We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



2

The use of the GLRT for revealing faults in atomic frequency standards

Emilia Nunzi
University of Perugia
Italy

1. Introduction

Atomic clocks are instruments widely employed in many synchronization systems. When such measurement instruments are inserted in complex systems as telecommunication networks, global satellite navigation systems, tests of fundamental physics of matter, an unexpected and anomalous behavior or a degradation of the performance of the clock may give rise to an error condition in the global system (Galysh et al., 1996); (Vioarsson et al., 2000). In some cases, as in the GPS system, the anomaly must be identified and a real-time alarm signal must be transmitted to user. Aerospace systems and navigation support systems used in the area of personal security, for both military and civil purposes, must satisfy strict requirements of the parameters relative to the integrity, reliability, availability and accuracy of the signal. The possible lacking of information concerning one of the above features, may imply a series of inefficiencies and system bugs for end-users.

Monitoring stability of atomic clock frequency data is important for guaranteeing the correct behavior of the electronic system where they are inserted. The principal application where the frequency stability monitoring is a challenging problem is in GNSS systems (like GPS or Galileo) where the overall system performance critically depends on performance of on–board clocks. When the clock behaves bad, thus the anomaly has to be detected fast in order to provide an adequate action for restoring the correct behavior of the clock. In the field of navigation system "integrity", most of the studies are related to satellite integrity (Bruce et al., 2000) and not specifically to that of the embedded clock, while recently a study on GPS clock integrity showed GPS clock strange behaviors (Weiss et al., 2006), asking for suitable new statistical tools for its characterization.

The scientific and industrial community has done a lot of efforts for the theoretical characterization of the behavior of atomic clocks. The purpose is to improve the accuracy and reliability of these instruments while reducing their size and cost. Since such objectives are often in contradiction, the scientific and industrial research is investigating innovative techniques to overcome the limits imposed by the technological development.

A common assumption when analyzing atomic clock data is that clock noise is stationary or that at least increments in the frequency values are stationary and thus data are examined by using stability analysis tools such as the Allan variance (IEEEstd, 1999); (D. W. Allan, 1987). The scientific literature has shown how this hypothesis cannot be always verified in reality, in particular in the application context of the satellite navigation or in experiments of fundamental physics of matter. If this hypothesis is not satisfied, the accuracy and reliability

of the classical fault detection techniques, and therefore of the whole system where the devices are inserted, are compromised (Vioarsson et al., 2000); (Bruce et al., 2000).
In this context, this book chapter collects and summarizes the last years proposals about the use of Generalized Likelihood Ratio test (GLRT) as a fault detection technique, complementary to the classical ones, for revealing faults from frequency data (E. Nunzi et al., 2007); (E. Nunzi et al., 2007); (E. Nunzi et al., 2008); (E. Nunzi & P. Carbone, 2008); (E. Nunzi & D. D'Ippolito, 2009); (E. Nunzi et al., 2009).
In particular, the GLRT (S. M. Kay, 1998), following the Neyman-Pearson (NP) approach is presented and its effectiveness is demonstrated when clock frequency data are subjected to jumps in the mean and/or in the dynamic range. This method is largely employed for revealing faults in industrial manufacturing processes or for supporting the decision making problems in many different applications fields.
Although the large number of scientific publications allows a simple interpretation of the GLRT outcome, the application of the GLRT to frequency data acquired from atomic clocks still need to be properly customized and metrologically characterized.
It follows that reasons for proposing the GLRT as an alternative method for revealing faults in atomic frequency can be summarized as it follows:

1. GLRT does not require the stationarity hypothesis on processed data;
2. GLRT gives a reliable outcomes also when the acquired data record is affected by missing data;
3. GLRT is easy to implement (since it based on the evaluation of the Maximum Likelihood Estimates (MLEs) of the data model parameters) (S. M. Kay, 1996);
4. GLRT functionalities can be extended for revealing also, in the meantime or separately, anomalous behavior other than mean and variance changes.

In order to give a comprehensive presentation of the problem and of the theory needed for applying the GLRT technique to frequency signals acquired from an atomic clock, the next section recalls the atomic clock frequency sample model adopted for the analysis of the fault occurrence and suitable for the application of the fault detection theory as indicated in (S. M. Kay, 1998). The evaluated GLRT detector is applied to both simulated and experimental data subjected to anomalous behavior in order to validate the presented statistical models and theory.

2. Mathematical model of frequency data

In this chapter, a single frequency sample, $y[\cdot]$, is modeled as a white Gaussian random variable with unknown mean and standard deviation, indicated with μ_0 and σ_0, respectively. When a N–length record of frequency data is collected, all samples are assumed to be independent and identically distributed (i.i.d.). Moreover, an anomalous behavior of the clock is defined as a change in the frequency size and/or dynamic range. As a consequence, the anomaly, when and if it occurs, can be modeled as a change in the mean and/or in the standard deviation of the statistical model (D. W. Allan, 1987); (E. Nunzi et al., 2007); (E. Nunzi et al., 2007); (IEEEstd, 1999).
The simple model assumed for atomic clock frequency behavior, allows the description of the parameters model change by means of two different statistical hypotheses indicated with $\mathcal{H}_0$ and $\mathcal{H}_1$. $\mathcal{H}_0$ denotes the assumption that the clock behavior respects the given model;

$\mathcal{H}_1$ represents the assumption that the mean and/or standard deviation have changed to μ_1 and/or σ_1. These hypotheses are formally described by the following equations:

$$\mathcal{H}_0 \quad : \quad y[n] \sim \mathcal{N}(\mu_0, \sigma_0), \quad n = 0, ..., N-1; \tag{1}$$

$$\mathcal{H}_1 \quad : \quad y[n] \sim \begin{cases} \mathcal{N}(\mu_0, \sigma_0), & n = 0, ..., n_0 - 1; \\ \mathcal{N}(\mu_1, \sigma_1), & n = n_0, ..., N-1, \end{cases} \tag{2}$$

where n_0 is the unknown sample number at which the change happens when $\mathcal{H}_1$ is true, and N is the number of analyzed data.
Let us indicate with K_μ and σ_{0f} the additive frequency jump and the standard deviation factor, respectively. When $\mathcal{H}_1$ is true, i.e. it is true that processed data include some $(N - n_0)$ anomalous samples, thus the following identities are true:

$$\mu_1 = \mu_0 + K_\mu \tag{3}$$

$$\sigma_1 = \sigma_0 \cdot \sigma_{0f}. \tag{4}$$

It should be noticed that there are many unknown parameters: μ_0 and σ_0 (both under $\mathcal{H}_0$ and $\mathcal{H}_1$), μ_1, σ_1 and n_0 when (and if) a fault occurs. On the other hand, N is a parameter of the data acquisition process.
In order to improve text readability, vectors of the unknown parameters when $\mathcal{H}_0$ and $\mathcal{H}_1$ are true are introduced and they are, respectively, $\boldsymbol{\theta}_{H0}$=$[\mu_0, \sigma_0]$ and $\boldsymbol{\theta}_{H1}$=$[\mu_0, \sigma_0, \mu_1, \sigma_1, n_0]$.

