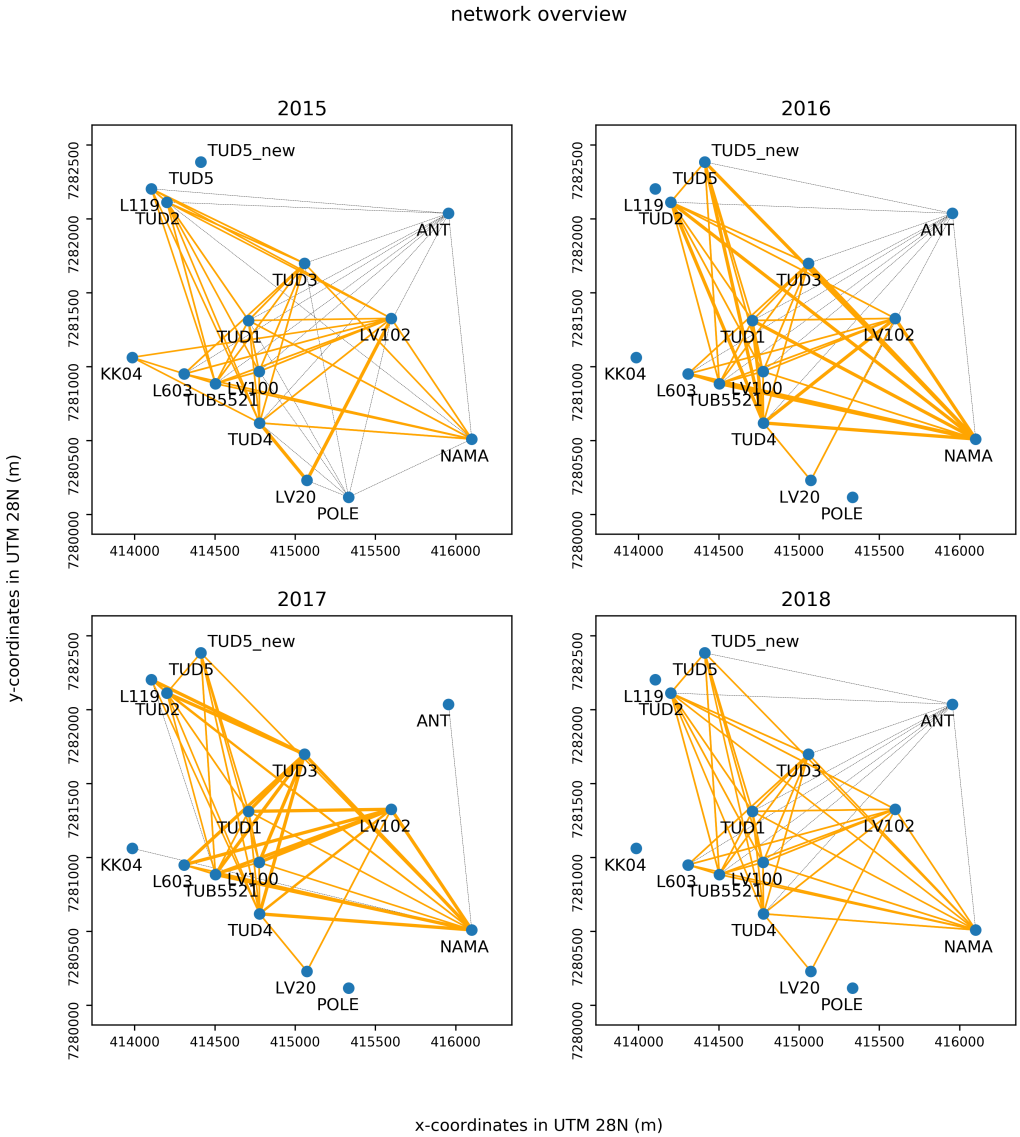


Figure 2.2: Tachymetric network in 2015, 2016, 2017 and 2018. Orange lines mean distance and directional measurements, black lines mean only directional measurements. The thickness of the line represents the number of repeated measurements.

From 2016 or 2017[1] onwards the total station *Leica 1201+* is used for the measurements, in combination with standard prisms (GPR1) and one 360°prism. Before 2016 or 2017 the total station *Leica 1101* is used for the measurements, in combination with standard prisms (GPR1).

The data has been processed using the software MOVE3, which performs a least squares adjustment on the measurements [10]. The software asks for input parameters of the geometry as well as input parameters of the stochastic model. Screenshots of these input parameters are included in appendix A, which can be used for the feature processing of measurements in the context of the educational fieldwork. Initial input parameters of the stochastic model are chosen empirically after experimenting with the data in MOVE3 (see Table 2.2).

year	std direction	std distance	std zenith angle
2015	0.005 gon	0.005 m + 2ppm	0.005 gon
2016	0.005 gon	0.005 m + 2ppm	0.005 gon
2017	0.003 gon	0.003 m + 1.5ppm	0.003 gon
2018	0.003 gon	0.003 m + 1.5ppm	0.003 gon

Table 2.2: Input parameters for the stochastic model, values for directions and zenith angles are ten times larger than as specified by the instrument manufacturer [7] [8].

2.2. Observation testing

The software MOVE3 performs an F-test (overall model test) and W-tests (data-snooping) in order to control the quality of the output. The critical value of the W-tests is calculated using a preset significance level α_0, indicating the probability of a false rejection of the null hypothesis. The different significance levels of the tests are connected to each other using the power of the test γ, indicating the probability of a correct rejection of the null hypothesis. For both tests the preset value for γ is the same. The level of significance of the F-test, α, is determined by α_0, γ and the redundancy of the network, using the B-method of testing as introduced by Baarda [2]. Connecting the levels of significance in this way results in that both the F-test and the W-test can detect an outlier with equal probability. The F-test can then be performed as overall model test indicating whether datasnooping (W-tests) are needed to detect outliers or not.

In the following paragraph (Section 2.2.1) the B-method of testing is described comprehensively, with illustrating figures using representative preset values for the level of significance for the W-test, α_0 (0.001) for the 'power' γ (0.8) and for the redundancy (296). This serves as illustration for an evaluation of the B-method of testing (Section 2.2.2), leading to a proposed alternative approach (Section 2.2.3).

2.2.1. B-method of testing

The W-test is performed by calculating a one dimensional test statistic:

$$T_1 = \frac{c^T Q_y^{-1} \hat{e}}{c^T Q_{\hat{r}} c} \tag{2.1}$$

where c is a vector with zeroes and a single one, $\hat{e} = y - A\hat{x}$ and $Q_{\hat{r}} = Q_y^{-1} - Q_y^{-1} A(A^T Q_y^{-1} A)^{-1} A^T Q_y^{-1}$.
If Q_y is diagonal, as is the case, T_1 can be also expressed as:

$$T_1 = \frac{\hat{e}_i^2}{\sigma_{\hat{e}_i}^2} \tag{2.2}$$

T_1 has a central $\chi^2(1)$-distribution under the null-hypothesis, so the squareroot of T_1 follows a standard normal distribution. The distribution of the alternative hypothesis is actually not known, but it is assumed to have a non central $\chi^2(1,\lambda)$-distribution, which means that the squareroot of T_1 follows a normal distribution with a non-zero mean equal to $\sqrt{\lambda}$. The value for the non-centrality parameter λ is calculated by using the preset values for α_0 and γ. The value of λ is not relevant for deciding whether to reject or accept the W-test, but it is needed in order to calculate the significance level of the F-test, α [9].

[1] Logsheets of 2016 are not available to verify what instrument has been used.

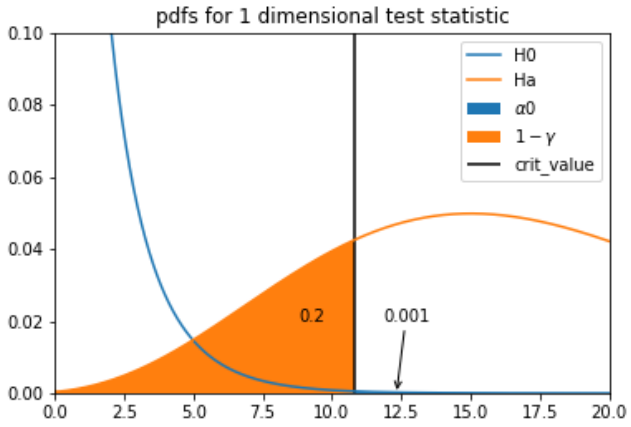


Figure 2.3: The one dimensional test statistic T_1 has a central $\chi^2(1)$-distribution under the null-hypothesis (blue). The alternative hypothesis H_a is defined by the value of 1 - γ. The test statistic under H_a follows a non central $\chi^2(1,\lambda)$-distribution, with non-centrality parameter λ (orange). This non-centrality parameter λ is found by finding a non-central distribution that satisfies $1-\gamma = 0.2$.

The test statistic of the F-test is:

$$T_q = \hat{e}_0^T Q_{yy}^{-1} \hat{e}_0 \tag{2.3}$$

where $\hat{e}_0$ is the least square residual of the full model, $\hat{e}_0 = y - A\hat{x}_0$, and Q_{yy} is the a-priori variance matrix.[2] The test statistic of the F-test follows a central $\chi^2(q)$-distribution, where q stands for the redundancy. This means that the test statistic represents a sum of the squares of q independent standard normal random variables. Under the alternative hypothesis this is not the case, one out of all the measurements is an outlier and therefore the test statistic does follow a non-central $\chi^2(q,\lambda)$ distribution. This means that the mean of the summed squared normal distributed variables is not equal to zero, but is related to the aforementioned noncentrality parameter by $\lambda = \sum_{i=1}^{q} \mu_i^2$ [6, Ch. 29]. The critical value, indicating whether to reject or accept the null-hypothesis is calculated by using the preset value γ. [9]

In the example the redundancy is rather high (296) and this causes the distributions of the test statistic under the null hypothesis and the alternative hypothesis to overlap largely. This leads to a very high value for the level of significance, α, meaning that the probability of rejecting the null-hypothesis while it is true (type I error) is very large.

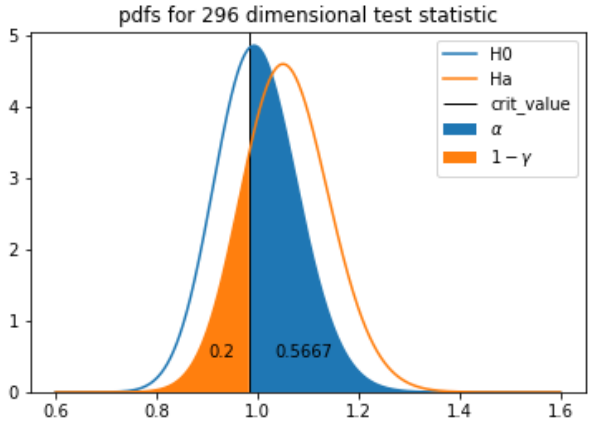


Figure 2.4: The test statistic under the alternative hypothesis H_a follows a non-central $\chi^2(q,\lambda)$ distribution (orange), which is scaled by the redundancy. The non-centrality parameter λ is the same non-centrality parameter as found in Figure 2.3. Under the null hypothesis H_0 the test statistic follows a central $\chi^2(q)$-distribution (blue). The power of the tests is required to be 0.8 and this requirement implies the level of significance α, and the corresponding critical value.

[2] For an elaborate explanation of the test statistics see Teunissen et al. [12] and Polman and Salzmann [9].

2.2.2. The B-method of testing evaluated

A level of significance α of 0.5667, found for the 296 dimensional test statistic in Figure 2.4 means that the probability of rejecting H_0 is larger than the probability of accepting H_0 when H_0 is true (type I error). This is the reason to evaluate the B-method of testing as performed by the MOVE3 software. As indicated by Teunissen [11], for larger redundancies the level of significance α of the overall model test increases monotonically, as depicted in Figure 2.5.

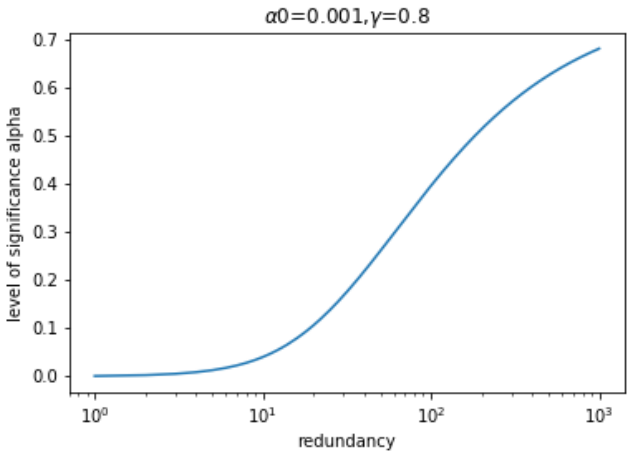


Figure 2.5: The redundancy vs. the level of significance alpha for the overall model test.

When α_0 is chosen smaller than 0.001, α will be smaller as well, but this does not offer a solution to the imposed problem, because α remains too large[3] for higher redundancy, as depicted in Figure 2.6. Besides that, choosing α_0 very small implies a large Minimal Detectable Bias (MDB). An error of the size of the MDB can be detected with probability γ, larger errors result in a larger probability of detection [12]. A large MDB is thus not desirable, because then the observation testing is insensitive for smaller biases.

