sciendo
ARTIFICIAL SATELLITES, Vol. 55, No. 1-2020
DOI: 10.2478/arsa-2020-0001

CHARACTERIZING POSITIONING ERRORS WHEN USING THE SECOND-GENERATION AUSTRALIAN SATELLITE-BASED AUGMENTATION SYSTEM

Mehdi KHAKI[1,2], Ahmed EL-MOWAFY[2]
[1] School of Engineering, University of Newcastle, Callaghan, New South Wales, Australia
[2] School of Earth and Planetary Sciences, Curtin University, Perth, Australia
e-mail: Mehdi.Khaki@Newcastle.edu.au

ABSTRACT. Fault detection and exclusion (FDE) is the main task for pre-processing of global navigation satellite system (GNSS) positions and is a fundamental process in integrity monitoring that is needed to achieve reliable positioning for applications such as in intelligent transport systems. A widely used method is the solution separation (SS) algorithm. The FDE in SS traditionally builds the models assuming positioning errors are normally distributed. However, in urban environments, this traditional assumption may no longer be valid. The objective of this study is to investigate this and further examine the performance of alternative distributions, which can be useful for FDE modelling and thus improved navigation. In particular, it investigates characterization of positioning errors using GNSS when the Australian satellite-based augmentation system (SBAS) test bed is used, which comprised different positioning modes, including single-point positioning (SPP) using the L1 global positioning system (GPS) legacy SBAS, the second-generation dual-frequency multi-constellation (DFMC) SBAS service for GPS and Galileo, and, finally, precise point positioning (PPP) using GPS and Galileo observations. Statistical analyses are carried out to study the position error distributions over different possible operational environments, including open sky, low-density urban environment, and high-density urban environment. Significant autocorrelation values are also found over all areas. This, however, is more evident for PPP solution. Furthermore, the applied distribution analyses applied suggest that in addition to the normal distribution, logistic, Weibull, and gamma distribution functions can fit the error data in various cases. This information can be used in building more representative FDE models according to the work environment.

Keywords: GNSS, SBAS, Error characteristics, Urban environment, Kinematic positioning

1. INTRODUCTION

In positioning, measurements that are noisy and may contain errors are used. Therefore, a fault detection and exclusion (FDE) process should be implemented, which usually relies on statistical hypothesis testing by applying a consistency check among all possible sets of observations that can provide a solution. Different methods are used for this purpose. A widely approach is the solution separation (SS) method, which is applied in the position domain. In this method, a position error bound is created for each possible fault mode that might be miss-detected by

accounting for the difference between the all-observation position solutions and a position solution unaffected by the fault (see Blanch et al., 2015; El-Mowafy and Yang, 2016; El-Mowafy, 2017).

The FDE check is applied such that SS statistic, standardized by the standard deviation, should be less than a threshold value. This threshold value is driven from the statistical distribution of the position error. It is computed as the inverse of the tail probability of this distribution computed at a chosen significance level, taken as the selected false alarm rate. In addition, because position errors are not known in practice, a bound on the maximum error with a certain integrity risk, known as the protection level (PL), is computed and is consistently checked against a threshold value, known as the alert limit, to alert the user that integrity of the system is unavailable if the PL exceeds that threshold.

In the general practice, and for simplicity of processing and because of the lack of studies characterizing positioning errors in different environments, current FDE and PL computation methods typically assume normal distribution of the position errors in the fault-free mode. This assumption may hold true in the open sky environment. However, in environments other than open sky, this assumption may not be valid (Imparato et al., 2018), particularly for land applications such as in intelligent transport systems (ITS), where the presence of high multipath and non-line-of-sight (NLOS) signals can alter the Gaussian distribution with zero-mean assumption, leading to a degradation in the detection ability of some anomalies. Thus, an investigation of the position errors in different environments is required to better represent FDE and PL models. Furthermore, the error distribution is critical for some groups of positioning solutions such as Kalman filtering, particularly for modeling the process noise when forming the dynamic models.

One should, however, note that determining the exact distribution of position error in each instantaneous individual environment is a very difficult task, especially in real-time with dynamic change of the work environment in the kinematic mode and the limited number of observations available for positioning, which cannot build a representative model. Therefore, the goal can be relaxed to finding a more generic positioning error distribution that better describes the actual distribution in most conditions, as much as possible, compared to the normal distribution. This is the goal of this contribution.

Furthermore, because of the complexity of studying all individual observations at each location with different directions and elevation angles, and the varying environmental conditions, where performance of global navigation satellite systems (GNSS) receivers can be easily affected by various error sources such as multipath effect, and signal interference (Braasch, 1996; Spilker et al., 1996), studying the positioning errors can provide insight into signal characteristics (Ahmad et al., 2014, Zhu et al., 2018).

The main objective of this study is to investigate the characterization of positioning errors over various environments in the presence of satellite-based augmentation system (SBAS). In this article, we restrict our focus on ITS applications, and because of varying possible environments, we subdivide the study environment into three categories, namely, open sky, low-density urban environment, and high-density urban environment. The low-density urban environment refers to the area with low elevation and a few high-rise buildings. The high-density urban environment, on the other hand, is dominated by the presence of high-rise buildings (on all or most sides of the road).