3. Generalized likelihood ratio Test (GLRT): theory

When there are only two different models, the scope of the detection theory is the determination of the optimal criterion for identifying which of the models is the most likely to be underlying the given experimental data and when, in case, the clock model parameters have changed their values. The commonly employed decision–rule is based on the so called likelihood ratio test (LRT), which maximizes the detection probability (PD) for a given false alarm probability (PFA). It should be noticed that, in this context, PFA is the probability of deciding for $\mathcal{H}_1$ when $\mathcal{H}_0$ is true, and PD is the probability of deciding for $\mathcal{H}_1$ when $\mathcal{H}_1$ is true.
The LRT technique is based on the evaluation of the likelihood ratio, i.e. the ratio between the likelihood functions calculated when the hypothesis $\mathcal{H}_1$ is true and when the hypothesis $\mathcal{H}_0$ is true. On the basis of the Neyman-Pearson (NP) theorem, if this ratio is sufficiently large, the hypothesis $\mathcal{H}_0$ is rejected.
From a practical point of view, the NP approach is applied to the set of available data $\mathbf{y} = \{y[0], ..., y[N-1]\}$ and it is based on the evaluation of the likelihood ratio of $\mathbf{y}$, $L_G(\mathbf{y}; [\boldsymbol{\theta}_{H0}, \boldsymbol{\theta}_{H1}])$, defined as the ratio between the likelihood function of $\mathbf{y}$ under $\mathcal{H}_1$, $p(\mathbf{y}; \boldsymbol{\theta}_{H1}, \mathcal{H}_1)$, and the likelihood ratio of $\mathbf{y}$ under $\mathcal{H}_0$, $p(\mathbf{y}; \boldsymbol{\theta}_{H0}, \mathcal{H}_0)$. If $L_G(\mathbf{y}; [\boldsymbol{\theta}_{H0}, \boldsymbol{\theta}_{H1}])$ is larger than a given value γ, i.e.:

$$L_G(\mathbf{y}; [\boldsymbol{\theta}_{H0}, \boldsymbol{\theta}_{H1}]) = \frac{p(\mathbf{y}; \boldsymbol{\theta}_{H1}, \mathcal{H}_1)}{p(\mathbf{y}; \boldsymbol{\theta}_{H0}, \mathcal{H}_0)} > \gamma, \tag{5}$$

thus, the NP approach decides for $\mathcal{H}_1$, otherwise $\mathcal{H}_0$ is assumed to be true. The value of γ is chosen on the basis of the target PFA value, α, that should be guaranteed a–priori and it can be evaluated by solving the following equation:

$$PFA = Pr\{L_G(\mathbf{y}; [\boldsymbol{\theta}_{H0}, \boldsymbol{\theta}_{H1}]) > \gamma; \mathcal{H}_0\} < \alpha. \tag{6}$$

It should be noticed that for evaluating the γ value that satisfies (6), the statistic of $L_G\left(\mathbf{y};[\boldsymbol{\theta}_{H0},\boldsymbol{\theta}_{H1}]\right)$ is needed.
$L_G\left(\mathbf{y};[\boldsymbol{\theta}_{H0},\boldsymbol{\theta}_{H1}]\right)$ is calculated by following the statistical model introduced in section 2, i.e. by considering that each measured value, $y[n]$, is the realization of an independent Gaussian random variable. It follows that the likelihood functions of the data vector $\mathbf{y}$ under $\mathcal{H}_0$ and $\mathcal{H}_1$ can be evaluated by multiplying the marginal likelihood function of each data sample, thus obtaining:

$$p\left(\mathbf{y};\boldsymbol{\theta}_{H0}\right)=\frac{1}{\left(2\pi\sigma_0^2\right)^{N/2}}\exp\left(-\frac{1}{2\sigma_0^2}\sum_{n=0}^{N-1}\left(y[n]-\mu_0\right)^2\right) \tag{7}$$

$$\begin{aligned}p\left(\mathbf{y};\boldsymbol{\theta}_{H1}\right)=&\frac{1}{\left(2\pi\sigma_0^2\right)^{n_0/2}\left(2\pi\sigma_1^2\right)^{(N-n_0)/2}}\\&\exp\left(-\frac{1}{2\sigma_0^2}\sum_{n=0}^{n_0-1}\left(y[n]-\mu_0\right)^2-\frac{1}{2\sigma_1^2}\sum_{n=n_0}^{N-1}\left(y[n]-\mu_1\right)^2\right)\end{aligned} \tag{8}$$

Since all values of the parameters vectors $\boldsymbol{\theta}_{H0}$ and $\boldsymbol{\theta}_{H1}$ are unknown and supposed to be deterministic, a GLRT technique, instead of a LRT, has been applied (S. M. Kay, 1998). Thus, parameter values in (5) have been replaced by their maximum likelihood estimates (MLEs), that will be indicated with a hat–sign on the symbol of the corresponding variable name. By indicating with $\widehat{\boldsymbol{\theta}}_{H0}=[\widehat{\mu}_{0_H0},\widehat{\sigma}_{0_H0}]$, the MLE of $\boldsymbol{\theta}_{H0}$ when $\mathcal{H}_0$ is true, with $\widehat{\boldsymbol{\theta}}_{H1}=[\widehat{\mu}_{0_H1},\widehat{\sigma}_{0_H1},\widehat{\mu}_{1_H1},\widehat{\sigma}_{1_H1},\widehat{n}_0]$ the MLE of $\boldsymbol{\theta}_{H1}$ when $\mathcal{H}_1$ is true, it follows that the GLRT test decides for $\mathcal{H}_1$ if $L_G\left(\mathbf{y};[\widehat{\boldsymbol{\theta}}_{H0},\widehat{\boldsymbol{\theta}}_{H1}]\right)>\gamma$, i.e.:

$$L_G\left(\mathbf{y};[\widehat{\boldsymbol{\theta}}_{H0},\widehat{\boldsymbol{\theta}}_{H1}]\right) \quad = \quad \frac{p\left(\mathbf{y};\widehat{\boldsymbol{\theta}}_{H0},\mathcal{H}_1\right)}{p\left(\mathbf{y};\widehat{\boldsymbol{\theta}}_{H1},\mathcal{H}_0\right)}>\gamma \tag{9}$$

where:

$$p\left(\mathbf{y};\widehat{\boldsymbol{\theta}}_{H0}\right) \quad = \quad \frac{1}{\left(2\pi\widehat{\sigma}_{0_H0}^2\right)^{N/2}}\exp\left(-\frac{1}{\widehat{\sigma}_{0_H0}^2}\sum_{n=0}^{N-1}\left(y[n]-\widehat{\mu}_{0_H0}\right)^2\right) \tag{10}$$

$$\begin{aligned}p\left(\mathbf{y};\widehat{\boldsymbol{\theta}}_{H1}\right) \quad = \quad &\frac{1}{\left(2\pi\widehat{\sigma}_{0_H1}^2\right)^{\widehat{n}_0/2}\left(2\pi\widehat{\sigma}_{1_H1}^2\right)^{(N-\widehat{n}_0)/2}}\cdot\\&\cdot\exp\left(-\frac{1}{2\widehat{\sigma}_{0_H1}^2}\sum_{n=0}^{\widehat{n}_0-1}\left(y[n]-\widehat{\mu}_{0_H1}\right)^2-\frac{1}{2\widehat{\sigma}_{1_H1}^2}\sum_{n=\widehat{n}_0}^{N-1}\left(y[n]-\widehat{\mu}_{1_H1}\right)^2\right).\end{aligned} \tag{11}$$

It can be shown that MLEs of the mean and of the standard deviation are, respectively, the sample mean and the sample standard deviation and that the MLE of n_0 is the sample number that maximizes the NP detector over the whole available data record (S. M. Kay, 1996). Thus, MLE estimates of the unknown parameters are:

$$\widehat{\mu}_{0_H0} = \frac{1}{N}\sum_{n=0}^{N-1} y[n] \tag{12}$$

$$\widehat{\mu}_{0_H1} = \frac{1}{\widehat{n}_0}\sum_{n=0}^{\widehat{n}_0-1} y[n] \tag{13}$$

$$\widehat{\mu}_{1_H1} = \frac{1}{N-\widehat{n}_0}\sum_{n=\widehat{n}_0}^{N-1} y[n] \tag{14}$$

$$\widehat{\sigma}_{0_H0} = \sqrt{\frac{1}{N}\sum_{n=0}^{N-1}\left(y[n]-\widehat{\mu}_{0_H0}\right)^2} \tag{15}$$

$$\widehat{\sigma}_{0_H1} = \sqrt{\frac{1}{\widehat{n}_0}\sum_{n=0}^{\widehat{n}_0-1}\left(y[n]-\widehat{\mu}_{0_H1}\right)^2} \tag{16}$$

$$\widehat{\sigma}_{1_H1} = \sqrt{\frac{1}{N-\widehat{n}_0}\sum_{n=\widehat{n}_0}^{N-1}\left(y[n]-\widehat{\mu}_{1_H1}\right)^2} \tag{17}$$

$$\widehat{n}_0 = \text{argmax}_{n_0=0,\ldots,N-1}\{L_G\left(\mathbf{y};[\widehat{\boldsymbol{\theta}}_{H0},\widehat{\boldsymbol{\theta}}_{H1}]\right)\}. \tag{18}$$

By analyzing and comparing firstly expressions (15) with (10), and thus equations (16)–(17) with (11), it follows that arguments of the exponential terms in (10) and (11) are both equal to $(-N/2)$. It follows that the condition to be tested by the GLRT can be simplified as:

$$\begin{aligned} L_G\left(\mathbf{y};[\widehat{\boldsymbol{\theta}}_{H0},\widehat{\boldsymbol{\theta}}_{H1}]\right) &= \frac{p\left(\mathbf{y};\widehat{\boldsymbol{\theta}}_{H0},\mathcal{H}_1\right)}{p\left(\mathbf{y};\widehat{\boldsymbol{\theta}}_{H1},\mathcal{H}_0\right)} \\ &= \frac{\left(\widehat{\sigma}^2_{0_H1}\right)^{\widehat{n}_0/2}\left(\widehat{\sigma}^2_{1_H1}\right)^{(N-\widehat{n}_0)/2}}{\left(\widehat{\sigma}^2_{0_H0}\right)^{N/2}} > \gamma \end{aligned} \tag{19}$$

By taking into account that the logarithm is a monotonically increasing function, thus inequality (19) does not change if the logarithm is taken on both inequality sides. Thus the NP approach can be further simplified by solving the following equivalent inequality:

$$\begin{aligned} T\left(\mathbf{y};[\widehat{\boldsymbol{\theta}}_{H0},\widehat{\boldsymbol{\theta}}_{H1}]\right) &= \log\left(L_G\left(\mathbf{y};[\widehat{\boldsymbol{\theta}}_{H0},\widehat{\boldsymbol{\theta}}_{H1}]\right)\right) = \\ &= \frac{N}{2}\log\left(\frac{\widehat{\sigma}^2_{0_H0}}{\widehat{\sigma}^2_{1_H1}}\right) - \frac{\widehat{n}_0}{2}\log\left(\frac{\widehat{\sigma}^2_{0_H1}}{\widehat{\sigma}^2_{1_H1}}\right) > \gamma' = \log\gamma \end{aligned} \tag{20}$$

This expression shows that the detector is function of the three different MLE variance estimators. In order to characterize this NP detector, and to customize the parameters of the

detection algorithm for revealing specifically faults of atomic clock frequency data, the statistical characterization of $T\left(\mathbf{y};[\widehat{\boldsymbol{\theta}}_{H0},\widehat{\boldsymbol{\theta}}_{H1}]\right)$ is needed. However, the three variance estimators are not mutually independent. In particular, $\widehat{\sigma}^2_{0_H0}$ is a function of $\widehat{\sigma}^2_{0_H1}$ and $\widehat{\sigma}^2_{1_H1}$, both under $\mathcal{H}_0$ and $\mathcal{H}_1$. In fact, by requiring the energy of the sequence $\{y[0],y[1],...,y[N-1]\}$ to be equal to the sum of the energies of the two complementary subsequences $\{y[0],y[1],...,y[\widehat{n}_0-1]\}$ and $\{y[\widehat{n}_0],y[\widehat{n}_0+1],...,y[N-1]\}$, the following relationship is always true:

$$\widehat{\sigma}^2_{0H0} = \frac{\widehat{n}_0}{N}\widehat{\sigma}^2_{0H1}+\frac{N-\widehat{n}_0}{N}\widehat{\sigma}^2_{1H1}+\frac{\widehat{n}_0}{N}\frac{N-\widehat{n}_0}{N}\left(\widehat{\mu}_{0H1}-\widehat{\mu}_{1H1}\right)^2, \tag{21}$$

both under $\mathcal{H}_0$ and $\mathcal{H}_1$.
This mutual relationship between variance estimators under $\mathcal{H}_0$ and $\mathcal{H}_1$ makes difficult task the evaluation of the probability density function (and thus of the likelihood function) of (20). It follows that the evaluation of the γ threshold from (6) is a difficult task.
In order to derive a criterion for evaluating the γ value, a further theoretical analysis is performed.
By substituting (21) in (20), the detector expression is formally obtained as function of MLE estimators of parameters under $\mathcal{H}_1$:

$$\begin{aligned} T\left(\mathbf{y};[\widehat{\boldsymbol{\theta}}_{H0},\widehat{\boldsymbol{\theta}}_{H1}]\right) &= \frac{N}{2}\log\left(\frac{\widehat{n}_0}{N}\frac{\widehat{\sigma}^2_{0H1}}{\widehat{\sigma}^2_{1H1}}+\frac{\widehat{n}_0}{N}\frac{N-\widehat{n}_0}{N}\frac{\left(\widehat{\mu}_{0H1}-\widehat{\mu}_{1H1}\right)^2}{\widehat{\sigma}^2_{1H1}}+\frac{N-\widehat{n}_0}{N}\right)- \\ &- \frac{\widehat{n}_0}{2}\log\frac{\widehat{\sigma}^2_{0H1}}{\widehat{\sigma}^2_{1H1}}. \end{aligned} \tag{22}$$

Equation (22) shows that the GLRT detector $T(\mathbf{y})$ depends only on the MLE estimates under $\mathcal{H}_1$ by means of a function $g(\cdot)$, i.e. $T(\mathbf{y})=g(\widehat{\theta}_{H1})$. By recalling the invariance property of the MLE (S. M. Kay, 1996), it follows that also $T(\mathbf{y})$ is an MLE estimate of a theoretical value, T_{teor} which can be evaluated by substituting the MLE estimates of parameters with the corresponding theoretical values. Thus, in order to evaluate expression of T_{teor}, MLE estimates in (22), $\widehat{\sigma}^2_{0H1}$, $\widehat{\sigma}^2_{1H1}$, $\widehat{\mu}_{0H1}$, $\widehat{\mu}_{1H1}$ and $\widehat{n}_0$, have been replaced by the corresponding theoretical value, i.e. by σ^2_0, σ^2_1, μ_0, μ_1 and n_0, respectively. Moreover, by exploiting relationships (3) and (4), it can be shown that the theoretical behavior of the GLRT detector is described by the following equation:

$$T_{teor} = \frac{N}{2}\log A+\left(\frac{N}{2}-\frac{n_0}{2}\right)\log\frac{1}{\sigma^2_{0f}} \tag{23}$$

where

$$A = \frac{K^2_\mu}{\sigma^2_0}\frac{(N-n_0)(n_0-1)}{(N-1)^2}+\left(\frac{n_0-1}{N-1}+\frac{N-n_0}{N-1}\sigma^2_{0f}\right). \tag{24}$$

This equation can be used in practical cases for evaluating a-priori the γ threshold in the comparison process, from the knowledge of: the length N of the data sequence, the initial standard deviation σ_0 value, the frequency jump, K_μ, the frequency variance change factor, σ^2_{0f}, and the number of anomalous samples, $(N-n_0)$, before the detection event.

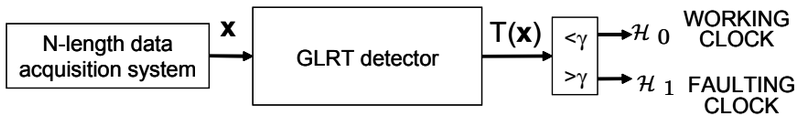


Fig. 1. Simplified block diagram of a GLRT scheme. $\boldsymbol{y}$ is the N–length available data sequence. $T(\boldsymbol{y})$ is the GLRT detector scalar output. γ is the threshold value used for assuming true $\mathcal{H}_0$ or $\mathcal{H}_1$.

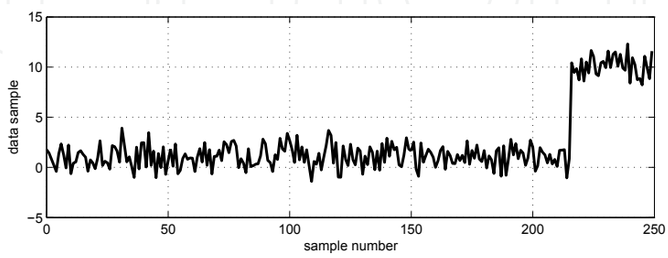


Fig. 2. Behavior of a normalized frequency data record, normally distributed, simulated by considering the following parameters: record length $M = 250$, $\mu_0 = 1$, $\sigma_0 = 1$, $K_\mu = 9$ (thus $\mu_1 = 10$), $\sigma_{0f} = 1$ (thus $\sigma_1 = \sigma_0$), $n_0 = 216$.

All these parameters values affect the GLRT performance and influence, in particular, the choice of the threshold value, γ, that is strictly dependent on the target PFA and that should be defined a–priori before the application of the GLRT. From a practical point of view, the use of the GRT is summarized in Fig.1: the available data sequence is processed by the GRT detector (20) and the corresponding scalar output is compared to a give $\gamma^{'}$ value in order to assume $\mathcal{H}_0$ or $\mathcal{H}_1$ true.

4. Simulation results

In this section simulation results are reported in order to validate theoretical results obtained in sec. 3. At first, theoretical formula (23) will be validated by means of Monte Carlo simulations and an application example is introduced in order to clarify its practical applicability. Moreover, the statistical characterization of the GLRT is presented by analyzing the behavior of PD versus PFA (i.e. the Receiver Operating Characteristics (ROCs)) . ROCs are evaluated by means of Monte Carlo simulations since the theoretical behavior of the detector is a difficult task, as already stated in sec.3.