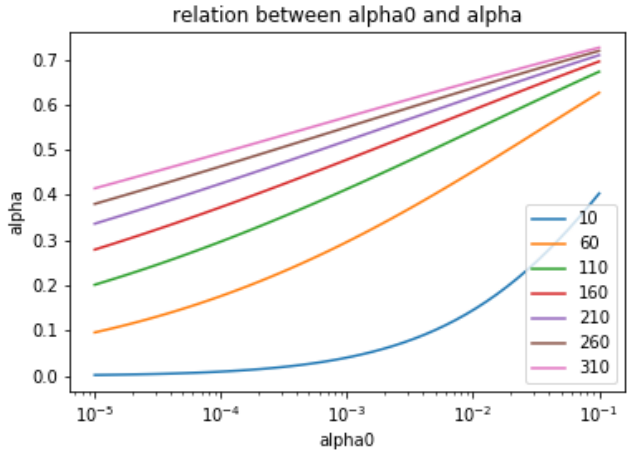


Figure 2.6: The values of α versus α_0, given that $\gamma = 0.8$ for different values for the redundancy.

Using the overall model test to detect the presence of outliers before performing data-snooping to identify the outliers can save time and reduce computational costs – namely, if the overall model test is accepted data-snooping can be skipped. If the value for α is too large, the overall model test does not detect the presence outliers properly. A workaround to this is to perform the adjustment in steps, as indicated by Baarda [2]. Performing the adjustment on a subset of the measurements reduces the redundancy and allows to detect outliers in the overall model test. If the test is accepted data-snooping on the subset of measurements can be

[3] indicated upper bound for α is 0.1 [2, p. 25]

skipped, time and computational costs can be saved or, if the equipment is still in the field, measurements can be repeated. However, nowadays data-snooping is not expensive and time-consuming anymore, the software MOVE3 performs it anyway, and in this specific case repeating measurements is not possible. Performing the adjustment in steps does not offer any advantage, therefore an alternative procedure is proposed.

2.2.3. An alternative approach

The F-test gets less sensitive for detecting outliers when the redundancy increases and for larger redundancies the F-test cannot be used as a detecting test. The maximum redundancy for which the F-test suffices as detecting test depends on α_0 (see Figure 2.6) and γ. With testing settings α_0 equal to 0.001 and γ equal to 0.8, redundancies larger than 20 result in α's larger than 0.1 (see Figure 2.5). The redundancies of the measurements of the tachymetric network are larger than 250, making the F-test not usable as detecting test. Therefore the "detection, identification, adaptation"-procedure as described by Teunissen et al. [12] and by Polman and Salzmann [9] as "toetsingsprocedure volgens de B-methode van toetsen" has to be performed differently, namely simultaneously in an interative process, leading to an adjustment in which all W-test values are below the critical value and the F-test is around 1, indicating that the stochastic model is representative. The detection and identification of outliers happens through the values of the W-tests. The highest value of the W-test is investigated, if the corresponding observation is a blunder (typo or event during measuring such as turning a switch to secure the instrument) possibly multiple measurements should be deselected. In case no clear reason for the outlier is found the value of the F-test is checked. If this value is lower than 1, the reason for detecting the outlier is not due to a stochastic model that is too strict and the measurement is deselected. If the value of the F-test is larger than 1, the stochastic model is too strict and should be loosened. This process is repeated until all values of the W-test are smaller than the critical value of 3.29 (significance level of 0.001). The value of the F-test should then be around 1[4]. If this is not the case the stochastic model is strictened and possibly more outliers are detected and the same procedure is repeated (see Figure 2.7).

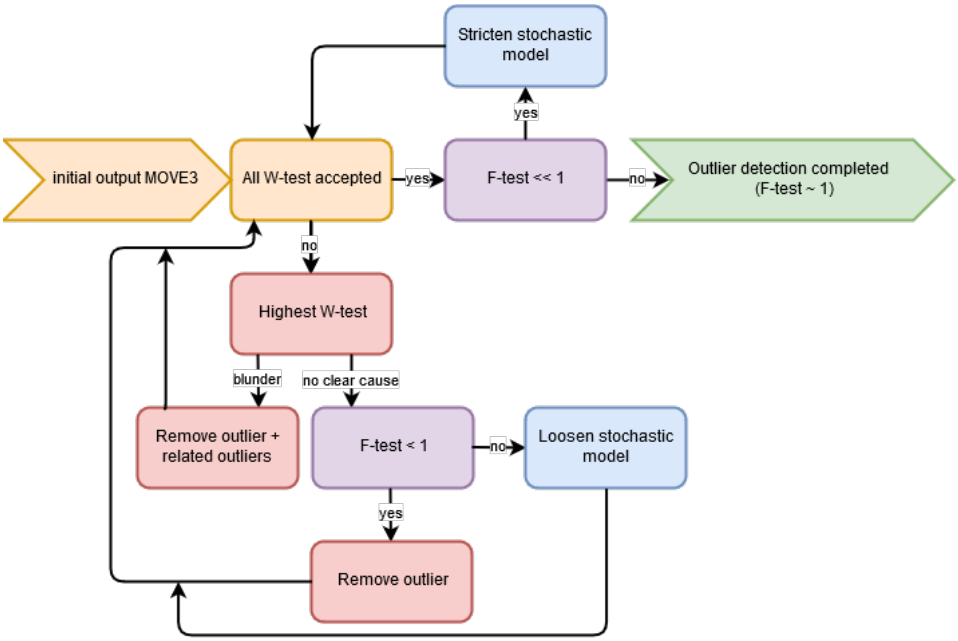


Figure 2.7: Flowchart of the iterative process to detect outliers and evaluate the stochastic model using the output of the Free network adjustment performed by the software MOVE3.

[4]One could set a level of significance for the F-test to quantify this.

3

Results

Yearly solutions of the tachymetric network are estimated using the software MOVE3. Firstly a free network adjustement is performed to detect outliers and evaluate the stochastic model as described in Section 2.2.3 (Section 3.1). After that, a weighted constrained adjustment is performed, in which GNSS data constrains the free network solution (Section 3.2). From the yearly solutions linear multiyear trends of the stations (Section 3.3) and relative velocities (Section 3.4) have been estimated using linear least squares.

3.1. Free network adjustment

In the free network solution the given coordinates of 'known' stations do not constrain the solution; the geometrical relations of the network only depend on the observations [10]. The testing procedure has been described in Section 2.2 (see Figure 2.7). Starting point is a 3D adjustment with the stochastic model as shown in Table 2.2 in Section 2.1. Final goal is a 2D adjustment with a value of circa 1 for the F-test. The final goal of the adjustment is limited to horizontal coordinates (2D), because for some years the zenith angle measurements are subject to a systematic bias and cannot be used for height estimation. Also, for some years there are no height measurements of the tripods. To have consistent results the final goal for all years is 2D, but the processing always initiated with a 3D adjustment. Starting in 3D allows to detect outliers in the zenith angle measurements, which influence the horizontal distances. This results in an additional mitigation for outliers: switching to a 2D geometry. The procedure per year is described in detail in appendix B, where per detected outlier I verified whether the measurement is an inexplicable outlier or a blunder, and in the latter case other involved measurements are simultaneously deselected. Detected blunders are related to typos (see for example Figure B.12 in Appendix B), mistakes in the measurement procedure (see for example Figure B.9), change in height of a station (see for example Figures B.10 and B.11 in Appendix B) and possibly the turning of a switch to secure the instrument or a change of observer or a minor impact on the instrument (see for example Figure B.1 in Appendix B). An instrumental index error (see Appendix B.2.1) has been discovered for the measurements of 2016 and 2018. The number of deselected measurements, the final value for the F-test and the final stochastic model are shown in Table 3.1.

year	deselec. meas.	final value F-test	std direction	std distance	std zenith angle
2015	12 (≈ 4%)	0.989	0.005 gon	0.006 m + 4ppm	0.005 gon
2016	12 (≈ 3%)	0.974	0.003 gon	0.003 m + 1.5ppm	0.003 gon
2017	14 (≈ 4%)	0.862	0.0036 gon	0.001 m + 0.5 ppm	0.003 gon
2018	8 (≈ 3%)	0.895	0.003 gon	0.0015 m + 0.7 ppm	0.003 gon

Table 3.1: Overview of the adjusted free network per year. Details of the processing is described in appendix B. Percentages of deselected measurements are based on the total number of direction and distance measurements.

In the yearly solutions 3 to 4 % of the measurements is rejected. Compared to the value of 1 % as mentioned by Sweco Nederland B.V. [10] this value is large, but this can be contributed to the inexperience of the

observers[1]. The consistency of the percentage of rejected observations over the years affirms the new approach for detecting outliers and tuning the stochastic model as described in Section 2.2.3. The values of the stochastic model for the directional and zenith angle measurements are much higher than specified by the instrument manufacturer (order 10).

3.2. Weighted constrained adjustment

After deselecting detected outliers and adjusting the stochastic model, a weighted least squares adjustment is performed, in which the 'known' stations are considered as measurements with their respective uncertainties [10] (see Figure 3.1). The 'known' coordinates are derived from the GNSS measurements.[2] Although the GNSS stations could act as validation for the trends estimated by the tachymetry network, they have been all used in the weighted adjustment in order to improve the precision of the estimates.

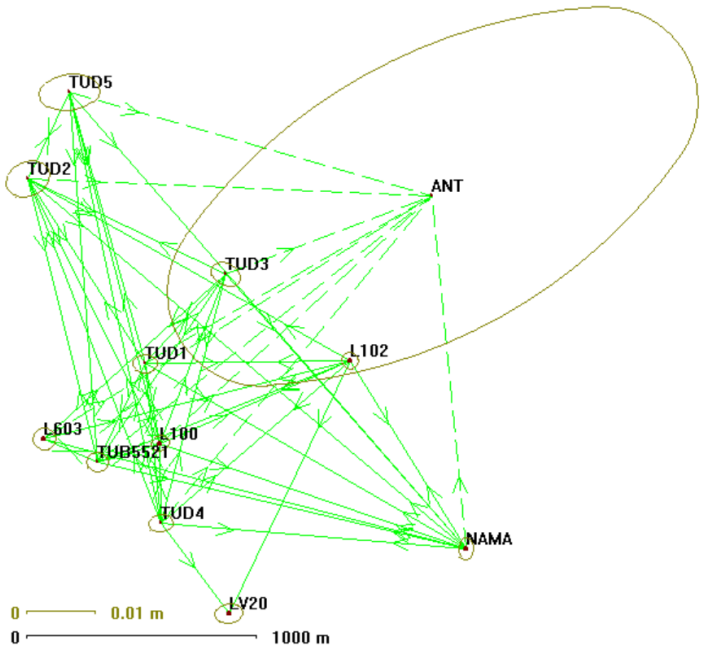


Figure 3.1: Weighted constrained adjustment results of 2018, performed by the software MOVE3. In green the measurements, where little arrows indicate the direction and dashed lines the absence of distance measurements. In brown the error ellipses of the estimates.

Figure 3.2 serves as illustration for the weighted constrained adjustment. The same adjustment has been performed for all years, resulting in four sets of coordinates and associated variance-covariance matrices. Station ANT has clearly a much larger error ellipse, because the station is an antenna without measurement setup to which only directions are measured.

3.3. Multiyear trends

Multiyear trends per station are estimated using linear least squares. The uncertainties of the estimates are calculated using

$$Q_{\hat{x}} = (A^T Q_y^{-1} A)^{-1} \tag{3.1}$$

where A is the design matrix and Q_y is a composite of the yearly variance-covariance matrices as estimated by the software MOVE3 in the weighted constrained adjustment. The estimate for TUD4 in 2015 has been left out of the evaluation since it is known that the wrong benchmark has been used that year. Also the estimated trend for station ANT has not been plotted, because the station is an antenna without measurement setup to

[1] In evaluating the percentage of rejected measurements one needs to consider also that false rejections are contributing to the percentage. The level of significance is set to 0.001, meaning that 0.1 % of the measurements is falsely rejected (type I error). Adjusting the level of significance will thus result in a slightly different percentage of rejected measurements.

[2] The used GNSS coordinates and a description of the sources and the processing is shown in appendix C.

which only directions are measured, resulting in large uncertainties (see Figure 3.1). The results are shown in Figure 3.2 and listed in Table 3.2.

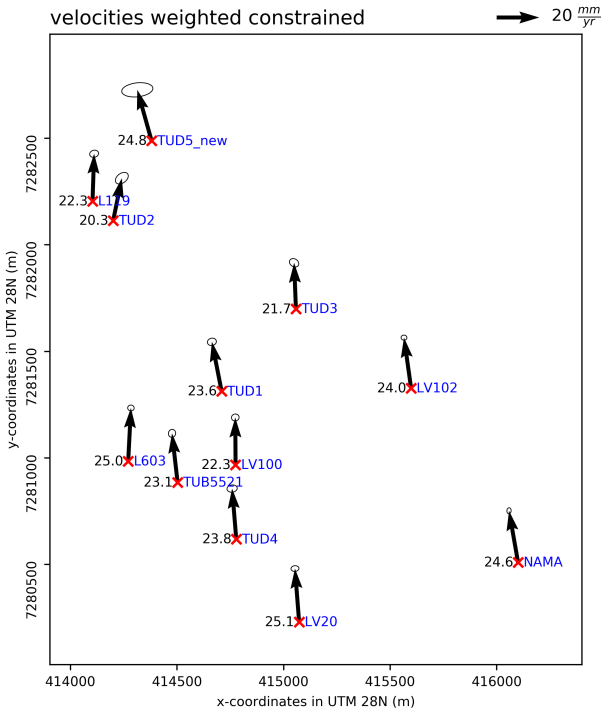


Figure 3.2: Estimated velocities and their respective 2σ-error ellipses, values are listed in Table 3.2. The arrows display the magnitude and direction of the velocity in a different scaling than the coordinates on the x- and y-axis. The benchmark locations are displayed with red crosses, the names of the benchmarks are annotated in blue and the value of the magnitude of the velocity in mm/year is written next to the arrow in black.