We used different positioning methods provided through the new Australian SBAS test bed, which form the basis for the upcoming operational system. SBAS has become more popular

in areas where accuracy and integrity are important such as in aviation. It also improves and extends other applications of GNSS, for example, in navigating vehicles, precision farming, and geodesy. There are several SBAS platforms currently in use to provide correction to the GNSS users around the world, for example, Wide Area Augmentation System (WAAS) over the United States, European Geostationary Navigation Overlay Service (EGNOS) over Europe, and MTSAT Satellite Augmentation System (MSAS) over Japan. Different from these systems, the Australian SBAS includes three augmentation services (enabling three positioning modes), namely, legacy L1 GPS SBAS, DFMC SBAS (using GPS and Galileo), and precise point positioning (PPP) (using GPS and Galileo). In addition to the legacy L1 GPS frequency (1,575.42 MHz) that is implemented in accordance with the radio technical commission for astronautics (RTCA) minimum operational performance standards (MOPS), the Australian SBAS includes the second-generation SBAS of dual-frequency multi-constellation (DFMC) service over the L5 frequency (1,176.45 MHz). This is provided as satellite orbits and clocks information transmitted to the entire footprint of the geostationary (GEO) SBAS satellite through messages for dual-frequency users to eliminate the dominant first-order term of the ionospheric delays by forming ionosphere-free linear combination of the two frequencies (Barrios et al., 2018). Furthermore, the Australian SBAS provides the float-ambiguity PPP, where GPS only corrections were transmitted over L1, and GPS together with Galileo corrections were transmitted over L5, supporting applications in different sectors besides aviation, including transport, maritime, mining, and precision agriculture (Dawson and Amos, 2018).

In the remainder of this study, the data sets used in characterizing the position errors are first presented in Section 2.1. We then discuss the error analysis approach in Section 2.2. Various statistical methods are used to comprehensively analyze the characteristics of position errors computed using different SBAS positioning modes. More importantly, the distribution of these errors is examined against multiple distributions families. The results and discussions are provided in Section 3.3, and concluding remarks are given in Section 4.

2. TEST DATA AND METHODS

In this section, the data used in the analysis are first presented, followed by a description of the error analysis approach.

2.1. GNSS data

Multiple data sets are used in this study to investigate the error characteristics. These include $\sim$60,000 epochs of positioning from actual measurements acquired from 10 different kinematic tests conducted in open, low-density, and high-density urban environments in Sydney and Wollongong, Australia, during the period from June to August 2018. The use of these different scenarios can provide insights into characterizing signals and positioning errors even though that error distribution results are largely location dependent. Efforts have been made to run these kinematic tests similarly (such as vehicle speed and time of data collection). The use of multiple tests can ensure the consistency of the measurements over each environment. This was checked by comparing the collected positioning data. For all these epochs, and as mentioned above, three different navigation solutions are used, each was independently applied using the whole data set. The first is SBAS legacy GPS L1, the second is the new-generation SBAS comprising DFMC GPS (L1 and L2) and Galileo (E1 and E5a) ionosphere-free combinations, both solutions are differential GNSS that use phase-smoothed pseudorange measurements. The third

positioning mode is a float-ambiguity PPP over L5 service using GPS and Galileo. The use of this new SBAS service facilitates the characterization of positioning errors for each of these solution methods.

All data sets were collected using the same set of equipment and settings comprising Topcon G3-A1 antenna (Topcon, Tokyo, Japan), Septentrio AsteRx-U receiver (Septentrio, Leuven, Belgium), GMV's Software Receiver SRX-10 Frontend to decode SBAS messages, and a signal splitter. Real-time PPP solutions were computed using GMV's magicAPK software, whereas magicGEMINI software was used to compute SBAS solutions (for details about the software, refer to Barrios et al., 2017; CRC.SI Technical report, 2017). The same data sets with similar satellite constellation and time of day were used in the three positioning modes, depending on the methodology of each mode. Nevertheless, despite using the same equipment and software (with the same settings) for analyzing the results over different environments, satellite constellation and time of data collection are (slightly) different. The solutions are summarized in Table 1. The position errors are calculated as the difference between each solution's output and reference coordinates obtained by processing the data using an independent ambiguity-fixed post-mission kinematic solution, which provided precision at 5 mm to 4 cm, where the rover observations were processed with observations from a nearby base station.

Table 1. Three different positioning solutions used in this study to assess their impacts on error distributions.

Solution	Acronym	Constellation	frequency of observations
Single frequency	L1	GPS	L1
Dual frequency multi-constellation	DFMC	GPS+Galileo	L1 and L2 (GPS), E1 and E5a (Galileo)
Precise point positioning	PPP	GPS+Galileo	L1 and L2 (GPS), E1 and E5a (Galileo)

2.2. Error analysis

In order to analyze the characteristic of position errors, various tests are carried out, including studying autocorrelation and distribution analyses, as well as goodness-of-fit tests (cf. Sections 2.2.1-2.2.4). These are briefly described in the following sections. The tests are applied to all error time series from the three solutions, that is L1 and DFMC SBAS, and PPP. These methods often need to assume that the errors are overbounded by a normal distribution. The objective is to investigate this and examine whether the positioning errors are random or follow a specific pattern.

2.2.1. Autocorrelation

Autocorrelation analysis is applied to each error time series to measure the correlation between errors at different times. To this end, Eq. (1) is used using various time lag (k) values,

$$r_k = \frac{\sum_{i=1}^{N-k}(x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=1}^{N}(x_i - \bar{x})^2}, \tag{1}$$

where r_k indicates the autocorrelation coefficient at lag k, x_i is the error value for $i = 1, 2, ..., N$ (N being the length of data), and $\bar{x}$ represents the mean of error values, respectively. The confidence level of 95% is considered to test against the null hypothesis that a true correlation is zero.