4.1 Validation of (23)

In order to validate theoretical formula (23), Monte Carlo simulations on NREC=10000 records of simulated data, each of length M=250, have been performed. Each data record has been generated by considering a M–length sequence of Gaussian distributed data with initial mean and standard deviation values equal to $\mu_0 = 1$ and $\sigma_0 = 1$, respectively, and by setting a frequency jump $K_\mu = 9$ on $n_0 = 216$. For clarification purposes, the behavior of one data record versus the data sample index is shown in Fig.2.
For each data record, the GLRT detector (20) has been applied consecutively NREC times to N=200 data samples by following a First In First Out (FIFO) data organization. The behavior

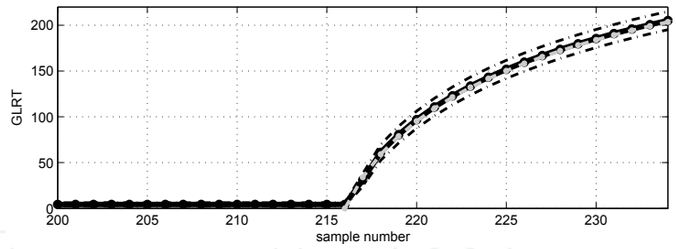


Fig. 3. Behavior of the averaged GLRT detector (solid black line) versus the index of the last acquired data sample, obtained by means of Monte–Carlo simulations. The average has been evaluated on $NREC = 10000$ GLRTs, each obtained by processing a simulated frequency data record with characterizing parameters equal to those of record shown in Fig.2 of data. The two dashed black lines represent the averaged GLRT plus and minus the corresponding sample standard deviation. Dashed gray line shows the behavior of T_{teor} given by (23) by using the same set of parameters values.

of the detector output, averaged over the number of record NREC, versus the index of the last acquired data sample, is shown in Fig.3 with a solid black line. Moreover, dashed black lines represent the averaged GLRT plus and minus the corresponding standard deviation evaluated on the same NREC data records. This figure shows that the GLRT detector is approximately equal to 0 if data are not affected by anomalies, i.e. until the last processed data sample is smaller than 216. When processed data include anomalous samples, thus the GLRT value increases.

For comparison purposes, the theoretical behavior of the detector described by eq.(23) versus the index of the last acquired data sample is also shown in Fig.3 with a dashed gray line for the same set of parameters values, i.e. $K_\mu = 9$, $\sigma_0 = 1$, $\sigma_{0f} = 1$, $n_0 = 216$. The theoretical behavior is close to the averaged GLRT output.

In order to give a quantitative characterization of the theoretical behavior, the relative displacement between the averaged GLRT evaluated by means of Monte Carlo simulations and the theoretical detector (i.e. between the solid–black and the dashed–gray lines in Fig.3) is shown in Fig.4(a) and the corresponding relative standard deviation is shown in Fig.4(b).

This figure shows that the error between the theoretical and simulated GLRT reduces as the number of samples with anomaly, and processed by the detector, increases. In particular, if $(N - n_0) > 2$ (i.e. sample index > 218), thus the error is smaller than 2% with a corresponding type A uncertainty, estimated by dividing the relative standard deviation by $\sqrt{NREC} = 100$, smaller than 0.1% (GUM, 1997). This result validates the expression of T_{teor}.

It follows that the theoretical expression (23) can be used for designing an accurate test procedure for revealing faults in data affected by frequency jump and/or by standard deviation variation. In particular, (23) gives the γ value to be used in the comparison process for a given set of: N, target readiness expressed in terms of $(N - n_0)$, target $K\mu$ and σ_{0f} .

4.2 An application example of (23)

In order to clarify how to use theoretical expression (23), let us consider an analysis problem on the simulated data sequence shown in Fig.2. In this case, the GLRT outcome applied to $N = 200$ data managed by using a FIFO strategy is shown in Fig.5. If the detector readiness is considered, in order to reveal a fault by using no more than $(N - n_0) = 4$ samples, by substituting in (23) values $N = 200$, $\sigma_0 = 1$, $n_0 = 196$, $\sigma_{0f} = 1$, $K_\mu = 9$, thus $T_{teor} = 95.37$. By

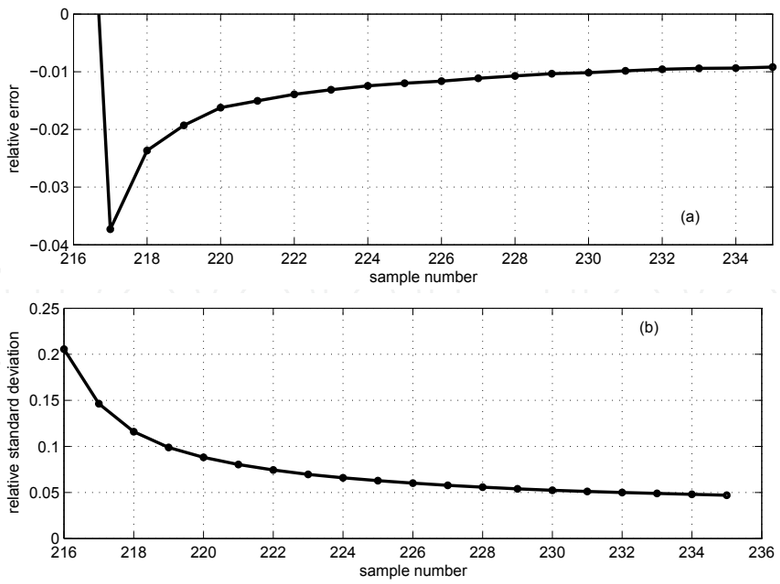


Fig. 4. Relative deviation (a) between the averaged GLRT obtained by means of Monte–Carlo simulations and T_{teor} shown in Fig.3 and corresponding relative sample standard deviation (b).

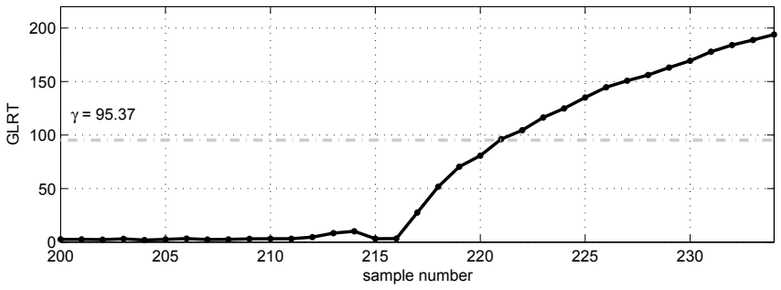


Fig. 5. Behavior of the GLRT detector applied to the data sequence shown in Fig.2 versus the index of the last acquired data sample.

setting $\gamma = T_{teor} = 95.37$, we can see in Fig.5 that the GLRT value is greater than γ for $n = 220$ and a warning signal should be emitted. This means that an additive frequency jump at least equal to 9 has occurred no more than 4 samples before the alarm signal emitted on the sample number 220.