Station	velocity x-direction	$\pm 2\sigma$ (mm/yr)	velocity y-direction	$\pm 2\sigma$ (mm/yr)
ANT	86.62	±43.39	-54.69	±32.89
L119	0.75	±2.09	22.30	±1.69
L603	1.28	±1.51	24.94	±1.38
LV100	-0.13	±1.79	22.25	±1.64
LV102	-3.34	±1.37	23.75	±1.20
LV20	-1.95	±1.87	25.07	±1.34
NAMA	-4.23	±1.00	24.19	±1.43
TUB5521	-2.61	±1.70	22.98	±1.88
TUD1	-4.56	±2.10	23.20	±1.74
TUD2	4.06	±3.03	19.85	±2.73
TUD3	-0.85	±2.19	21.64	±1.98
TUD4	-1.95	±2.40	23.70	±1.71
TUD5_new	-6.70	±7.31	23.90	±3.36

Table 3.2: Velocities and their respective uncertainties as estimated using linear least squares.

Looking at the absolute velocities as presented in Figure 3.2, a clear northward trend is visible. The mean of the estimated velocities (Table 3.2, not considering station ANT) in North-direction and East-direction are 23.15 mm/year and -1.69 mm/year. These numbers are based on the GNSS reference frame ITRF2014 (see Appendix C). If the Eurasian plate is assumed to be stable, the area is moving with a mean velocity of 6.4 mm/year northwards and -8.6 mm/year eastwards with respect to Eurasian plate. The area is located in the

transition zone of the plate spreading and this is the largest contributor to the mean velocity. Another, small, contribution is from the glacial isostatic adjustment (GIA). Furthermore the northward movement South of Krafla, as well as other deformation signals in the region can be modelled well by two deep Mogi sources, one inflating North of Krafla, one deflating below the Krafla caldera. The reason for these deformations is not clear, it could be readjustments at depth following the Krafla rifting episode or migration of magma at depth or the setting of a new equilibrium pressure in the ductile crust [3][3].

3.4. Relative velocities

Subtracting the aforementioned mean velocities (23.15 mm/year in North-direction, -1.69 mm/year in East-direction) from the estimates results in relative velocities which tell us how stations are moving with respect to each other, see Figure 3.3. The estimated relative trends in East- and North-direction are shown in Figures 3.4 and 3.5.

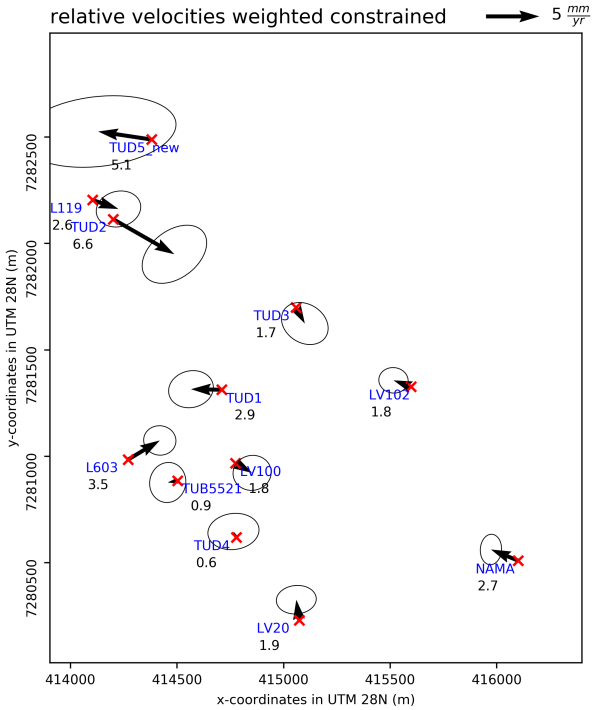


Figure 3.3: Estimated relative velocities and their respective 2σ-error ellipses. The arrows display the magnitude and direction of the relative velocity in a different scaling than the coordinates on the x- and y-axis. The benchmark locations are displayed with red crosses, the names of the benchmarks are annotated in blue and the value of the magnitude of the relative velocity in mm/year is written next to the arrow in black.

[3]The article of Drouin et al. [3] is limited to the period 2008-2014. The processes mentioned are slow and therefore assumed to be qualitatively valid for the period 2015-2018.

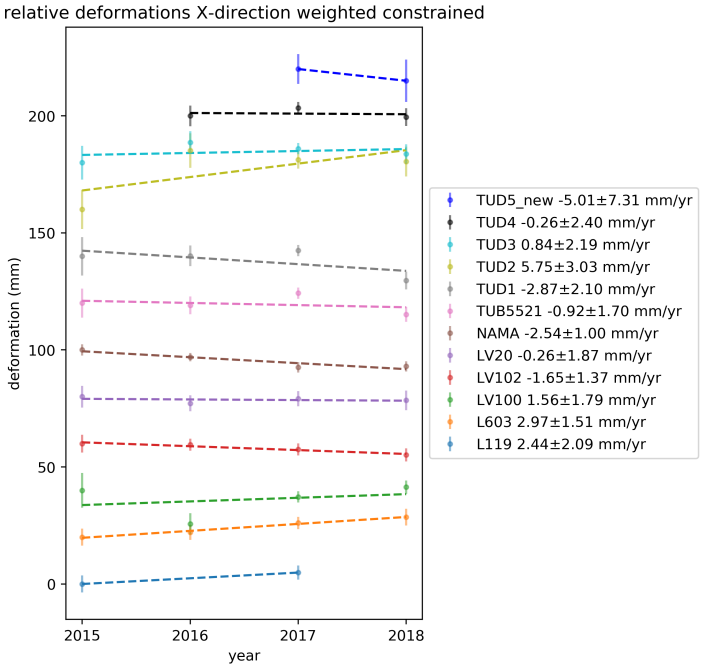


Figure 3.4: Estimated trend in East-direction plotted together with relative coordinates and their respective 2σ-error bars. In the legend the 2σ-interval of the trend is shown.

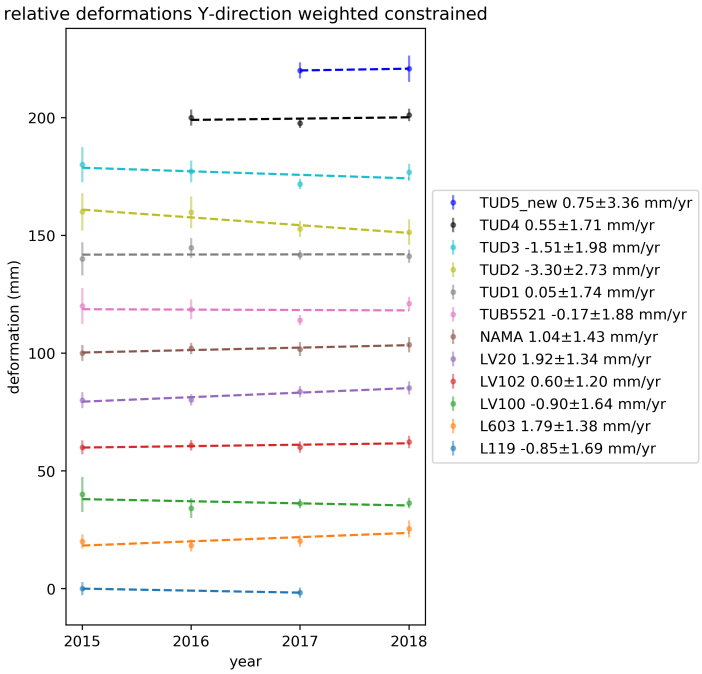


Figure 3.5: Estimated trend in North-direction plotted together with relative coordinates and their respective 2σ-error bars. In the legend the 2σ-interval of the trend is shown.

The relative signal is very small and the error ellipses are rather large. The question is whether the estimates are good enough to interpret them, which is evaluated qualitatively (visually) as well as quantitatively in Section 3.5.

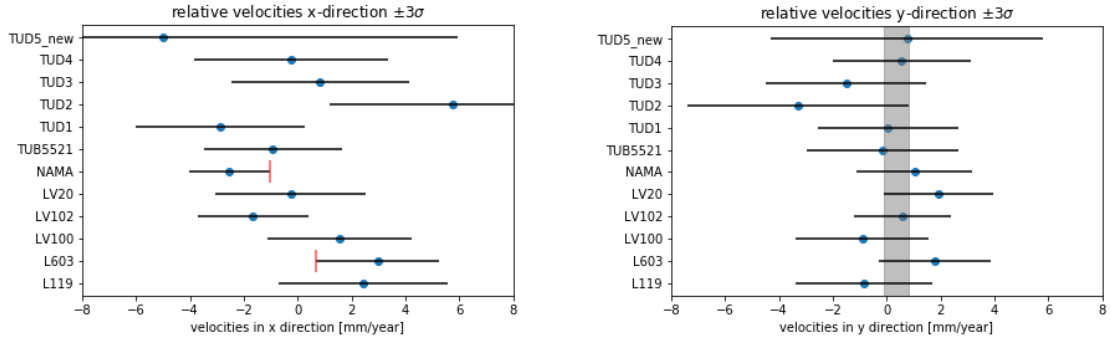
Main contributor to the relative velocities is the contraction in the area of Bjarnarflag due to natural geothermal processes and/or due to extraction for the geothermal powerplant. Other contributors are the

spatial variations of the aforementioned broadband deformations (plate spreading, GIA and two deep Mogi sources) [3]. The small scale deformations have been modelled by Drouin et al. [3] as three shallow deflating Mogi sources. Globally the results of this study agree, but a quantitative comparison has not been performed.

3.5. Analysis

The values of the estimated relative trends and their respective 3σ-confidence intervals (99.73%) are shown in Figure 3.6. For the northward trends (Figure 3.6b, y-direction) all values overlap, indicated with a gray region. All relative velocities in North-direction could thus be zero. For the eastward trends (Figure 3.6a, x-direction) there is no overlapping region. Benchmark L603 is clearly moving westwards with respect to benchmark NAMA, marked with two vertical red lines in Figure 3.6a. Looking at Figure 3.4, showing the relative trends in East-direction, this is confirmed.

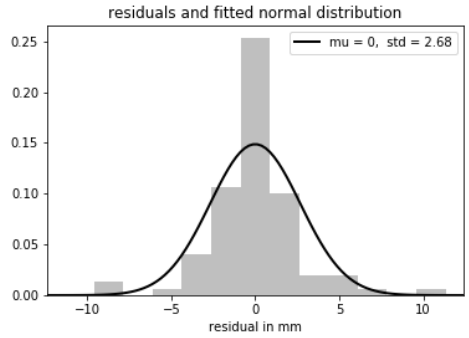
The data used to estimate the relative trends in east- and north-direction (shown in Figures 3.4 and 3.5) shows visually detectable outliers, for example in east-direction TUD2 in 2015, or TUD1 and TUB5521 in 2017. The visually detectable outliers in east-direction do not coincide clearly with the visually detectable outliers in north-direction (correlation between residuals is 0.53, see Figure 3.7b). There can be confidence in the yearly solutions, because these have been thoroughly checked for outliers and the measurements are redundant, but possibly some tripods were not perfectly centered above the benchmark. This would not influence the yearly adjustment of the network, but it causes outliers in the multiyear data. The visually detectable outliers can thus be due to a wrong measurement set up or due to nonlinear movements, in which case the assumed model is wrong. In the first case, outliers can be identified using statistical tests. These have not been performed, because of the small number of observations, maximum 4 per year, 86 in total. In the second case, if the movements are nonlinear, the estimated residuals of the least-squares fit of the model would not be normally distributed. This is also hard to verify quantitatively due to the small number of observations. Evaluating the values visually (see Figure 3.7) shows that the residuals are following a normal distribution quite well, for the north-direction the spread of the residuals is smaller than for the east-direction.



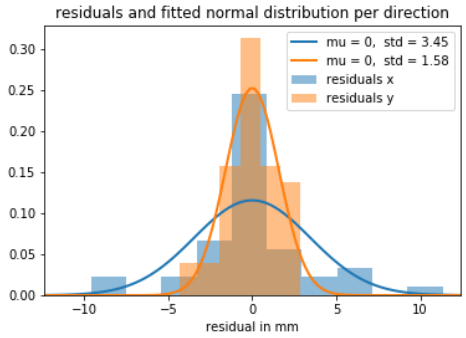
(a) x-direction. Benchmark L603 is clearly moving westwards with respect to benchmark NAMA, marked with two vertical red lines.

(b) y-direction.

Figure 3.6: Overview of relative velocities in x- and y-direction and their respective 3σ-intervals. Gray area shows overlapping values.



(a) All (except 'ANT').



(b) Per direction (except 'ANT').

