engineering proceedings

MDPI

Proceeding Paper

On the Quantification of the GNSS Signals' Quality for RFI Assessment †

Naveed Ahmed

Department of Engineering Cybernetics, Norwegian University of Science and Technology, 7034 Trondheim, Norway; naveed.ahmed@ntnu.no

† Presented at the European Navigation Conference 2023, Noordwijk, The Netherlands, 31 May–2 June 2023.

Abstract: The performance of radio-frequency interference (RFI) detection and localization could be improved if they are applied on good quality inputs. RFI assessment is an important aspect of evaluating overall signal quality. The worthiness of a Global Navigation Satellite System (GNSS) receiver output can be generally quantified by a number of parameters readily available from a receiver. However, such discrete parameters do not give a detailed picture of the quality of the received GNSS signals. Statistical treatment of the received signals both in the absence and presence of interference gives some interesting insight about the data. In this paper, we study if the baseband data from the front-end Analog to Digital Converter (ADC) are normally distributed and if the presence of interference affects the statistical behavior of the distribution, often characterized by its probability density function (PDF) and other related parameters, such as skewness and kurtosis. In the second part of the paper, we study the feasibility of the Shapiro–Wilk (SW) test as a method to study the effect of interference on the GNSS signal while also serving as a potential approach to assess RFI. Skewness and kurtosis are statistical measures used to examine the shape of the distribution of a set of data. The implementation of the Shapiro–Wilk test is also studied, which is a normality test used to check whether a set of data follows a normal distribution. The above approaches have been evaluated using an experiment, where an RTLSDR is used as a reference GNSS receiver and simulated noise is added in the real signals. The data have been logged both in the presence and absence of wideband interference. The obtained results show the potential of the techniques presented for both the quantification of GNSS signal quality and the RFI assessment, alike.

Keywords: RFI assessment; goodness of fit; Shapiro–Wilk (SW) test; skewness; kurtosis

Citation: Ahmed, N. On the Quantification of the GNSS Signals' Quality for RFI Assessment. *Eng. Proc.* **2023**, *54*, 19. https://doi.org/10.3390/ENC2023-15440

Academic Editors: Tom Willems and Okko Bleeker

Published: 29 October 2023

1. Introduction

A compromise in the safety of the GNSS dependent applications has been widely reported in recent years [1–3]. This leads to an escalation in the demand for more robust, resilient, and secure navigation systems for such navigation-reliant applications. The impact of RFI on the receiver operations has been observed at both pre-correlation and post-correlation stages of the GNSS receiver [4].

The subject of RFI detection has been widely covered in the literature, where both statistical and non-statistical approaches have been proposed. The non-statistical approaches often encompass the monitoring of observables in the received signal such as C/N_0, AGC, and spectrum in terms of power spectral density (PSD), pseudorange, etc., and aim to observe the abnormalities in the behavior of these observables. In [5], interference is detected using day to day C/N_0 monitoring, where the observation of distinct patterns declares the presence of RFI. The Wavelet-Hough transform approach is proposed in [6], which is a combination of wavelet transform, which breaks down a signal into different frequency components extracting local spectral and temporal information in the signal, and Hough transform, which is a popular feature extraction technique that is widely used in image processing; here, it is used for a template-matching purpose. A joint Time-Frequency (TF) analysis has been proven to be an instrumental technique [7]. These techniques work as the interference exhibits difference behavior in the time and frequency domain, and even a very short duration interference that might not be visible in the time domain could be identified in the frequency domain. The statistical approaches work around analyzing signals in the time domain or the pre-correlation domain and exploit the established statistical concepts and principles for RFI detection purposes [8]. In [9], a Chi-square-based goodness-of-fit method is proposed. The in-phase (I) and quadrature (Q) IF samples from the RF front end are investigated if they follow the normal distribution. From the literature, there is no attempt to apply normality tests for detecting distortions in the GNSS signals. It is important to clarify the difference between the normality testing and goodness-of-fit (GoF) methodologies. Apart from the obvious mathematical differences, the normality testing methods take a sample from the population and check whether its population follows certain normal distribution, whereas the GoF methods aim to fit the available dataset in order to follow the normal distribution.

The rest of the paper is organized as follows: Section 2 explains the statistical techniques used in the paper for RFI detection, Section 3 presents the SW-based distribution analysis technique, Section 4 presents the scenario considered for validating the proposed methods, Section 5 discusses the obtained results, and finally, Section 6 concludes the paper.

2. Statistical Characteristics of Data Distribution

In our analysis, we underscore the importance of the distribution modeled by the selected data both in the absence and presence of noise. In the current case, we are interested in two important parameters, skewness and kurtosis, that potentially give a good picture of how the distribution might look like.

2.1. Skewness

Skewness is a measure of the asymmetry of a distribution. It describes how much the shape of the distribution deviates from the symmetry around its mean.

If the distribution is symmetric, then the skewness is zero. If the distribution has a longer tail on the right side (i.e., the positive side of the x-axis), then it is said to be positively skewed, and the skewness value will be positive. Conversely, if the distribution has a longer tail on the left side (i.e., the negative side of the x-axis), then it is said to be negatively skewed, and the skewness value will be negative. Mathematically speaking, it is computed as

$$skewness(x) = \frac{3\overline{x} - M_d}{s} \tag{1}$$

where $\overline{x}$ is the arithmetic mean, M_d is the median of the distribution, and s is the standard deviation of the distribution.

The skewness is generally interpreted as:

1. If the skewness is between -0.5 and 0.5, the data are nearly symmetrically distributed around the mean.
2. If the skewness is between -1 and -0.5 (negative skewed) or between 0.5 and 1 (positive skewed), the sample data are representative of a slightly skewed distribution.
3. If the skewness is lower than -1 (negative skewed) or greater than 1 (positive skewed), the sample data would represent an extremely skewed distribution.

2.2. Kurtosis

The kurtosis is a numerical quantification of the tailedness of a distribution. The standard normal distributions have a kurtosis of around 3, where most of the data are concentrated around its mean. If the kurtosis is greater than 3 or, in other words, the excess kurtosis is positive, it means most of the data are concentrated around its mean. Such types of distributions are termed as a leptokurtic distribution. On the contrary, if the kurtosis is less than 3 or, in other words, the excess kurtosis is negative, it means that most of the data are concentrated around its ends/tails. Such types of distributions are termed as platykurtic. Mathematically, kurtosis is defined as

$$kurtosis(x) = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^4}{\left[\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2\right]^2} \tag{2}$$

where n is the number of samples, x_i is the ith value, and $\overline{x}$ is the sample mean.

3. Distribution Analysis for RFI Assessment Techniques

The literature covers many normality testing and goodness-of-fit methods that are suitable for certain applications [10]. Some of the methods are applied on the empirical distribution formed by the dataset [11,12], whereas others are applied on the subset of the whole dataset that is assumed to be the representative of the whole dataset. In the latter category, two popular methods for normality testing are the Shapiro–Wilk (SW) [13] and Shapiro–Francia (SF) [14] methods. Generally, both methods exhibit equally good performance; however, in some cases, the SW method shows some sensitivity to the sample size and the SF method tends to be powerful in alternate hypothesis testing [15]. For the sake of generality, we do not discriminate the two methods in terms of performance.

In this section, we discuss the normality testing using the SW method and present them in context of the GNSS RFI assessment. We initiate the analysis by formulating the hypotheses test for the presence of interference in the system.

$$H_0(RFI\ absent)\text{: the sample is taken from a normal distribution} \tag{3}$$

$$H_1(RFI\ present)\text{: the sample is not taken from a normal distribution} \tag{4}$$

Mathematical Formulation of the Shapiro–Wilk Test

Let us suppose x to be a sample of n GNSS observations derived from a population with an unknown distribution, where the sample is characterized by its mean ($\overline{x}$) and standard deviation (σ).

$$x = \{x_1, x_2, x_3, \ldots, x_n\} \tag{5}$$

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{6}$$

$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 \tag{7}$$

We need to reorder the sample values in an ascending order, i.e., from smallest to largest.

$$x(1) \leq x(2) \leq \cdots \leq x(n) \tag{8}$$

The mean or the expected values of the sorted sample values under normality can be computed and the standardized values for the SW test can be calculated.

$$m(i) = E[x(i)] = \frac{i - 0.375}{n + 0.25},\ \forall i = 1, 2, \ldots, n \tag{9}$$

$$z(i) = (x(i) - \overline{x})/s,\ \forall i = 1, 2, \ldots, n \tag{10}$$

Finally, we calculate the coefficients a of the normal distribution, the test statistic W, and the p-value.

$$a(i) = \frac{m(i) - \overline{m}}{s_m},\ \forall i = 1, 2, \ldots, n \tag{11}$$

where

$$\overline{m} = \frac{1}{n}\sum_{i=1}^{n} m(i) \tag{12}$$

$$s_m^2 = \frac{1}{n-1}\sum_{i=1}^{n}(m(i) - \overline{m})^2 \tag{13}$$

$$W = \frac{(\sum_{i=1}^{n} a(i)z(i))^2}{\sum_{i=1}^{n}(x(i) - \overline{x})^2} \tag{14}$$

$$p = P(W > \alpha) \tag{15}$$

where α is the significance level, that is set to be 0.05; W is the observed value of the test statistic; and P is the probability function of the null hypothesis, which states that the sample comes from a normally distributed population.

4. Experimental Setup

A static test is performed using a multiple purpose and low-cost RTL-SDR that is mainly used as a TV tuner [16] (See Figure 1). However, due to a wide-range demodulator, the reception of GPS L1 signals is also possible .

Figure 1. RTL-SDR (sixth generation) USB Dongle showing antenna connector (left) and USB port to connect with PC (right).

The RTL-SDR design is based on the RTL2832U demodulator chip and the raw I/Q data can be accessed directly, which allows the DVB-T TV tuner to be used as a wide-band software defined radio via a custom software driver. Some important technical specifications, and the settings used to record the dataset, are given in Table 1.

Table 1. RTL-SDR Technical Specifications [9].

Parameter	Value
Tune Low (MHz)	24
Tune Max (MHz)	1766
RX Bandwidth (MHz)	3.2/2.56 (Stable)
ADC Resolution (Bits)	8
ADC Sampling Frequency (MHz)	2.048
Intermediate Frequency (Hz)	0

5. Results and Discussion

The frequency spectrum of the GPS L1 signals at the baseband level is shown in Figure 2. The effect of varying noise levels on the spectrum is evident in the figures. The wideband noise or white noise has a constant power spectral density across all frequencies, where the noise energy is spread out across the entire spectrum of the signal, effectively masking or obscuring any underlying signal components. The same observation can be seen that the noise adds many unwanted frequencies to the original signal spectrum, making it more difficult to identify and analyze the desired signal.

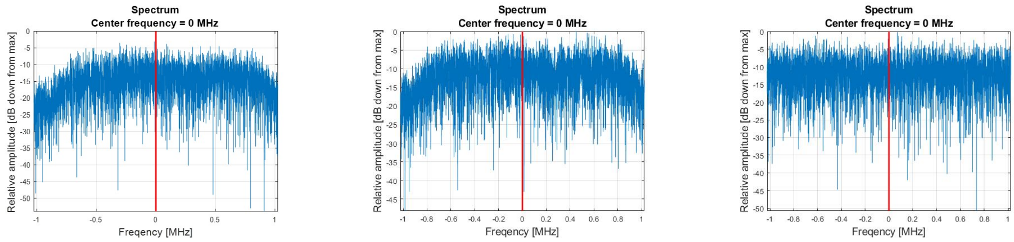

Figure 2. Power Spectrum Density (PSD) of clean (**left**) and noisy GNSS signals at −20 dB (**center**) and −60 dB (**right**).

5.1. Effect of Noise on the Statistical Properties

The test results obtained using different noise characteristics have been summarized in Table 2.

Table 2. Results showing the impact of noise on parameters of PDF (Sample Size = 4096).

GNSS Signal and Noise	Sample Size	Skewness	Kurtosis
L1 with no noise	1024	1.43	5.66
L1 with white noise @ −20 dB	1024	1.30	5.29
L1 with white noise @ −40 dB	1024	0.56	3.25
L1 with white noise @ −60 dB	1024	0.72	3.48
L1 with white noise @ −80 dB	1024	0.58	3.00
L1 with no noise	2048	1.52	6.67
L1 with white noise @ −20 dB	2048	1.39	6.15
L1 with white noise @ −40 dB	2048	0.67	3.25
L1 with white noise @ −60 dB	2048	0.61	3.27
L1 with white noise @ −80 dB	2048	0.6	3.00
L1 with no noise	4096	1.26	5.17
L1 with white noise @ −20 dB	4096	1.16	4.96
L1 with white noise @ −40 dB	4096	0.62	3.21
L1 with white noise @ −60 dB	4096	0.60	3.11
L1 with white noise @ −80 dB	4096	0.70	3.48
L1 with no noise	8192	1.39	6.00
L1 with white noise @ −20 dB	8192	1.25	5.52
L1 with white noise @ −40 dB	8192	0.62	3.14
L1 with white noise @ −60 dB	8192	0.62	3.26
L1 with white noise @ −80 dB	8192	0.66	3.32

We observe some very interesting trends from the above results. In comparison with the reference parameters of the PDF computed using clean signals, we observe that the distributions tend to become less skewed and so does the kurtosis approach to limiting the kurtosis value to 3.0. A slight increase for the added noise of intensity −60 dB can be ignored here as it does not have a substantial impact on the data distribution. This means the higher noise tends to overcome the real GNSS signals, leaving to follow an even greater corruption, whereas the lower intensity noise tends to normalize the signals to follow a moral and normal distribution-like behavior.

As far as the impact of sample size on the distribution shape is concerned, the values in the table indicate that there is no drastic change in the shape of the bell curve characterized by the corresponding skewness and kurtosis values, as shown in Figure 3.

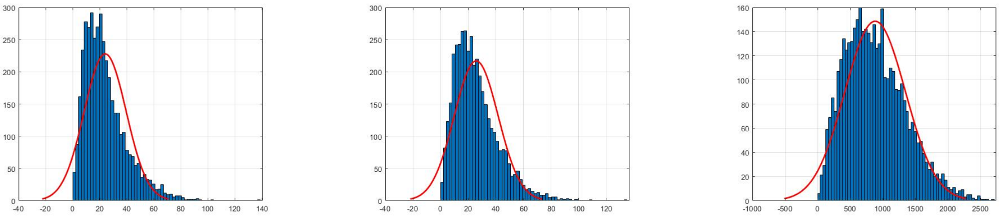

Figure 3. Histogram and distribution fitting for the samples taken from the clean (**left**) and noisy GNSS signals at −20 dB (**center**) and −60 dB (**right**), where the sample size is 4096.

5.2. Shapiro–Wilk Test Results

Typically, the test results are interpreted based on the statistics value and p-value that are compared against the significance level, which is considered to be 0.05. If the p-value is greater than the significance level (usually 0.05), then there is not enough evidence to reject the null hypothesis that the data are normally distributed. This means that the sample is likely to be normally distributed. If the p-value is less than or equal to the significance level, then there is evidence to reject the null hypothesis and conclude that the data are not normally distributed. This means that the sample is unlikely to be normally distributed.

Table 3 tabulates the SW test results for various test cases. It can be seen that the null hypothesis is not rejected in any case, meaning that the injection of noise does disturb the normal behavior of the distribution. However, the parameters pertaining to the SW test paint an interesting picture. Firstly, the decrease in the noise level in the signal decreases the value of the test statistic, and this trend is visible for each sample size in the table. Secondly, the impact of sample size is also evident, as having a bigger sample size increases the value of the test statistic, leading to the conclusion that the SW, in the current case, is sensitive to the sample size.

Table 3. Shapiro–Wilk Test Results performed for various noise levels.

GNSS Signal	Sample Size	Test Statistics Value (w)	p-Value	SK-Test Inference
L1 with white noise @ −20 dB	1024	0.04	1	Null Hypothesis not rejected
L1 with white noise @ −40 dB	1024	0.004	1	Null Hypothesis not rejected
L1 with white noise @ −60 dB	1024	4.4×10^{-5}	1	Null Hypothesis not rejected
L1 with white noise @ −80 dB	1024	4.4×10^{-7}	1	Null Hypothesis not rejected
L1 with white noise @ −20 dB	2048	0.09	1	Null Hypothesis not rejected
L1 with white noise @ −40 dB	2048	0.009	1	Null Hypothesis not rejected
L1 with white noise @ −60 dB	2048	8.8×10^{-5}	1	Null Hypothesis not rejected
L1 with white noise @ −80 dB	2048	8.5×10^{-7}	1	Null Hypothesis not rejected
L1 with white noise @ −20 dB	4096	0.13	1	Null Hypothesis not rejected
L1 with white noise @ −40 dB	4096	0.016	1	Null Hypothesis not rejected
L1 with white noise @ −60 dB	4096	1.79×10^{-4}	1	Null Hypothesis not rejected
L1 with white noise @ −80 dB	4096	1.78×10^{-6}	1	Null Hypothesis not rejected
L1 with white noise @ −20 dB	8192	0.32	1	Null Hypothesis not rejected
L1 with white noise @ −40 dB	8192	0.03	1	Null Hypothesis not rejected
L1 with white noise @ −60 dB	8192	3.48×10^{-4}	1	Null Hypothesis not rejected
L1 with white noise @ −80 dB	8192	3.5×10^{-6}	1	Null Hypothesis not rejected

6. Conclusions

The statistical methods of monitoring the signal quality and assessing the interference in the GNSS signals have been presented in the paper. The assessment covers the subject from two different loosely related methods. The first method is based on the deriving the statistical properties of the distribution formed by the dataset and comparing it with the same using the clean dataset. The second method is based on the SW-based normality testing method that identifies if the selected sample from the complete dataset is representative of the complete dataset by checking if the sample exhibits the properties of a normal distribution. In both cases, we studied the effect of the noise level and sample size on the desired observables. Such assessment could be instrumental in an offsite assessment of the dataset. Even though the preliminary results help us explain the phenomenon, the implementation of the presented techniques should be implemented on a wider range of scenarios with different receiver types, signals, and noise sources.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing is not applicable to this article

Acknowledgments: The author would like to acknowledge the support of Imtiaz Nabi at the Institute of Space Technology, Islamabad, Pakistan, for assisting with real data collection using the RTL-SDR.

Conflicts of Interest: The author declares no conflict of interest.

References

1. U.S. Coast Guard Navigation Center. GPS Navigation Problems. Available online: https://www.navcen.uscg.gov/?pageName=GPSProblems (accessed on 20 March 2023).
2. Nandra, R.; Sridevan, S.; Hancke, R. GPS Spoofing: The Next Generation. *IEEE Commun. Mag.* **2018**, *56*, 180–186.
3. Humphreys, T.E.; Ledvina, B.M. Assessing the Spoofing Threat: Development of a Portable GPS Civilian Spoofer. *J. Field Robot.* **2013**, *30*, 531–548.
4. Dovis, F. *GNSS Interference Threats and Countermeasures*; Artech House: Norwood, MA, USA, 2015.
5. Thompson, R.J.; Wu, J.; Balaei, A.T.; Dempster, A.G. Detection of RF interference to GPS using day-to-day C/N_0 differences. In Proceedings of the International Symposium on GNSS, Taipei, Taiwan, 26–28 October 2010.
6. Sun, K.; Zhang, T. A New GNSS Interference Detection Method Based on Rearranged Wavelet–Hough Transform. *Sensors* **2021**, *21*, 1714. [CrossRef] [PubMed]
7. Sun, K.; Jin, T.; Yang, D. An Improved Time-Frequency Analysis Method in Interference Detection for GNSS Receivers. *Sensors* **2015**, *15*, 9404–9426. [CrossRef] [PubMed]
8. Balaei, A.T.; Dempster, A.G. A statistical inference technique for GPS interference detection. *IEEE Trans. Aerosp. Electron. Syst.* **2009**, *45*, 1499–1511. [CrossRef]
9. Motella, B.; Pini, M.; Presti, L.L. GNSS interference detector based on Chi-square Goodness-of-fit test. In Proceedings of the 2012 6th ESA Workshop on Satellite Navigation Technologies (Navitec 2012) and European Workshop on GNSS Signals and Signal Processing, Noordwijk, The Netherlands, 5–7 December 2012; pp. 1–6. [CrossRef]
10. Specht, M. Statistical Distribution Analysis of Navigation Positioning System Errors—Issue of the Empirical Sample Size. *Sensors* **2020**, *20*, 7144. [CrossRef] [PubMed]
11. Lilliefors, H.W. On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown. *J. Am. Stat. Assoc.* **1967**, *62*, 399–402. [CrossRef]
12. Smirnov, N. Table for Estimating the Goodness of Fit of Empirical Distributions. *Ann. Math. Stat.* **1948**, *19*, 279–281. [CrossRef]
13. Shapiro, S.S.; Francia, R. An Approximate Analysis of Variance Test for Normality. *J. Am. Stat. Assoc.* **1972**, *67*, 215–216. [CrossRef]
14. Shapiro, S.S.; Wilk, M.B. An Analysis of Variance Test for Normality (Complete Samples). *Biometrika* **1965**, *52*, 591–611. [CrossRef]
15. Ahmad, F.; Khan, R.A. A power comparison of various normality tests. *Pak. J. Stat. Oper. Res.* **2015**, *11*, 331–345. [CrossRef]
16. RTL-SDR. Available online: https://www.rtl-sdr.com/about-rtl-sdr/ (accessed on 20 March 2023).