Theoretical expression T_{teor} can be used also for analyzing the detector accuracy versus the number of employed anomalous samples $N - n_0$. In particular, Fig.6 shows the theoretical behavior of the GLRT with the $N - n_0$ when a data sequence of length $N = 200$ affected by a frequency jump equal to $K_m = 9$ is analyzed. This behavior clearly shows that the frequency jump is detectable also by using just one anomalous sample. The large value of the GLRT detector ($T_{teor} = 34$) when $N - n_0 = 1$ is mainly due to the large jump size considered in this application example and to the large $N = 200$ employed. It should be noticed that a large N value is due to the hypothesis that initially the clock is working properly and that a fault

occurs after that the monitoring process starts. A large N value increases the accuracy of the MLE estimates of μ_0 and σ_0 since they are asymptotic Gaussian with N (S. M. Kay, 1996).

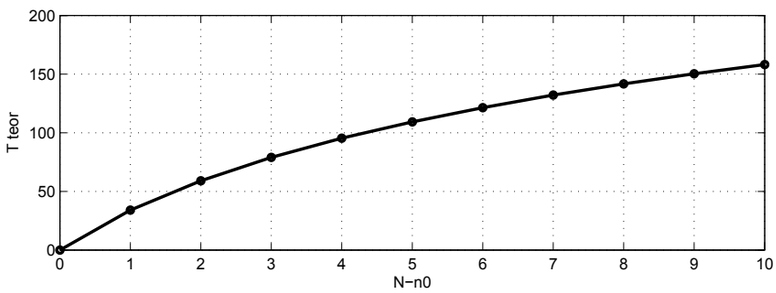


Fig. 6. Behavior of T_{teor} versus $N - n_0$ given by (23) obtained by setting characterizing parameters equal to those of record shown in Fig.2 of data.

4.3 Statistical characterization of the NP detector

In order to illustrate the properties of the NP detector when it is applied to data with parameters values typically employed in frequency standards and, in particular, to relate reasonable values for the threshold parameter γ to the corresponding PFA value, the test has been characterized by calculating the receiver operating characteristic (ROC), i.e. the behavior of PD versus PFA, for many γ and n_0 values, by means of Monte Carlo simulations (S. M. Kay, 1996); (S. M. Kay, 1998).

In particular, two sets of ROCs are presented here: the first one for analyzing the GLRT sensitivity (i.e.the detection capability by using no more than a given number of faulty samples), the other one for evaluating the GLRT readiness (i.e. the number of employed faulty samples for a given fault).

GLRT SENSITIVITY ANALYSIS

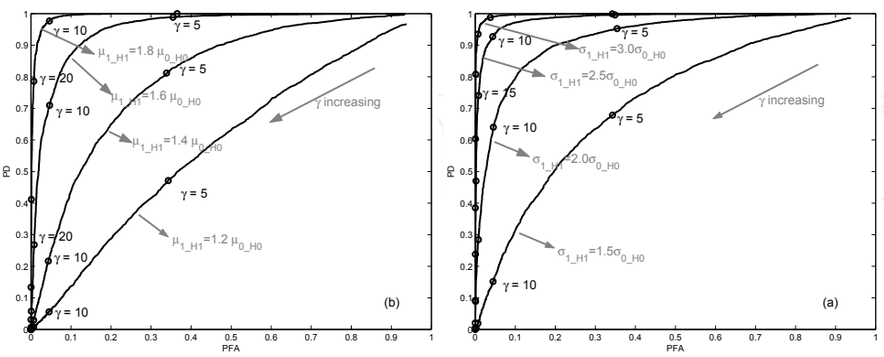


Fig. 7. ROCs obtained by considering 100 data samples and by setting $(N - n0) = 15$. (a) $\mu_1 = \mu_0$ and $\mu_1 = \{1.5\sigma_0, 2\sigma_0, 2.5\sigma_0, 3\sigma_0\}$. (b) $\mu_1 = \{1.2\mu_0, 1.4\mu_{0x}, 1.6\mu_0, 1.8\mu_0\}$ and $\sigma_1 = \sigma_0$.

Fig. 7(a) and (b) shows the ROCs when $\mu_1 = \mu_0$ and $\sigma_1 = \{1.5\sigma_0, 2\sigma_0, 2.5\sigma_0, 3\sigma_0\}$, and $\mu_1 = \{1.2\mu_0, 1.4\mu_0, 1.6\mu_0, 1.8\mu_0\}$ and $\sigma_1 = \sigma_0$, respectively, with $\mu_0 = 2.36 \cdot 10^{-11}$,

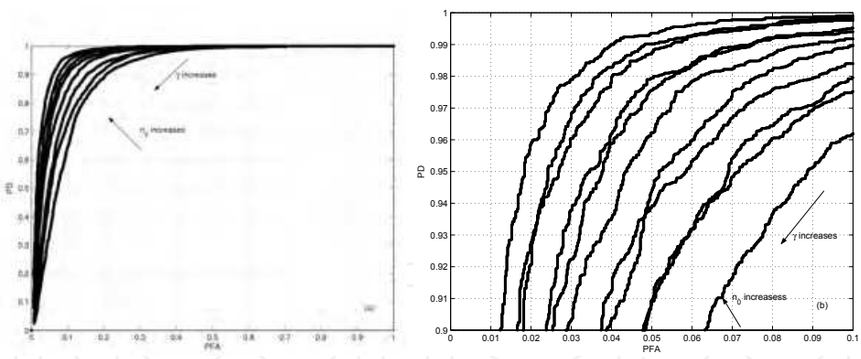


Fig. 8. Each line represents a ROC corresponding to a particular value of n_0 obtained by using a Monte Carlo approach based on 5000 pairs of data record each of length N=100 when: σ_1/σ_0=3, γ varies from 0 to 10. (a): n_0 ranges from 13 to 20. (b): n_0 ranges from 20 to 30.