Figure 3.7: Histogram of residuals of least squares regression (as shown in Figures 3.4 and 3.5)

The significance of the estimated relative velocities can be tested quantitatively using the F-statistic in which the residual and regression sums of squares of two models are compared:

$$F_0 = \frac{(\text{RegSS}_1 - \text{RegSS}_0)/q}{\text{RSS}_1/(n-k-1)} \tag{3.2}$$

where RegSS = TSS - RSS with TSS = $\sum(Y_i - \bar{Y})^2$ and RSS = $\sum(Y_i - \hat{Y}_i)^2$. q is the dimension of the subset of slopes that are defined as zero in the null hypothesis, n is the number of observations, k the number of unknown slopes in the alternative hypothesis. The F-statistic follows a $F(q, n-k-1)$ distribution under the null hypothesis [5].

Using a level of significance of 5% three different hypotheses are tested consecutively against the same alternative hypothesis, namely that all **relative** velocities are non-zero. From these tests (see Table 3.3), it can be concluded that the relative velocities in east-direction are significant. The relative trends in north-direction are not significant when tested separately (first null hypothesis), but they are significant when tested simultaneously with the relative trends in east-direction (third null hypothesis). The trends in Figure 3.5 could thus possibly be horizontal lines, and this would increase the root mean square error slightly in north-direction, from 10.35 mm to 14.87 mm. The root mean square error in east-direction is 22.62 mm.

null hypothesis	value F_0	critical value F-statistic	conclusion
All relative y-trends zero, all relative x-trends nonzero	0.94	1.91	null hypothesis accepted
All relative x-trends zero, all relative y-trends nonzero	2.80	1.91	null hypothesis rejected
All relative trends zero	1.87	1.70	null hypothesis rejected

Table 3.3: Values for the F-statistic and conclusions on the different null hypotheses.

4

Conclusions and recommendations

The tachymetric measurements of the network can be used to investigate mm/year deformation of the earth's surface in the area of Bjarnarflag. In order to do so the yearly measurements need to be tested for outliers and the stochastic model needs to be verified. For this an alternative method has been used. The estimates of the multiyear trends of the benchmarks are not easy to interpret, because of the relatively large error ellipses. Recommendations for improving the estimates in the coming years are given.

The consistency over the years of the percentage of rejected observations of the yearly solutions affirms the new testing procedure as described in Section 2.2.3, where the Detection, Identification and Adaptation steps are performed simultaneously in order to delete outlying measurements and tune the stochastic model. The values of the tuned stochastic model for the directional and zenith angle measurements are much higher than specified by instrument manufacturer (order 10). An explanation lies partly in the non-constant circumstances and the inexperienced observers, but an evaluation of the instrument settings and possibly a test setup is recommended to investigate this further. Also a calibration of the instrument is recommended, as this would eliminate the index error. Saving the measurements digitally would eliminate typo's and save processing time.

With deformation rates of the tachymetric network movements in between GPS stations could be captured. This could answer the question whether the GPS values are representative for the whole area, or whether the benchmarks are moving relatively to each other. This relative signal in the Bjarnarflag area can be due to extraction of water or steam by the geothermal powerplant, due to the instability of the benchmarks or due to a natural process as described in Chapter 1. The estimated relative velocities are significant, but show large uncertainty intervals. Collecting more data in the coming years would increase the precision of the estimates of the (relative) velocities (see Table 4.1). Longer timeseries would allow meaningful tests for outliers and gives a better understanding of whether the benchmarks are moving linearly or not and thus whether the estimated trends are accurate or not.

number of years	mean std x (mm)	reduction in %	mean std y (mm)	reduction in %
4	2.4	-	1.8	-
5	1.6	67	1.3	69
6	1.2	50	0.9	52
7	0.9	39	0.8	41

Table 4.1: Estimated increase of precision. First line shows the mean standard deviation of the trends as shown in Figures 3.4 and 3.5 (so excluding 'ANT'). To estimate the predicted values for the mean std it is assumed that the network and the variance-covariance matrix resulting from the network adjustment for the coming years are the same as in 2018, and the network has been measured yearly.

To accelerate the increase of precision of the estimated velocities the precision of the yearly adjustment must be improved. This can be done by reevaluating the stochastic model of the measurements, by adding more GNSS measurements, and by reconsidering the network design. Extra (GNSS-)benchmarks can be

included to densify and fortify the network.

The (improved) estimated horizontal velocities of deformation can be used in combination with remote sensing data such as InSAR, impairing the limitations of both techniques. InSAR is insensitive for deformations in the north-south direction and has therefore been combined with GPS data by Drouin et al. [4], using the tachymetry data as well for this can improve the results due to an increase in the number of in-situ measurements. The limitations of tachymetry (poor spatial resolution, poor temporal resolution) can be impaired using InSAR data, which gives insight into the global deformation patterns of the area and, with the higher temporal resolution, insight into the linearity of the deformation. Comparing the tachymetric data to InSAR data can also reveal information on the stability of the individual benchmarks. For research on horizontal deformation due to a specific extraction or injection well of the geothermal powerplant or due to a specific local natural process the design of the network should be reconsidered in order to capture a signal needed to confirm or invalidate a hypothesis.

Estimating horizontal deformation velocities in an area using tachymetry asks for repeated measurement campaigns of GNSS and tachymetry, which is costly and time-consuming. The reached precision of the estimated coordinates measured by tachymetry is lower than the GNSS-estimates, but if the number of GNSS receivers is limited and equipment for tachymetric measurements is available, the spatial resolution of the GNSS measurements can be improved using tachymetry. Measures need to be taken, such as setting up prisms on tripods, no changing of observer during a measurement cycle, measuring under stable conditions and calibrating the instrument, to reach the best precision of the estimated coordinates in order to capture a mm/year-signal.

Bibliography

[1] J.E. Alberda. *Inleiding Landmeetkunde*. Delftse Uitgevers Maatschappij b.v., Delft, fifth edition, 1978.

[2] W. Baarda. A testing procedure for use in geodetic networks. *Netherlands geodetic commission, publications on geodesy, new series*, 2(5), 1968.

[3] Vincent Drouin, Freysteinn Sigmundsson, Benedikt G. Ófeigsson, Sigrún Hreinsdóttir, Erik Sturkell, and Páll Einarsson. Deformation in the northern volcanic zone of iceland 2008–2014: An interplay of tectonic, magmatic, and glacial isostatic deformation. *J. Geophys. Res. Solid Earth*, 122:3158–3178, 2017.

[4] Vincent Drouin, Freysteinn Sigmundsson, Sandra Verhagen, Benedikt G. Ófeigsson, Karsten Spaans, and Sigrún Hreinsdóttir. Deformation at krafla and bjarnarflag geothermal areas, northernvolcanic zone of iceland, 1993–2015. *Journal of Volcanology and Geothermal Research*, 344:92 – 105, 2017. ISSN 0377-0273. doi: https://doi.org/10.1016/j.jvolgeores.2017.06.013. URL `http://www.sciencedirect.com/science/article/pii/S0377027317303712`. Volcano Geodesy: Recent developments and future challenges.

[5] John Fox. *Applied regression analysis and generalized linear models*. SAGE Publications, second edition, 1997.

[6] Norman Lloyd Johnson, Samuel Kotz, and N. Balakrishnan. *Continuous univariate distributions, Volume 2*. Wiley Sons, second edition, 1995.

[7] Leica Geosystems AG. *Leica TPS1100 Professional Series*. 2002.

[8] Leica Geosystems AG. *Leica TPS1200+ User Manual*. 2008.

[9] J. Polman and M. Salzmann. *Handleiding voor de technische werkzaamheden van het Kadaster*. Kadaster, cop. 1996, Apeldoorn, 1996.

[10] Sweco Nederland B.V. Move3 user manual, version 4.4.

[11] P.J.G. Teunissen. *Testing theory: an introduction*. Delft University Press, Delft, 2000.

[12] P.J.G. Teunissen, D.G. Simons, and C.C.J.M. Tiberius. *Probability and Observation Theory, Lecture notes*. Department of Geoscience and Remote Sensing, Faculty of Civil Engineering and Geosciences, Delft University of Technology, 2009.

A

MOVE3 Input

This appendix can be used in addition to the file 'MOVE3_instructions_2017' for the feature processing of measurements in the context of the educational fieldwork.

Processing the data in MOVE3 is done in three steps:

- Import data
- Perform free network adjustment
- Perform weighted constrained adjustment

The free network adjustment is performed in order to verify the stochastic model and detect and deselect possible outliers[1]. Coordinates of two stations need to be given as input, these can be chosen arbitrarily. Starting from a 3D adjustment allows to identify outliers in the zenith angles. During the free network adjustment the dimension has to be set to 2D, as has been done for the measurements of previous years, because the height estimates were not reliable. When all W-tests are accepted and the F-test value is approximately 1 the weighted constrained adjustment can be performed.

The weighted constrained adjustment uses GPS coordinates and their respective standard deviations to calculate the actual coordinates. For transforming the GPS coordinates from geocentric XYZ coordinates as downloaded from `http://icelandsupersite.hi.is/gps/ts/NVZ.html` to UTM28 coordinates I used the python script 'GPS_MOVE3_XYZtoUTM28.ipynb', which needs input generated by the matlab script 'Save_std_neu.m'. The weighted constrained adjustment returns a '.out2' file containing among other things the adjusted coordinates of the benchmarks and a '.var' file containing the full covariance matrix (if the 'additional output' in 'general options -> adjustment' has been set correctly as shown in Figure A.2). These files can be used to estimate multiyear trends using the python script 'Multiyear_trends_MOVE3out2files.ipynb'.

The file 'MOVE3_instructions_2017' gives concise instructions on how to use the software. I reccomend to use slightly different project settings, as shown in Figures A.1, A.2, A.3 and A.4.

[1] An iterative scheme for this process is included in section 2.2.3 (Figure 2.7). The iterative process for the measurements of 2015, 2016, 2017 and 2018 is described in appendix B

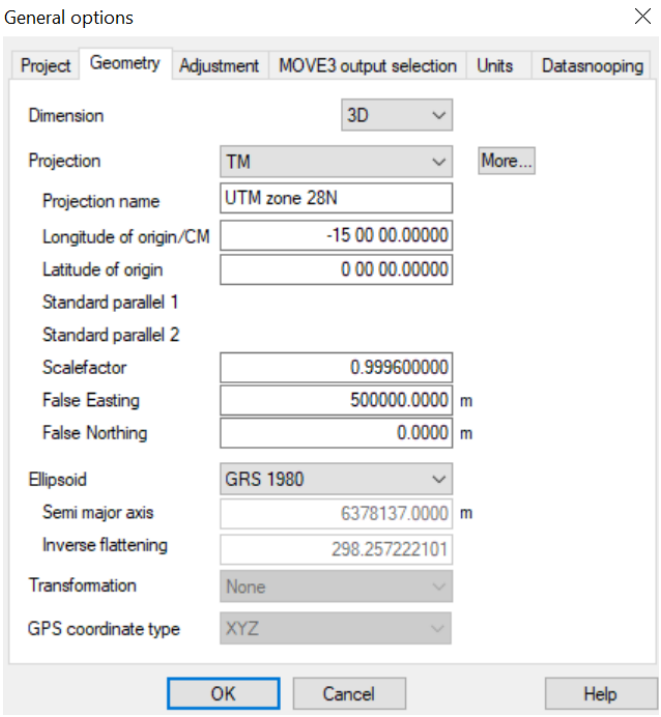


Figure A.1: Initial project settings, later the dimension is set to 2D. UTM zone 28N is chosen instead of a local projection in order to connect the adjusted coordinates to other years and other types of data.

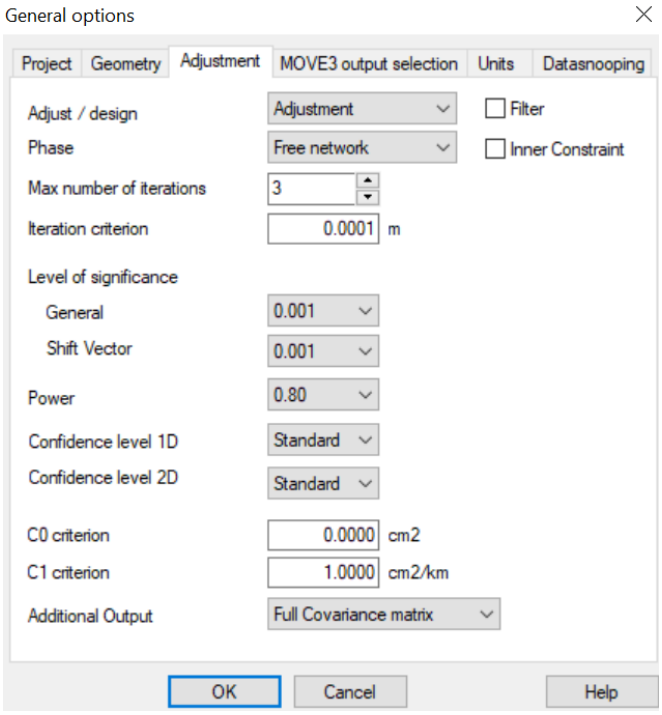


Figure A.2: Initial project settings, later the phase is set to "Weighted Constrained" and the additional output "Full Covariance matrix" is generated.

