

On the probability distribution of GNSS carrier phase observations

Xiaoguang Luo · Michael Mayer · Bernhard Heck

Received: 25 February 2010 / Accepted: 22 November 2010 / Published online: 23 December 2010
© Springer-Verlag 2010

Abstract When processing observational data from global navigation satellite systems (GNSS), the carrier phase measurements are generally assumed to follow a normal distribution. Although full knowledge of the probability distribution of the observables is not required for parameter estimation, for example when using the least-squares method, the distributional properties of GNSS observations play a key role in quality control procedures, such as outlier and cycle-slip detection, in ambiguity resolution, as well as in the reliability assessment of estimation results. In addition, when applying GNSS positioning under critical observation conditions with respect to multipath and atmospheric effects, the validity of the normal distribution assumption of GNSS observables certainly comes into doubt. This paper illustrates the discrepancies between the normal distribution assumption and reality, based on a large and representative data set of GPS phase measurements covering a range of factors, including multipath impact, baseline length, and atmospheric conditions. The statistical inferences are made using the first through fourth sample moments, hypothesis tests, and graphical tools such as histograms and quantile–quantile plots. The results show clearly that multipath effects, in particular the near-field component, produce the dominant influence on the distributional characteristics of GNSS observables. Additionally, using surface meteorological data, considerable correlations between distributional deviations from normality on the one hand and atmospheric relative humidity on the other are detected.

Keywords GNSS · Residual analysis · Probability distribution · Sample moments · Hypothesis tests

Introduction

Global navigation satellite systems (GNSS) like GPS serve as an efficient and reliable tool for a wide range of geodetic applications in the industrial, commercial, and cadastral sectors. The ongoing system modernization and expansion, as well as new developments in GNSS equipment will significantly enhance the performance of satellite-based positioning techniques. However, the rising demands for more accurate positions and realistic interpretation of the associated quality measures cannot be solely fulfilled by hardware improvements, in light of deficiencies in the mathematical models applied within GNSS data processing.

These mathematical models consist of functional and stochastic components. Although the functional model formulating the mathematical relations between observations and unknown parameters has been investigated in considerable detail, it nevertheless contains deficiencies in modeling site-specific effects and atmospheric influences. In contrast to the highly developed functional model, the stochastic model characterizing the statistical properties of GNSS observations is still a controversial research topic. In fact, the stochastic model is an essential issue in high-accuracy positioning applications and in the realistic quality interpretation of estimated parameters, such as ambiguities (Teunissen et al. 1998; Teunissen 2000), tropospheric delays (Jin and Park 2005; Luo et al. 2008), and site coordinates (Howind 2005, p. 80, 91; Schön and Brunner 2008a). The specification of a realistic stochastic model requires fundamental knowledge about the probability distribution of the observables.

X. Luo (✉) · M. Mayer · B. Heck
Geodetic Institute, Karlsruhe Institute of Technology (KIT),
Englerstr. 7, 76131 Karlsruhe, Germany
e-mail: luo@kit.edu

With respect to ambiguity resolution, probabilistic inferences concerning the integer ambiguity estimators are commonly based on the probability distribution of the original observations. Although many of the results and properties of ambiguity resolution hold true for a larger class of distributions than only Gaussian (Teunissen 1999a), for example, elliptically contoured distributions, the integer rounding, bootstrapping, and least-squares (LS) ambiguity estimators are unbiased if the probability density function of the float ambiguity estimators is symmetric, for instance, a member of the family of multivariate normal distributions (Teunissen 2002; Verhagen and Teunissen 2006). In the context of quality control like outlier and cycle-slip detection, statistical tests performed during GNSS data processing directly rely on the statistical distributions of the observables (Teunissen 1998). The cycle-slip validation procedure presented in Kim and Langley (2001) makes use of the normal distribution assumption of triple differences to characterize the probability distribution of the discrimination test statistic. Hence, in the interest of accurate ambiguity resolution and reliable data quality assessment, a realistic verification of the normality-based statistical model of GNSS observations is indispensable.

The contributions to verify the probability distribution of GNSS observables can be distinguished according to the investigated measures, i.e., either residuals resulting from GNSS data processing or original observations. Using LS residuals at one-second rate, Tiberius and Borre (1999) analyzed the distribution of GPS code and phase observations from a zero, a short (3 m), and an average baseline (13 km) in the field. Evaluating the first through fourth sample moments and applying different statistical hypothesis tests on the empirical distribution function, the normal distribution assumption seemed to be reasonable for the data from the zero and short baselines. However, deviations from normality arose for the 13 km baseline and were attributed to multipath effects and unmodeled differential atmospheric delays. Based on the theory of directional statistics (Mardia and Jupp 1999), Cai et al. (2007) applied the von Mises distribution to describe the distributional properties of the fractional parts of GPS phase double differences. For the data collected from short baselines (2–3 km), the von Mises distribution showed considerably better fitting results in comparison with the Gaussian distribution.

Within the above-mentioned verification studies, zero and short baselines were used in order to mitigate the influences of multipath and differential atmospheric delays. However, under practical circumstances where longer baselines (≥ 30 km) are processed and the impact of these error sources cannot be sufficiently reduced, the real probability distribution of the observables deviates from

the assumed normality. These deviations at the observation level can be retrieved in the residuals resulting from the LS evaluation. This paper illustrates the discrepancies between the Gaussian distribution assumption and reality using representative 1 Hz GPS phase residuals resulting from processing baselines ranging between 30 and 200 km under variable atmospheric conditions. Applying methods from both descriptive and inferential statistics, distributional analyses are performed by means of sample moments, probability plots, and hypothesis tests. The following two sections provide a brief summary of the sample moments and the employed normal distribution tests. Afterward, incorporating surface meteorological data, a case study of statistical inferences is presented. Finally, some concluding remarks provide an outlook on future research work.

Sample moments

The analysis of sample moments does not consider the probability density in detail but only from the perspective of some of its characteristics. For a one-dimensional real-valued random variable X with n independent realizations x_1, \dots, x_n , the r -th central sample moment is given by

$$m_r = \frac{1}{n} \sum_{i=1}^n (x_i - m_1)^r, \quad r \geq 2 \quad (1)$$

where m_1 denotes the sample mean representing an unbiased estimator for the mean μ defined as the expectation of X with $\mu = E(X)$. In the case of $r = 2$, replacing the denominator n by $n - 1$ in Eq. 1 due to the unknown μ , an unbiased estimator s^2 for the variance σ^2 of X given by $\sigma^2 = E((X - \mu)^2)$ can be obtained. Although the normal distribution is completely characterized by μ and σ^2 , the third and fourth standardized moments, i.e., skewness (S) and kurtosis (K), allow the asymmetry and peakedness of the probability density function to be investigated. If $S > 0$ ($S < 0$), then the bulk of the distribution is concentrated on the left (right) part of the probability density function, while a normal distribution has a skewness of zero ($S = 0$), implying a symmetric distribution. Kurtosis measures the degree of peakedness of the probability density of a real-valued random variable. A distribution with a sharp peak around the mean is termed leptokurtic ($K > 3$), a lower peak around the mean with wider tails is referred to as platykurtic ($K < 3$), and the normal distribution is called mesokurtic ($K = 3$). In distribution analysis, the consideration of higher-order moments is necessary, because no probability distribution could be legitimately described as Gaussian normal unless both skewness and excess kurtosis ($K - 3$) are equal to zero (Fiori and Zenga 2009).

