Article

Normality Analysis for RFI Detection in Microwave Radiometry

Jose Miguel Tarongi * and Adriano Camps

Remote Sensing Lab (RSLAB), Department of Signal Theory and Communications (TSC), Universitat Politècnica de Catalunya (UPC), IEEC/CRAE-UPC, Campus Nord, Building D, 08034 Barcelona, Spain; E-Mail: camps@tsc.upc.edu

*   Author to whom correspondence should be addressed; E-Mail: jose_miguel_tarongi@tsc.upc.edu; Tel.: +34-934-017-361; Fax: +34-93-4017-232.

Abstract: Radio-frequency interference (RFI) present in microwave radiometry measurements leads to erroneous radiometric results. Sources of RFI include spurious signals and harmonics from lower frequency bands, spread-spectrum signals overlapping the "protected" band of operation, or out-of-band emissions not properly rejected by the pre-detection filters due to its finite rejection. The presence of RFI in the radiometric signal modifies the detected power and therefore the estimated antenna temperature from which the geophysical parameters will be retrieved. In recent years, techniques to detect the presence of RFI in radiometric measurements have been developed. They include time- and/or frequency domain analyses, or time and/or frequency domain statistical analysis of the received signal which, in the absence of RFI, must be a zero-mean Gaussian process. Statistical analyses performed to date include the calculation of the Kurtosis, and the Shapiro-Wilk normality test of the received signal. Nevertheless, statistical analysis of the received signal could be more extensive, as reported in the Statistics literature. The objective of this work is the study of the performance of a number of normality tests encountered in the Statistics literature when applied to the detection of the presence of RFI in the radiometric signal, which is Gaussian by nature. A description of the normality tests and the RFI detection results for different kinds of RFI are presented in view of determining an omnibus test that can deal with the blind spots of the currently used methods.

Keywords: microwave radiometers; radio frequency interference; detection; mitigation

1. Introduction

The performance of microwave radiometers can be seriously degraded by the presence of radio-frequency interference (RFI). This is specially important in populated land areas [1–3]. When RFI power is larger than the variance of the measured thermal noise, a simple time-domain algorithm can be implemented to detect this power variation, and the data is eliminated ("blanking"). However, when RFI is comparable or smaller than this variance, it becomes more difficult to detect, and the estimated power may be erroneous [3]. Furthermore, RFI may be present even in the calibration, producing a systematic error in the whole data set. Nowadays, several RFI detection and mitigation methods have been developed. The algorithms more widely used to solve this problem are: time, frequency or statistical analyses of the received signal. Time analysis consists of the detection of power peaks in the received signal that are larger than the variance of the measured thermal noise in the absence of RFI. These values are blanked (discarded) and only the RFI-free data are left [4,5]. However, since the detected power is a smoothed (averaged) version of the instantaneous one, if the duration of the RFI peaks is shorter than the integration time they may pass undetected. Frequency analysis is based on the study of the received signal spectrum, and consists of discarding sub-bands with power higher than the variance of the thermal noise [6]. Finally, statistical analysis is based on the fact that the radiometric signal (thermal noise) follows a determined probability distribution, which in the absence of RFI, is a zero-mean random Gaussian variable. Therefore, the probability density function and the statistical parameters, such as the moments, are perfectly known. The best known time domain statistical analysis in microwave radiometry is the Kurtosis (ratio of the fourth moment and the square of the second moment) which must be equal to 3 in RFI-free conditions [7–9]. The sixth order moment [10], and other algorithms [11−14] have also been studied.

In this research, the suitability of several normality tests for RFI detection in microwave radiometry has been analyzed. The normality tests used in this study are: Jarque-Bera (JB), Shapiro-Wilk (SW), Chi-square (CHI2), Anderson-Darling (AD), Lilliefors-Smirnov-Kolmogorov (L), Lin-Muldhokar (LM), Agostino-Pearson (K2), Cramer-von Mises (CM), in addition to the Kurtosis (K) and Skewness (S) statistical parameters to detect signals nonnormality. The ultimate objective is to compare these normality tests to obtain an omnibus test to detect RFI, or at least, the best normality test for a determined type of RFI. Extensive Monte Carlo simulations have been used to compare tests performance.

A brief description of these tests is given in Section 2. Section 3 analyzes the validity of these tests in relation to its probability of false alarm. Section 4 shows the results obtained in the different analyses of a thermal noise signal (Gaussian) contaminated with different RFI signals. Finally Section 5 summarizes the conclusions obtained in this study.

2. Normality Tests

The rationale behind the use of normality tests to detect RFI in microwave radiometry is the fact that the thermal noise signal measured by radiometers, follows a zero-mean Gaussian distribution, while in general, man-made RFI are not Gaussian.

Normality tests used in this study have been widely used in statistical literature and are described in this section. Some of these tests can be used with probability distributions different from the normal distribution.

The radiometric signal must be sampled following the Nyquist theorem, which states that the sampling frequency f_s must be at least twice the signal's bandwidth B. In this study, a sample is defined as the value of the received signal amplitude obtained by an Analog-to-Digital Converter (ADC) every $T_s = 1/f_s$ seconds. Thus, the number of samples or sample size is the total number of values obtained in the sampling process of the signal.

A brief summary of the statistical tests used in this study is provided below:

Kurtosis test: Kurtosis is a statistical parameter related to the shape of the probability density function (PDF) of a random variable. Kurtosis of a Gaussian random variable is always 3 independently of its mean and variance. Assuming a random process X, the Kurtosis (K) follows:

$$K = \frac{\mu_4}{\sigma^4} = \frac{E\left[\left(X - E[X]\right)^4\right]}{\left(E\left[\left(X - E[X]\right)^2\right]\right)^2} \tag{1}$$

$$\hat{K} = \frac{\frac{1}{N}\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^4}{\left(\frac{1}{N}\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^2\right)^2} \tag{2}$$

where N is the sample size and $\overline{X}$ represents the sample mean of X. Equation (2) defines the Kurtosis estimator $\hat{K}$ used in this study. In the specific case that the random process is a zero-mean Gaussian process, the value of $\hat{K}$ tends to 3 as the sample size increases. The Kurtosis test consists of comparing the estimated Kurtosis value of the received signal with tabulated values of the cumulative distribution function (CDF) of the Kurtosis of a Gaussian random variable of N samples. The Kurtosis parameter has been used in microwave radiometry RFI detection, although it exhibits some problems in detecting some particular signals [7–9,11].

Skewness test: Skewness is a statistical parameter related to the asymmetry of the PDF of a random variable. In this case, the Skewness of a Gaussian random variable is always 0. Assuming a zero-mean random process X, the Skewness (S) follows:

$$S = \frac{\mu_3}{\sigma^3} = \frac{E\left[\left(X - E[X]\right)^3\right]}{\left(E\left[\left(X - E[X]\right)^2\right]\right)^{3/2}} \tag{3}$$

$$\hat{S} = \frac{\frac{1}{N}\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^3}{\left(\frac{1}{N}\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^2\right)^{3/2}} \tag{4}$$

Equation (4) defines the Skewness estimator $\hat{S}$ used in this study. In the specific case that the random process is Gaussian $\hat{S}$ tends to 0 as the sample size increases. The Skewness test is based on

comparing the estimated Skewness value of the received signal with tabulated values of the Skewness of a Gaussian random variable. Kurtosis and Skewness CDF tables have been computed from 2^{16} Monte-Carlo simulations.

Jarque-Bera test (JB): JB test is a normality test based on the Skewness and the Kurtosis of the process, e.g. analyses the normality of a process taking into account both the Kurtosis and the Skewness of this process. JB test is defined as:

$$\text{JB} = \frac{N}{6}\left(\hat{S}^2 + \frac{\left(\hat{K}-3\right)^2}{4}\right) \tag{5}$$

where N is the sample size, $\hat{S}$ is the Skewness estimator of the process, and $\hat{K}$ is the Kurtosis estimator of the process. In case of normality $\hat{S}$ and $\hat{K}$ are asymptotically independent, and hence, the JB test asymptotically has a chi-square distribution with two degrees of freedom. This fact unfortunately leads to an error measurement when the sample size is low [15–17].

D'Agostino K-squared test (K2): K^2 test, like JB test, is also based on the Skewness and the Kurtosis of the process, with the particularity that the Skewness and the Kurtosis of the process are first transformed to avoid the error measurements present in the JB test when the sample size is low. In case of normality K^2 test follows a chi-squared distribution even with a low sample size. The definition of the K^2 test is complex and it can be consulted in [18] for the interested readers.

Kolmogorov-Smirnov (KS) and Lilliefors (L) tests: KS test is based on the empirical distribution function (EDF); given N ordered values of a sample X the EDF is defined as:

$$\hat{F}_N\left(x\right) = \frac{\#\left\{X_i \mid X_i \leq x\right\}_{\forall i=1..N}}{N} = \frac{1}{N}\sum_{i=1}^{N} I\left(X_i \leq x\right) \tag{6}$$

where $I(\cdot)$ is the indicator of the event, X_i is the ith element of the sample to be tested, whose values must be ordered from the lowest to the highest, and $\hat{F}_N\left(x\right)$ is a step function that increases by $1/N$ at the value of each ordered data point. KS test correlates the empirical distribution function with the normal distribution function, with a determined mean and variance that must be known. Since the mean and the variance are usually not known parameters, this test is replaced by the L test to avoid the errors introduced by a wrong variance estimation. The L test is a slight modification of the KS test in which the mean and variance of the normal distribution are obtained from the sample X [19]. The L test is defined as:

$$\text{L} = \max_{1 \leq i \leq N}\left| F\left(Y_i\right) - \hat{F}\left(X_i\right)\right| \tag{7}$$

where $\hat{F}\left(X_i\right)$ is the value of the ith element of the EDF of X, and $F(Y_i)$ is the value of the ith element of the normal distribution function with (L test case) mean and variance (σ_Y^2) equal to:

$$\overline{Y} = \frac{1}{N}\sum_{i=1}^{N} X_i \tag{8}$$

$$\sigma_Y^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^2 \tag{9}$$

L confidence values are obtained from the CDF of the L test result when applied to a Gaussian distribution. Hence, these values represent the result of the test in case of normality [19]. A total of 2^{16} Monte-Carlo simulations have been performed to obtain these confidence values, which are tabulated as the previous Kurtosis and Skewness values.

On the other hand, the L test has still one limitation, it tends to be more sensitive near the center of the distribution, than at the tails. In general, the probability of detection will be set as high as possible, making more important the tails of the distribution than the center.

Anderson-Darling test (AD): The AD test is a modification of the L test that gives more weight to the tails than the L test. As this test is also based on the comparison of distribution functions, the values of the sample to test must be ordered. This test consists of:

$$AD^{*2} = -N - \frac{1}{N} \sum_{i=1}^{N} (2i-1)\left(\ln \Phi(Y_i) + \ln\left(1 - \Phi(Y_{N+1-i})\right) \right) \tag{10}$$

with:

$$Y_i = \frac{X_i - \overline{X}}{\sigma_X} \tag{11}$$

$\Phi(\cdot)$ represents the standard normal cumulative distribution function (CDF) operator. As it is described in [20], AD^{*2} must be adjusted for the sample size as follows:

$$AD^2 = AD^{*2}\left(1 + \frac{0.75}{N} + \frac{2.25}{N^2} \right) \tag{12}$$

Critical values can be consulted from tables [20], although an empirical development of these critical values for the normal case is presented in [21].

Shapiro-Wilk test (SW): The SW test belongs to the EDF comparison group of tests. Again, samples must be sorted from the lowest to the highest values in order to be able to use this normality test. The SW test is defined as:

$$SW = \frac{\left(\sum_{i=1}^{N} a_i X_i \right)^2}{\sum_{i=1}^{N} \left(X_i - \overline{X} \right)^2} \tag{13}$$

The main part of the SW test is the vector of coefficients a_i, $i = 1...N$. These coefficients are tabulated in [22] for the case of less than 50 samples, or can be analytically calculated [23]. Furthermore, to ease the application of the SW test, in [24] the SW test has been transformed to have a normal distribution in the case of normality of the tested signal. A drawback of this test is the limitation of the sample size to a maximum of 2,000 values [24]. Longer sample lengths can be tested by dividing it in several shorter length sets of samples and calculating the SW test on each set, and averaging the results as they are normally distributed [13,14].

Cramer-von Mises test (CM): The CM test is a variation of the L test. It is defined as:

$$CM = \frac{1}{12N} + \sum_{i=1}^{N} \left(\frac{2i-1}{2N} - \Phi(Y_i) \right)^2 \tag{14}$$

where Y_i has been already defined in (11). Confidence values are obtained following the same methodology as with the L test.

Lin-Mudholkar test (LM): The LM test is based on the fact that the mean and variance of a random sample are independently distributed, if and only if, the parent population is normal (a simple test for normality against asymmetric alternatives). The LM test of a sample X is defined as [25,26]:

$$LM = 0.5 \log\left(\frac{1+R}{1-R}\right) \qquad (15)$$

being R the cross-correlation between the sample X and the sample Y:

$$R = \frac{\sum_{i=1}^{N}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{N}(X_i - \overline{X})^2 \sum_{i=1}^{N}(Y_i - \overline{Y})^2}} \qquad (16)$$

$$Y_i = \left[\frac{1}{N}\left[\left(\sum_{i \neq j} X_j^2\right) - \frac{\left(\sum_{i \neq j} X_j\right)^2}{N-1}\right]\right]^{1/3} \qquad (17)$$

The LM parameter presents a normal distribution, however it is not normalized neither in the mean, nor in the variance. In [26] it is specified the procedure to normalize the LM test. This test is sensitive only to departures from normality due to Skewness. As shown in [27], this procedure is generally much more powerful at detecting Skewness than the Skewness coefficient itself, although it has little power in detecting nonnormal symmetrical distributions. As an example, uniform or platykurtic distributions will pass this test easily, and therefore other tests such as AD test have to be also used in conjunction to it.

Chi-square test (CHI2): The strong point of the chi-squared goodness-of-fit test is that it can be used to test if a data sample belongs to a process with a determined distribution. However, in our application this is not an advantage as the only distribution to be analyzed is the normal one. On the other hand the CHI2 test requires a sample size high enough for the chi-square approximation to be valid. CHI2 test of the sample X is defined as:

$$CHI2 = \sum_{i=1}^{N}\frac{(O_i - E_i)^2}{E_i} \qquad (18)$$

where CHI2 represents the result of the test, which asymptotically approaches to a χ^2 distribution, O_i is the frequency of the ith possible outcome of the sample (assuming O to be an histogram of the sample X), E_i is the theoretical frequency (which in our case is the histogram of the zero mean unit variance random normal process), and N is the number of possible outcomes of each event.

3. Validation of Normality Tests in the Absence of RFI

In order to evaluate the performance of the different tests listed in section 2, two error types must be introduced first:

Type I error: rejection of a true hypothesis. In our context, type I error is also known as probability of false alarm (P_{fa}). This error is produced when, in the absence of RFI sources, the algorithm "detects" the presence of RFI in a determined sample, leading to the blanking (elimination) of correct data, thus reducing the total integration time.

Type II error: acceptance of a false hypothesis. In our context, type II error is known as probability of missed detection (P_{miss}). This error is produced when a RFI is present in the signal, but it is not detected, leading to an erroneous measurement, but it is assumed to be correct. Probability of detection (P_{det}) used in this work is defined as $1 - P_{miss}$.

It is obvious that the objective is to obtain a low probability of false alarm and a high probability of detection; but both types of errors have a strong correlation, where the setting of the value of one of these parameters determines the value of the other. This way, if it is desired to minimize the probability of false alarm, the RFI detection threshold must be set to a relative high value, leading to a low probability of detection. On the other hand, if it is desired to maximize the probability of detection, the RFI detection threshold must be set to a relative low value, leading to a high probability of false alarm. A good way to evaluate the compromise between P_{det} and P_{fa}, is the calculation of the so-called Receiver Operating Characteristic (ROC) curves $P_{det}(P_{fa})$ [11,13,14]. In addition to ROC plots of the Interference to Noise Ratio (INR) as a function of the sample size and other parameters are shown.

Validation of the normality tests is performed in order to minimize the errors in the threshold calculation for a determined pair of values P_{det} and P_{fa}. Normality test errors introduced by the different tests in the P_{det} and P_{fa} values must be acceptable. The method followed to check the normality tests is the calculation of the ROC curve by means of 2^{15} Monte Carlo simulations of a Gaussian signal in the absence of RFI, for every test, and varying determined parameters (sample size, quantization level). In this study a test is considered valid when the error between the ROC curve of the test and the RFI free case ROC curve ($P_{det} = P_{fa}$) is less than the 5% (e.g. for $P_{fa} = 0.1 \rightarrow 0.095 < P_{det} < 0.105$).

Quantization has been modeled by varying the number of discretization levels (determined by the number of bits), and assuming a dynamic range of the ADC of $\pm 8\sigma$, where σ^2 is the RFI-free noise power, to avoid signal clipping.

Figures 1a and 1b represent the ROC curves for the SW test in the case of calculating the test in blocks of 2,048 and 4,096 samples respectively, and averaging the results to obtain a Gaussian distribution [24]. Test validation has been performed to these two sizes of the sets to obtain the largest sample size that can be used with the SW test. For the 4,096 samples case, performance of the SW test decreases with the sample length, while in the case of 2,048 samples the SW test does not have significant variations with respect to the ideal behavior. Explanation of this behavior can be found in [23], where the performance of the SW test is guaranteed for a maximum of 2,000 samples, but not for larger sample lengths. The number of bits has been set to 20 ($2^{20}-1$ quantization levels) in both cases to avoid quantization errors.

Figure 1. Comparison between the performance of Shapiro-Wilk test with block size set of 2,048 (Figure 1a) and 4,096 (Figure 1b). 20 quantization bits have been used to neglect quantization errors. For 4,096 samples (>2,000 [23]) the performance degrades, while does not for 2,048 samples.
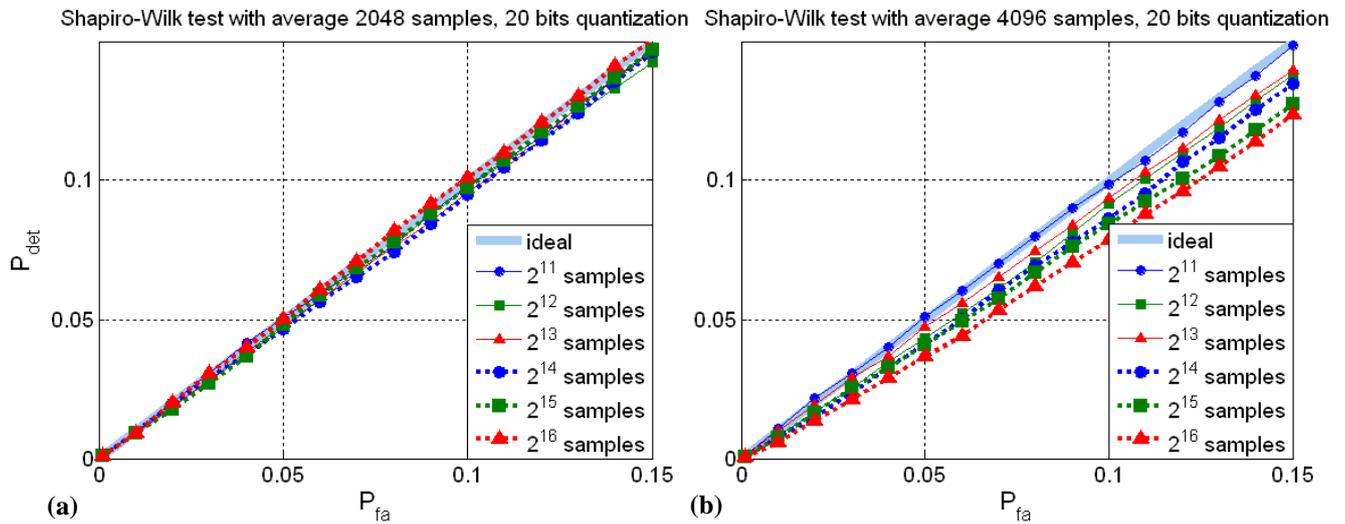

Figure 2 represents the validity of the SW test, as a function of the sample length and the quantization level. Solid lines represent the cases where the test is considered to be valid (error <5%), while dotted and dash-dotted lines represent invalid test cases. As it can be appreciated, the longer the sample size, the more the quantization levels are required since the quantization process introduces a discretization error of the PDF. Quantization makes the normal distribution to become similar to a binomial distribution, which is detected as non-normal. Therefore, as the sample size increases, the number of quantization levels must increase to avoid this "change of distribution" from normal to binomial.

Figure 2. Shapiro-Wilk validity test as a function of the sample length and the quantization level. Dotted and dash-dotted lines represent invalid test cases (error >5%), solid lines represent cases where the test is valid (error <5%). Lower number of bits in lines of the same color (same sample size) lead to sample sets "detected" as interference.
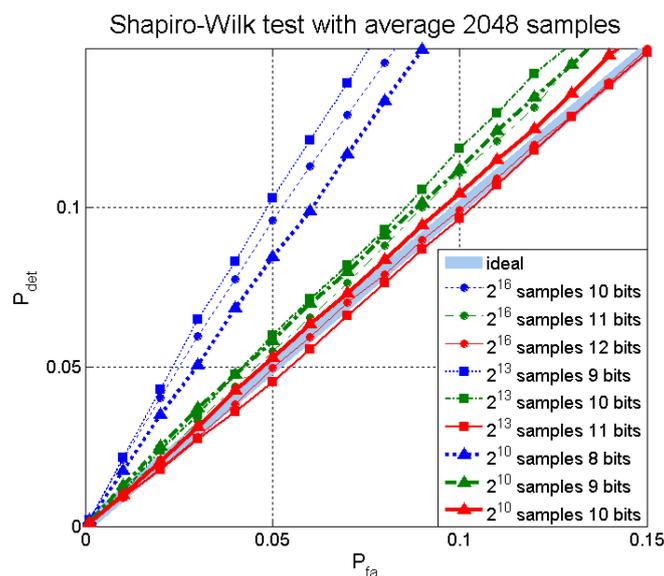
Figure 3 shows the validity of the normality tests as a function of the quantization bits (and the sample size (the lower the better, except for K2 and JB test). Actually, the validity of the tests is more influenced by the number of quantization bits than by sample size, provided it is high enough.

Figure 3. Minimum number of bits to neglect the quantization error, ECDF based test need more quantization bits as the sample size increase. As chi2 test, K2 test and JB tests are asymptotic, do not work properly for a low sample size. K, S and LM tests do not appear in graph as performance of these tests is acceptable for less than 6 quantization bits even for a sample size of 2^{16}.
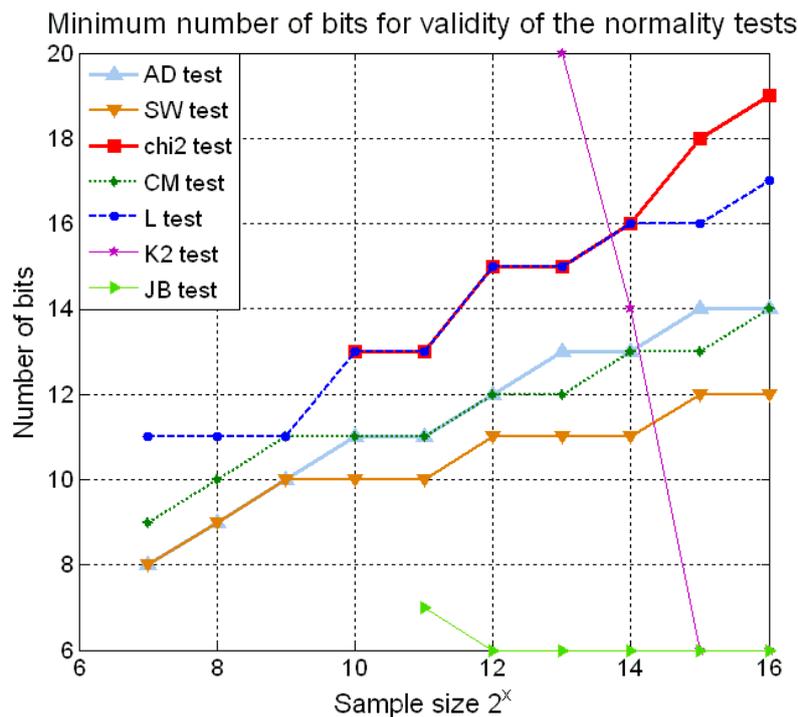
4. Performance of Normality Tests for Different RFI Types

The performance of each normality test is calculated for different types of RFI representative of the ones actually encountered:

Pulsed sinusoidal signal: this signal has been extensively studied in microwave radiometry [7–11,13,14], as it is a common interference signal (e.g. a radar signal or a third order intermodualtion product spurious signal). The signal model of this RFI is described as:
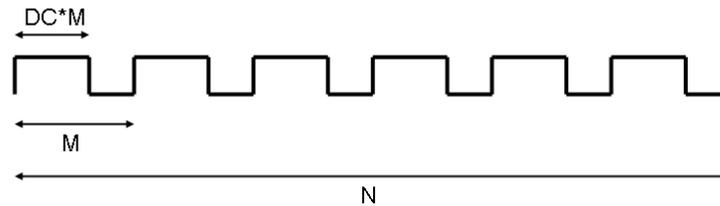
$$PS[i] = A\cos\left(2\pi f_0 T_s i + \varphi_0\right) H[i] \qquad i = 1..N \tag{19}$$

where *PS[i]* is the sampled pulsed sinusoidal signal, *A*, *f₀* and *φ₀* are the amplitude, frequency, and initial phase of the RFI respectively, T_s is the sampling period, and *H[i]* is a train of pulses described as:

$$H[i] = \begin{cases} 1 & \mathrm{mod}_M(i) \leq M \cdot DC \\ 0 & \mathrm{otherwise} \end{cases} \qquad i = 1..N \tag{20}$$

where *N* is the sample length, *M* is the pulse length, and *DC* is the duty cycle factor of every pulse. Hence the RFI is a train of *N/M* pulsed sinusoidal signals of *DC* duty cycle factor.

Figure 4. Representation of *H*[*i*] function, where *N* is the total sample size, *N/M* is the total number of pulses and *DC**M is the pulse length (in samples).


Chirp signal: A chirp consists of a linearly varying frequency sinusoidal signal. The signal model of the RFI chirp is described as:

$$CH[i] = A\cos\left(\left(2\pi f_0 + \pi\beta T_s i\right)T_s i + \varphi_0\right)H[i] \qquad i=1..N \tag{21}$$

where *CH*[*i*] is the sampled pulsed chirp signal, β is the chirp rate of the linear frequency modulation, A, f_0 and φ_0 are the amplitude, initial frequency, and initial phase of the RFI respectively, and *H*[*i*] is the train of pulses function described in Equation (20).

Pseudo-Random Noise (PRN) signal: a PRN signal is a signal that satisfies one or more standard tests for statistical randomness. This signal consists of a deterministic sequence of pulses with (−1 and 1 values) that repeats itself after a period, which is usually very long, leading to a spread spectrum behavior of the signal. Without loss of generality in this work, the firsts 10,230 output bits of a maximum length sequence generator of 14 stages (Figure 5) are used as the deterministic sequence of the PRN interfering signal.

Figure 5. Maximum length sequence generator from which the PRN interfering signal used in this work is obtained. PRN interfering signal is composed by the repetition of the first 10,230 output bits.
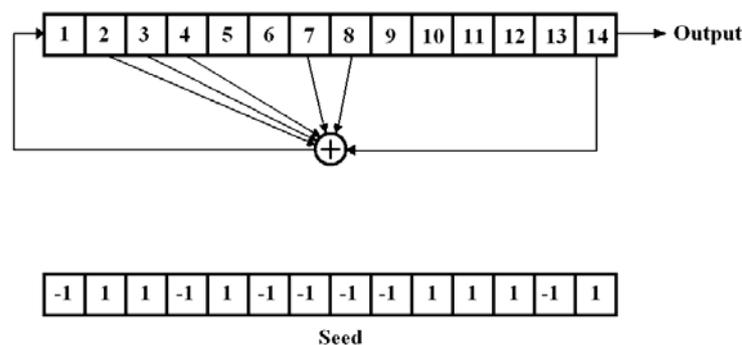

Telegraphic signal: this signal is a baseband digital amplitude modulated signal (*i.e.* ASK or OOSK). Telegraphic signal consist of a sequence of pulses (which values can be 0 and 1 or −1 and 1) where every bit of the message is modulated as one independent pulse. Duration of the pulse will determine the transmission speed. This signal is usually scrambled as originally it contents a high entropy value.

4.1. Pulsed Sinusoidal Signal

It is widely known [7–11] that a pulsed sinusoidal interfering signal of 0.5 duty cycle cannot be detected by the Kurtosis test as this signal has a Kurtosis equal to 3, independently of the frequency of the interfering signal. Hence, a study of different alternatives to detect non-Gaussian signals is performed. Figure 6a shows the performance of different normality tests for sample sizes of 1,024 samples (dotted lines) and 16,384 samples (solid lines). The performance is measured in terms of the required INR to obtain a ROC curve with a $P_{det} = 0.9$ for a $P_{fa} = 0.1$. In order to get reliable results, performance has been calculated as the average of 2^{15} Monte-Carlo simulations. In the following figures, some tests are not plotted due to its poor performance, as it is usually the case of Skewness, Lin-Mudholkar and chi-square tests.

Figure 6. Normality test performance in the detection of a pulsed sinusoidal interference of 1,024 samples (dotted line) and 16,384 samples (solid line) (Figure 6a), and a chirp signal of 16,384 samples (Figure 6b) as a function of signal's duty cycle. Both graphs represent the INR value required to obtain a ROC curve with a $P_{fa} = 0.1$ for $P_{det} = 0.9$. SW test has the best performance around duty cycle = 0.5, followed by AD test for 1,024 samples, while AD tests presents the best performance for 16,384 samples since SW test performance is degraded due to averaging. For low or high duty cycles the K test outperforms. JB and K2 tests results can not be used with 1,024 samples. Parameters f_0 and φ_0 are selected at random for the pulsed sinusoidal interference, while are defined as $f_0 = 2 \times 10^{-5}$ and $f_N = 0.15$ with $N = 16,384$; and $\varphi_0 = 0°$ for the pulsed chirp interference. Results obtained from a Monte-Carlo set of 2^{15} simulations.
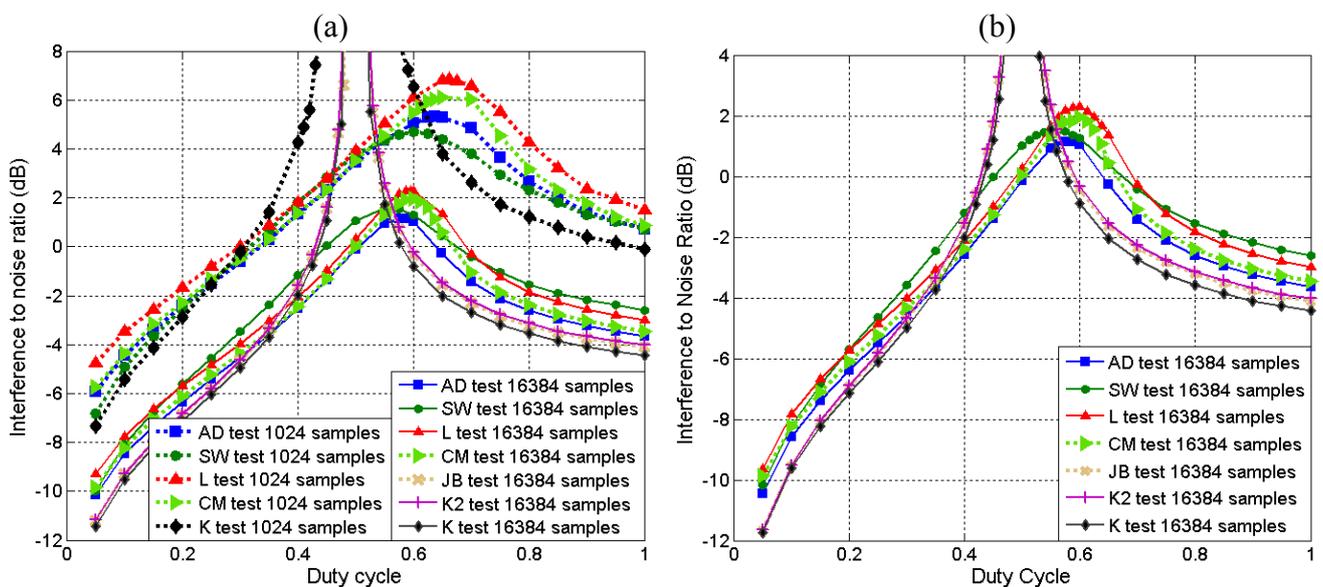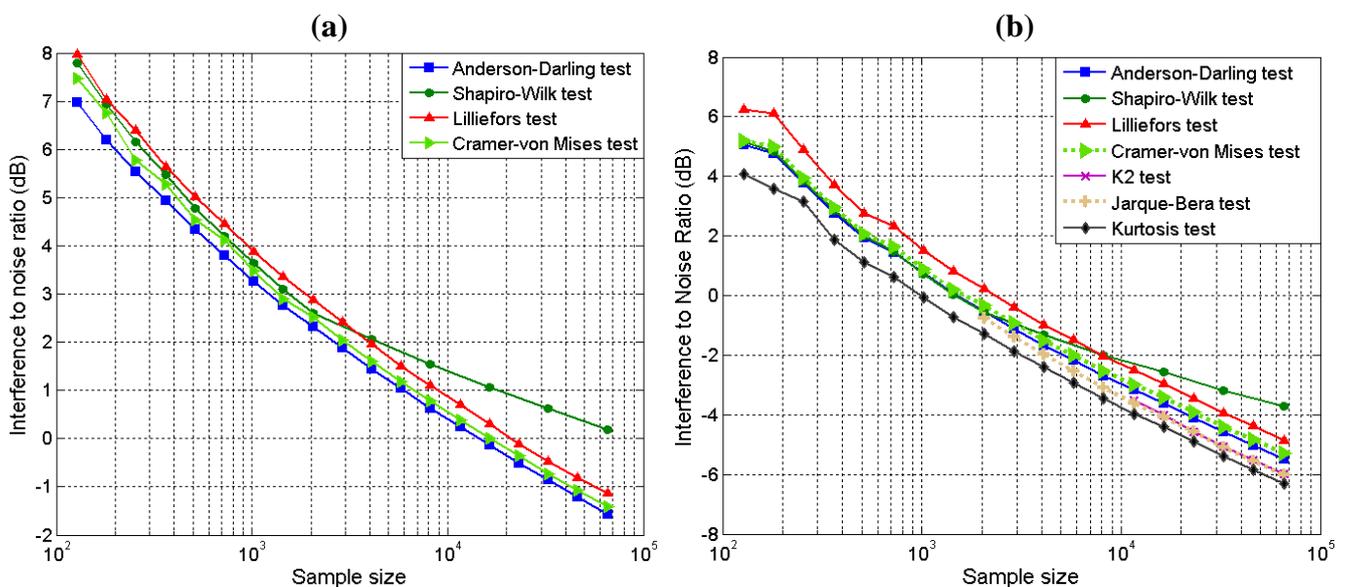

Best normality tests to detect pulsed sinusoidal interfering of duty cycle around 0.5 are the AD, the SW, the CM and the L tests, with improving performance as the sample size increases. These four ECDF based normality tests perform in a similar manner, having a better performance the SW test for

shorter sample sizes and the AD test for longer sample sizes. The CM and L tests have a performance in between the other tests for large duty cycle.

For duty cycle values different from 0.5 and nearby values, however, Kurtosis test and Kurtosis-related tests outperform. In Figure 6a Kurtosis and Kurtosis related tests for have a peak around 0.5 that narrows with increasing sample size. Performance of JB and K2 test are worse than the Kurtosis alone, since these two tests depend also on the Skewness parameter which is zero in the analyzed signal.

In Figure 7a the performance of the AD, SW, L and CM normality tests in the detection of a pulsed sinusoidal signal of exactly 0.5 duty cycle is compared as a function of the sample size. K, JB and K2 tests are not present in Figure 8 since they cannot detect sinusoidal signals of 0.5 duty cycle. AD, L and CM tests follow almost the same trend while SW test has a different trend for sample lengths of 4,096 and above since blocks of 2,048 samples have to be averaged to ensure a good performance of the test.

Figure 7. Normality test performance in the detection of a 0.5 duty cycle pulsed sinusoidal interference (Figure 7a) and a chirp interfering signal (Figure 7b) as a function of the signal's sample size. Both graphs represent the INR value required to obtain a ROC curve with a $P_{fa} = 0.1$ for $P_{det} = 0.9$. In the 7a graph Kurtosis test is not present, and AD test has the best performance as sample size increases while in the 7b graph Kurtosis algorithm has the best performance followed by the Kurtosis based normality tests (JB and K2). In both cases SW test performance does not improve as fast as the others above 2,048 samples due to averaging and CM test has a slightly worse performance than AD test. Parameters f_0 and φ_0 are selected at random for the pulsed sinusoidal interference, while are defined as $f_0 = 2 \cdot 10^{-5}$ and $f_N = 0.15$ with $N = 16,384$; and $\varphi_0 = 0°$ for the pulsed chirp interference. Results obtained from a Monte-Carlo set of 2^{15} simulations.
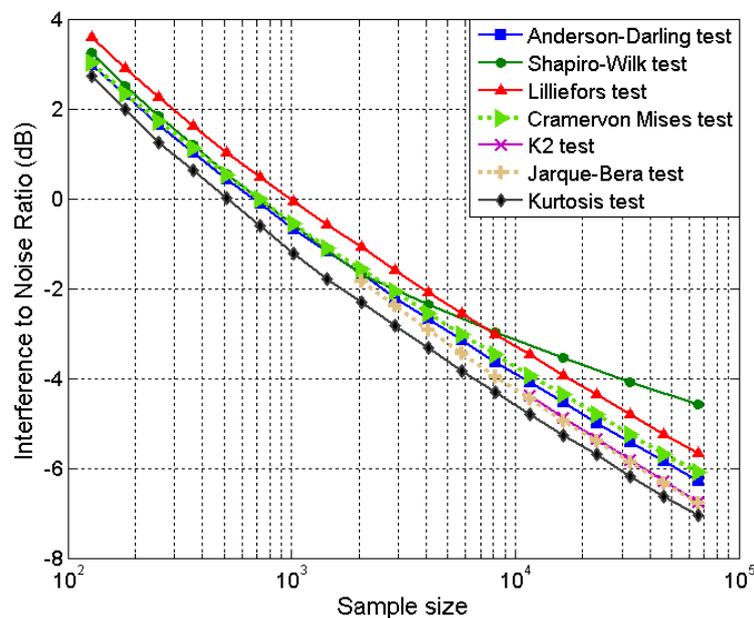

4.2. Chirp Interfering Signal

The performance of all the normality tests in the detection of an interfering chirp signal with a variable duty cycle is very similar to the case of the detection of a pulsed sinusoidal signal. In fact,

a 0.5 duty cycle blind spot is still present for Kurtosis algorithm and Kurtosis-based normality tests. Figure 6b presents K, K2, JB, AD, SW, L, and CM test performance in the detection of a pulsed chirp signal as a function of the duty cycle, showing the similarities commented previously. Furthermore Figure 6a (solid lines) and 6b shows similar trends for the different normality tests and duty cycles, even the INR values of the tests are almost the same in both figures.

In the case that the signal has a duty cycle of 1, the best detection algorithm is the Kurtosis test followed by the Kurtosis-based and Skewness-based tests. Figure 10 shows the INR trend Vs. sample size: The AD, L, CM, K, JB and K2 tests follow almost the same trend, while SW test has a different trend for sample length of 4,096 and above, as in Figure 8. JB and K2 tests start taking values of INR at 2,048 and 11,585 number of samples respectively as these tests are not considered valid for a lower sample size (see Section 3). The K, JB and K2 tests are shown (since as the duty cycle of the chirp is 1 and not 0.5 as in Figure 7a).

Figure 8. Normality test performance in the detection of a PRN interference as a function of the signal's sample size. Graphs represent the INR value required to obtain a ROC curve with a $P_{fa} = 0.1$ for $P_{det} = 0.9$. Kurtosis test has the best performance, followed by JB and K2 tests. Best ECDF test performance is obtained with the AD test followed by the CM test. SW test performance does not improve as the rest for large sample size due to averaging. Results obtained from a Monte-Carlo set of 2^{15} simulations.


4.3. Pseudo-Random Noise Interfering Signal

Figure 8 shows the performance of the different tests Vs the sample size. The Kurtosis test achieves the best performance in the detection of this kind of interfering signal, followed by JB and K2 tests. Hence, Kurtosis-based tests perform better than ECDF-based tests (AD, L, CM, SW) in the detection of PRN signals. Performance of ECDF-based tests is quite similar to the sinusoidal and chirp interfering signals, obtaining the best results with the AD and the SW tests for lower sample sizes, and with the AD and CM tests for higher sample sizes.

Figure 9 shows the ROC curves of the normality tests performance in the detection of a PRN signal of 16,384 number of samples; in this figure, Kurtosis gets a value of P_{det} = 0.9 for P_{fa} = 0.1. In this figure it is clearly shown that Kurtosis outperforms for the same INR as the rest of normality tests, followed by the two Kurtosis-based and Skewness-based tests (JB and K2 tests) which have almost the same behavior. The four ECDF based tests (AD, CM, L and SW tests) have a worse performance. The worst performance is obtained by the S, LM and CHI2 tests. S and LM tests fail as the interfering signal Skewness is zero. CHI2 test has a lesser performance in the RFI detection than the ECDF-based tests and Kurtosis based tests, therefore it is not recommended in the RFI detection.

Figure 9. Normality tests performance in the detection of a PRN interference signal. Sample size is 16,384 and INR is −5.2 dB. It is appreciated the difference in performance between the Kurtosis and the rest of tests. Also, bad performance of the Skewness, LM and CHI2 tests is shown.
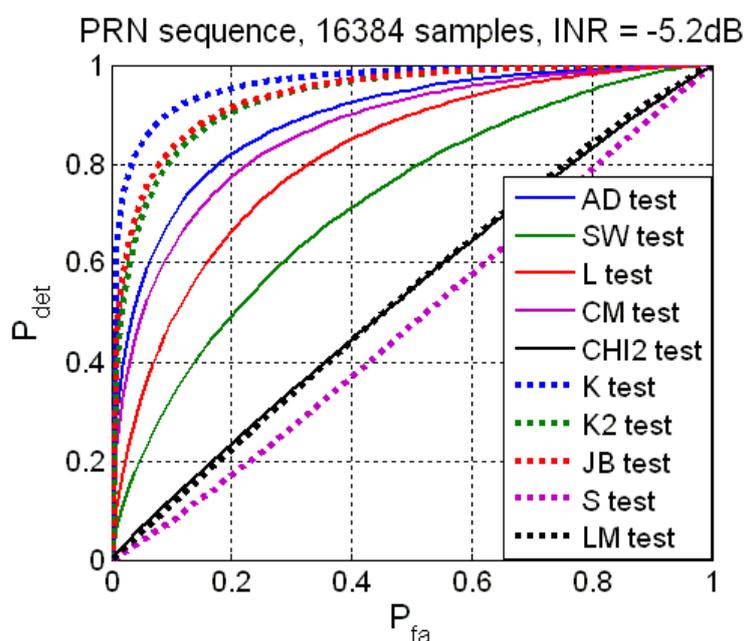
4.4. Telegrafic Interfering Signal

For this type of interfering signal, depending on the actual message transmitted the performance of the different normality test is quite variable. In this study, three different interfering signals have been used, called messages 1, 2 and 3. Message 1 is a plain text file (very low randomness), message 2 is a zipped file and message 3 is a jpg file (high level of randomness, as redundancy is eliminated). In Figures 10 and 11 ROC curves of message 1 are presented as a function of the sample size and the INR, respectively. As for other kinds of interfering signals increasing the sample size increases the probability of RFI detection, and the normality tests performance degrades with decreasing INR (Figure 11). In Figures 10 and 11 it is shown that Kurtosis and Skewness-based tests are far better than AD, CM and L tests; and that the SW test has a better performance than the rest of the ECDF tests even for a large sample size.

In Figures 12 and 13 ROC curves of the three different messages are presented. Figure 12 presents the normality test performance in the detection of the three different messages with the same INR and

the same sample size. Message 1 is relatively easy to detect by any normality test due to its statistical nature (plain text file with high redundancy) while messages 2 and 3 are undetectables for INR = −20dB due to the low redundancy of these messages (compressed data). Figure 13 presents the normality test performance in the detection of the messages 1, 2 and 3 with a higher INR = −5.2 dB, showing that the best normality tests to detect this low redundant telegraphic signal is again the Kurtosis. In fact, the relative performance of the normality tests is exactly the same as the PRN case.

Figure 10. Normality tests performance in the detection of a telegraphic interference signal. ROC curves of the different tests are presented as a function of the sample size, setting the INR to −16 dB. The best normality tests for detecting message 1 varies depending of the sample size, for 2,048 samples is the SW test, while for 1,024 samples are the JB and the K2 tests (highest P_{det} with a low P_{fa}), as both values of Kurtosis and Skewness are non-zero. Results obtained from a Monte-Carlo set of 2^{15} simulations.
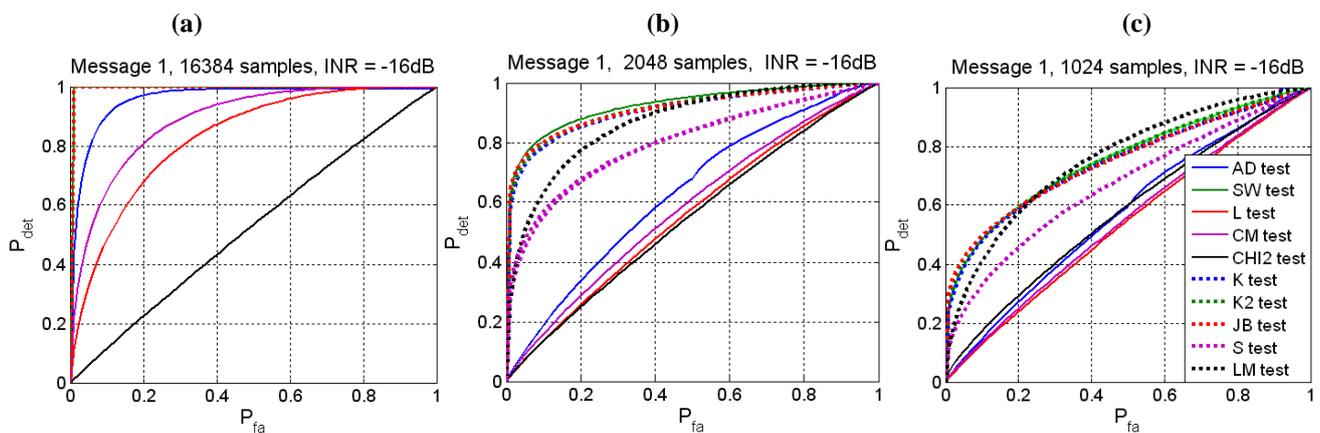

Figure 11. Normality test performance in the detection of a telegraphic interference signal (Message 1). ROC curves of the different tests are presented as a function of the INR, setting the sample size to 2,048. Like in Figure 10, the best normality tests for detecting message 1 is the SW test (highest P_{det} with a low P_{fa}). Results obtained from a Monte-Carlo set of 2^{15} simulations.
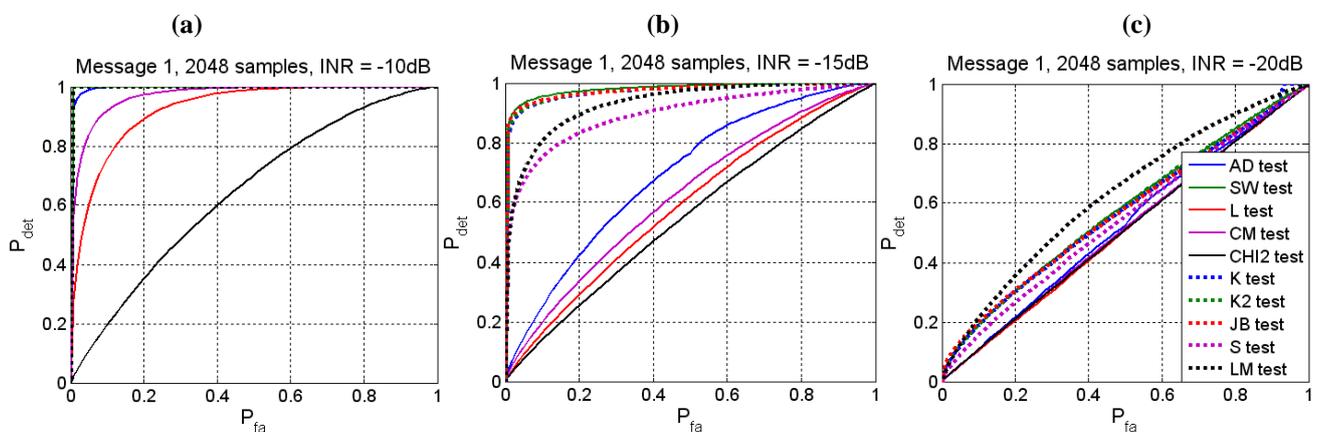
Figure 12. Normality test performance in the detection of a telegraphic interference signal. ROC curves of the different tests are presented in function of 3 different messages, setting the sample size to 16,384 and INR to −20 dB. The best normality tests for detecting message 1 is JB and K2 tests as both values of Kurtosis and Skewness are non-zero. In the cases of message 2 and 3, both compressed data (zip and jpg archives), −20 dB is a very low INR to detect these messages so, in Figure 13, INR is set to -5.2dB. Results obtained from a Monte-Carlo set of 2^{15} simulations.
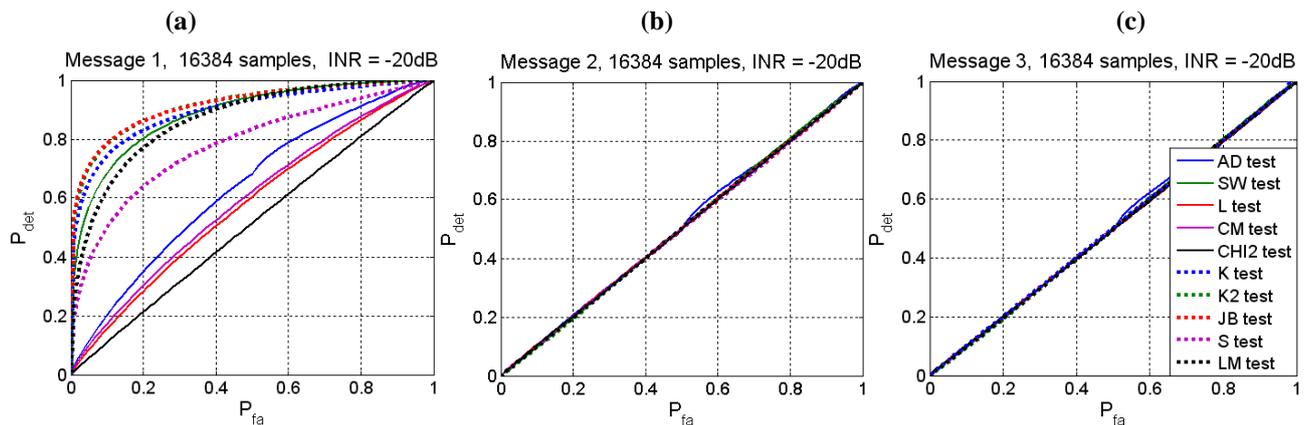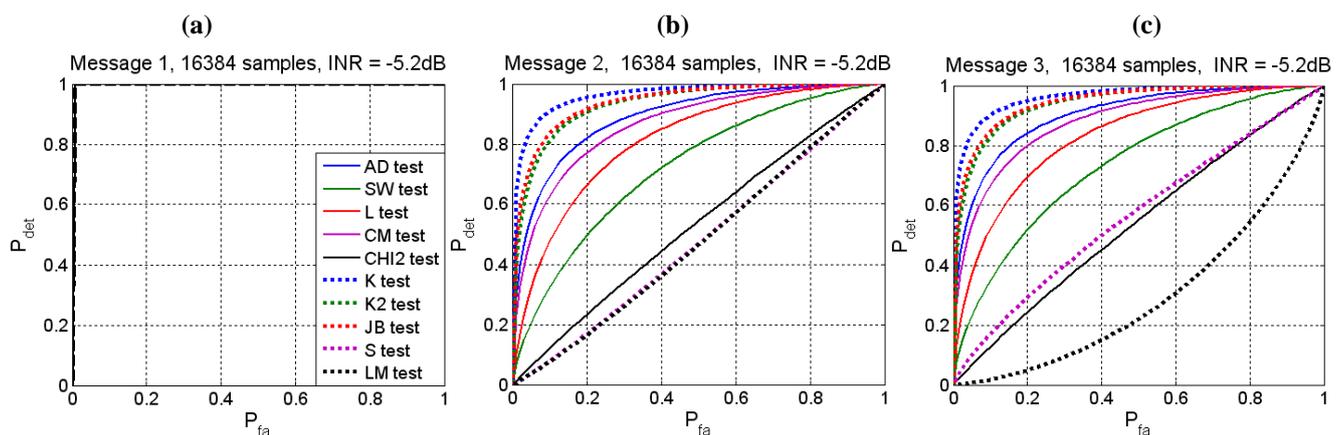

Figure 13. Normality test performance in the detection of a telegraphic interference signal. ROC curves of the different tests are presented in function of 3 different messages, setting the sample size to 16,384 and INR to −5.2 dB. Message 1, is easily detectable for all the normality tests due to its elevated Messages 2 and 3, both compressed data (a zip and a jpg archives), Kurtosis seems to be the best detection algorithm. Results obtained from a Monte-Carlo set of 2^{15} simulations.

To simulate message scrambling and encryption, all 3 messages are scrambled by means of an XOR operation between original message and the PRN code previously studied in this work. Results obtained are shown in Figure 14 which has to be compared with the result obtained in Figure 9. The high similarity between Figure 9 and the result of the test detection of the scrambled signal of all three messages and the PRN signal (Figure 14), shows that the scrambling process usually employed in communications makes the detection of RFI more difficult.

Figure 14. Normality test performance in the detection of a telegraphic interference signal. ROC curves of the different tests are presented in function of 3 different messages scrambled with PRN signal shown in Figure 9, setting the sample size to 16,384 and INR to −5.2 dB. It can be observed that all ROC curves are almost equal for each interfering signal, deducing that if the telegraphic signal is encrypted or scrambled, it can be treated as a spread spectrum signal. Results obtained from a Monte-Carlo set of 2^{15} simulations.

(a) (b) (c)
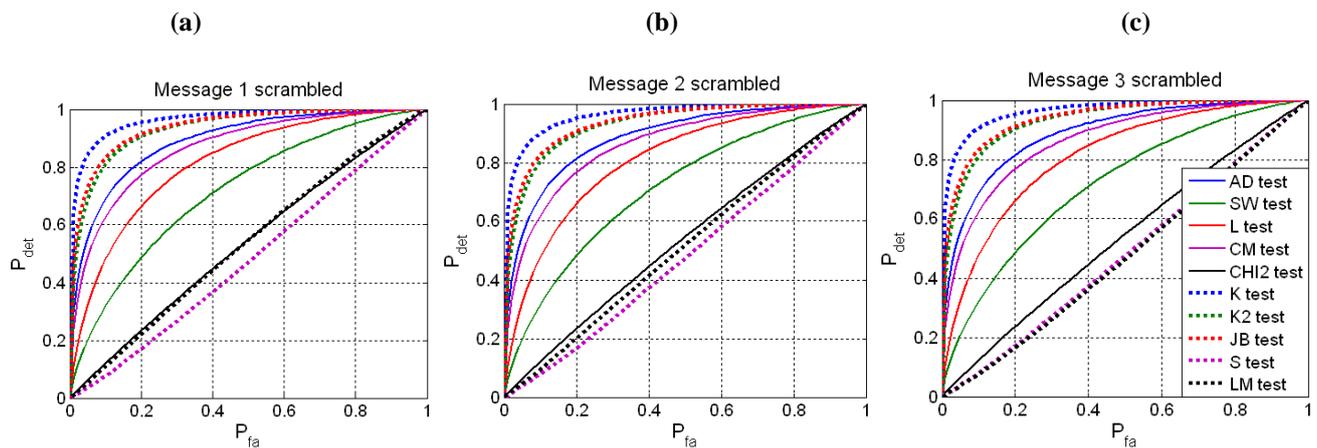
5. Conclusions

In this study the performance of ten different normality tests has been analyzed in terms of their capability to detect radio-frequency interference in microwave radiometry. These tests have been first validated in terms of sequence length and number of quantization bits in the absence of interference. Their capability to detect sinusoidal, chirp, PRN and telegraphic signals has then been analyzed.

It has been shown that the Kurtosis is the best RFI detection algorithm for almost all kinds of interfering signals, although it is known that it has a blind spot for sinusoidal and chirp interfering signals of 0.5 duty cycle.

Skewness-based algorithms (S and LM) usually have a lower performance than other tests as the Skewness of PRN, sinusoidal and chirp interfering signals is almost zero. However nonscrambled telegraphic signals present a higher Skewness parameter leading to a better performance of the S and LM tests than in case of sinusoidal, chirp, PRN and telegraphic scrambled signals.

Kurtosis-based normality tests (JB and K2 tests) have a good performance if both the Kurtosis and the Skewness are high enough. However, since Skewness is usually almost zero, both tests have a performance slightly worse than the Kurtosis test, except in the case of nonscrambled telegraphic interfering signals. Their performance is very similar, although the JB test has always a slightly better performance. In any case, as K, both tests present a blind spot detection for sinusoidal and chirp interfering signals of 0.5 duty cycle.

The four Empirical Distribution Function based normality tests: Anderson-Darling (AD), Lilliefors (L), Cramer-von Mises (CM), and Shapiro-Wilk (SW) tests have a similar performance for PRN, sinusoidal and chirp interfering signals. For a low sample size AD and SW tests work better that the CM and L test, but as sample size increases, SW test performance degrades in front of CM and L tests, as SW test must be averaged above 2,048 sample size to obtain a correct performance.

As compared to the other normality tests, CHI2 normality test has a poor performance, therefore is not recommended to use in RFI detection algorithms.

In summary, the Kurtosis is the best RFI detection algorithm for almost all kinds of interfering signals, although it has a blind spot for sinusoidal (chirp) signals of 0.5 of duty cycle. The AD test is a complementary normality test that covers this blind spot, and has a very good performance for all the studied sample sizes. The combination of the Kurtosis and the AD tests seems capable to detect most types of RFI. The performance of the detection tests improves with the sample size and depends on the duty cycle of the pulsed RFI.

Future research will be devoted to the optimum combination of these statistical analysis with time and frequency blanking methods, since these methods outperform statistical analysis in some specific cases for example low duty cycle pulsed sinusoidal signals (short pulses are easily detected in time domain), or high duty cycle pulsed sinusoidal signals or a continuous wave (a tone is easily detected in frequency domain).

Acknowledgements

References

1. Njoku, E.G.; Ashcroft, P.; Chan, T.K.; Li L. Global survey and statistics of radio-frequency interference in AMSR-E land observations. *IEEE Trans. Geosc. Rem. Sens.* **2005**, *43*, 938–947.
2. Ellingson, S.W.; Johnson, J.T. A polarimetric survey of radio-frequency interference in C- and X-Bands in the continental United States using Windsat radiometry. *IEEE Trans. Geosc. Rem. Sens.* **2006**, *44*, 540–548.
3. Younis, M.; Maurer, J.; Fortuny-Guasch, J.; Schneider, R.; Wiesbeck, W.; Gasiewski, A.J. Interference from 24-GHz automotive radars to passive microwave earth remote sensing satellites. *IEEE Trans. Geosc. Rem. Sens.* **2007**, *42*, 1387–1398.
4. Ellingson, S.W.; Hampson, G.A.; Johnson, J.T. Design of an L-band microwave radiometer with active mitigation of interference. *IEEE Intern. Geosc. Rem. Sens. Symp.* IGARSS **2003**, *3*, 1751–1753.
5. Guner, B.; Johnson, J.T.; Niamswaun, N. Time and frequency blanking for radio frequency interference mitigation in microwave radiometry. *IEEE Trans. Geosc. Rem. Sens.* **2007**, *45*, 3672–3679.
6. Johnson, J.T.; Gasiewski, A.J.; Guner, B.; Hampson, G.A.; Ellingson, S.W.; Krishnamachari, R.; Niamsuwan, N.; McIntyre, E.; Klein, M.; Leuski, V.Y. Airborne radio-frequency interference studies at C-band using a digital receiver. *IEEE Trans. Geosc. Rem. Sens.* **2006**, *44*, 1974–1985.
7. Ruf, C.S.; Gross, S.M.; Misra, S. RFI detection and mitigation for microwave radiometry with an agile digital detector. *IEEE Trans. Geosc. Rem. Sens.* **2006**, *44*, 694–706.

8. De Roo, R.D.; Misra, S.; Ruf, C.S. Sensitivity of the Kurtosis statistic as a detector of pulsed sinusoidal RFI. *IEEE Trans. Geosc. Rem. Sens.* **2007**, *45*, 1938–1946.

9. Misra, S.; Ruf, C.S. Detection of radio-frequency interference for the aquarius radiometer. *IEEE Trans. Geosc. Rem. Sens.* **2008**, *46*, 3123–3128.

10. De Roo, R.D.; Misra, S. Effectiveness of the sixth moment to eliminate a kurtosis blind spot in the detection of interference in a radiometer. *IEEE Intern. Geosc. Rem. Sens. Symp.* **2008**, *2*, 331–334.

11. Johnson, J.T.; Potter, L.C. Performance study of algorithms for detecting pulsed sinusoidal interference in microwave radiometry. *IEEE Trans. Geosc. Rem. Sens.* **2009**, *47*, 628–636.

12. Piepmeier, J.R.; Mohammed, P.N.; Knuble, J.J. A double detector for rfi mitigation in microwave radiometers. *IEEE Trans. Geosc. Rem. Sens. Sensing* **2008**, *46*, 458–465.

13. Guner, B.; Frankford, M.T. Johnson, J.T. On the Shapiro–Wilk test for the detection of pulsed sinusoidal radio frequency interference. *IEEE Intern. Geosc. Rem. Sens. Symp.* **2008**, *2*, II-157–II-160.

14. Guner, B.; Frankford, M.T. Johnson, J.T. A study of the Shapiro-Wilk test for the detection of pulsed sinusoidal radio frequency interference. *IEEE Trans. Geosc. Rem. Sens.* **2009**, *47*, 1745–1751.

15. Jarque, C.M.; Bera, A.K. A Test for normality of observations and regression residuals. *Intern. Statist. Rev.* **1987**, *55*, 163–172.

16. Urzua, C.M. On the correct use of omnibus tests for normality. *Econom. Let.* **1996**, *53*, 247–251.

17. Poitras, G. More on the correct use of omnibus tests for normality. *Econom. Let.* **2006**, *90*, 304–309.

18. D'Agostino, R.B.; Belanger, A.; D'Agostino, R.B., Jr. A suggestion for using powerful and informative tests of normality. *Amer. Statist.* **1990**, *44*, 316–321.

19. Lilliefors, H.W. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *J. Amer. Statist. Assoc.* **1967**, *64*, 387–389.

20. D'Agostino, R.B.; Stephens, M.A. *Goodness-of-Fit Techniques*, 1st ed.; Marcel Dekker: New York, NY, USA, 1986; pp. 122–133.

21. Trujillo-Ortiz, A.; Hernandez-Walls, R.; Barba-Rojo, K.; Castro-Perez, A. AnDartest: Anderson-Darling test for assessing normality of a sample data. *MATLAB file*, 2007.

22. Shapiro, S.S.; Wilk, M.B. An analysis of variance test for normality (complete samples). *Biomet.* **1965**, *52*, 591–611.

23. Royston, J.P. Approximating the Shapiro-Wilk W-test for non-normality. *Statist. Comput.* **1992**, *2*, 117–119.

24. Royston, J.P. An extension of Shapiro and Wilk's W test for normality to large samples. *Appl. Statist.* **1982**, *31*, 115–124.

25. Tong, H. *Non-Linear Time Series: a Dynamical System Approach*, 1st ed.; Oxford University Press: New York, NY, USA, 1990; pp. 324–325.

26. Thode, H.G., Jr. *Testing for Normality*, 1st ed.; Marcel Dekker: New York, NY, USA, 2002; pp. 88–89.

27. Lin, C.C.; Mudholkar, G.S. A simple test for normality against asymmetric alternatives. *Biomet.* **1980**, *67*, 455–461.