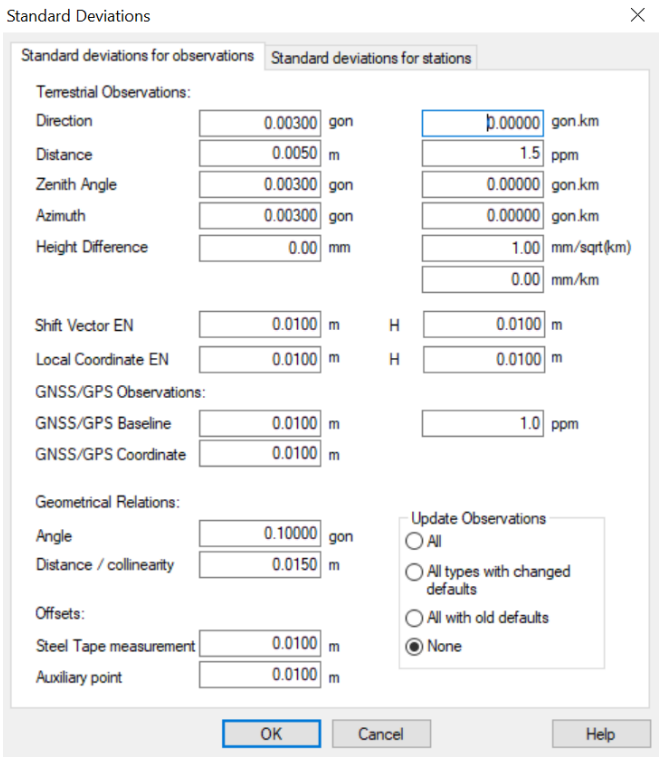


Figure A.3: Initial settings for standard deviations of observations, these values differ from the values used from 2017 onwards (see section 2.1).

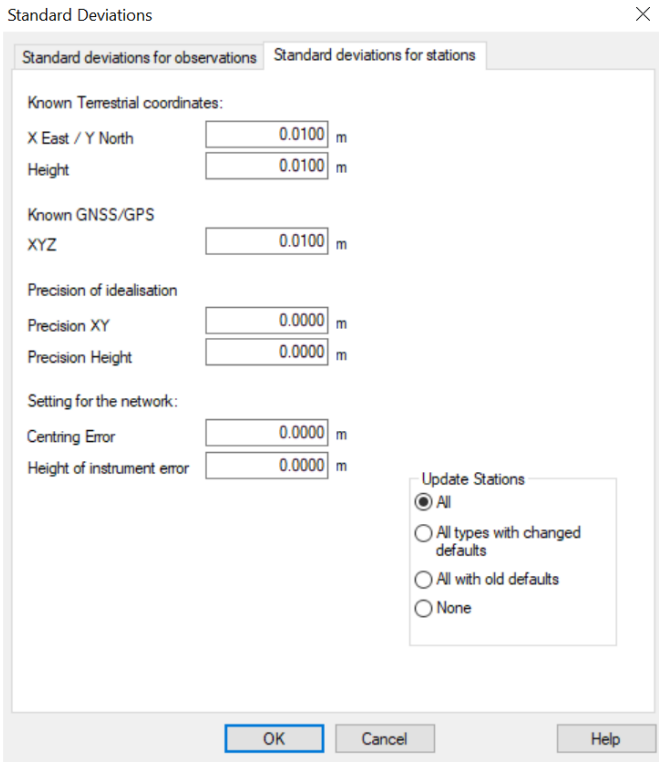


Figure A.4: Initial settings for standard deviations of stations. The given standard deviations of the GPS observations will be used later for the weighted constrained adjustment and can be specified per station via view -> stations -> double click on the station you want to edit.

B

Iterative process per year

In this appendix the testing procedure as described in section 2.2 (see Figure 2.7) is applied to the measurements per year. Per detected outlier I verified whether the measurement is an inexplicable outlier or a blunder, and in the latter case other involved measurements are simultaneously deselected.

B.1. Iterations free network adjustment in 2015

Forward and backward distance from TUD2 to LV102 There is no distance measurement from LV102 to TUD2 to compare. Compared to 2016 there is a large difference of approximately 2 meters (1605.46 in 2016 versus 1603.39 in 2015). The directional measurements are similar to 2016, leading to the conclusion that it must be an error in the notation of the distance measurement. Distance measurements deselected.

Backward direction from LV100 to ANT Antenna is actually an antenna to which is aimed during measuring, not a very thin and easy goal, prone to errors. Directional measurement deselected.

Backward direction from L119 to TUB5521, TUD4 and LV102 Probably something happened during measuring, a change of observer or a movement of the instrument (see figure B.1). No influence on vertical angles and distances, differences between the forward and backward measurements smaller than the 1-σ interval as indicated by the instrument manufacturer (see figures B.2 and B.3). Directional measurements deselected.

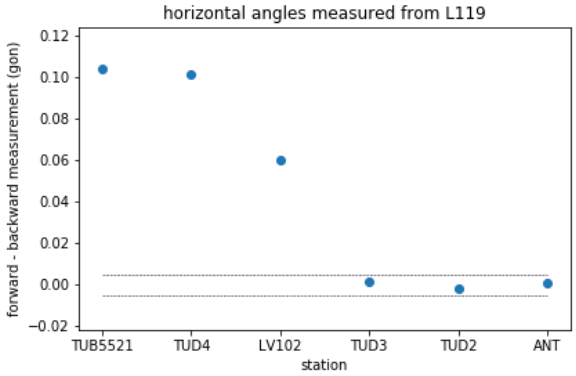


Figure B.1: Differences between the forward and the backward measurements of directions. The order of measurements is kept in the figure, forward measurements were from TUB5521 to ANT and backward measurements from ANT to TUB5521. In dashed black lines the 1-σ interval as specified in table 2.2 in chapter 2 (0.005 gon). From this it is clear that something happened either between LV102 and TUD3 in the forward measurements or between TUD3 and LV102 in the backward measurements.

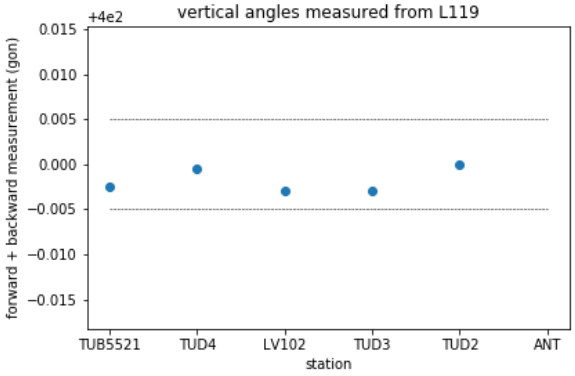


Figure B.2: Summation of forward and backward measurement of vertical angles. In dashed black lines the 1-σ interval as specified in table 2.2 in chapter 2 (0.005 gon). Here no change visible between LV102 and TUD3.

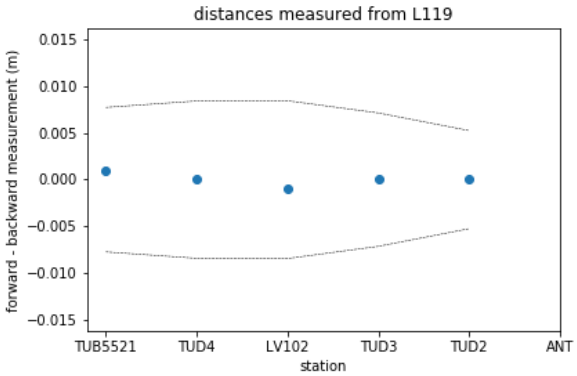


Figure B.3: Difference of forward and backward measurement of distances. In dashed black lines the 1-σ interval as specified in table 2.2 in chapter 2 (0.005 m + 2 ppm). Here no change visible between LV102 and TUD3.

Forward direction from TUB5521 to KK04, L603 and ANT Something happened in forward measurement cycle between TUD2 and KK04 (see B.4), the resulting constant offset between forward and backward measured horizontal angles is not a problem, since relative angles are of interest. No influence on vertical angles and distances, differences between the forward and backward measurements smaller than the 1-σ interval as indicated by the instrument manufacturer (see figures B.5 and B.6). Directional measurements deselected.

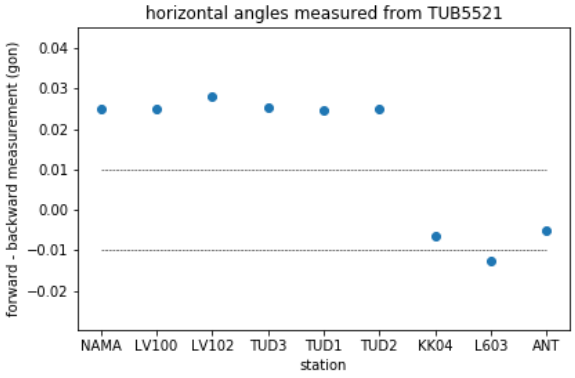


Figure B.4: Differences between forward and backward measurements of directions. In dashed black lines the 2-σ interval as specified in table 2.2 in chapter 2 (0.01 gon). It shows that something happened between TUD2 and KK04 in the forward measurements.

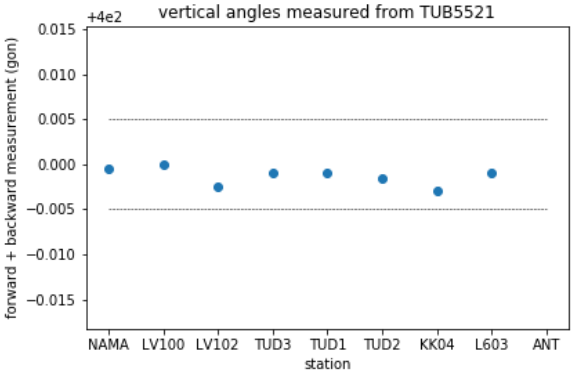


Figure B.5: Summation of forward and backward measurement of vertical angles. In dashed black lines the 1-σ interval as specified in table 2.2 in chapter 2 (0.005 gon). Here no change visible between TUD2 and KK04.

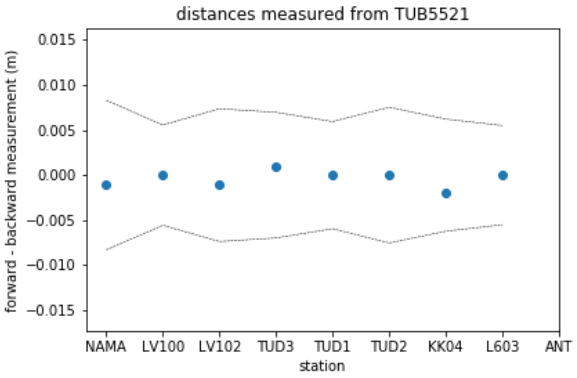


Figure B.6: Difference of forward and backward measurement of distances. In dashed black lines is the 1-σ interval as specified in table 2.2 in chapter 2 (0.005 m + 2 ppm). Here no change visible between TUD2 and KK04.

Forward direction LV100 to TUD4 There is no clear cause for this outlier looking at the difference of the forward and backward measurements of horizontal angles, see figure B.7. Loosening the stochastic model is not an option, because the overall model test shows a value below 1 (0.898), indicating that the stochastic model is not too strict. Forward direction deselected.

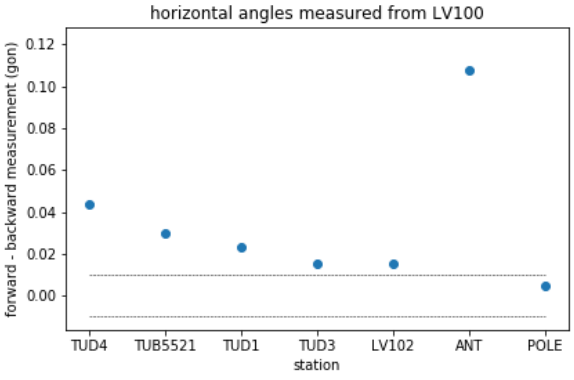


Figure B.7: Differences between forward and backward measurements of directions. In dashed black lines the 2-σ interval as specified in table 2.2 in chapter 2 (0.01 gon). Reason for the decaying difference between forward and backward measurements is not clear. Backward reading of direction from LV100 to ANT has already been removed in an earlier stage.

Forward and backward zenith angle from TUD4 to LV20 Forward and backward zenith angle do agree (summation is 399.999), most likely explanation is that the prism height of LV20 has changed. Zenith angle

measurements are deselected.

Forward direction from TUD3 to POLE POLE is a little pole. The precision of idealization is set to 0.1 m.

Forward direction TUD3 to NAMA Differences in horizontal angles show a lot of variance (see figure B.8). Largest difference to POLE is accounted for in the precision of idealization of station POLE. No clear cause for the variance. No deselection, but geometry changed from 3D to 2D.

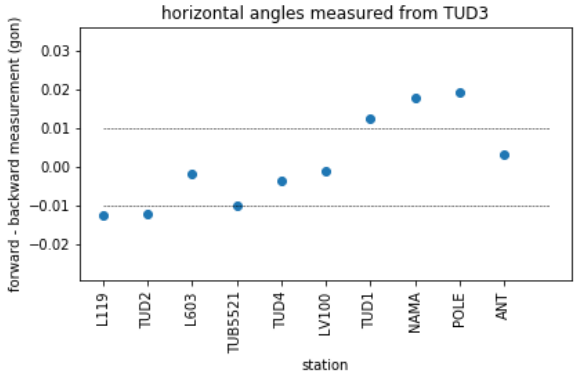


Figure B.8: Differences between forward and backward measurements of directions. In dashed black lines the 2-σ interval as specified in table 2.2 in chapter 2 (0.01 gon). Reason for the difference between forward and backward measurements is not clear.

Forward and backward distances between NAMA and L603 The four measurements between the two stations are within one standard deviation of 0.005 m + 2 ppm. Stochastic model for distance measurements is loosened to 0.06 m + 4 ppm. No detected outliers anymore, 0.989 for F-test.

B.2. Iterations free network adjustment 2016

L603 and LV102 used as base stations, because no measurements to L119 and LV20 is not well located in the network. Station BATH is left out, because it is a large and unclear target.

Measurements related to LV20 Forward distance and vertical angle from TUD4 to LV20 is a typo, values restored from raw data. Measurements to and from LV20, TUD4 and LV102 form a triangle (see figure B.9), angles do not sum up to 200 gon. Angle TUD4-LV20-LV102 is measured 120 gon, visually detected as erroneous. Confirmed by note in raw data of station LV20: "Done in 2 seperate days". Angle LV102-LV20-ANT is correct. Directional measurements from LV20 to TUD4 deselected.

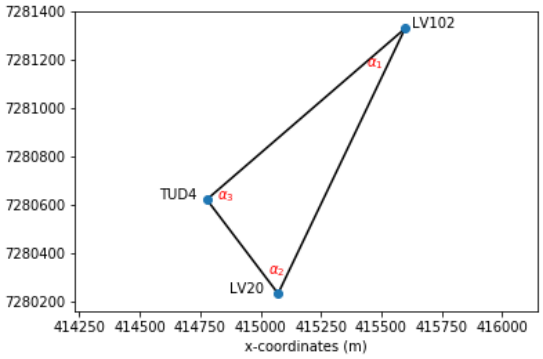


Figure B.9: Aprroximate coordinates of LV102, LV20 and TUD4. Angles making up the triangle are $\alpha_1 \approx 26.2$ gon, $\alpha_2 \approx 120.0$ gon, $\alpha_3 \approx 104.0$ gon, summation is 249.2 gon instead of expected 200 gon. The angle α_2 is wrong, it should be a value of approximately 70.8 gon.

Zenith angles related to L603 Differences between all zenith angles per station pair (see figure B.10 and figure B.11) shows that probably station L603 has been set up again on a different height during the measurement of the network. No deselection, but geometry changed from 3D to 2D.

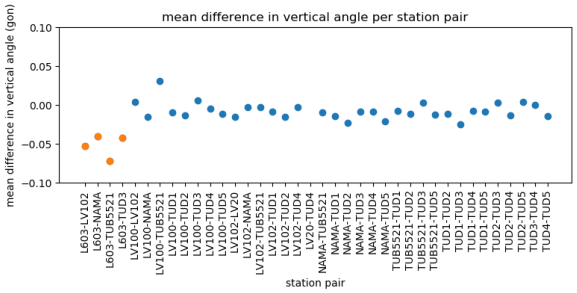


Figure B.10: Difference in vertical angle per station pair has been calculated by comparing the vertical angles measured from A to B with the vertical angles measured from B to A. A positive value means that the highest station has moved downwards with respect to the lowest station. Stations pairs including L603 are shown in orange. Change in vertical angle results in larger vertical displacement if the distance between the two stations is large. Vertical displacement due to the change in vertical angle is shown in figure B.11.

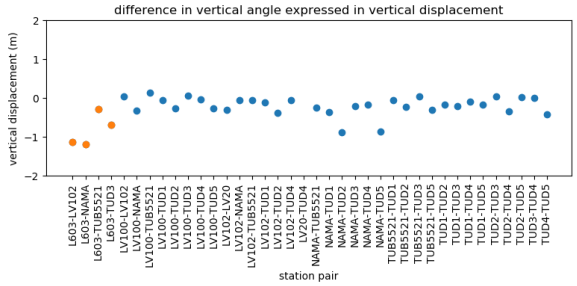


Figure B.11: Vertical displacement due to the change in vertical angle, calculated by $vert\_displacement = slant\_distance \times sin(\alpha)$, where α is the change in vertical angle. A positive value means that the highest station has moved downwards with respect to the lowest station. The change in vertical angle results in a larger vertical displacement if the distance between the two stations is large. Stations pairs including L603 are shown in orange.

Measurements related to LV100 Large difference between forward and backward direction to ANT (see figure B.12). Backward direction to ANT deselected.
Also difference between forward and backward direction to NAMA and to TUD4 (see figure B.13). No influence on vertical angle and distances (see figures B.14 and B.15). Forward and backward direction to NAMA and to TUD4 deselected.
Maybe the station was not level or out-centered, because larger variance of directional measurements and detected outlier in forward and backward distance from TUB5521 to LV100. Precision of idealization set to 0.005 m.

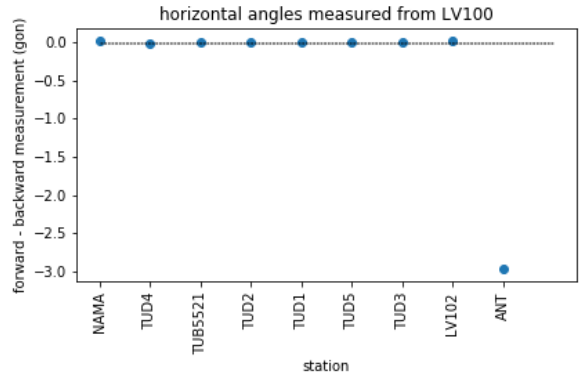


Figure B.12: Differences between forward and backward measurements of directions. In dashed black lines the 2σ interval as specified in table 2.2 in chapter 2 (0.01 gon). Clear problem with measurement to ANT, same data with different scale in figure B.13

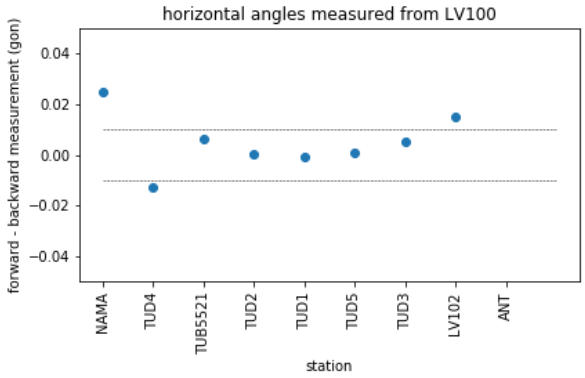


Figure B.13: Differences between the forward and the backward measurements of directions. In dashed black lines the 2-σ interval as specified in table 2.2 in chapter 2 (0.01 gon). Measurements to NAMA, TUD4 and LV102 show large differences.

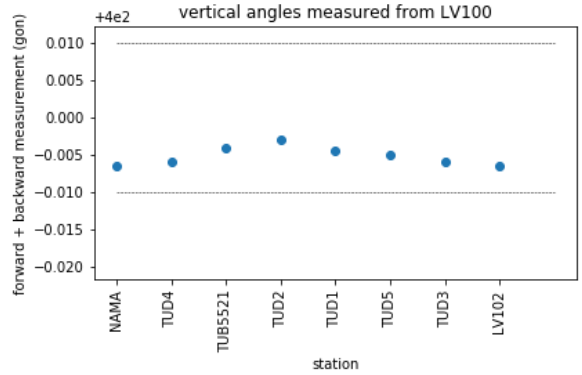


Figure B.14: Summation of forward and backward measurement of vertical angles. In dashed black lines the 2-σ interval as specified in table 2.2 in chapter 2 (0.005 gon). Mean of summations is slightly smaller than 400 gon (see B.2.1)

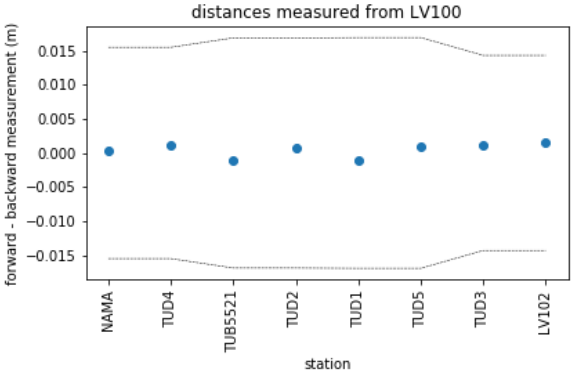


Figure B.15: Difference of forward and backward measurement of distances. In dashed black lines the 2-σ interval as specified in table 2.2 in chapter 2 (0.01 m + 4 ppm).

Stochastic model F-test is 0.608, stochastic model is too loose. Changed to values of instrument used from 2017 on, 0.003 gon for the directions and zenith angles, 0.003 m + 1.5 ppm for the distances. New outliers detected.

Forward direction from TUB5521 to L603 Constant offset in directions between forward and backward measurements except for L603 (see figure B.16), relative angle between L603 and other measurements is influenced. No clear cause. Distances not influenced (see figure B.17). Forward direction deselected.

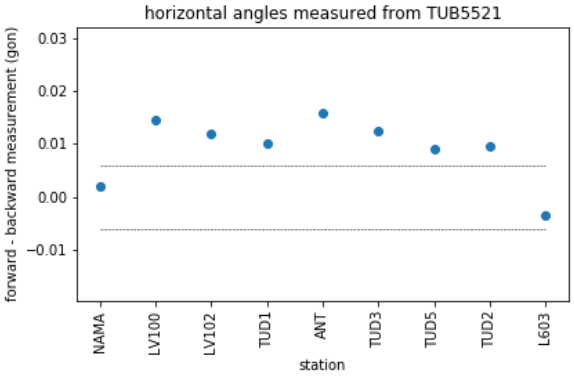


Figure B.16: Differences between forward and backward measurements of directions. In dashed black lines the adjusted 2-σ interval (0.006 gon). L603 lies well in the interval centered around zero, but relative angles are relevant.

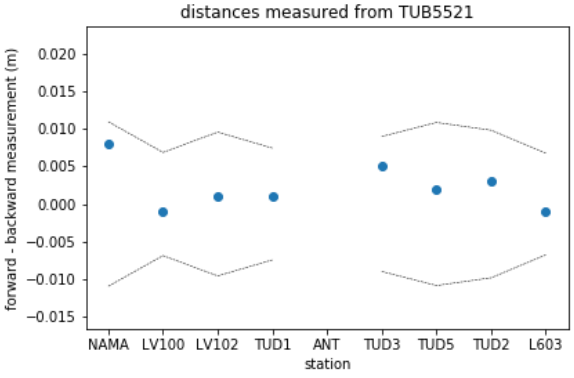


Figure B.17: Differences between forward and backward measurements of distances. In dashed black lines the adjusted 2-σ interval (0.006 m + 3 ppm).

Backward direction from TUD1 to ANT Antenna is actually an antenna to which is measured, not a very thin and easy goal, prone to errors. Precision of idealization set to 0.1 m.

Forward and backward distance from L603 to LV102 Large difference between measurements from L603 to LV102 and measurements from LV102 to L603 (see figure B.18). No clear reason. Forward and backward distance from L603 to LV102 deselected.

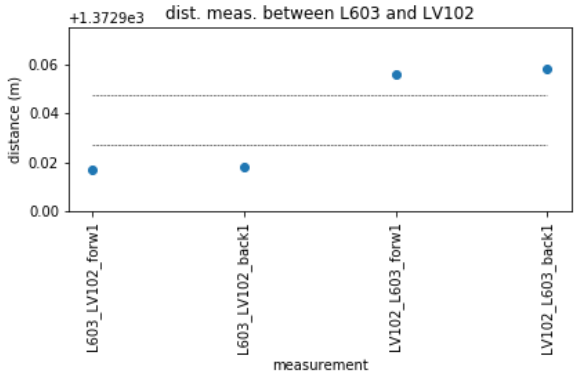


Figure B.18: Slant distance measurements between L603 and LV102. Black dashed line shows adjusted 2-σ adjusted interval (0.006 m + 3 ppm) around the mean.

Forward distance from L603 to NAMA Difference between forward and backward distance measurement from L603 to NAMA is large (see B.19). No clear reason. Forward distance deselected.

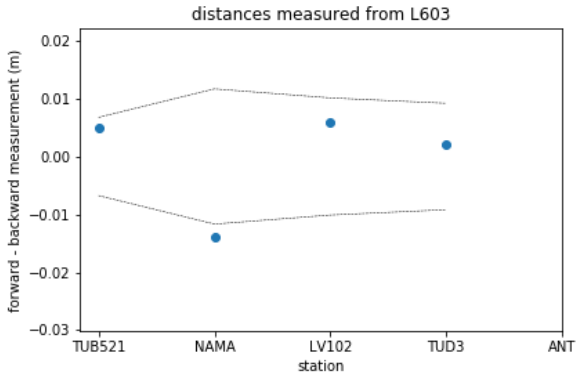


Figure B.19: Difference in distance measurements from L603. Black dashed line shows adjusted 2-σ interval (0.006 m + 3 ppm)

Forward direction from NAMA to TUD4 Differences in horizontal angles between forward and backward measurements show that something happened between TUD1 and TUD2 (see figure B.20). During first measurement cycle (blue) the conditions were rainy, repeated measurement cycle (orange) was performed in sunny conditions, which can be traced back in the differences in distances (see figure B.21), still acceptable. Forward directions to TUD4, TUB5521, L503 and TUD1 deselected.

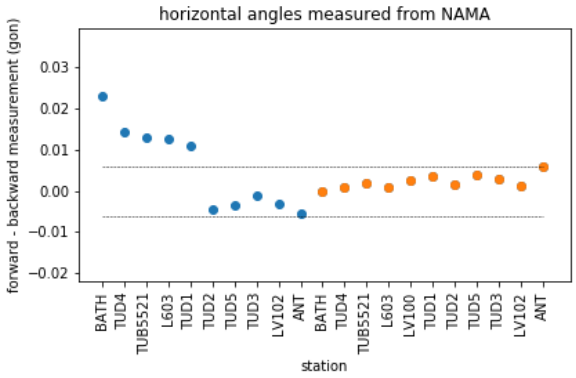


Figure B.20: Differences between forward and backward measurements of directions. In dashed black lines the adjusted 2-σ interval (0.006 gon). The first measurement cycle is indicated in blue, the second in orange. It is clear that something happened between TUD1 and TUD2 in the first measurement cycle.

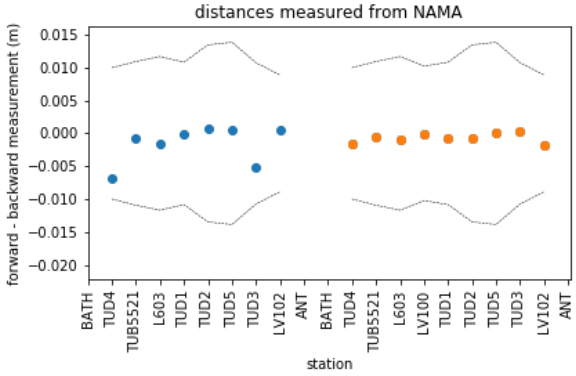


Figure B.21: Differences between forward and backward measurements of distances. In dashed black lines the adjusted 2-σ interval (0.006 m + 3 ppm). The first measurement cycle is indicated in blue, the second in orange. The spread in the first measurement cycle is larger, due to rainy conditions.

Forward direction from LV100 to TUD1 and backward direction from LV100 to TUD2 Looking again at figure B.13, problem seems to be measurement from LV100 to LV102. Forward direction from LV100 to LV102 deselected. No detected outliers anymore, 0.974 for F-test.

B.2.1. Instrumental index error

When making the plots I found it remarkable that the mean difference in vertical angle seems to be slightly under 400 gon, both in figure B.11 and, more clearly visible, in figure B.14. Looking more carefully at all measurements of 2016 it revealed that all sums of vertical angles of the forward and backward measurements are slightly under 400, see figure B.22. This indicates that the instrument has an index error that has not been accounted for. It is possible to account for this in a total station, however, it needs to be calibrated and that has not been done. Also, accounting for it afterwards is possible by using that the error δ can be calculated by $vertical\_angle = 400 + 2\delta$. [1]. This is not done in this case, because the solution is in 2D and therefore the influence on the heights estimates is eliminated and the influence on the distance observations is negligible.

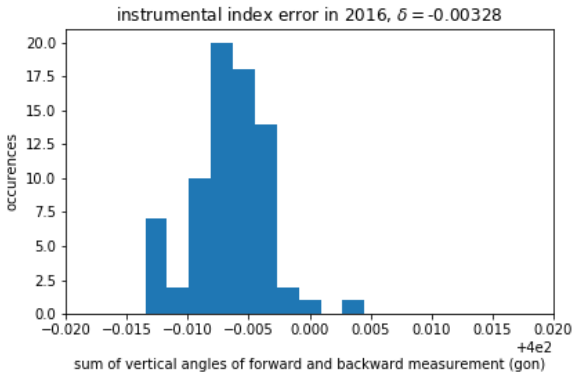


Figure B.22: Histogram plot showing the summations of vertical angles measured in the forward and backward measurements.

B.3. Iterations free network adjustment 2017

Stations KK04 and ANT not used, because not enough measurements to estimate locations.

Second forward distance from NAMA to TUD5_new Notational error, 262.9507 m compared to 2624.0388 m in the backward measurement. Value restored to 2624.0400 m using logsheets.

Forward distance and vertical angle from L102 to NAMA Measurement differs from other measurements of station pair (see figure B.23 and B.24), resulting in different horizontal distance (see figure B.25). Also backward measurement from L102 to NAMA slightly different, but not detected as outlier. Forward distance and vertical angle deselected.

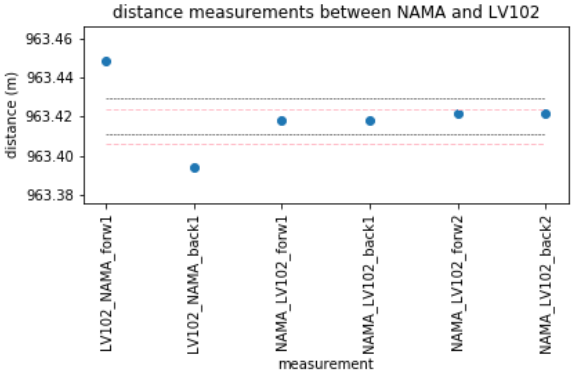


Figure B.23: Slant distance measurements between LV102 and NAMA. Black dashed line shows 2-σ interval around mean, as indicated by the instrument manufacturer (0.006 m + 3ppm). Pink dashed line shows 2-σ interval around mean without detected outlier, forward distance measurement from L102 to NAMA.

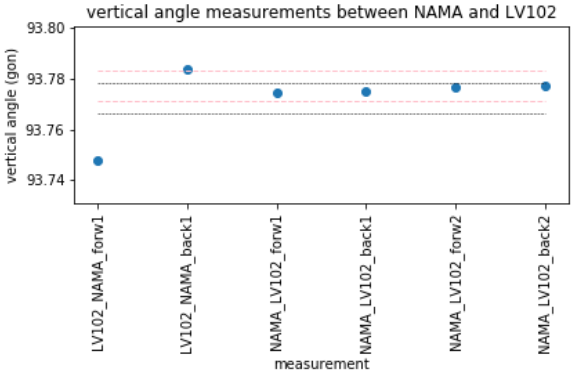


Figure B.24: Vertical angle measurements between LV102 and NAMA. Black dashed line shows 2-σ interval around the mean, as indicated by the instrument manufacturer (0.006 gon). Pink dashed line shows 2-σ interval around the mean without detected outlier, forward vertical angle measurement from L102 to NAMA.

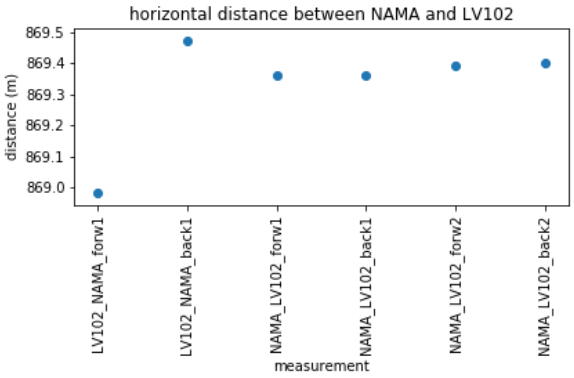


Figure B.25: Horizontal distances between LV102 and NAMA, computed by $horizontal\_distance = slant\_distance \times (1 - cos(\frac{vertical\_angle}{400} 2\pi)$. First value clearly deviates from other values.

Directions from TUD3 Measurments from TUD3 are repeated, presumably because discrepancies between forward and backward measurements were noted when equipment was still there. In first cycle of measurements something happened between TUD1 and LV603 (see figure B.26), vertical angles and distances not influenced (see figures B.27 and B.28). Forward measurements of first cycle to TUD4, LV100, TUB5521 and TUD1 deselected.
In the second cyle of measurements, TUD2 and L119 are detected as outliers. Probably something happened after first three forward measurements and something happened again after first three backward measurements (see figure B.26), vertical angles and distances not influenced (see figures B.27 and B.28). Forward measurements of second cycle to NAMA, TUD4 and LV100 deselected, backward measurements of second cycle to L119, TUD2 and L603 deselected.

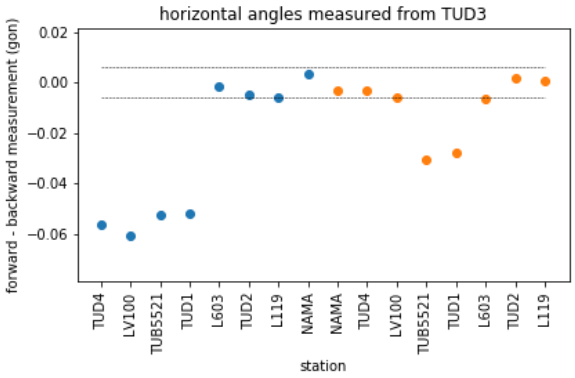


Figure B.26: Differences between forward and backward measurements of directions. In blue the first cycle of measurements, in orange the second. In dashed black lines the 2-σ interval as specified in table 2.2 in chapter 2 (0.006 gon). In the first cycle something happened between TUD1 and L603. In the second cycle something happened during forward measurements after TUD4 and during backward measurements after L603.

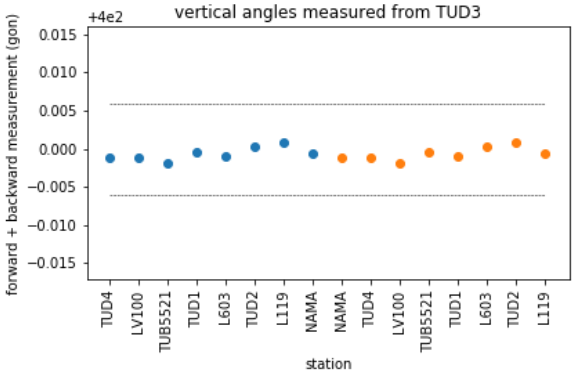


Figure B.27: Summation of forward and backward measurement of vertical angles. In dashed black lines the 2-σ interval as specified in table 2.2 in chapter 2 (0.006 gon).

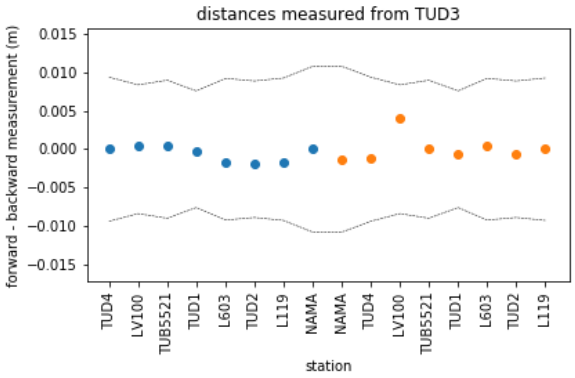


Figure B.28: Difference of forward and backward measurement of distances. In dashed black lines the 2-σ interval as specified in table 2.2 in chapter 2 (0.006 m + 3 ppm).

Distances from and to NAMA Horizontal distances per measurement pair agree better than slant distances (see figure B.29), notably for measurement pairs involving NAMA. Geometry set to 2D.

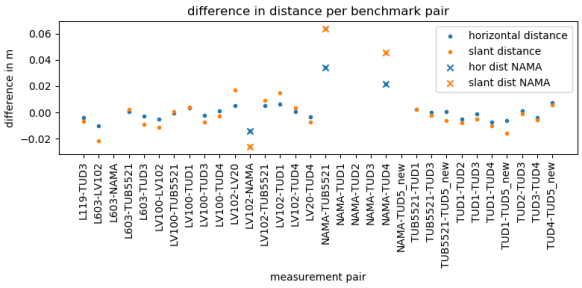


Figure B.29: Difference between mean distance from station A to station B and mean distance from station B to station A, both slant (orange) and horizontal (blue). Foward measurement from LV102 to NAMA is not considered (outlier). Horizontal distance differences are closer to zero than slant distance differences, especially for the measurement pairs involving NAMA (crosses).

Forward direction from TUD2 to TUD3 Variance of difference between forward and backward angle is high (see B.30), vertical angles and distances not influenced (see figure figures B.31 and B.32). No explanation. Forward direction deselected.

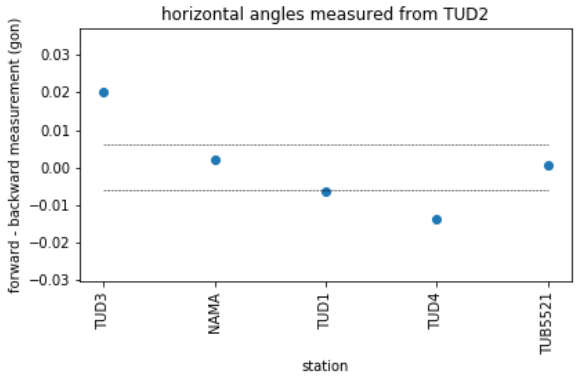


Figure B.30: Differences between forward and backward measurements of directions. In dashed black lines the 2-σ interval as specified in table 2.2 in chapter 2 (0.006 gon). Variance of the measurements is high, no clear reason.

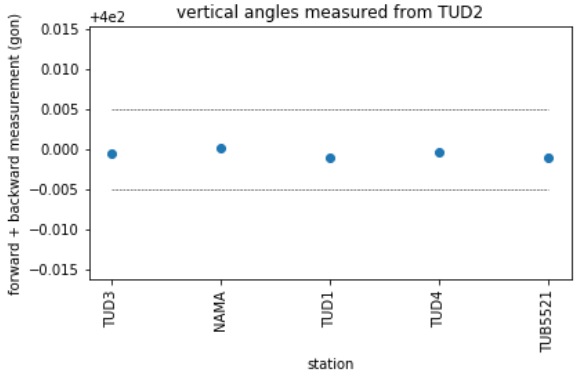


Figure B.31: Summation of forward and backward measurement of vertical angles. In dashed black lines the 2-σ interval as specified in table 2.2 in chapter 2 (0.006 gon).

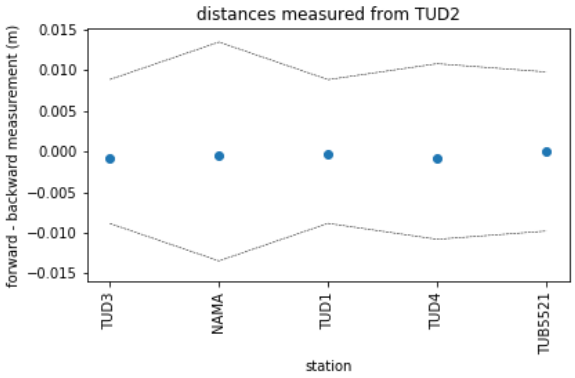


Figure B.32: Difference of forward and backward measurement of distances. In dashed black lines the 2-σ interval (0.006 m + 3 ppm) as specified in table 2.2 in chapter 2.

Directions from LV102 Difference betweeen forward and backward measurements does not show discrepancies (see figure B.33). Stochastic model looser for directions (0.0036 gon).

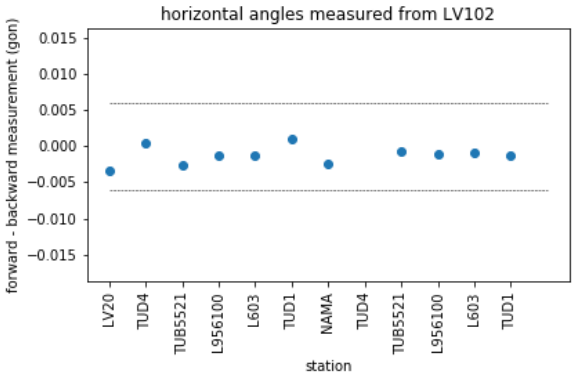


Figure B.33: Differences between forward and backward measurements of directions. In dashed black lines the 2-σ interval as specified in table 2.2 in chapter 2 (0.006 gon).

F-test value too low Stochastic model stricter for distances (0.0010 m + 0.5 ppm). F-test closer to 1, variance componenet analysisis in balance.

Forward direction from LV102 to TUD4 No explanation (see figure B.33). Directional measurement deselected. No detected outliers anymore, 0.862 for F-test.

B.4. Iterations free network adjustment 2018

Backward distance and zenith angle from NAMA to LV100 Probably typo in zenith angle on logsheet (see figures B.34 and B.35). Slant distance is deduced from horizontal distance (as returned by instrument) and vertical angle and therefore also erroneous (see figure B.36). Distance and zenith angle deselected.

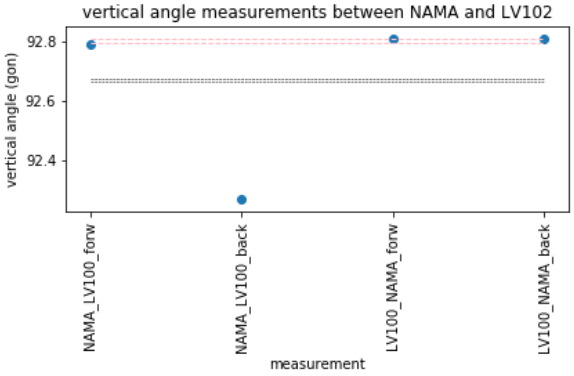


Figure B.34: Vertical angle measurements between NAMA and LV100. Black dashed line shows the 2-σ interval around the mean, as indicated by the instrument manufacturer (0.003 gon). The pink dashed line shows the 2-σ interval around the mean without backward vertical angle measurement from NAMA to LV100.

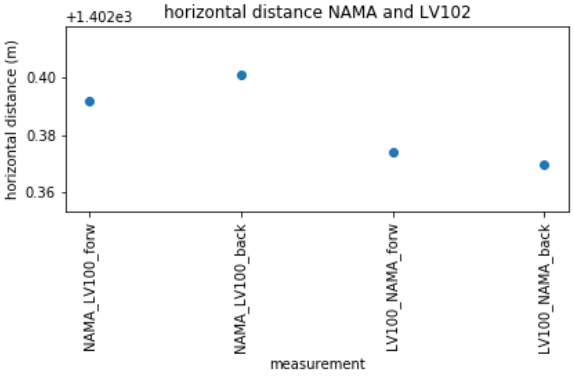


Figure B.35: Horizontal distance measurements between NAMA and LV100, as returned by instrument. Variance of the horizontal distance is rather large, presumably because very windy conditions, distance are still acceptable.

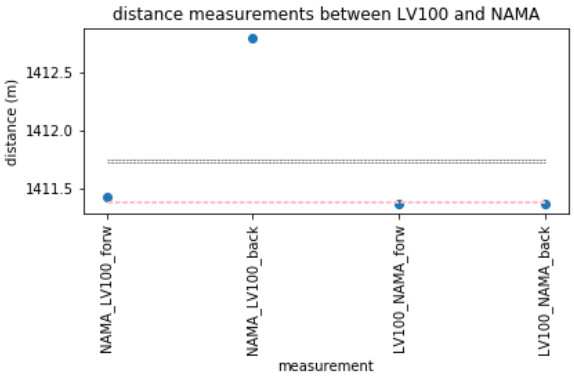


Figure B.36: Slant distance measurements between NAMA and LV100, deduced from the horizontal distance measurements and the vertical angle. Black dashed line shows the 2-σ interval around the mean, as indicated by the instrument manufacturer (0.006 m + 3ppm). The pink dashed line shows the 2-σ interval around the mean without backward vertical angle measurement from NAMA to LV100.

Zenith angles Related to index error as described in appendix B.2.1. Estimated value for δ is large (see figure B.37), 0.00125 gon $\approx$ 42% of σ_{zenith_angle}, which equals a vertical distance of 2 cm when the slant distance is 1000 m. Geometry set to 2D

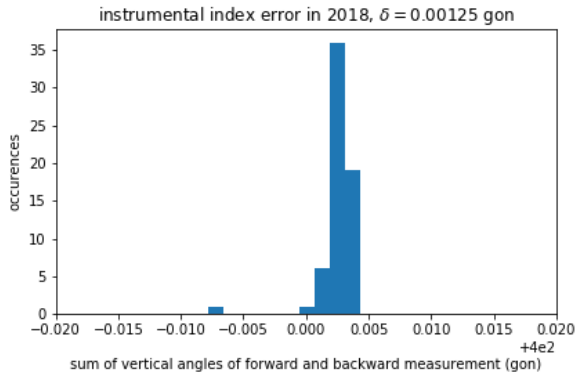


Figure B.37: Histogram plot of summations of vertical angles measured in the forward and backward measurements.

Distance from TUD1 to TUD3 Bias in distance measurements from TUD1 or from TUD3 (see B.38), because distance from TUD1 to TUD3 is detected as outlier, presumably the bias lies there. No explanation. Forward and backward distance measurements from TUD1 to TUD3 deselected.

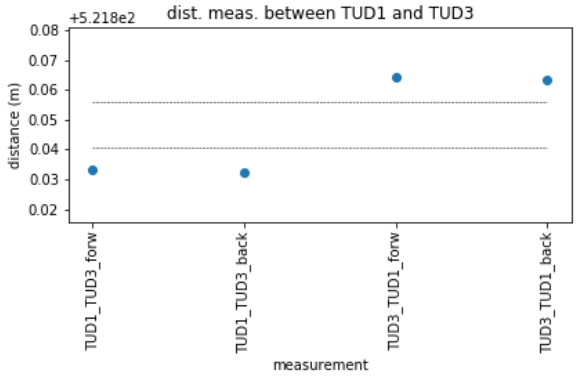


Figure B.38: Distance measurements between TUD1 and TUD3. Clear bias in the first two measurements.

Distances from TUB5521 to TUD1 and to TUD3 Bias in distance measurements from TUB5521 to TUD1 and to TUD3 (see figures B.39 and B.40). Because distance from TUB5521 to TUD1 and to TUD3 are detected as outliers, presumably the bias lies there. No explanation. Forward and backward distance measurements from TUB5521 to TUD1 and to TUD3 deselected.

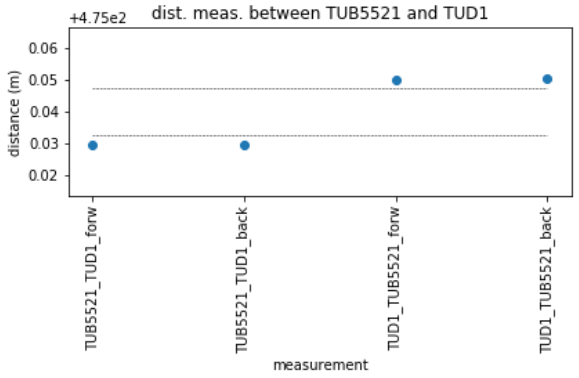


Figure B.39: Distance measurements between TUB5521 and TUD1. Clear bias in the first two measurements.

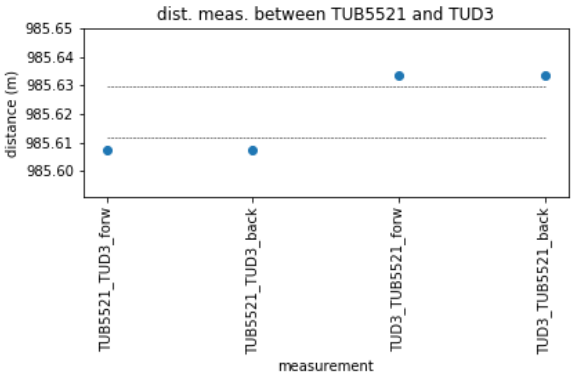


Figure B.40: Distance measurements between TUB5521 and TUD3. Clear bias in the first two measurements.

F-test value too low (0.639) Stochastic model stricter for distances (0.0015 + 0.7 ppm). No detected outliers anymore, 0.895 for F-test.

C

GNSS Data

All GNSS solutions except for the station NAMA are retrieved from the website `http://icelandsupersite.hi.is/gps/ts/NVZ.html`, where the results of analyzed GPS data using the GAMIT/GLOBK 10.6 software at University of Iceland are available. The sites positions were downloaded as geocentric XYZ coordinates in the ITRF2014 reference frame, with no reference plate. They have been converted to UTM28 coordinates using the python interface to PROJ.4. The uncertainties are converted to local NEU coordinates using a matlab script by Hans van der Marel.

The GNSS solutions for NAMA, marked with an asterix in table C.1, are re-processed by TU Delft, using precise point positioning. The geocentric XYZ coordinates and uncertainties in the IGb08 reference frame up to and including 2016 and in the IGS14 reference frame from 2017 have been converted to UTM28 coordinates in the same way as the other GNSS measurements.

Station	year	x (m)	y (m)	z (m)	std_x (m)	std_y (m)	std_z (m)
LV20	2015	415072.7418	7280230.3685	398.7835	0.0024	0.0018	0.0075
LV20	2016	415072.7374	7280230.3920	398.7953	0.0018	0.0013	0.0057
LV20	2017	415072.7358	7280230.4195	398.7776	0.0022	0.0016	0.0071
LV20	2018	415072.7370	7280230.4423	398.7777	0.0026	0.0018	0.0079
L102	2015	415598.1539	7281326.8840	455.8124	0.0022	0.0016	0.0068
L102	2016	415598.1503	7281326.9075	455.8289	0.0017	0.0012	0.0051
L102	2017	415598.1480	7281326.9323	455.8061	0.0023	0.0016	0.0076
L102	2018	415598.1484	7281326.9547	455.8012	0.0023	0.0016	0.0072
L119	2015	414103.4744	7282204.1897	407.4861	0.0018	0.0014	0.0056
L119	2016	414103.4734	7282204.2123	407.4788	0.0014	0.0012	0.0045
L119	2017	414103.4755	7282204.2334	407.4816	0.0017	0.0012	0.0055
L119	2018	414103.4761	7282204.2567	407.4717	0.0019	0.0014	0.0059
L603	2015	414270.8681	7280984.9729	384.8315	0.0020	0.0015	0.0063
L603	2016	414270.8693	7280984.9951	384.8142	0.0018	0.0014	0.0058
L603	2017	414270.8706	7280985.0204	384.8478	0.0023	0.0016	0.0075
L603	2018	414270.8732	7280985.0474	384.8166	0.0026	0.0022	0.0079
BF19	2016	414774.5434	7280967.2079	391.0517	0.0025	0.0017	0.0084
BF19	2017	414774.5441	7280967.2309	391.0735	0.0024	0.0017	0.0078
BF19	2018	414774.5431	7280967.2556	391.0306	0.0021	0.0015	0.0064
NAMA*	2015	416099.8574	7280510.2263	550.2063	0.0012	0.0017	0.0035
NAMA*	2016	416099.8528	7280510.2510	550.2055	0.0009	0.0013	0.0027
NAMA*	2017	416099.8470	7280510.2719	550.2065	0.0012	0.0018	0.0037
NAMA*	2018	416099.8445	7280510.3004	550.1970	0.0012	0.0018	0.0035

Table C.1: GNSS data per year in UTM 28N, gray values are not used. Naming is inconsistent with the tachymetry benchmarks, 'L102' equals 'LV102' and 'BF19' equals 'LV100'