$\sigma_0 = 1.046 \cdot 10^{-11}$, $N = 100$, and by setting $(N - n_0) = 15$. Circles indicate the γ value used for evaluating the corresponding PFA and PD. Fig. 7 can be employed to design a detection test that can track changes in the mean and in the standard deviation. As an example, one can see that a PD $= 93\%$ is obtained when the change in the mean is at least equal to a factor 1.8, and that the condition $T\left(\mathbf{y}; [\widehat{\boldsymbol{\theta}}_{H0}, \widehat{\boldsymbol{\theta}}_{H1}]\right) > \gamma = 10$ guarantees a PFA $< 8\%$. The same threshold value $\gamma = 10$ can be employed to detect a change in the standard deviation by a factor at least equal to 3.0, with PD $> 97\%$ and PFA $< 8\%$.

GLRT READINESS ANALYSIS

To this purpose, for each value of n_0 varying from 13 to 20, 5000 pairs of N–length data records, normally distributed, have been synthesized. For each record pair, the first sequence has been created by employing mean and standard deviation values respectively equal to $\overline{\mu}_0 = 2.3650 \cdot 10^{-11}$ and $\overline{\sigma}_0 = 1.0462 \cdot 10^{-11}$, that is by assuming $\mathcal{H}_0$ true. The second data set presents the first n_0 samples equal to those of the first record, while the last $N - n_0$ samples have been generated as normally distributed with mean $\overline{\mu}_0$ and $\sigma_1 = 3\sigma_0$. Thus, the second data record meets with $\mathcal{H}_1$. Both data records have been employed for calculating the corresponding value of $T\left(\mathbf{y}; [\widehat{\boldsymbol{\theta}}_{H0}, \widehat{\boldsymbol{\theta}}_{H1}]\right)$ when $\mathcal{H}_0$ is true, i.e. by applying (20) to the first record, and when $\mathcal{H}_1$ is true, i.e. by applying (20) to the second record.
Simulation results obtained by assuming N=100 are shown in Figs.8(a) and (b). In particular, each line in Fig.8(a) is the ROC corresponding to a particular n_0 value, ranging from 13 to 20, with γ varying from 0 to 15, as indicated by the figure label, when $\mu_0 = \mu_1 = \overline{\mu}_0$ and $\sigma_0/\sigma_1 = 3$ with $\sigma_0 = \overline{\sigma}_0$. Fig.8(b) shows the same ROC reported in Fig.8(a) but detailed for PFA$<$10%, PD$>$90% and n_0 ranging from 20 to 30. This figure shows that by processing a set of data of length $N = 100$, the fault detection occurs with PFA$<$5% and PD$>$95% only for n_0 larger than 25. Equivalently, under the same assumptions, it is possible to detect a fault within 25 samples from its occurrence. Moreover, for each n_0 value, this figure gives information on the range of the threshold values γ which can be used for revealing a fault.

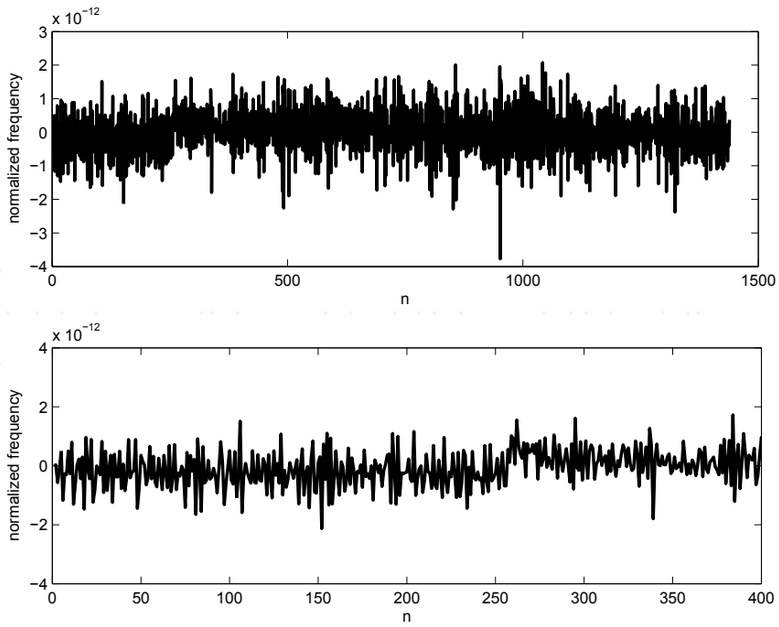


Fig. 9. (a): Behavior of normalized frequency data of atomic clock on satellite GPS 22 (downloaded by IGS). (b) Zoom of figure (a) on the first 400 experimental data samples.

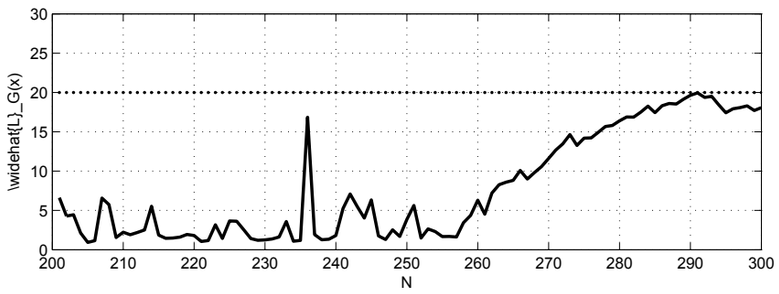


Fig. 10. Behavior of the GLRT detector applied to data shown in Fig. 9 (a) with $N = 200$.

5. Experimental results

Figure 9(a) shows the behavior of experimental normalized frequency data of GPS satellite number 22. Data have been obtained by the International GNSS Service (IGS) – Formerly the International GPS Service (J.M. Dow et al., 2005). The GLRT has been applied to these data by setting a processing window length $N = 200$. Thus, data samples have been processed by following a FIFO strategy and the processing result is shown in Fig.10. In particular, each time a new data sample is processed, the detector outcome is updated. The GLRT has a a spike on sample numbered as 236 while presents an increasing behavior after sample index 260. By requiring a small threshold value, i.e. $\gamma = 10$, warning signals are emitted on sample indexes 237 and 266. The first one is probably a false alarm, while the second is a true false alarm. If the γ value is increased (i.e. $\gamma = 18$), thus the false alarm probability is obviously reduced but the detection process requires a larger number of faulty samples in order to detect the fault

occurrence (Fig.10 shows that at least 22 faulty samples are needed since $\gamma = 18$ implies a fault detection on sample index 282).
A-posteriori analysis of data parameters confirm the validity of the GLRT technique. In fact, by zooming the behavior of the experimental frequency data on the first 400 samples (see Fig.9(b)) it is clear (and confirmed) that the first warning is a false alarm, while the second warning signal a fault detection signal and that the fault has occurred on sample numbered as 260 (as argued by a visual inspection of the GLRT behavior). In particular, the fault can be classified as a mean change. In fact, the arithmetic mean on the first 260 sample is equal to $\mu_0 = -2.21 \cdot 10^{-13}$ and on samples from 261 to 461 is equal to $\mu_0 = 1.96 \cdot 10^{-13}$. The corresponding standard deviation values are almost the same and equal to $\sigma_0 = \sigma_1 = 6 \cdot 10^{-13}$. Thus, we can asume $K_\mu = 4 \cdot 10^{-13}$ and $\sigma_{0f} = 1$.
By substituting these parameters values in the theoretical detector (23) and by requiring the fault detection within no more than 6 faulty samples, i.e. by using $N - n_0 = 6$, thus (23) gives $T_{teor} = 9.14$, which is a threshold value congruent with the analysis of the GLRT outcome shown in Fig.10.