2.2.2. Distribution

The data are experimentally tested against multiple possible probability distributions to investigate their level of agreement. These distributions with their references are listed in Table 2, which are selected because they are widely used in data analysis. The comparison can demonstrate whether the positioning errors follow the typically assumed Gaussian pattern or they better suit other types of popular probability distribution functions (pdf). The comparison is performed both visually and quantitatively. The visual comparison is performed using the time series histograms, curve fitting, and Q-Q (quantile-quantile) plots, and the numerical comparison is carried out using goodness-of-fit tests (see Section 2.2.4).

Table 2. Various distribution used in this study. These are selected because of their popularity in the analytical field.

Distribution name (alphabetically)	References
Beta	Johnson et al. (1995)
Birnbaum-Saunders	Birnbaum and Saunders (1969)
Exponential	Balakrishnan (1996)
Gamma	Jambunathan (1954)
Generalized Pareto	Hosking and Wallis (1987)
Logistic	deCani and Stine (1986)
Lognormal	Johnson et al. (1994)
Nakagami	Laurenson (1994)
Normal	Lukas (1942)
Poisson	Frank (1967)
Rayleigh	Siddiqui (1964)
Rician	Rice (1945)
t location-scale	Meyer (1987)
Weibull	Weibull (1951)

2.2.3. Skewness and Kurtosis

To further investigate whether the data follow the Gaussian distribution, skewness and kurtosis tests are applied. These can help us to indicate the departure of the error distributions from a Gaussian distribution (with a skewness 0 and a kurtosis 3). The distribution shape can be quantified by kurtosis (i.e., heavy-tailed or light-tailed, in comparison to normal distribution), whereas skewness measures the distribution asymmetry (Khaki et al., 2017; Joanes and Gill, 1998). Skewness (s) and kurtosis (k) are calculated using the following equations, respectively,

$$s = \frac{\mathbb{E}(x-\mu)^3}{\sigma^3}, \tag{2}$$

$$k = \frac{\mathbb{E}(x-\mu)^4}{\sigma^4}, \tag{3}$$

where μ and σ are the mean and standard deviation of x, respectively, and $\mathbb{E}(.)$ denotes the expected value.

2.2.4. Goodness-of-fit

In order to statistically measure how well the statistical models fit the observation errors, the goodness-of-fit tests are carried out. These tests describe the misfit between the data and the values expected under the assumed distributions. There are various tests available, and the choice of a specific one depends on the expected kind of possible distortion of the theoretical distribution. In this context, three mostly used tests, including Anderson–Darling (Stephens, 1974), Kolmogorov–Smirnov (Chakravarti et al., 1967), and Chi-squared (Snedecor et al., 1989), are used to indicate whether two samples fit each other: the first is the computed position errors and the second is a tested distribution pdf. Multiple tests are selected to better investigate the goodness of fit between different distributions. Confidence can be achieved if the tests obtain similar results. On the other hand, discrepancies between the results can show less agreement with a specific distribution.

3. RESULTS AND DISCUSSIONS

Characteristics of different positioning solutions errors using different SBAS methods and PPP in various environments are explored in this section. This is important for improving our understanding of GNSS errors and their distribution when SBAS signals are available, which can be very helpful for various positioning objectives such as integrity monitoring. Here, we first analyze the normality of the positioning errors. This is performed by studying multiple various aspects of the errors, including their empirical cumulative distribution function (cdf), skewness and kurtosis, as well as their autocorrelation. We then present potential distribution functions as an alternative to Gaussian distribution and assess their performance using various goodness-of-fit methods. All these tests are applied for three test environments and for various positioning solutions (cf. Section 2.1).

3.1. Normality tests

We first show that the traditional assumption that the position errors follow a Gaussian distribution cannot be generalized as mostly done to simplify the models. Figure 1 shows the histogram plot of the position error time series for the three different solutions (L1, DFMC, and PPP) over low- and high-density urban environments. It is clear from the figure that the distribution varies among different methods. However, the bell-shaped graphs of the normal distribution can be seen in some instances such as horizontal and vertical error histograms of L1, PPP, and to a lesser degree DFMC. This is more evident for the high-density urban environment. Such different histogram patterns demonstrate that the assumption that all positioning errors are normally distributed does not present actual results. This can also be seen when studying the empirical cdf (Cox and Oakes, 1969; Lawless, 2003) of the error time series (blue lines) and comparing them with the Gaussian (normal) cdf (red lines), as shown in Figure 2.

The empirical cdf graphs of the positioning errors of the solutions shown in Figure 2 over the open sky and low- and high-density urban environments depict disagreements with the normal cdf graphs for different solutions. The average difference between the two distributions in each graph is also indicated in the figure. The PPP results have the largest misfit compared to L1 and DFMC solutions. These disagreements are clearer in urban environments and especially high-density urban environment, particularly because of multipath, where all solutions are found to deviate from normal cdf. In open sky and low-density urban environments, DFMC results show noticeable discrepancies to normal cdf. L1 solutions, however, better match the normal cdf in open sky, whereas the misfit can be observed in the low-density urban environment. These graphs explain the non-Gaussian histogram patterns of different solutions in Figure 1

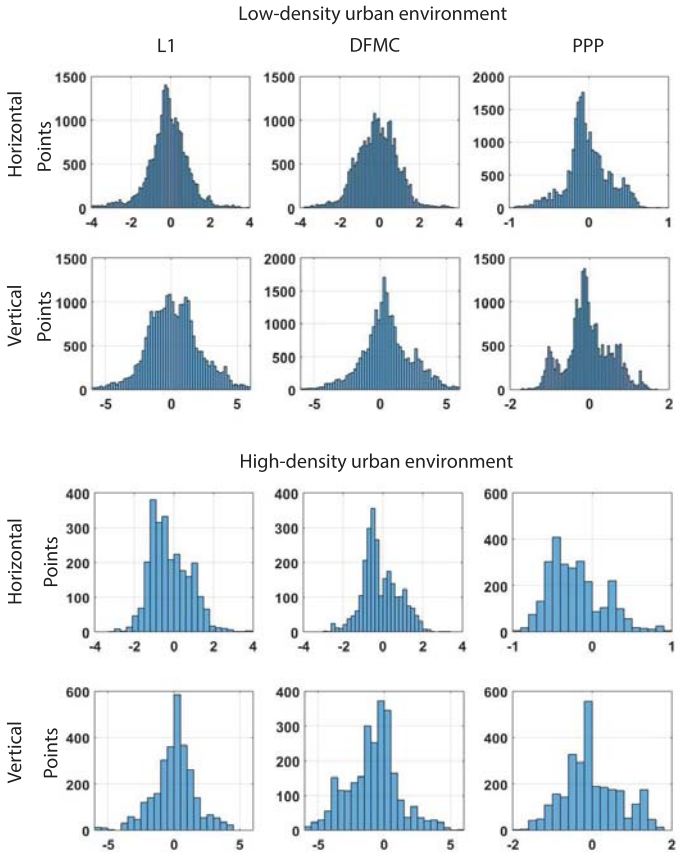

Figure 1. Histogram plots of average error time series for three different solutions over low- and high-density urban environments in horizontal (average of East and North error components) and vertical directions. Note that x axis in the plots represents the error values and the y axis indicates frequency. For a better presentation, only results of the urban environments are shown because of their larger discrepancy from a Normal distribution.

over various environments. This can further be clarified through studying skewness and kurtosis factors.

As mentioned earlier, the skewness and kurtosis measure the departure of the error distributions from a Gaussian distribution (which has a skewness of 0 and a kurtosis of 3). We calculate these two factors for L1, DFMC, and PPP positioning errors over various environments. Figure 3 demonstrates the estimated skewness and kurtosis in the horizontal and vertical directions. It is clear from the figure that errors in vertical direction are more disperse from a Gaussian distribution. It can be seen that L1 and DFMC reach the closest values to the zero in the open sky environment. This, however, is no longer the case over the low- and high-density urban environments, especially for DFMC. Negative values for the skewness indicate the presence of data that are skewed left, and positive values for the skewness indicate the presence of data that are skewed right. From the figure, in general, DFMC obtains smaller factors than PPP. Larger kurtosis in low- and high-density urban environments indicate larger discrepancies from normal distribution. Positive kurtosis indicates a heavy-tailed distribution, and negative kurtosis indicates a light-tailed distribution. PPP errors further depict more disagreement with a Gaussian distribution in terms of the distribution asymmetry. These results confirm the presented outcomes in Figure 2, where PPP and, to a lesser degree DFMC deviate from Gaussian cdf, particularly over urban environments.

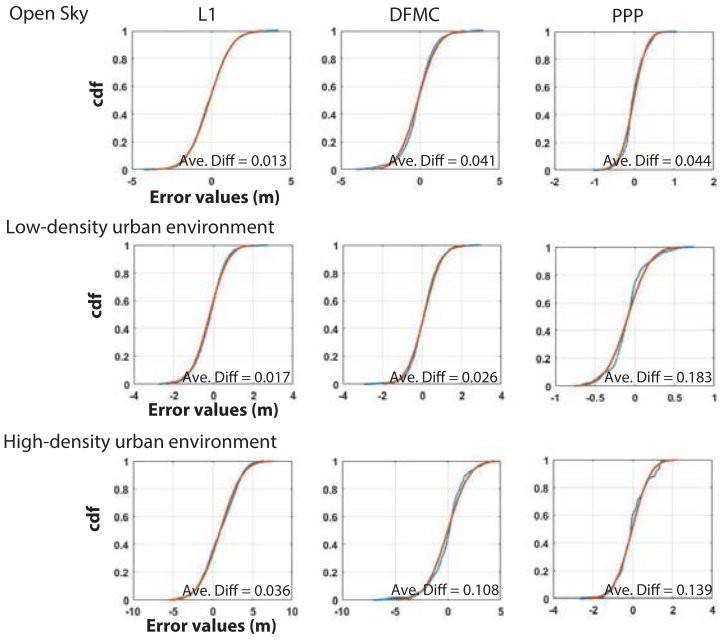

Figure 2. Empirical cdf (blue) against normal cdf (red) for the (L1, DFMC, and PPP) solutions over open sky and low- and high-density urban environments estimated using the average horizontal and vertical errors.

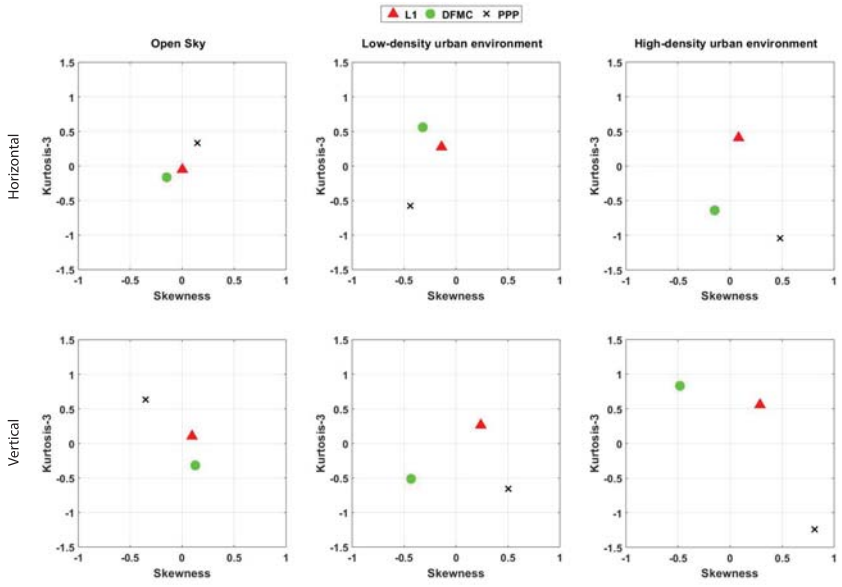

Figure 3. The estimated skewness and kurtosis, averaged for horizontal and vertical errors, for the different experiments applied.

3.2. Autocorrelation analysis

Studying autocorrelation of the positioning errors is of interest because it shows the time lag over which the data can be considered independent, and thus shows the frequency (i.e., time interval) required to sample the data when a set of independent data points are required to build a distribution. The degree of autocorrelation of position errors can be impacted by different factors such as, type of observations, processing algorithm, work environment and the form of the trajectory. Moreover, autocorrelation analyses on different time series over different situations can provide useful information about the dependency of errors caused by time-correlated corrections transmitted through SBAS signals, particularly for PPP. This information can be important for different filtering techniques such as Kalman filtering.

Figure 4 shows the autocorrelation values of the position errors in each mode and over different test environments considered in this study using 95% and 99% confidence intervals (shown as blue and red lines, respectively). The autocorrelation values outside the confidence intervals are considered to be statistically significant. The results show that, contrary to the open sky environment that has a short initial lag and mostly insignificance autocorrelation, considerable autocorrelation values are found for position errors of all solutions in the urban environments. This is more obvious for PPP, especially in urban environments. Positioning using L1 SBAS gave much smaller values of correlation time, of a few seconds, over open sky. However, this correlation time increases significantly over low- and high-density urban environments. Positioning using DFMC SBAS, on the other hand, shows higher autocorrelated values even in the open sky environment. It should be noted that L1 and DFMC SBAS positioning use snapshot least squares processing (but with phase-smoothed data), thus a low time correlation could be observed. On the other hand, PPP uses Kalman filtering, which, because of the use of dynamic models in the error states, may introduce correlated positioning errors in the urban environment in the presence of multipath and NLOS.

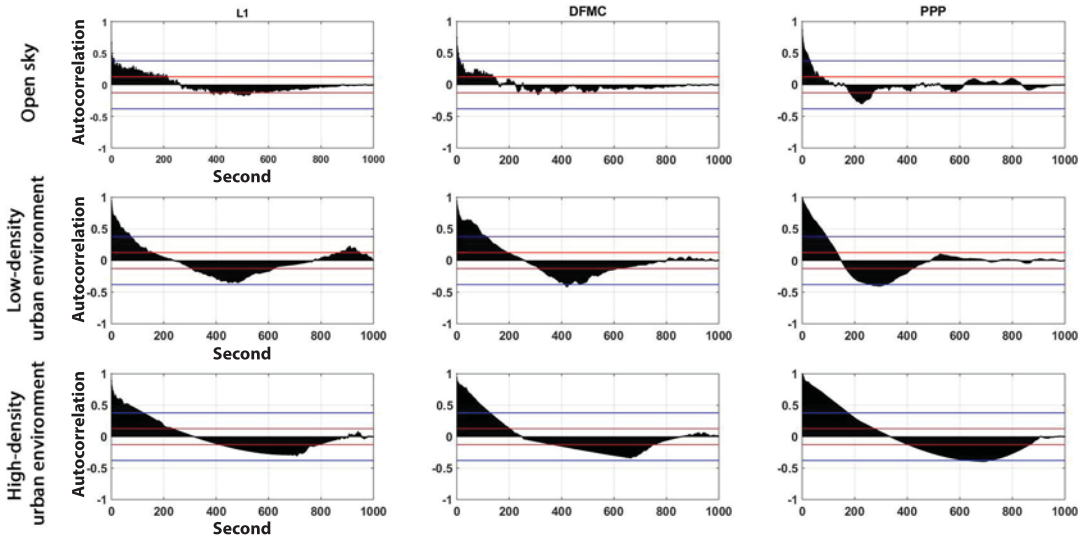

Figure 4. The average autocorrelation results (for horizontal and vertical errors) at different lags (in seconds) over open sky and low- and high-density urban environments.

The results presented in Figure 4 illustrate the non-Gaussian behavior of the error time series in the presence of significant autocorrelation values, which can affect the normal distribution assumption, especially in the widely used Kalman filtering approach. Useful information can

be found from the correlation time values, which can show the period in which time series can be assumed temporarily independent. The calculated lag values for each solution type, over open sky and low- and high-density urban environments, for the large data sets used in this study are presented in Table 3. The table shows that the independent lags are generally close for open sky and low-density urban environment for different solution types. This, however, is different over high-density urban environment, where significantly larger time lags for all solutions are observed. Among the solutions, PPP has the largest time lags (correlation time), especially over high-density urban environment. The position error time series were largely autocorrelated within 6 minutes interval with correlations higher than 0.70 for $\sim$1.5 minutes. L1 and DFMC SBAS positioning errors also show large correlation values ($>$0.5 at 0.99 significant level) over the urban environments for approximately 1 minute. The results indicate that for all cases, autocorrelation exist and it is essential to be accounted for in the positioning models. Overall, it can be inferred from Table 3 that similar to the previous results, PPP errors deviates more from Gaussian pattern followed by L1 and DFMC SBAS positioning errors over different environments, especially in high-density urban environments.

Table 3. Correlation time (in seconds) from the autocorrelation analysis of the time series of L1, DFMC, and PPP positioning errors over open sky and low- and high-density urban environments.

		Confidence intervals		**Correlation time** $> \alpha$		
Test environment	Solution	95%	99%	α = 0.5	0.7	0.9
Open sky	L1 SBAS	47	11	7	2	1
	DFMC SBAS	66	14	9	3	1
	PPP	84	38	32	8	3
Low-density	L1 SBAS	125	74	36	9	2
	DFMC SBAS	169	97	71	14	4
	PPP	237	115	96	45	8
High-density	L1 SBAS	172	124	63	11	3
	DFMC SBAS	198	131	98	42	4
	PPP	315	178	156	87	13

3.3. Error distribution analysis

In this section, we review multiple popular distributions against the tested positioning error time series to assess which distributions better fit the data. Recall that the collected sample size is $\sim$60,000, which gives relatively representative sample. This assessment is performed both visually, based on a histogram analysis, and numerically, using the goodness-of-fit investigations. Similar to Figure 1, we calculate the histograms of all solutions over different environments and compare them with all the studied pdfs (cf. Section 2.2.2). Afterwards, the best fits are investigated among all distributions applied, which include normal, logistic, Weibull, gp, generalized Pareto, and generalized extreme value (gev). In order to help this process, Q-Q plots (quantile of data vs. quantile of the attempted distribution) of all distributions are generated and analyzed

to find the best matches. Figures 5 and 6 exemplify the best-achieved results, in terms of fitted distributions, over low-density urban environment for the histogram and Q-Q plot, respectively, which agree with the skewness and kurtosis conclusions. The pdfs of logistic, Weibull, and somewhat the "Normal" are observed to better match the data as shown in Figures 5 and 6. These three distributions are also found to better fit the positioning errors over open sky and, to a lesser degree, high-density urban environments. Among the three, however, logistic and Weibull distributions match better the data than the normal distribution.

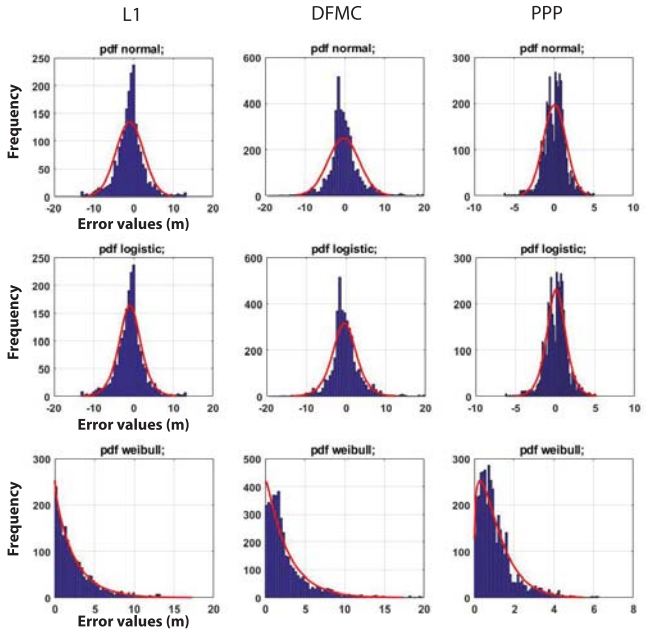

Figure 5. Three best-fitted pdfs (red) plotted on corresponding histogram plots of average horizontal and vertical error time series for different solutions.

To better assess the relevance of different distributions against the solutions positioning errors, the goodness-of-fit analysis is undertaken. As mentioned earlier, three different tests of Anderson-Darling, Kolmogorov-Smirnov, and Chi-squared are applied on error time series of L1, DFMC, and PPP positioning errors over open sky and low- and high-density urban environments. Table 4 summarizes the results, where it can be seen that similar to the results in Figures 5 and 6 logistic, Weibull, and normal distributions are found to be the best matches to the data. For each test, p-value is also reported, which indicates the probability of observing a test statistic as extreme as, or more extreme than, the observed value under the null hypothesis. In other words, small p-values cast doubt on the validity of the null hypothesis, that is a doubt that the data follow the tested distribution. In Table 4, larger p-values are also obtained for logistic, Weibull, and normal distributions. Overall, smaller p-values for all distributions are achieved over urban environments compared to open sky region. Moreover, smaller p-values are achieved for PPP than the DFMC SBAS results. L1 SBAS positioning errors generally better follow the indicated distributions.

Furthermore, Table 4 proves that the normal distribution was appropriate in the open environment, but had a degraded performance in the urban environment. The logistic and Weibull distributions, t-location-scale, gamma and exponential distributions can also be considered as candidates to fit the data according to the positioning method used and environment, as shown

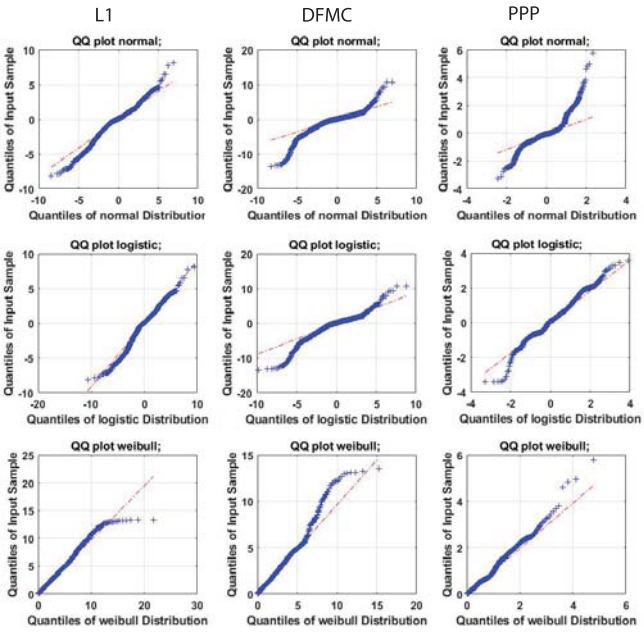

Figure 6. Q-Q plots of error time series and the distributions used in Figure 5.

in the table. On the basis of these results, one can re-confirm that the normal distribution assumption, usually made in most literature, cannot be generally considered valid over urban environments. Besides, the non-Gaussianity of the noises can be increased in ionosphere-free combination by using the second frequency (as shown by the DFMC SBAS and PPP results). Table 4 suggests that based on the observation types used for positioning and also the test environment, different distributions can be considered for modeling the positioning errors. Overall, the three distributions found in the tests that better match the data, are the logistic, Weibull, and normal distributions.

Table 4. The goodness-of-fit analysis results using Anderson–Darling, Kolmogorov–Smirnov, and chi–squared tests applied on various experiments, averaged over horizontal and vertical error components. The highest p-value obtained by each test for each experiment is also reported.

		Anderson–Darling		**Kolmogorov–Smirnov**		**Chi–squared**	
Test environment	Solution	Distribution	p-value	Distribution	P-value	Distribution	p-value
Open sky	L1	logistic	0.77	normal	0.51	weibull	0.31
	DFMC	logistic	0.72	normal	0.49	gamma	0.27
	PPP	tlocationscale	0.45	logistic	0.26	normal	0.16
Low-density	L1	normal	0.46	weibull	0.62	logistic	0.15
	DFMC	logistic	0.58	gamma	0.32	normal	0.26
	PPP	weibull	0.13	logistic	0.22	exponential	0.11
High-density	L1	weibull	0.37	logistic	0.30	normal	0.16
	DFMC	gamma	0.21	weibull	0.44	gamma	0.28
	PPP	logistic	0.19	tlocationscale	0.12	gp	0.09

4. CONCLUSION

This study investigated the characterization of GNSS positioning errors. The main objective is to find the best distribution for defining the threshold value in the fault detection testing using the solution separation algorithm in ITS, and to give insight on the behavior of positioning errors using some possible methods for these applications and under different environments. The positioning methods considered are based on the second-generation Australian SBAS test bed and include L1 and DFMC SBAS, and PPP. Testing was performed in the kinematic mode using a vehicle traveling in various environments including open sky, low-density urban environments, and high-density urban environments. The data size is $\sim$60000 epochs acquired from 10 different kinematic tests, where the L1, DFMC SBAS, and PPP methods were independently applied for the whole data set. Different statistical tests were used to first investigate the common assumption of normality of error distributions as well as explore other distributions that may better fit the positioning errors. Our results indicated that, in most of the cases, the error time series do not follow a normal distribution, especially in urban environments.

Significant autocorrelation values were also found over different areas for different solutions applied, especially for PPP. The autocorrelation values for PPP errors were larger than 0.70 (at 0.99 significant level) for $\sim$1.5 min. These results along with autocorrelation values for other solutions applied are important for modeling the temporal dependency of the data. In another effort, different probability density functions (pdf) were tested against the positioning errors to find the best fits. According to the goodness-of-fit test results using Anderson–Darling, Kolmogorov–Smirnov, and chi-squared methods, it was found that some distribution functions such as logistic, and Weibull can better fit the distribution of the errors than normal distributions, especially in the urban environments.

Acknowledgments. Transport for New South Wales is acknowledged for its help in data gathering. The positioning data was acquired during the Australian test-bed Project in 2018, which is developed and funded by Geoscience Australia and administrated by FrontierSI. Both are acknowledged. The first author is grateful to Curtin University for partially supporting the study and providing the resources.

REFERENCES

Ahmad, K.A.B., Sahmoudi, M., Macabiau, C., (2014). *Characterization of GNSS Receiver Position Errors for User Integrity Monitoring in Urban Environments*. ENC-GNSS 2014, European Navigation Conference, Apr 2014, Rotterdam, Netherlands. $< hal - 01160130 >$.

Balakrishnan, K. (1996). *Exponential Dist: Theory, Methods and Applications*. CRC Press.

Barrios, J., Caro, J., Calle, J.D., Carbonell, E., Rodriguez, I., Romay, M.M., Jackson, R., Reddan, P.E., Bunce, D., Soddu, C., (2017). *Australian and New Zealand Second Generation Satellite Positioning Augmentation System Supporting Global SBAS Concept*, Proceedings of the 30th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2017), Portland, Oregon, September 2017, pp. 979-996. https://doi.org/10.33012/2017.15243

Barrios, J., Caro, J., Calle, J.D., Carbonell, E., Pericacho, J.G., Fernandez, G., Esteban, V.M., Fernandez, M.A., Bravo, F., Torres, B., Calabrese, A., Diaz, A., Rodriguez, I., Lainez, M.D., Romay, M.M., Jackson, R., Reddan, P.E., Bunce, D., Soddu, C., (2018). *Update on Australia and New Zealand DFMC SBAS and PPP System Results*. In proceedings of ION GNSS+ 2018, Miami, Florida, September 24-28, 2018, pp. 1038-1067.

Birnbaum, Z.W., Saunders, S.C. (1969), *A new family of life distributions*, Journal of Applied Probability, 6 (2): 319-327, doi:10.2307/3212003.

Blanch, J., Walter, T., Enge, P., et al., (2015). *Baseline advanced RAIM user algorithm and possible improvements*, IEEE Transactions on Aerospace and Electronic Systems. 51(1):713-732.

Braasch, M.S., (1996). *Multipath effects, in Global Positioning System: Theory and Applications*, vol. 1, B. W. Parkinson and J. J. Spilker, Jr., Eds. Washington, DC, USA: AIAA, ch. 14, pp. 547-568.

Chakravarti, Laha, Roy, (1967). *Handbook of Methods of Applied Statistics*, Volume I, John Wiley and Sons, pp. 392-394.

Cox, D.R., Oakes, D. (1984). *Analysis of Survival Data*, Chapman & Hall.

deCani, J.S., Stine, R.A. (1986). *A note on deriving the information matrix for a logistic distribution. The American Statistician*. American Statistical Association. 40: 220-222. doi:10.2307/2684541.

CRC.SI Technical report (2017). *Technical Specifications Document for Satellite-Based Augmentation System (SBAS) Testbed*, The Australia and New Zealand Cooperative Research Centre, Revision 5 10 November 2017, https://www.crcsi.com.au/assets/Program-1/SBAS-Project/Australia-NZ-Testbed-Technical-Specifications-Rev05.pdf.

Dawson, J., Amos, M., (2018). *Progress Towards a Regional SBAS Service for Australia and New Zealand*, Proc. Of the Institute of Navigation, ION GNSS+, Miami, FL, 24-28 Sept, 1-11.

El-Mowafy, A., Yang, C., (2016). *Limited Sensitivity Analysis of ARAIM Availability for LPV-200 over Australia using real data*. Advances in Space Research, 57(2): 659-670.

El-Mowafy, A., (2017). *Advanced Receiver Autonomous Integrity Monitoring Using Triple Frequency Data with a Focus on Treatment of Biases*. Advances in Space Research, 59(8): 2148-2157.

Frank, A.H. (1967). *Handbook of the Poisson Distribution*. New York: John Wiley & Sons.

Hosking, J.R.M., Wallis, J.R. (1987). *Parameter and Quantile Estimation for the Generalized Pareto Distribution*. Technometrics. 29 (3): 339-349. doi:10.2307/1269343.

Imparato, D., El-Mowafy, A., Rizos, C. (2018). *Positioning Integrity Monitoring: from Aviation to Land Applications. Multifunctional Operation and Application of GPS*. In Tech Publisher, UK, Chapter 2, 23-43.

Jambunathan, M.V. (1954). *Some Properties of Beta and Gamma Distributions*. Ann. Math. Stat. 25, 401-405.

Joanes, A.H., Gill, C.A. (1998). *Comparing Measures of Sample Skewness and Kurtosis*. The Statistician 47(1): 183-189.

Johnson, N.L., Kotz, S., Balakrishnan, N. (1994). 14: *Lognormal Distributions, Continuous univariate distributions*. Vol. 1, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics (2nd ed.), New York: John Wiley & Sons, ISBN 978-0-471-58495-7, MR 1299979.

Johnson, N.L., Kotz, Samuel, B.N. (1995). Chapter 21: *Beta Distributions*. Continuous Univariate Distributions Vol. 2 (2nd ed.). Wiley. ISBN 978-0-471-58494-0.

Khaki, M., Hoteit, I., Kuhn, et al., (2017). *Assessing sequential data assimilation techniques for integrating GRACE data into a hydrological model*, Advances in Water Resources, Volume 107, Pages 301-316, ISSN 0309-1708, http://dx.doi.org/10.1016/j.advwatres.2017.07.001.

Laurenson, D. (1994). *Nakagami Distribution. Indoor Radio Channel Propagation Modelling by Ray Tracing Techniques*. Retrieved 2007-08-04.

Lawless, J.F. (2003). *Statistical Models and Methods for Lifetime Data*, 2nd ed., Wiley.

Lukas, E. (1942) *A characterization of the normal distribution*. Annals of Mathematical Statistics 13: 91-93.

Meyer, J. (1987). *Two-Moment Decision Models and Expected Utility Maximization*. American Economic Review. 77 (3): 421-430.

Rice, S.O. (1945). *Mathematical Analysis of Random Noise*. Bell System Technical Journal 24, 46-156.

Siddiqui, M.M. (1964). *Statistical inference for Rayleigh distributions*, The Journal of Research of the National Bureau of Standards, Sec. D: Radio Science, 68D(9): 1005-1010.

Snedecor, G.W., Cochran, William, G. (1989). *Statistical Methods, Eighth Edition*, Iowa State University Press.

Spilker, J.J., Axelrad, P., Parkinson, B.W., et al., (1996). *Interference effects and mitigation techniques*, in Global Positioning System: Theory and Applications, vol. 1, B. W. Parkinson and J. J. Spilker, Jr., Eds. Washington, DC, USA: AIAA, 1996, ch. 20, pp. 717-771.

Stephens, M.A. (1974). *EDF Statistics for Goodness of Fit and Some Comparisons*, Journal of the American Statistical Association, 69, 730-737.

Weibull, W. (1951). *A statistical distribution function of wide applicability* (PDF), J. Appl. Mech.-Trans. ASME, 18 (3): 293-297.

Zhu, N., Marais, J., Betaille, D., Berbineau, M. (2018). *GNSS Position Integrity in Urban Environments: A Review of Literature*, in IEEE Transactions on Intelligent Transportation Systems. 19(9):2762-2778. doi: 10.1109/TITS.2017.2766768

Received: *2019-11-06*

Reviewed: *2020-01-20 (by L. Jaworski) and 2020-01-20*

Accepted: *2020-01-21*