Table 1 Overview of the first through fourth central sample moments with the corresponding distributions and confidence intervals

Sample moment	Definition	Distribution	Confidence interval
Mean	$m_1 = \frac{1}{n} \sum_{i=1}^n x_i$	$\frac{\sqrt{n}(m_1 - \mu)}{\sigma} \sim N(0, 1)$	$\left[\mu - \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2}, \mu + \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2} \right]$
Variance	$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - m_1)^2$	$f \cdot \frac{s^2}{\sigma^2} \sim \chi_f^2$	$\left[\frac{\sigma^2 \chi_{2/f}^2}{f}, \frac{\sigma^2 \chi_{1-2/f}^2}{f} \right]$
Skewness	$g_1 = \frac{m_3}{m_2^{3/2}}$	$g_1 \xrightarrow{D} N\left(0, \frac{6}{n}\right)$	$\left[-\sqrt{\frac{6}{n}} z_{1-\alpha/2}, \sqrt{\frac{6}{n}} z_{1-\alpha/2} \right]$
Kurtosis	$g_2 = \frac{m_4}{m_2^2}$	$g_2 \xrightarrow{D} N\left(3, \frac{24}{n}\right)$	$\left[3 - \sqrt{\frac{24}{n}} z_{1-\alpha/2}, 3 + \sqrt{\frac{24}{n}} z_{1-\alpha/2} \right]$

If X is normally distributed with $X \sim N(\mu, \sigma^2)$, the confidence intervals for the sample moments can be analytically derived or asymptotically approximated (Snedecor and Cochran 1989, p. 79; Niemeier 2002, p. 72). Table 1 gives an overview of the first through fourth central sample moments with the corresponding distributions and confidence intervals, where z_p denotes the p -quantile of the standard normal distribution, $\chi_{p,f}^2$ is the p -quantile of the chi-square distribution with f degrees of freedom ($f = n - 1$), and α is the significance level corresponding to the probability of committing a type I error. The estimators g_1 and g_2 for skewness and kurtosis are biased. Additionally, the normal distributions of g_1 and g_2 represent asymptotical approximations. For a more detailed discussion on the sample moments, the reader is referred to textbooks on statistics, such as Wuensch (2005).

Hypothesis tests for normal distribution

A total of five well-known statistical hypothesis tests for normal distribution are employed within this study. The null hypothesis H_0 specifies that the independent samples x_1, \dots, x_n represent n realizations of a normally distributed random variable X with mean μ and variance σ^2 . In the case of unknown values of μ and σ^2 , the sample statistics m_1 and s^2 can be used within the test procedure (Table 1). Under this circumstance, H_0 is referred to as a composite hypothesis. The following text describes the core characteristics, as well as the relative strengths and weaknesses, of the applied hypothesis tests.

The Jarque–Bera (JB) test statistic (Jarque and Bera 1980) provides a goodness-of-fit measure of departures from a normal distribution using the sample skewness and kurtosis (Table 1). As is well known, the sample moments are very sensitive to outliers. Thus, the JB test statistic is sensitive to blunders and extreme observations. Using a robust measure of variance, Gel and Gastwirth (2008) suggested an advanced JB test which is more resistant to outliers and provides equal or higher statistical power than the standard JB test. Furthermore, the chi-square

approximation of the JB test statistic is poorly valid for small sample sizes. This leads to a large rate of wrong rejections of H_0 (type I error). In the MATLAB® Statistics Toolbox™ (MST), a table of critical values computed using a Monte Carlo simulation is applied for $n < 2000$.

The Kolmogorov–Smirnov (KS) test statistic is a single distance measure defined as the supremum of the absolute difference between the empirical and theoretical cumulative distribution functions (CDF) (Chakravarti et al. 1967, p. 392). For small sample sizes ($n \leq 20$), the critical values for the KS test statistic are tabulated in Miller (1956). If $n > 20$, the critical values can be derived using the quantile values of the Kolmogorov distribution given in Teusch (2006, p. 104) or by the analytical approximation implemented in MST. The KS test can only be applied to continuous distributions, and the distribution to be tested must be completely specified. In the case of unknown characteristic parameters such as the mean and variance, the Lilliefors (LF) test is preferred. This test uses exactly the same test statistic as the KS test, but more appropriate critical values computed using a Monte Carlo simulation for instance. More detailed information on the LF test can be found in Lilliefors (1967) and in Abdi and Molin (2007, p. 540).

The chi-square (CS) goodness-of-fit test verifies whether the frequency distribution of an observed sample is consistent with the expected theoretical one (Lehmann and Romano 2005, p. 590). The CS test statistic follows asymptotically a chi-square distribution with $(m - u)$ degrees of freedom, where m denotes the number of bins and u the number of unknown characteristic parameters plus one (e.g., $u = 3$ for a normal distribution). The CS test can be applied to both discrete and continuous distributions. However, the CS test statistic is sensitive to the choice of bins. According to Reißmann (1976, p. 359), m between 10 and 15 with a bin width of approximately $s/2$, where s is the sample standard deviation, seems to be reasonable in practice. Furthermore, the expected counts in each bin should not be less than 5. Therefore, the CS test requires a sufficiently large sample size for reliable test results.

Table 2 Critical values for the test statistics discussed in the text at different significance levels ($n = 3600$)

Hypothesis test	Notation	MATLAB [®] command	Critical value at α			Source
			1%	5%	10%	
Jarque–Bera	JB	jbttest	9.485	5.972	4.552	Analytical approximation in MST
Kolmogorov–Smirnov	KS	kstest	0.027	0.023	0.020	
Lilliefors	LF	lillietest	0.018	0.015	0.014	
Chi-square (e.g., $m = 10$)	CS	chi2gof	18.475	14.067	12.017	
Anderson–Darling	AD	–	1.035	0.752	0.631	Stephens (1986)

The Anderson–Darling (AD) test is based on a weighted (higher weight to the tails) overall distance measure between the empirical and theoretical CDF (Anderson and Darling 1952). For the modified AD test statistic that is adjusted with respect to sample size, the critical values for the hypothesis of normality are given in Stephens (1986, Table 4.9). Although the AD test is restricted to continuous distributions, Stephens (1974) found the AD test statistic to be one of the best empirical CDF statistics for detecting most departures from a normal distribution. Additionally, the AD test makes use of the specified distribution in calculating the critical values. This results in the advantage of allowing more sensitive tests on the one hand, while on the other, the disadvantage of needing to compute individual critical values for each kind of distribution to be tested.

All hypothesis tests discussed above are one-sided right-tailed tests, indicating that the null hypothesis of normality is rejected if the test statistic is larger than the corresponding critical value. In addition, the JB, KS, LF, and CS test statistics are available in MST. As an example, for a sample size of 3600, the critical values at different significance levels are shown in Table 2.

Case study

Using representative GPS phase observations with respect to multipath impact, baseline length, and atmospheric conditions, this section presents a case study dealing with the influences of different factors on the probability distribution of the observables. Due to deficiencies in the mathematical models, the distributional disturbances at the observation level can be observed in the residuals resulting from the GNSS data processing. Taking so-called sidereal lags into account, the residual database is first homogenized and investigated for outliers. Subsequently, distributional assessments are made by evaluating the sample moments, by applying normal distribution tests and by utilizing graphic tools such as histograms and quantile–quantile (Q–Q) plots for visual inspection.

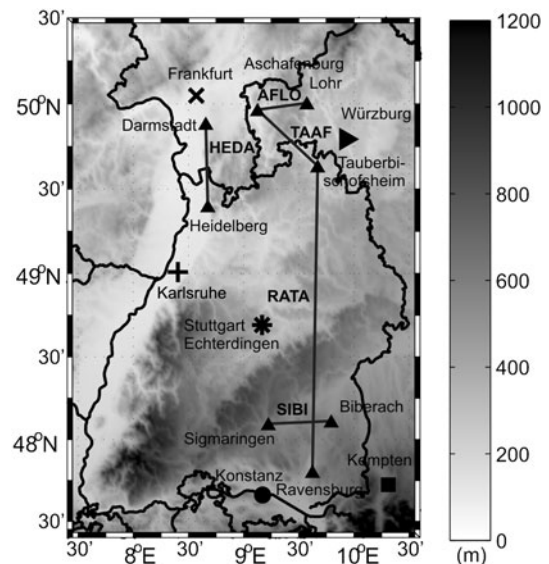


Fig. 1 Selected SAPOS[®] sites (symbol: filled triangle) and the available meteorological stations (symbol: multiplication sign, filled right point triangle, plus sign, asterisk, filled circle, filled square) in the region of investigation with the corresponding relief model (ETOPO1) of the Earth's surface (Amante and Eakins 2009)

Database

A total of 21 days of 1 Hz GPS phase measurements from the SAPOS[®] (Satellite Positioning Service of the German State Survey) network in the area of the state of Baden–Württemberg in the southwest of Germany have been processed using the Bernese GPS Software 5.0 (Dach et al. 2007) in post-processing mode for the selected baselines visualized in Fig. 1. According to the results presented in Knöpfner et al. (2010), the baseline HEDA is the one most strongly affected by multipath effects, and the longest baseline RATA is about 200 km long. In Table 3, some important specifications of the GNSS data processing carried out in this work are listed.

Surface meteorological data, such as air pressure p , temperature T , and relative humidity rh , from the available meteorological stations are used to characterize the near-ground atmospheric conditions during the processing

Table 3 Some important specifications of the GNSS data processing carried out in this work using the Bernese GPS Software 5.0

Parameter	Characteristic
Observations	1 Hz GPS phase double differences
Processing period	DOY2007: 161–181, UT: 15–18 h
Observation weighting	$w = f(\text{SNR}^a)$ (Luo et al. 2008)
Elevation cutoff angle	3 degrees
Satellite orbits, Earth orientation parameters	Precise final IGS ^b products
Ionospheric model for L5 ambiguity resolution	Precise final CODE ^c products
Tropospheric a priori model	Saastamoinen (Saastamoinen 1973)
Mapping function	Niell (dry, wet; Niell 1996)
Time span for tropospheric parameters	15 min
Ambiguity resolution	SIGMA strategy (L5, L3; Dach et al. 2007, p. 175)
Receiver antenna correction	Individual absolute calibration

^a SNR Signal-to-Noise Ratio, ^b IGS International GNSS Service, ^c CODE Center for Orbit Determination in Europe

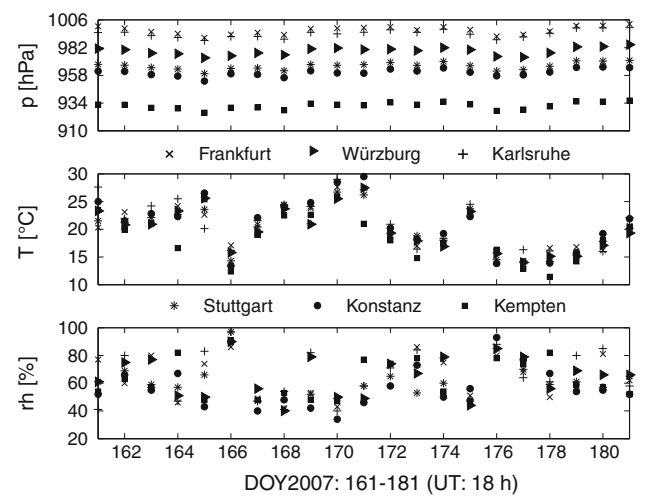


Fig. 2 Selected surface meteorological parameters at 18 h (UT) during the period of investigation (see Fig. 1 for the location of the stations represented by the symbols)

period (Figs. 1, 2). The meteorological data provided by the Deutscher Wetterdienst are free of charge and have a temporal resolution of 6 h. Considering the time span of the GNSS data processing (UT: 15–18 h), the p , T , rh values at 18 h are incorporated into this study. The p values visualized in Fig. 2 show strong dependencies on the sites’ altitudes, and the T , rh data illustrate highly variable atmospheric conditions during the 21-day period. In comparison with T , greater spatial variability in rh is clearly visible. Analyzing the corresponding ionospheric index I95 (Wanninger 2004) and using the ionosphere-free linear combination (L3), the remaining ionospheric effects in the

L3 residuals are negligible. Hence, in the following text, atmospheric effects are mainly attributed to tropospheric influences.

The residuals resulting from processing the daily 3-h GPS data set are related to the same UT time interval. However, due to the sidereal or geometry-repeat lag of approximately 3 min 56 s (236 s), the double difference residuals of a given identical baseline and satellite pair on different days are obtained under dissimilar satellite geometries. In order to guarantee the comparability of the residuals on a daily basis, within this case study the satellite-specific sidereal lags are empirically determined by shifting the satellite azimuth and elevation angles between two consecutive days. Preserving a high reproducibility of a few seconds, the evaluated absolute sidereal lags vary from satellite to satellite between 240 and 263 s. For a given satellite pair, the arithmetical mean of the related sidereal lags is used to specify the time windows, so that from the 3-h residual database, a total of 285 1 h or 3600-epoch residual time series with almost identical satellite geometry over the entire 21-day period are extracted for the subsequent distribution analysis.

As mentioned before, outliers degrade the performance of statistical inferences. Thus, in the primary stage of the distribution analysis, outliers in the 1-h residual database must be detected and appropriately handled. Under the assumption of independent and identically distributed LS residuals v_i , the studentized double difference residuals (SDDR; Cook and Weisberg 1982, p. 18) given by

$$r_s(i) = \frac{v_i}{\hat{\sigma}_0 \sqrt{q_{vv}(i, i)}}, \quad i = 1, \dots, n \quad (2)$$

are homoscedastic with a constant variance of 1 (Howind 2005, p. 39) and follow Pope’s τ -distribution (Pope 1976), where $\hat{\sigma}_0$ denotes the a posteriori standard deviation and $q_{vv}(i, i)$ are the diagonal elements in the cofactor matrix of \mathbf{v} . Due to the extremely large degree of freedom induced by the high redundancy of GNSS data processing in static mode, Pope’s τ -distribution can be well approximated by Student’s t -distribution, which approaches the standard normal distribution (Heck 1981). The outlier detection is performed using the 3-sigma strategy. As a result, 172 SDDR time series possess samples beyond the 3-sigma region with $|r_s(i)| \geq 3$. Based on the fact that the sample variance is more sensitive to outliers than the sample mean, the F -test is employed to assess the impact of the detected 3-sigma outliers on the sample variance (Niemeier 2002, p. 91). Figure 3a illustrates the test results at $\alpha = 1\%$. For most 1 h SDDR time series, the identified 3-sigma outliers appear to insignificantly affect the sample variance. For the case of time series whose sample variances are significantly influenced (black dots in Fig. 3a), high correlation between the F -test statistic (thick gray line) and the number

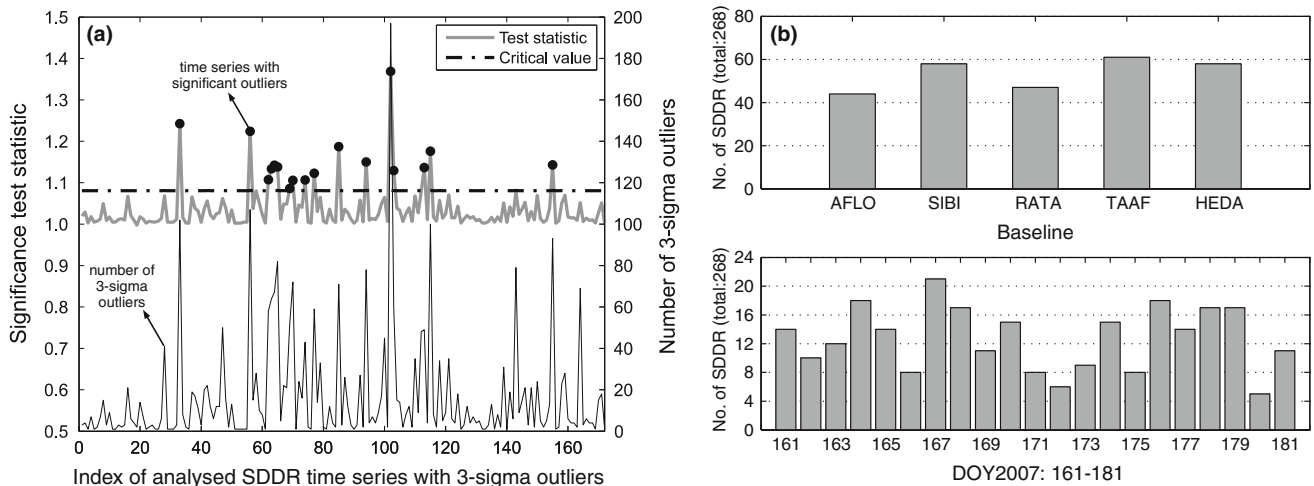


Fig. 3 Results of outlier detection (a) and final residual data distribution with respect to baseline and day of year (DOY) (b)

of 3-sigma outliers (thin black line) is obviously present. Nearly half of the identified 17 SDDR time series with significant outliers are related to the longest baseline RATA. These complete 17 time series (i.e., not only the identified single 3-sigma outliers) are excluded from the following distribution analysis, leading a total of 268. In Fig. 3b, the spatial and temporal distributions of the final 1 h SDDR database are displayed in the form of histograms.

Distribution analysis using sample moments

Using the formulas given in Table 1 and substituting $\mu = 0$ and $\sigma^2 = 1$ for the theoretical mean and variance, the sample moments and the associated confidence intervals of SDDR are computed for $\alpha = 1\%$. Regarding the remaining atmospheric effects in SDDR as random, the baseline-related average sample moments over the whole period of investigation are analyzed to highlight the influences of multipath (MP) and baseline length. Table 4 gives the baseline-related statistical characteristics for the first through fourth sample moments, such as arithmetical mean, standard deviation (STD), and percentage (absolute number) of the SDDR time series whose sample moments are located within the corresponding confidence bounds.

In comparison with other baselines, the HEDA and RATA sample means possess larger biases, which can be interpreted as being due to unmodeled multipath effects and the remaining differential atmospheric delays, which, for long baselines, cannot be sufficiently eliminated by differencing. Furthermore, the large variations in the sample moments and the considerable deviations in the average sample variance from the theoretical value one indicate the unrealistic nature of assuming independent GNSS observations, with the largest deviations from the

unit variance found for the shortest baseline AFLO. Being less critical compared to the sample mean and variance, all baseline-related averages of the sample skewness and kurtosis are within the corresponding confidence intervals. Larger variations in the skewness and kurtosis are detected in the AFLO and HEDA samples. The AFLO-related mean kurtosis is slightly larger than three, which indicates a leptokurtic probability density with a sharper peak and heavier tails than a normal distribution. Nonetheless, the null hypothesis of normality cannot be rejected for nearly half of the data at $\alpha = 1\%$.

Distribution analysis using hypothesis tests

In order to cope with the deficiencies in the mathematical models applied within GNSS data processing, the normal distribution tests are employed in the composite hypothesis case using sample statistics m_1 and s^2 (Table 1). Considering the residual database as a whole, Fig. 4a compares the non-rejection rates of normality at different significance levels. Due to the inappropriate critical values for the case of composite hypotheses, the KS test provides obviously over-optimistic results that can be effectively corrected by the LF test (Table 2). Additionally, except for the KS test, the other four statistical tests deliver generally consistent outputs. Looking at the test decision for each SDDR time series at $\alpha = 1\%$ displayed in Fig. 4b, broad similarities can also be found. In consequence of the inadequate results, the KS test is excluded from further discussion in this work.

Considering aspects of baseline length and multipath impact, Fig. 5 provides the baseline-related presentation of the test results at a significance level of $\alpha = 1\%$, where the absolute number of SDDR data series that cannot be rejected as normally distributed are displayed. The influence of

Table 4 Baseline-related statistical characteristics for the first through fourth sample moments of the SDDR time series ($\alpha = 1\%$)

Baseline	AFLO	SIBI	TAAF	HEDA	RATA	All baselines
Length [km]	32.4	42.5	53.7	54.1	203.7	
MP impact	Weak	Weak	Weak	Strong	Weak	
Number of SDDR series	44	58	61	58	47	268
Sample mean						
Mean	0.02	0.03	0.00	-0.05	0.04	0.00
STD	0.36	0.28	0.27	0.27	0.25	0.29
[-0.04, 0.04]	14% (6)	7% (4)	20% (12)	12% (7)	11% (5)	13% (34)
Sample variance						
Mean	0.64	0.76	0.81	0.81	0.75	0.76
STD	0.25	0.29	0.30	0.29	0.28	0.29
[0.94, 1.06]	7% (3)	10% (6)	16% (10)	16% (9)	11% (5)	12% (33)
Sample skewness						
Mean	0.02	0.02	0.00	-0.01	0.03	0.01
STD	0.32	0.22	0.18	0.32	0.20	0.25
[-0.11, 0.11]	36% (16)	48% (28)	44% (27)	19% (11)	53% (25)	40% (107)
Sample kurtosis						
Mean	3.12	2.85	2.90	2.93	2.88	2.93
STD	0.46	0.31	0.26	0.50	0.29	0.38
[2.79, 3.21]	36% (16)	43% (25)	59% (36)	38% (22)	57% (27)	47% (126)

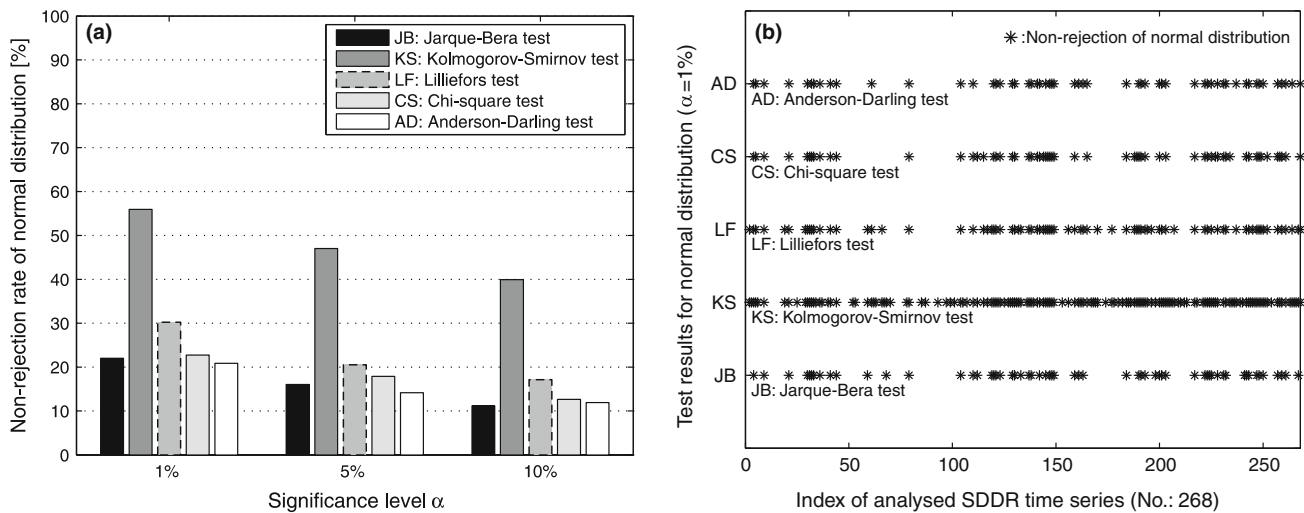


Fig. 4 Comparison of the test results applying different normal distribution tests. **a** The non-rejection rate of normal distribution as found by each hypothesis test, **b** SDDR time series for which the null hypothesis of normality cannot be rejected

baseline length on the probability distribution is analyzed by comparing the AFLO- and SIBI-related results with those of RATA. For the longer baseline RATA, even more SDDR data cannot be rejected as normally distributed. This phenomenon may be explained by the relatively weakly correlated GNSS observations due to the large baseline length. For long baselines, the assumption of independent observables made in the stochastic model is better fulfilled. In comparison with the baseline length, the influence of

different multipath effects on the test results is clearly stronger. Comparing the baselines TAAF and HEDA which have similar baseline lengths, significant disturbances in the distributional properties due to increased multipath effects can be postulated.

Multipath effects, as a major factor limiting GNSS positioning quality, can be subdivided into far-field and near-field components. Within the context of probabilistic distribution, far-field multipath introducing short-periodic

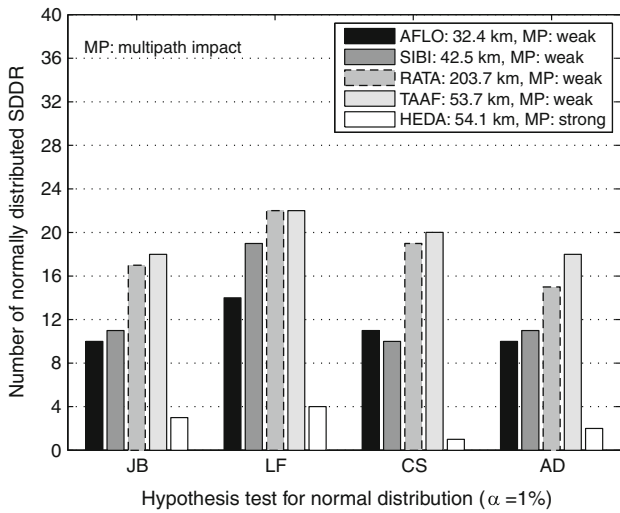


Fig. 5 Baseline-related presentation of the test results dealing with baseline length and multipath impact ($\alpha = 1\%$)

(up to half an hour; Seeber 2003, p. 317) and zero mean effects may primarily affect the distribution form, while near-field multipath showing long-periodic (up to several hours; Wübbena et al. 2006) and non-zero mean characteristics could interfere with both location and form parameters. Applying histograms and Q–Q plots, representative examples of SDDR time series are presented in Fig. 6. The elevation angles of the satellite pair PRN 18, 26 vary between 20° and 30° during the 1 h time span. DOY 170 in 2007, which had the lowest atmospheric variability, is selected to minimize the influences of differential atmospheric delays (Fig. 2).

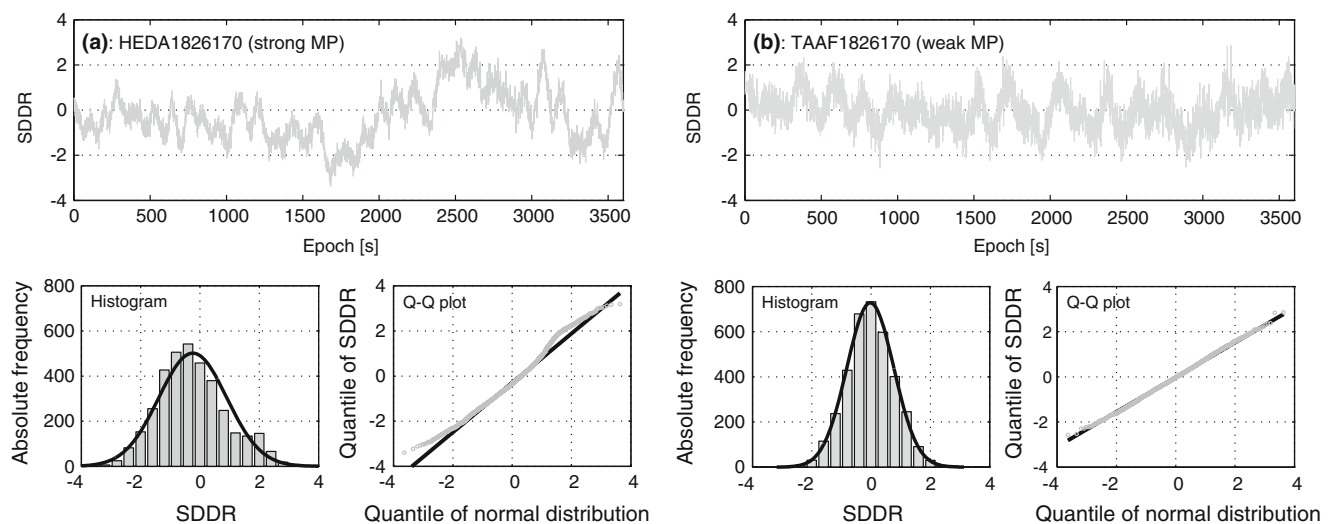


Fig. 6 Influences of increased multipath effects on the probability distribution. **a** HEDA-related example with strong MP, **b** TAAF-related example with weak MP; upper graphics SDDR time series,

lower graphics empirical distribution in gray, hypothesized normal distribution in black

For static receivers, the period (frequency) of multipath errors is inversely (proportionally) related to the antenna–reflector distance, i.e., near-field reflections cause slowly varying errors, whereas distant reflectors create rapidly oscillating quasi-periodic signals (Georgiadou and Kleusberg 1988; Wanninger 2000, p. 23). As Fig. 6a shows, the HEDA-related SDDR time series exhibits stronger near-field multipath effects which considerably affect both location and form of the corresponding histogram. The deviations from normality of the HEDA-related data and the validity of normality for the TAAF-related example can be assessed conveniently based on the corresponding Q–Q plots. However, obvious quasi-periodic signals with periods of several minutes are visible in the TAAF-related SDDR time series (Fig. 6b), indicating the presence of far-field multipath effects. Compared to zero mean far-field multipath, near-field multipath tends to play a more important role in distributional deviations of GNSS observations.

In order to validate the relationship between distributional deviations and variable atmospheric conditions, surface meteorological data are incorporated into this case study. According to Hopfield (1969), the tropospheric delay of GNSS signals can be subdivided into a dry and a complementary wet component. In contrast to the dry delay term, the wet component is very difficult to handle due to the high temporal and spatial variability of atmospheric water vapor. Thus, atmospheric humidity has a large potential to disturb the probability distribution of the observables. In Fig. 7, the daily median test statistics are plotted together with the corresponding mean atmospheric relative humidity.

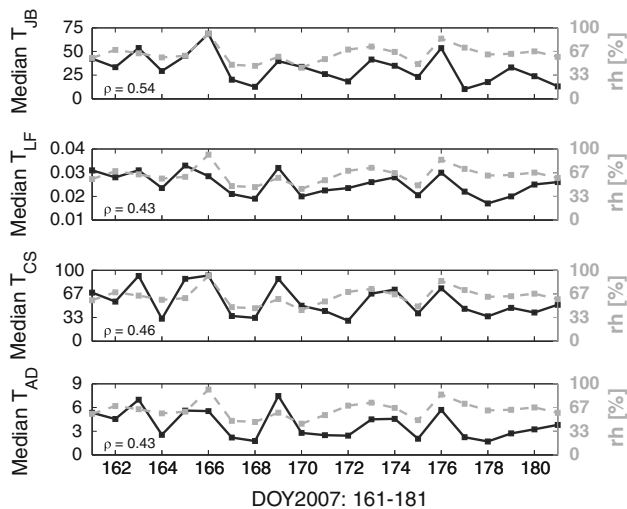


Fig. 7 Comparison of the daily median test statistics (*black line*) with the corresponding mean relative humidity values at 18 h (*dashed gray line*)

All normal distribution tests applied within this study are one-sided. Therefore, larger test statistics imply stronger departures from normality. As Fig. 7 illustrates, the daily median test statistics are obviously positively correlated with the corresponding mean atmospheric relative humidity (correlation coefficient of approximately 0.5), which indicates the considerable impact of wet atmosphere on the probability distribution of GNSS observables. Such a finding provides a motivation for improved tropospheric modeling, particularly for the wet delay component, by using, for example, representative meteorological data with high temporal and spatial resolution, as well as advanced mapping functions (Boehm et al. 2006).

Concluding remarks

This paper illustrates the influences of different factors on the probability distribution of GNSS observables by analyzing representative studentized double difference residuals (SDDR) of GPS phase observations. The consideration of satellite-specific sidereal lags and the employment of variance-based outlier detection guarantees consistency and high quality in the data. Apart from GNSS residuals, surface meteorological data are incorporated to characterize the prevailing atmospheric conditions. Statistical inferences are made using sample moments, hypothesis tests, and probabilistic plots.

The sample mean and variance illustrates considerably large variations and more significant deviations from normality than the sample skewness and kurtosis. Applying hypothesis tests, only approximately 20% of the SDDR data cannot be rejected as normally distributed at a

significance level of $\alpha = 1\%$. In addition, when incorporating surface meteorological data, considerable correlations between distributional discrepancies and relative humidity are detected.

The strong distributional deviations of SDDR from normality can be attributed to the deficiencies in modeling multipath effects and differential atmospheric delays, as well as to the unrealistic assumption of independent observations. The correlation behavior of GNSS phase measurements is affected by both multipath effects (Nahavandchi and Joodaki 2010) and prevailing atmospheric conditions (Schön and Brunner 2008b). Multipath, in particular the long-periodic and non-zero mean near-field component, significantly disturbs the distributional properties of the observables. Long baselines have the disadvantage of non-zero mean SDDR due to the remaining differential atmospheric delays, but the advantage of weakly correlated observations because of large separation distances. Nevertheless, the relatively high standard deviations in all sample moments and the considerably large biases in the sample variance indicate the unrealistic assumption of independence made in the stochastic model. Within this case study, the residuals are obtained after the ambiguities are resolved. Teunissen (1999b) stated that the ambiguity-resolved real-valued parameter estimators are not Gaussian, even if the data are Gaussian. This important aspect motivates further statistical inferences based on residuals produced prior to ambiguity resolution.

Future research work will focus on improving the mathematical models of GNSS observations to substantiate the physical causes for distributional deviations illustrated above. To mitigate site-specific influences, in particular multipath effects, sidereal filtering and stacking techniques will be employed (Choi et al. 2004; Larson et al. 2007). In order to cope with the deficiencies in the stochastic model, Luo et al. (2010) proposed autoregressive moving average (ARMA) processes to characterize the temporal correlations of GNSS observations. Within the framework of the international project GURN (GNSS Upper Rhine Graben Network), larger data sets will be available for future distribution analyses. Moreover, high-resolution meteorological data will be incorporated into GURN to produce reliable four-dimensional atmospheric water vapor fields (Knöpfler et al. 2010).

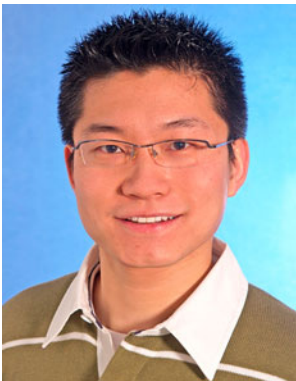
Acknowledgments We would like to thank the state survey office of Baden-Württemberg for providing the GNSS data and absolute antenna calibration values. The German Research Foundation (DFG) is gratefully acknowledged for supporting the research project “Improving the stochastic model of GPS observations by modeling physical correlations”. We also appreciate very much the professional comments from two anonymous reviewers as well as the valuable suggestions from the editorial office.

References

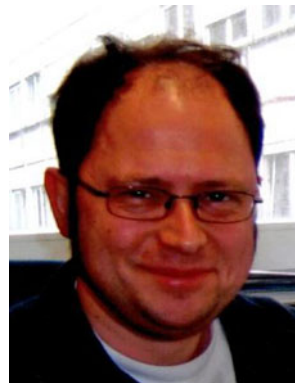
- Abdi H, Molin P (2007) Lilliefors test of normality. In: Salkind NJ (ed) *Encyclopedia of measurement and statistics*, vol 2. Thousand Oaks, CA, USA
- Amante C, Eakins BW (2009) ETOPO1 1 Arc-Minute Global Relief Model: procedures, data sources and analysis. NOAA Technical Memorandum NESDIS NGDC-24. Boulder, CO, USA
- Anderson TW, Darling DA (1952) Asymptotic theory of certain “goodness-of-fit” criteria based on stochastic processes. *Ann Math Stat* 23:193–212. doi:10.1214/aoms/1177729437
- Boehm J, Werl B, Schuh H (2006) Troposphere mapping functions for GPS and very long baseline interferometry from European Centre for Medium-Range Weather Forecasts operational analysis data. *J Geophys Res* 111:B02406. doi:10.1029/2005JB003629
- Cai J, Grafarend E, Hu C (2007) The statistical property of the GNSS carrier phase observations and its effects on the hypothesis testing of the related estimators. In: *Proceedings of ION GNSS 2007*, Fort Worth, TX, USA, Sept 25–28, 2007, pp 331–338
- Chakravarti IM, Roy J, Laha RG (1967) *Handbook of methods of applied statistics*, vol 1. Wiley, New York
- Choi K, Bilich A, Larson KM, Axelrad P (2004) Modified sidereal filtering: implications for high-rate GPS positioning. *Geophys Res Lett* 31:L22608. doi:10.1029/2004GL021621
- Cook R, Weisberg S (1982) *Residuals and influence in regression*. Chapman and Hall, New York
- Dach R, Hugentobler U, Fridez P, Meindl M (2007) *Bernese GPS Software Version 5.0*. Astronomical Institute, University of Berne, Berne
- Fiori AM, Zenga M (2009) Karl Pearson and the origin of kurtosis. *Int Stat Rev* 77(1):40–50. doi:10.1111/j.1751-5823.2009.00076.x
- Gel YR, Gastwirth JL (2008) A robust modification of the Jarque-Bera test of normality. *Econom Lett* 99:30–32. doi:10.1016/j.econlet.2007.05.022
- Georgiadou Y, Kleusberg A (1988) On carrier signal multipath effects in relative GPS positioning. *Man Geod* 13:172–179
- Heck B (1981) Der Einfluß einzelner Beobachtungen auf das Ergebnis einer Ausgleichung und die Suche nach Ausreißern in den Beobachtungen. *Allgemeine Vermessungs-Nachrichten (AVN)* 88:17–34
- Hopfield H (1969) Two-quartic tropospheric refractivity profile for correcting satellite data. *J Geophys Res* 74(18):4487–4499. doi:10.1029/JC074i018p04487
- Howind J (2005) *Analyse des stochastischen Modells von GPS-Trägerphasenbeobachtungen*. Deutsche Geodätische Kommission, DGK C584, Munich, Germany
- Jarque CM, Bera AK (1980) A test for normality of observations and regression residuals. *Int Stat Rev* 55(2):163–172
- Jin SG, Park PH (2005) A new precision improvement in zenith tropospheric delay estimation by GPS. *Current Science* 89(6):997–1000
- Kim D, Langley RB (2001) Quality control techniques and issues in GPS applications: Stochastic modelling and reliability testing. In: *Proceedings of tutorial and domestic session, international symposium on GPS/GNSS*, Jeju, Korea, Nov 7–9, 2001, pp 76–85
- Knöpfler A, Masson F, Mayer M, Ulrich P, Heck B (2010) GURN (GNSS Upper Rhine Graben Network)—status and first results. FIG congress 2010, facing the challenges—building the capacity, Sydney, Australia, April 11–16, 2010
- Larson KM, Bilich A, Axelrad P (2007) Improving the precision of high-rate GPS. *J Geophys Res Solid Earth* 112:B05422. doi:10.1029/2006JB004367
- Lehmann EL, Romano JP (2005) *Testing statistical hypotheses*, 3rd edn. Springer, New York
- Lilliefors HW (1967) On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *J Am Stat Ass* 62:399–402
- Luo X, Mayer M, Heck B (2008) Improving the stochastic model of GNSS observations by means of SNR-based weighting. In: Petrov BN, Csaki F (eds) *Observing our changing Earth. Proceedings of the 2007 IAG general assembly*, Perugia, Italy, July 2–13, 2007, IAG Symposia, vol 133, pp 725–734. doi:10.1007/978-3-540-85426-5_83
- Luo X, Mayer M, Heck B (2010) Analysing time series of GNSS residuals by means of AR(1)MA processes. In: *Proceeding of VII Hotine-Marussi symposium*, Rome, Italy, July 6–10, 2009, IAG Symposia (in print)
- Mardia KV, Jupp PE (1999) *Statistics of directional data*, 2nd edn. Wiley, New York
- Miller LH (1956) Table of percentage points of Kolmogorov statistics. *J Am Stat Ass* 51(273):111–121
- Nahavandchi H, Joodaki G (2010) Correlation analysis of multipath effects in GPS-code and carrier phase observations. *Surv Rev* 42(316):193–206. doi:10.1179/003962610X12572516251808
- Niell AE (1996) Global mapping functions for the atmosphere delay at radio wavelengths. *J Geophys Res* 101:3227–3246. doi:10.1029/95JB03048
- Niemeier W (2002) *Ausgleichsrechnung*. Walter de Gruyter, Berlin
- Pope AJ (1976) *The statistics of residuals and the detection of outliers*. NOAA Technical Report NOS 65 NGS 1, Rockville, MD
- Reißmann G (1976) *Die Ausgleichsrechnung*, 5th edn. VEB Verlag für Bauwesen, Berlin
- Saastamoinen J (1973) Contribution to the theory of atmospheric refraction. *Bull Geod* 107(1):13–34. doi:10.1007/BF02522083
- Schön S, Brunner FK (2008a) A proposal for modelling physical correlations of GPS phase observations. *J Geod* 82(10):601–612. doi:10.1007/s00190-008-0211-3
- Schön S, Brunner FK (2008b) Atmospheric turbulence theory applied to GPS phase data. *J Geod* 82(1):47–57. doi:10.1007/s00190-007-0156-y
- Seeber G (2003) *Satellite geodesy*, 2nd edn. Walter de Gruyter, Berlin
- Snedecor GW, Cochran WG (1989) *Statistical methods*, 8th edn. Iowa State University Press, Ames
- Stephens MA (1974) EDF statistics for goodness of fit and some comparisons. *J Am Stat Ass* 69:730–737. doi:10.2307/2286009
- Stephens MA (1986) Test based on EDF statistics. In: D’Agostino R, Stephens M (eds) *Goodness-of-fit techniques*. Vol 68 of *Statistics: textbooks and monographs*. Marcel Dekker Inc., New York
- Teunissen PJG (1998) Quality control and GPS. In: Teunissen PJG, Kleusberg A (eds) *GPS for Geodesy*, chapter 7, 2nd edn, Springer, Berlin
- Teunissen PJG (1999a) An optimality property of the integer least-squares estimator. *J Geod* 73(11):587–593. doi:10.1007/s001900050269
- Teunissen PJG (1999b) The probability distribution of the GPS baseline for a class of integer ambiguity estimators. *J Geod* 73(5):275–284. doi:10.1007/s001900050244
- Teunissen PJG (2000) The success rate and precision of GPS ambiguities. *J Geod* 74(3–4):321–326. doi:10.1007/s001900050289
- Teunissen PJG (2002) The parameter distributions of the integer GPS model. *J Geod* 76(1):41–48. doi:10.1007/s001900100223
- Teunissen PJG, Jonkman NF, Tiberius CCJM (1998) Weighting GPS dual frequency observations: Bearing the cross of cross-correlation. *GPS Solut* 2(2):28–37. doi:10.1007/PL00000033
- Teusch A (2006) *Einführung in die Spektral- und Zeitreihenanalyse mit Beispielen aus der Geodäsie*. Deutsche Geodätische Kommission, DGK A120, Munich, Germany

- Tiberius C, Borre K (1999) Probability distribution of GPS code and phase data. *Zeitschrift für Vermessungswesen (ZfV)* 124(8): 264–273
- Verhagen S, Teunissen PJG (2006) On the probability density function of the GNSS ambiguity residuals. *GPS Solut* 10(1):21–28. doi:10.1007/s10291-005-0148-4
- Wanninger L (2000) Präzise Positionierung in regionalen GPS-Referenzstationsnetzen. Deutsche Geodätische Kommission, DGK C508, Munich, Germany
- Wanninger L (2004) Ionospheric disturbance indices for RTK and network RTK positioning, In: *Proceeding of ION GNSS 2004*, Long Beach, CA, USA, Sept 21–24, 2004, pp 2849–2854
- Wübbena G, Schmitz M, Boettcher G (2006) Near-field effects on GNSS sites: analysis using absolute robot calibrations and procedures to determine corrections. In: *Proceeding of the IGS workshop 2006, perspectives and visions for 2010 and beyond*, ESOC, Darmstadt, Germany, May 8–12, 2006
- Wuensch KL (2005) Descriptive statistics/graphical procedures. In: Everitt BS, Howell DC (eds) *Encyclopedia of statistics in behavioral science*. Wiley, Chichester

Author Biographies



Xiaoguang Luo is currently a Ph.D. student and works at the Geodetic Institute at Karlsruhe Institute of Technology (KIT), Germany. He received his diploma in geodesy from the Karlsruhe University in 2006. He is interested in the analysis of the stochastic model, and atmospheric and site-specific effects of GNSS observations.



Michael Mayer received his Ph.D. in geodesy from the Karlsruhe University in 2005. He is currently head of the GNSS working group as the Chair of Physical and Satellite Geodesy in the Geodetic Institute at KIT. He is interested in analysis and mitigation of GNSS atmospheric and site-specific effects with a special focus on continuously operating reference stations.



Bernhard Heck is a professor at the Geodetic Institute at KIT. He currently holds the chair of Physical and Satellite Geodesy, where he has been teaching and conducting research since 1991. He received his Ph.D. and habilitation in geodesy from the Karlsruhe University in 1979 and 1984. He has worked extensively in the fields of geometrical geodesy, deformation analysis, gravity field analysis and GNSS positioning. In addition, he has been the Director of the Black Forest Observatory (BFO) since 1999. He is an active member of different special groups of the International Association of Geodesy (IAG) and was President of IAG Section IV (General Theory and Methodology) during 1999–2003.