6. Conclusion and future works

The GLRT detection algorithm for revealing faults from atomic clock frequency data has been analyzed. It has been shown that the GLRT algorithm efficiently detects anomalies when applied to atomic clock frequency data (E. Nunzi & P. Carbone, 2008)– (E. Nunzi et al., 2009). However, the mathematical expression of the detector depends on many parameters characterizing acquired data and previously analysis of the parameters dependencies have been based on ROCs evaluated by means of Monte–Carlo simulations. It follows that a proper customization of the GLRT technique to the analysis of data coming from atomic clocks is necessary and useful for designing the test procedure.
The information given by the GLRT method can be employed to detect the non stationary change in the clock behavior.
Future works are focused on the theoretical statistical characterization of the GLRT detector and on the designing of the test procedure for effectively reveals anomaly in real–time (i.e. while the data acquisition process is going on). Moreover, the MLE estimates used for evaluating the GLRT outcome could be employed for deriving information also on the identification of the occurred anomaly, although it is known MLE estimators are not optimal when a small number of faulty samples are employed (as required for this specific application).

Acknowledgment

This work has been supported by the Italian Ministry PRIN 2007 program (ARCADIA project: http://arcadia.diei.unipg.it).

7. References

D. W. Allan (1987) , "Time and frequency (time-domain) characterization, estimation and prediction of precision clocks and oscillators," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*,vol. UFFC-37, no. 6, pp. 647–654, Nov. 1987.

A. S. Bruce & A. J. Van Dierendonck & A. Jakab, J. Wiebe & B. Townsend (2000), "Detection of GPS Satellite Signal Failures in Satellite Based Augmentation Systems," Proc. of the *ION GPS Meeting*, 19-22 September 2000, Salt Lake City, Utah, USA (Institute of Navigation, Alexandria, Virginia), pp. 189-198.

J.M. Dow, R.E. Neilan, G. Gendt (2005), "The International GPS Service (IGS): Celebrating the 10th Anniversary and Looking to the Next Decade," *Adv. Space Res. 36* vol. 36, no. 3, pp. 320-326, 2005. doi:10.1016/j.asr.2005.05.125.

I. J. Galysh & D. M. Craig & W. G. Reid & J. A.Buisson (1996), "Performance Analysis Of The GPS Monitor Station Timing Subsystem Enhancement Program At The Naval Research Laboratory", Proc. *Precise Time and Time Interval*, 1996, Washigton, DC.

(1999) IEEE Standard Definitions of Physical Quantities for Fundamental Frequency and Time Metrology. Random Instabilities, IEEE Standard 1139, 1999.

S. M. Kay (1998), "Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory," *Prentice Hall*, 1998.

S. M. Kay (1998), "Fundamentals of Statistical Signal Processing, Volume 1: Estimation Theory," *Prentice Hall*, 1998.

E. Nunzi & P. Carbone (2007), "The generalized likelihood ratio test for detecting anomalous behaviors of atomic clocks," (Invited paper) Proc. of *SPIE Noise and Fluctuations in Photonics, Quantum Optics, and Communications Conference*, 21-24 May 2007, Florence, Italy, vol. 6603, pages: 660310-1 - 660310-9.

E. Nunzi, L. Galleani, P. Tavella & P. Carbone (2007), Detection of Anomalies in the Behavior of Atomic Clocks, *Instr. and Meas., IEEE Transactions on*, Vol.XX, Apr. 2007, pp. 523–528.

E. Nunzi & P. Carbone (2008), Monitoring Signal Integrity of Atomic Clocks by Means of the GLRT, *Metrologia*, Vol. 45, Dec - 2008, S103-S107.

E. Nunzi and P. Carbone and P. Tavella (2008), "Fault detection in atomic clock frequency standards affected by mean and variance changes and by an additive periodic component: the GLRT approach," *Instr. and Meas. Tech. Conf.*, May 2008.

E. Nunzi & D. D'Ippolito (2009), "A Novel Theoretical Analysis of Fault Detection for Atomic Clock," *IEEE International Workshop on Advanced Method for Uncertainty Estimation in Measurement - AMUEM*, Bucharest, Romania, July 2009.

E. Nunzi & G. Barchi & U. Bartoccini (2009), "Methods and Tools for Frequency Jump Detection," *IEEE International Workshop on Advanced Method for Uncertainty Estimation in Measurement - AMUEM*, Bucharest, Romania, July 2009.

L. Vioarsson & S. Pullen & G. Green & P. Enge (2001), Satellite Autonomous Integrity Monitoring and its Role in Enhancing User Performance, Proc. of the *ION GPS Meeting*, 11-14 September 2001, Salt Lake City, Utah, USA (Institute of Navigation, Alexandria, Virginia), pp. 690-702.

M. Weiss, P. Shome, R. Beard (2006), "GPS signal integrity dependencies on atomic clocks," Proc. *38th Annual Precise Time and Time Interval (PTTI) Meeting*, Washington DC, Dec 2006, pp. 439-448.

(1997) Guide to the Expression of Uncertainty in Measurement. ISO 1997



Fault Detection

Edited by Wei Zhang

ISBN 978-953-307-037-7
Hard cover, 504 pages
Publisher InTech
Published online 01, March, 2010
Published in print edition March, 2010

In this book, a number of innovative fault diagnosis algorithms in recently years are introduced. These methods can detect failures of various types of system effectively, and with a relatively high significance.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Emilia Nunzi (2010). The Use of the GLRT for Revealing Faults in Atomic Frequency Standards, Fault Detection, Wei Zhang (Ed.), ISBN: 978-953-307-037-7, InTech, Available from: http://www.intechopen.com/books/fault-detection/the-use-of-the-glrt-for-revealing-faults-in-atomic-frequency-standards

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2010 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License, which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen