Detection of anomalous radioxenon concentrations: A distribution-free approach

Published Version:
Scagliarini, M., Gualdi, R., Ottaviano, G., Rizzo, A. (2023). Detection of anomalous radioxenon concentrations: A distribution-free approach. ENVIRONMETRICS, 34(7 (November)), 1-23 [10.1002/env.2804].

Availability:
This version is available at: https://hdl.handle.net/11585/945417 since: 2023-10-18

Published:
DOI: http://doi.org/10.1002/env.2804

TITLE

Detection of Anomalous Radioxenon Concentrations: A Distribution-Free Approach

AUTHORS

Michele Scagliarini ORCID: 0000-0003-3209-4745
Department of Statistical Sciences, University of Bologna, Italy
Via delle Belle Arti, 41, 40126, Bologna, Italy
Email: michele.scagliarini@unibo.it

Rosanna Gualdi
Department of Statistical Sciences, University of Bologna, Italy
Via delle Belle Arti, 41, 40126, Bologna, Italy
Email: rosanna.gualdi@studio.unibo.it

Giuseppe Ottaviano
ENEA, Bologna Research Centre
Via Martiri di Monte Sole, 4, 40129, Bologna, Italy
Email: giuseppe.ottaviano@enea.it

Antonietta Rizzo
ENEA, Bologna Research Centre
Via Martiri di Monte Sole, 4, 40129, Bologna, Italy
Email: antonietta.rizzo@enea.it

Abstract

The detection of anomalous atmospheric radioxenon concentrations plays a key role in detecting both underground nuclear explosions and radioactive emissions from nuclear power plants and medical isotope production facilities. For this purpose, the CTBTO's International Data Centre uses a procedure based on descriptive thresholds. In order to supplement this procedure with a statistical inference-based method, we compared several non-parametric change-point control charts for detecting shifts above the natural radioxenon background. The results indicate that the proposed methods can provide valuable tools for the institutions responsible for the verification and classification of anomalous radioxenon concentrations.

1 Introduction

The Comprehensive Nuclear Test-Ban Treaty Organization (CTBTO) supports and promotes the entry into force of the Comprehensive Nuclear Test-Ban Treaty (CTBT), a treaty that bans nuclear tests (CTBT, 1996, United Nations, 1998). The core of the CTBTO is a verification regime based on the International Monitoring System (IMS), the International Data Centre (IDC) and On-Site Inspections (OSI).

The IMS consists of 321 planned worldwide stations to detect radioactivity, seismic, infrasound and hydroacoustic signals from nuclear explosions occurring in the atmosphere, underwater and underground (Figure 1). Among these, 80 stations are able to capture and measure fission product particulate in the atmosphere (radionuclide particulate stations), 40 also being equipped to measure fission product noble gases (noble gas stations). About 90 percent of the facilities are already in operation.

Figure 1: The International Monitoring System (Courtesy of CTBTO)

Among all possible fission products, 88 have been selected as relevant in detecting nuclear explosions due to their high fission yield and suitable decay times.
The noble gases isotopes, namely ^{131m}Xe, ^{133m}Xe, ^{133}Xe and ^{135}Xe, generally known as radioxenon, are of particular interest since they are the fission products that are most likely to escape from an underground explosion. They can therefore be useful in revealing violations of the CTBT Treaty (Kalinowski et al., 2010).
However, these radioxenon fission products can also be emitted by civil sources such as medical isotope production facilities, nuclear power plants and nuclear research reactors.
Drawing conclusions about a suspected underground nuclear explosion or atmospheric release from civil sources is a very complex task (Kalinowski et al., 2010). The first step in achieving this goal is to measure and properly characterise both the radioxenon atmospheric background levels and the possible anomalous radioxenon values above the average atmospheric background level for each station. To carry out this important activity, the IMS radioxenon data are reviewed by the analysts of the CTBTO's International Data Centre in Vienna and independently analysed by analysts at the Italian National Data Centre – Radionuclides (NDC-RN) at the ENEA Bologna Research Centre, in order to search for signals that can be reasonably traced back to civil sources or nuclear tests. The NDC-RN is the Scientific Support to the Italian National Authority, the Italian Ministry of Foreign Affairs.
This screening procedure has to be carried out on the basis of scientifically proven and robust criteria, and for this purpose the IDC has defined a classification system in which a measured activity concentration is ranked as anomalous if it is above the "abnormal activity concentration" value.
The abnormal activity concentration value, or "abnormal limit", used by the IDC, is defined as:

$$L_A = Q_2 + \lambda_A \left[Q_3 - Q_1 \right] \quad (1)$$

where Q_1 is the first quartile, Q_2 the median, Q_3 the third quartile and λ_A=3. The radioxenon concentrations are measured daily and the L_A value is constantly updated after each observation, since the last 365 observations are used for computing the quartiles.
In summary, an activity concentration is classified as anomalous if it is above the monitoring station's "typical background" established by applying an inter-quartile filter to the data, updated after each observation.
By using this method, experts at NDC-RN of the ENEA Bologna Research Centre can concentrate their efforts only on anomalous concentrations. These anomalous radioactivity measurements should be cross-checked with information on the activities of the civil sources operating in the area of the

monitoring station, in order to strengthen the discrimination efforts and achieve a correct classification.

The interquartile filter algorithm in use by the IDC is based on descriptive thresholds. Therefore, it could be suitably supplemented with an inference-based method able to provide some complementary results and to identify multiple change points in the sequence of observations. This in order to assign priorities and establish which anomalous values should be analysed first.

To this end, we compare and discuss several distribution-free methodologies that could be useful for distinguishing between the atmospheric radioxenon background and anomalous values above the background level. Within this framework we need methods that: can be used generally (not based on particular distributional assumptions); can be used on a stream of single observations (no subgroups); are able to estimate the presence of change points in the process location and/or process scale; work with a desired false alarm probability, or a given ARL(H_0); and offer post-signal diagnostics to interpret the results.

For our purposes, aware of being non-exhaustive, we consider the following nonparametric change-point control charts:

• a control chart based on Recursive Segmentation and Permutation, RS/P, (Capizzi and Masarotto 2013);

• a control chart based on the Mann–Whitney statistic (Hawkins and Deng, 2010);

• two control charts based on the Kolmogorov-Smirnov and Cramer-von Mises statistics (Ross and Adams, 2012)

The paper is structured as follows. In Section 2, the data are introduced and described. In Section 3, the distribution-free control charts considered for our purpose are described in detail. In Sections 4 and 5, the main results are explained. In Section 6 we study and compare the performance of the monitoring algorithms considered by means of two simulation studies. Finally, Section 7 provides an in-depth discussion, our concluding remarks and directions for further research.

2. The data

Among the twenty-five currently operational IMS noble gas stations, the station SEX63 located in Stockholm-Sweden (59.40°N, 17.93°E) and the station MNX45 located in Ulaanbaatar-Mongolia (47.89N, 106.33E) have been selected. Both the monitoring stations were certified for regular operation in 2013.

The station SEX63 is operated by the Swedish Defence Research Agency. It is a SAUNA-II fully automatic radioactive Xenon analyser system based on a beta-gamma coincidence particle system detector (Ringbom et al., 2003), which performs sampling, processing and measurement of radioactive Xenon isotopes in 12-hour pumped samples of ambient air.

The station MNX45 is operated by “Commissariat a l'Energie Atomique, Département Analyse Surveillance Environnement” (France). This is a fully automatic SPALAX (Fontaine et al. 2004) radioactive xenon analysis system, based on a high-resolution γ-spectrometer, which takes measurements every 24 hours.

It is worth noting that the choice of these monitoring stations is not accidental. The two stations have different measurement technologies (Sauna and Spalax), are located in contrasting geographical and industrial contexts and, as will be highlighted in this paragraph, the amount of data available varies.

The motivation for this choice is to compare the control charts in distinct scenarios in order to obtain the most complete picture of their performance and the results obtainable using these methodologies.

The monitoring station SEX63 is inserted in a highly industrialized region and it is relatively close to civil radioxenon sources. The data considered in this study consist of ^{133}Xe activity concentrations (mBq/m^3) measured from 5th December 2013 to 1st August 2018. The number of days in the time period of interest is 1701. The sampling frequency of the SEX63 measurements is 12 hours, so 3402 observations should be available.

In our case, data are completely missing on 285 days (16.75%). Of the remaining 1416 days, 371 days (26.2%) have only one observation. Therefore, there are a total of 2461 valid data (72.34%) and the percentages of valid data in each year are 92.59% in 2013, 83.84% in 2014, 68.49 in 2015, 70.08 in 2016, 64.52 in 2017 and 73.94 in 2018, respectively.

Figures 2 and 3 show the time series plot and the histogram, respectively, of the ^{133}Xe data at our disposal, while Table 1 provides some summary statistics.

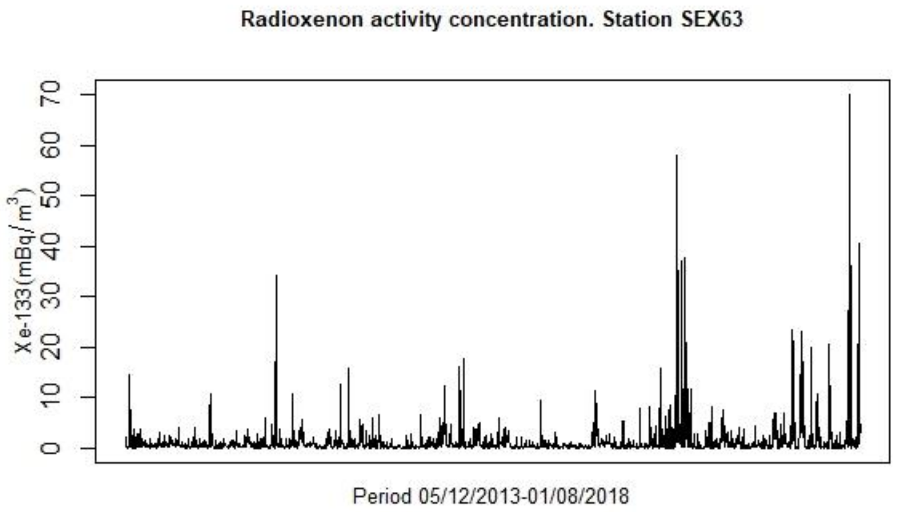

Figure 2. Time sequence plot of the available ^{133}Xe activity concentrations in the period 5/12/13-01/08/18

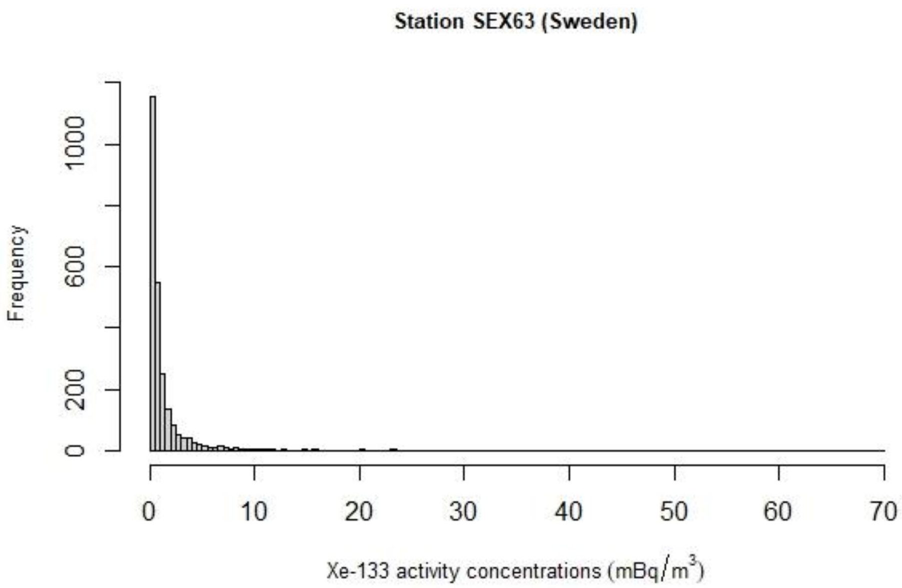

Figure 3. Histogram of the available ^{133}Xe activity concentrations

Min.	First Qu.	Median	Mean	Third Qu.	Max.	Variance	Skewness	Kurtosis
0.0911	0.2720	0.5500	1.3914	1.2200	70.0000	11.8431	9.9139	141.2442

Table 1. Descriptive statistics of ^{133}Xe activity concentrations (*mBq/m*^{3}) concentration for the SEX63 station

The monitoring station MNX45 is quite remote from civil sources, although it may be marginally affected by emissions from China and the Russian Federation. The data available for this study consist of ^{133}Xe activity concentrations (mBq/m^3) measured from 4th March 2013 to 9th June 2018. The number of days in the time period of interest is 1924. The sampling frequency of the MNX45 measurements is 24 hours. In our case, data are missing on 1474 days (76.61%). Overall, there are therefore 450 valid data (23.39%), while the percentages of valid data in each year are 13.20% in 2013, 18.36% in 2014, 28.77 in 2015, 18.58 in 2016, 34.52 in 2017 and 27.50 in 2018, respectively. Figures 4 and 5 show the time series plot and the histogram, respectively, of the ^{133}Xe data at our disposal, while Table 2 provides some summary statistics.

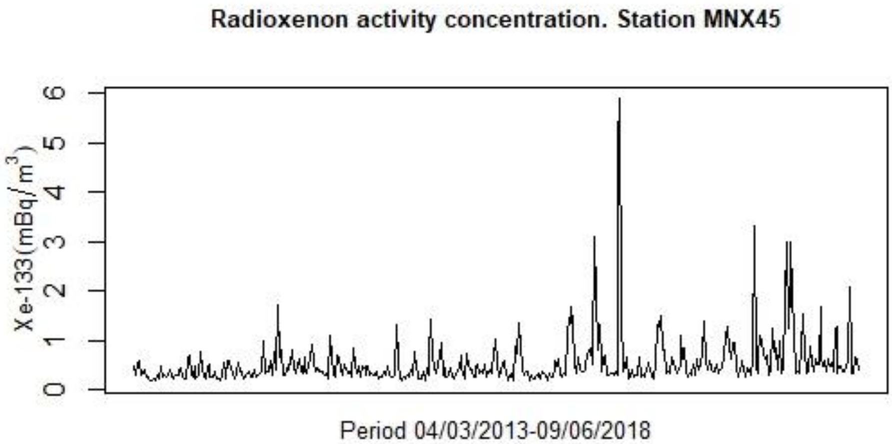

Figure 4. Time sequence plot of the available ^{133}Xe activity concentrations in the period 4/3/13-09/06/18

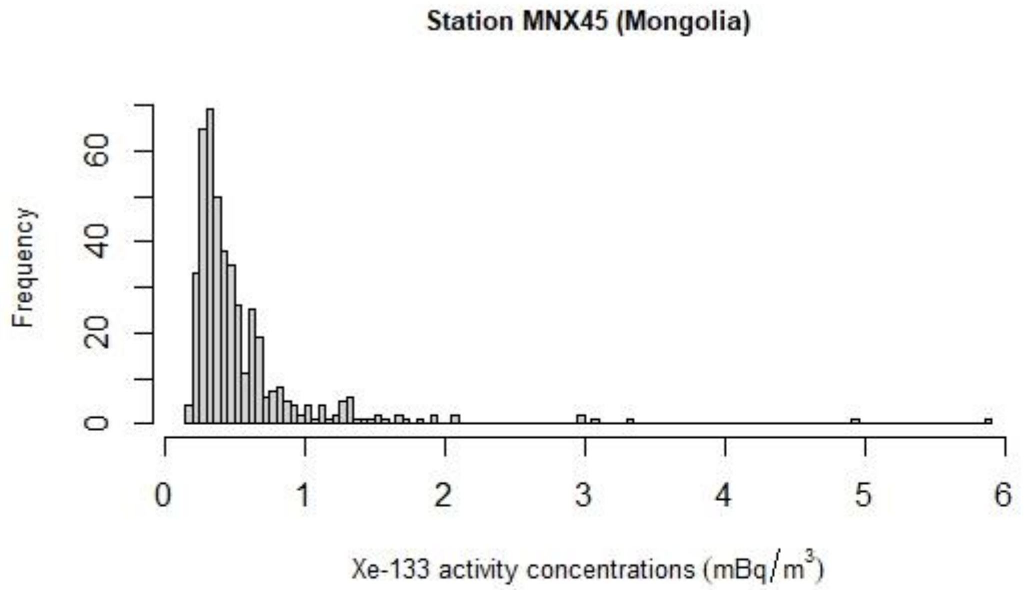

Figure 5. Histogram of the available ^{133}Xe activity concentrations

Min.	First Qu.	Median	Mean	Third Qu.	Max.	Variance	Skewness	Kurtosis
0.1840	0.3063	0.4045	0.5547	0.6078	5.8800	0.2702	5.2300	42.2403

Table 2. Descriptive statistics of ^{133}Xe activity concentrations (mBq/m^3) concentration for the MNX45 station

3. Methods

Statistical process control monitoring algorithms, such as control charts, are effective for monitoring process data in order to detect whether there has been a change point in which the data distribution significantly deviates from a baseline distribution (Montgomery, 2009).

Since its first applications in laboratories (Rodbard et al., 1968), statistical process control has received increasing attention in various non-manufacturing fields. Some recent examples include, but are not limited to, the monitoring of healthcare services (Erfanian et al., 2021), the monitoring and classification of environmental data (Paroissin et al., 2016) and the monitoring of customer waiting times (Suzuki et al., 2021).

As noted in the previous section, we are dealing with data characterised by a high level of skewness, which can lead to misleading results if it is not taken into due account. In this situation, or more generally in all the cases where the underlying distributional form of the monitored data cannot be assumed to be known, it is important to consider approaches that are as general as possible. Within this framework, distribution-free methods may be suitable candidates, as they offer the advantage of not depending on any distributional assumption.

Literature concerning distribution-free control charts includes a significant amount of interesting research. To name but a few: Chakraborti (2011), Chakraborti and Graham (2019), Jones-Farmer et al. (2009), Qui and Li (2011), Qui (2013), Zou et al. (2014). In these contributions, the importance of nonparametric control charts in modern and non-conventional applications of statistical process control is highlighted.

Furthermore, an aspect to be carefully considered is the fact that in the non-parametric setting, not all the monitoring schemes can be implemented with individual observations (n=1). However, in our case, there is a continuous stream of individual data from the monitoring station that should be analysed.

Numerous and interesting scientific contributions are available to address the change point detection problem in univariate data streams (Niu et al. 2016). For our purposes, aware of being non-exhaustive, we consider: the nonparametric change-point control chart proposed by Hawkins and Deng (2010) which is based on the Mann–Whitney statistic; the distribution-free control chart, based on recursive segmentation and permutation, RS/P, proposed by Capizzi and Masarotto (2013); the two non-parametric control charts based on the Kolmogorov-Smirnov and Cramer-von-Mises statistics proposed by Ross and Adams (2012).

In the Change Point Model (CPM) framework, it is assumed that the time ordered observations $x_1, x_2, ...$ are generated by the random variables $X_1, X_2, ...$ with unknown distribution functions $F_1, F_2, ...$, respectively. A change point occurs at instant τ when $F_\tau \neq F_{\tau+1}$ and it is usually assumed that the observations are independent and identically distributed between every pair of change points. Therefore, the distribution of the sequence can be described by the following model:

$$X_i \sim \begin{cases} F_0 \text{ if } 0 < i \leq \tau_1 \\ F_1 \text{ if } \tau_1 < i \leq \tau_2 \\ \ldots \\ F_k \text{ if } \tau_k < i \leq m \end{cases} \tag{2}$$

where $0 < i < \tau_1 < \tau_2 < \cdots < \tau_k < m$ denote k unknown change points.

Within this framework, it is of interest to test $H_0 : k = 0$, versus $H_1 : k > 0$, to estimate the number of change points, k, and to estimate their locations $\hat{\tau}_1, \hat{\tau}_2, ..., \hat{\tau}_k$.

3.1 Recursive Segmentation and Permutation CPM Control Chart

It has been proposed for Phase I analysis and is able to detect location (mean) and scale (variability) shifts in sub-grouped or individual data. The most relevant methodological aspects of the methodology are summarised below, while complete details can be found in Capizzi and Masarotto (2013) and Capizzi (2015).

The RS/P control chart consists of three steps. Initially, a set of elementary test statistics, designed to detect the possible shifts are calculated. These statistics depend on the type of shift to be detected. For detecting isolated shifts in the location of the monitored variable the test statistic is

$$T_0 = \max_{i=1,\dots,m} \left| \bar{x}_i - \bar{\bar{x}} \right|, \tag{3}$$

where $\bar{x}_i = \frac{1}{n}\sum_{j=1}^{n} x_{ij}$ and $\bar{\bar{x}} = \frac{1}{n}\sum_{i=1}^{m} \bar{x}_i$. It is important to note that detecting isolated shifts is only possible for n>1. For this reason, the computation of T_0 is omitted in the case of individual observations (n=1).

The statistics T_k, k=1, 2, …, K, designed for detecting step shifts, are computed using a forward recursive-segmentation approach:

- The algorithm starts with k=0 and proceeds in K successive stages. Here, K represents the maximum number of change points that the algorithm will search for. This parameter can be defined by the user, but it is suggested to set $K = \max\left(3, \min\left(50, \left[\frac{m}{15}\right]\right)\right)$ where $\left[x\right]$ denotes the integer closest to x (Capizzi and Masarotto 2013);
- At the beginning of each stage k, the interval [1, m] is partitioned into k subintervals, each having a length greater than or equal to l_{MIN}; l_{MIN} is also a user-modifiable parameter, and the choice l_{MIN}=5 improves the detection capability of the test without increasing the false alarm probability (Capizzi and Masarotto 2013);
- At stage k, conditionally on the results of the previous stages, one of these subintervals is split, adding a new potential change point; the identification of potential change points is achieved by maximizing a function f_k. The control statistic T_k is equal to the attained maximum value of function f_k. The function f_k has a formulation that depends on the goal to be achieved. For detecting shifts in the process location, the function f_k is defined as

$$f_k = \sum_{i=1}^{k+1} \left(\hat{\tau}_i - \hat{\tau}_{i-1}\right)\left(\bar{x}\left(\hat{\tau}_{i-1}, \hat{\tau}_i\right) - \bar{\bar{x}}\right)^2, \tag{4}$$

where $0=\hat{\tau}_0<\hat{\tau}_1<...<\hat{\tau}_k<\tau_{k+1}=m$ are the boundaries of the new partition and $\overline{x}(a,b)=\frac{1}{b-a}\sum_{i=a+1}^{b}\overline{x}_i$.

In the second step, the statistics $T_0,\ldots T_K$ are standardized and aggregated. In such a way, it is possible to obtain the overall control statistic

$$W_{obs}=\max_{r=0,\ldots,K}\frac{T_r-E\left(T_r\,|\,\mathbf{Y}_{(\cdot)}\right)}{\sqrt{\operatorname{var}\left(T_r\,|\,\mathbf{Y}_{(\cdot)}\right)}}, \tag{5}$$

where $\mathbf{Y}_{(\cdot)}$ denotes the ordered pooled sample. Therefore, the *p*-value can be computed by means of a permutation approach (Pesarin, 2001, Good, 2005):

- If we assume that, when the process is in control, the observations are *iid*, then it can be shown that (Lehmann and Romano 2005)

$$P\left\{\mathbf{Y}=(a_1,...,a_N)\,|\,\mathbf{Y}_{(\cdot)}\right\}=\begin{cases}\frac{1}{N!} \text{ if } (a_1,...,a_N) \text{ is a permutation of } \mathbf{Y}_{(\cdot)}\\ 0 \text{ otherwise}\end{cases}, \tag{6}$$

  where N=$m\cdot n$, then equation (6) does not depend on the unknown distribution $F_0(\cdot)$. Thus given a test statistic, the correspondent *p*-value can be calculated, conditionally to $\mathbf{Y}_{(\cdot)}$, as the proportion of permutations under which the statistic value exceeds or is equal to the statistic computed from the original sample.

In the third step, a time-varying mean estimate is calculated for diagnostic purposes. For detecting changes in scale, the steps are the same as for the process location, but the function f_k used for the recursive segmentation is different (Capizzi and Masarotto 2013).

The RS/P method is implemented in the *R* package *dfphase1* (R Core Team, 2020, Capizzi and Masarotto, 2017). To facilitate the interpretation of the results, the package *dfphase1* provides some graphical tools in which the following are displayed:

- The original observations;
- The test statistics;
- The possible time-varying estimates $\hat{\mu}_i$ and $\hat{\sigma}_i$, of process level and scale, respectively;
- The *p*-values adjusted using the Holm-Bonferroni correction (Holm, 1979) so that comparing both the *p*-values with the same threshold α leads to an overall first error probability (or false alarm probability) $\leq\alpha$;
- The estimated change points.

3.2 Mann-Whitney CPM Control Chart

It has been proposed for detecting location changes and uses a *U*-statistic based on the Mann–Whitney two-sample test (Hawkins and Deng 2010)

$$U_{k,n}=\sum_{i=1}^{k}\sum_{j=k+1}^{n}D_{ij}\quad 1\le k\le n-1, \tag{7}$$

where

$$D_{ij}=\operatorname{sgn}\left(X_i-X_j\right)=\begin{cases}1 \text{ if } X_i>X_j\\ 0 \text{ if } X_i=X_j\\ -1 \text{ if } X_i<X_j\end{cases}. \tag{8}$$

The variance of $U_{k,n}$ depends on the split point *k*, so for this reason it is suitably standardized, thus obtaining the statistic $T_{k,n}$ (Hawkins and Deng 2010). Therefore, the test for the presence of a change point and the estimate of its time of occurrence are given by maximizing $T_{k,n}$ over *k*,

$$T_{\max,n}=\max_{1\le k\le n-1}\left|T_{k,n}\right|. \tag{9}$$

Note that in this formulation, the data set size *n* is no longer fixed but grows as long as the process is monitored and every instant of time is analysed as a potential change point by carrying out a two-sample Mann-Whitney test between the observations before and after that time. If $T_{\max,n}$ exceeds a specified control limit $h_{n,\alpha}$, the method signals that a change has occurred and estimates the time of change as the maximizing split point:

$$\hat{\tau}_T=\arg\max_{1\le k\le n-1}\left|T_{k,n}\right|. \tag{10}$$

The thresholds $h_{n,\alpha}$ are such that the conditional probability of a false alarm at any *n* is constant and equal to α and are established by simulation (Hawkins and Deng 2010).

The MW control chart is implemented in the *R* package *cpm* (Ross, 2015).

3.3 Cramer-von Mises and Kolmogorov-Smirnov CPM Control Charts

The Cramer-von Mises (CvM) and Kolmogorov-Smirnov (KS) control charts are designed to detect arbitrary changes in the process distribution (Ross and Adams, 2012).

Both the control charts test for a change point immediately following any observation x_k by partitioning the observations into two samples $S_1 = \left(x_1, ..., x_k\right)$ and $S_2 = \left(x_{k+1}, ..., x_t\right)$, and comparing the corresponding empirical distribution functions $\hat{F}_{S_1}\left(x\right)$ and $\hat{F}_{S_2}\left(x\right)$.

The CvM test uses a statistic based on the square of the average distance between the empirical distributions

$$W_{k,t} = \int_{-\infty}^{\infty} \left|\hat{F}_{S_1}\left(x_i\right) - \hat{F}_{S_2}\left(x_i\right)\right|^2 dF_t\left(x\right). \tag{11}$$

A change point is detected if $W_{k,t} > {}_W h_{k,t}$.

The KS test uses a statistic defined as the maximum difference between the empirical distributions

$$D_{k,t} = \sup_x \left|\hat{F}_{S_1}\left(x_i\right) - \hat{F}_{S_2}\left(x_i\right)\right|, \tag{12}$$

and in this case a change point is also detected if $D_{k,t} > {}_D h_{k,t}$.

Whenever a new observation x_t is received, the sequence $\left(X_1, ..., X_t\right)$ is considered as being a fixed-size sample. If H_0 is rejected, the estimate $\hat{\tau}$ of the change point location is the value of k that maximizes $W_{k,t}$ (or $D_{k,t}$). The thresholds ${}_W h_{k,t}$ (or ${}_D h_{k,t}$), such as to guarantee the desired ARL_0, are determined by means of the Monte Carlo simulation.

For the sake of completeness, it is worth noting that both $W_{k,t}$ and $D_{k,t}$ have variances which depend on *k*. Consequently, the test is performed on the standardized versions of the statistics $W_{k,t}$ and $D_{k,t}$. Full details can be found in Ross and Adams (2012). The CvM and KS control charts are implemented in the *R* package *cpm* (Ross, 2015).

Please note that to allow fair comparisons in Section 4 we will set Mann-Whitney, Cramer-von Mises and Kolmogorov-Smirnov CPM Control Charts to the same value of ARL_0 and, as a reasonable choice, we have selected ARL_0=200.

4. Results for the monitoring station SEX63

For the monitoring station SEX63 we considered the 365 valid observations from April 19, 2017 to January 26, 2018. Figure 6 shows the ^{133}Xe activity concentrations along with the abnormal limit computed, on the basis of the available data, by the IDC (the red continuous line at 2.59 mBq/m^3).

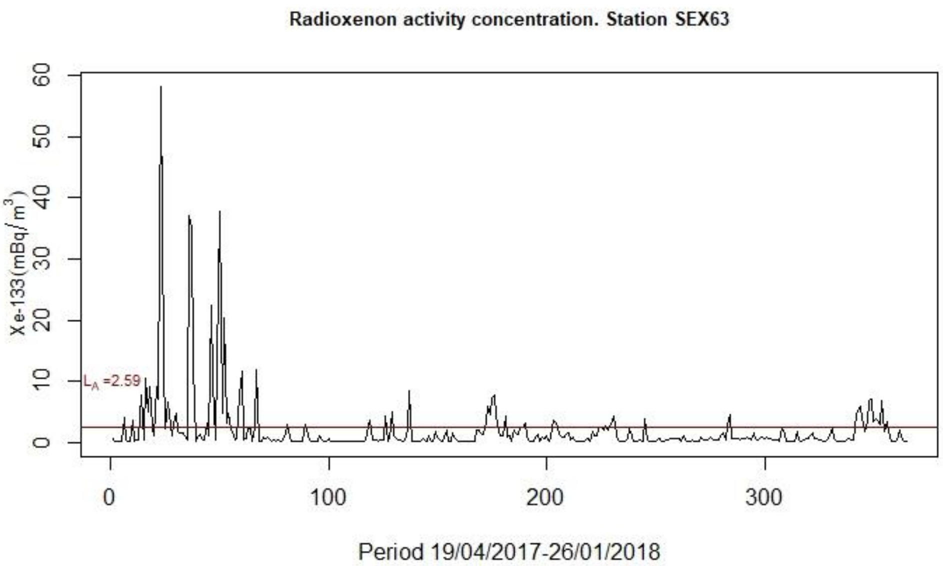

Figure 6. Station SEX63: time sequence plot of the available ^{133}Xe activity concentrations and L_A value for the period 19/04/2017-26/01/2018

In this case there are sixty-six observations above the L_A (observations 6, 10, 14, 16-18, 21-24, 26-27, 29-30, 36-38, 44, 46-47, 49-54, 59-60, 67, 126, 129, 137, 172-177, 181, 285, 342-345, 347-355). Among these, twenty-nine occur in the initial period up to the 67th observation, while the other thirty-seven exceedances are distributed over the remaining time period.

Before applying the CPM control charts, the independence assumption was checked. Figures 7 and 8 show the estimated global and partial autocorrelation functions, respectively.

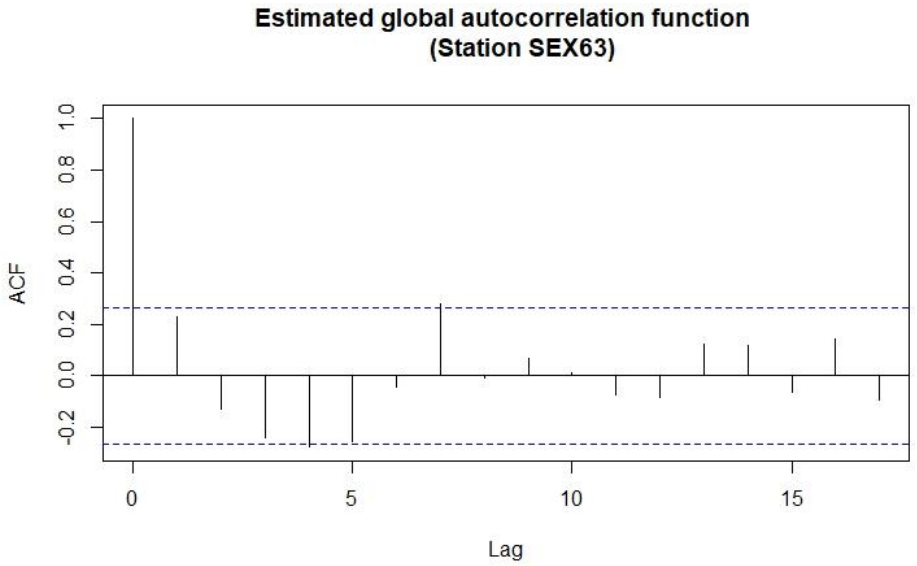

Figure 7. Station SEX63: estimated global autocorrelation function

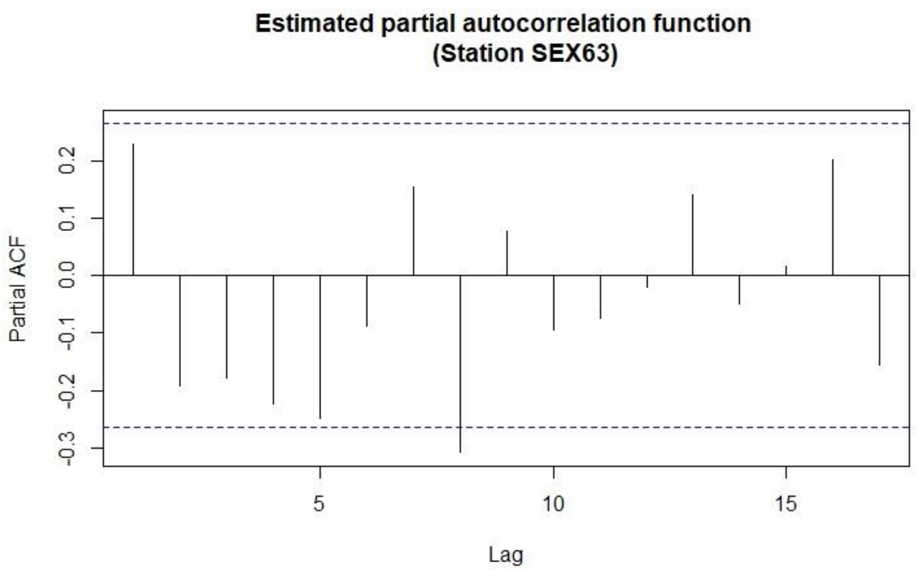

Figure 8. Station SEX63: estimated partial autocorrelation function

The autocorrelation analysis does not show any clear pattern in the estimated global and partial autocorrelation functions. Therefore, it is reasonable to assume that the observations are in fact independent.

4.1 RS/P Control Chart

Figure 9 shows the results obtained with the RS/P control chart for testing the stability of the ^{133}Xe mean (location) over time. The Figure shows the original observations (continuous line), the abnormal limit (the same as Figure 6), the estimated time-varying process means (dashed line) and the estimated change points. In the case in question the hypothesis of stability of the ^{133}Xe location is rejected and one change point is estimated at observations 61.

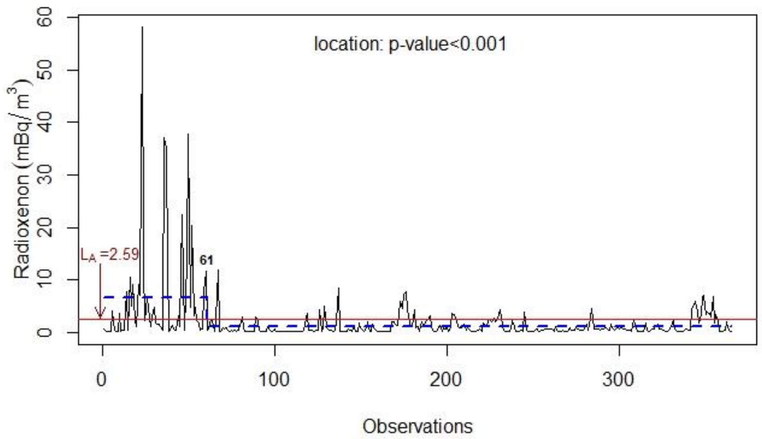

Figure 9. Station SEX63: results of RS/P control chart for detecting location shifts

Figure 10 illustrates results for monitoring the scale of the ^{133}Xe activity concentrations. Please note that we have made some changes to the output of the *dfPhaseI* package. To monitor variability, the solid line usually plotted by the RS/P procedure corresponds to either $s_i = \sqrt{\sum_{j=1}^{n}\left(x_{ij} - \hat{\mu}_i\right)^2 / n}$ for sub-grouped data or $\left|x_i - \hat{\mu}_i\right|$ for individual data. Here, in order to better understand the behaviour of the monitored phenomenon, we decided to plot the original observations instead of $\left|x_i - \hat{\mu}_i\right|$.

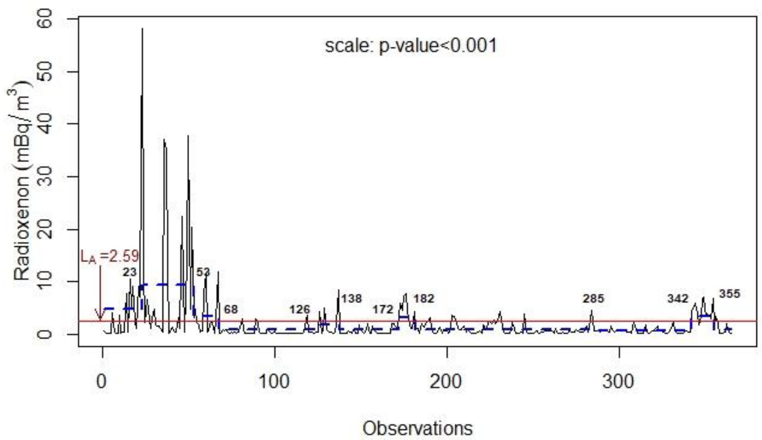

Figure 10. Station SEX63: results of RS/P control chart for detecting scale shifts

Therefore, in Figure 10 the original observations and the abnormal limit are the same as those shown in Figure 9, whereas the dashed line is an estimate of the possible time-varying process variability. It can be noted that the hypothesis of stability over time of the ^{133}Xe variability is rejected and ten possible change points are estimated at observations 23, 53, 68, 126, 138, 172, 182, 285, 342 and 355 respectively.

The results indicate that the RS/P control chart estimates change points in location and scale often corresponding to observations which exceed the abnormal limit threshold. As an example, there is an initial period, up to observation 68, in which one shift in the mean (Figure 9) and three changes in the variability (Figure 10) are estimated. In the same period, there are twenty-nine observations above the L_A threshold (observations 6, 10, 14, 16-18, 21-24, 26-27, 29-30, 36-38, 44, 46-47, 49-54, 59-60, 67).

In order to formally and objectively evaluate the relationship between the series of the estimated change points and the series of the abnormal limit exceedances, we used the Event Coincidence Analysis (ECA) (Donges et al., 2016). ECA considers the general viewpoint of possibly coupled point processes and compares the empirical co-occurrence frequencies with those one may expect under the hypothesis of uncoupled Poisson processes.

In the present case, there are two sequences of events of distinct types:

- events A (exceedances of L_A threshold) that occur at times $t_i^A \ \left(i = 1, ..., N_A\right)$;
- events B (the estimated change points) that occur at times $t_j^B \ \left(j = 1, ..., N_B\right)$.

Event coincidence analysis is based on counting coincidences between events of different types with B events that precede A events. A coincidence occurs if two events at t_i^A, t_j^B with $t_j^B < t_i^A$ are closer in time than a temporal tolerance or coincidence interval ΔT, i.e.

$$t_i^A - t_j^B \le \Delta T \ . \tag{13}$$

The precursor coincidence rate (14) is used to quantify the intensity of the coincidences between event time series A and B

$$r_p\left(\Delta T, \tau\right) = \frac{1}{N_A} \sum_{i=1}^{N_A} \Theta\left[\sum_{j=1}^{N_B} 1_{[0,\Delta T]}\left(\left(t_i^A - \tau\right) - t_j^B \right) \right]. \tag{14}$$

The rate $r_p(\Delta T, \tau)$ measures the fraction of A-type events that are preceded by at least one B-type event (note that multiple B-type events within the coincidence interval are counted only once). In (14), $\Theta(x) = 0$ for $x \leq 0$ and $\Theta(x) = 1$ otherwise, $1_I(\cdot)$ is the indicator function of the interval I (defined as $1_I(x) = 1$ for $x \in I$ and $1_I(x) = 0$ otherwise) and $\tau \geq 0$ is a time lag parameter which allows lagged relationships between event series to be taken into account. Event coincidence analysis is implemented in the *R* package *CoinCalc* (Siegmund et al., 2017). In addition to the computation of $r_p(\Delta T, \tau)$, the package *CoinCalc* provides a selection of significance tests in order to cope with specific properties of the event series under study. Under the assumption that the A and B-type events are randomly distributed and mutually independent, an analytical test based on a Poissonian approximation can be used. If conditions (events distributed independently and uniformly) for this approximation do not hold, *CoinCalc* provides an approximated significance test (Shuffle test) based on the generation of an ensemble of surrogate event series where only the numbers of events in both series (N_A and N_B) are prescribed and the actual event times are selected uniformly at random from the time interval of observations. In what follows we will show the results obtained with both tests.
For our purpose, we set as a reasonably coincidence interval $\Delta T = 10$. As far as the time lag parameter is concerned, we set $\tau = 0$ in such a way that only the instantaneous coincidences within the coincidence interval ΔT were considered.
Therefore, considering the eleven points estimated by the RS/P control chart in total, the result is $r_p(\Delta T, \tau) = 0.49$ with *p-value*=$0.9 \cdot 10^{-4}$ (Poisson) and *p-value*=0.001 (Shuffle).

4.2 Mann-Whitney Control Chart

Figure 11 shows the results obtained with the Mann-Whitney control chart. Please note that to obtain graphical outputs comparable with those of the RS/P control chart we made some changes in the output of the *cpm* package. Figure 11 reports the estimated change points, the original observations and the averages of the observations between every pair of change points (dashed line). In this case, twenty-two change points are estimated and the precursor coincidence rate is $r_p(\Delta T, \tau) = 0.61$ with *p-value*=0.01 (Poisson) and *p-value*=0.041 (Shuffle).

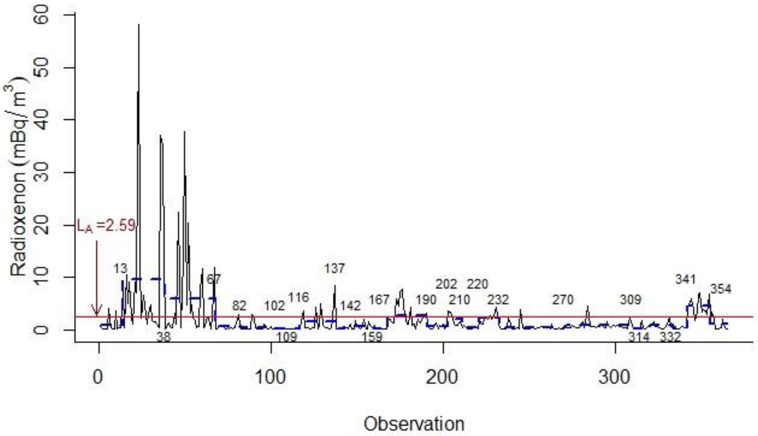

Figure 11. Station SEX63: results of MW control chart

4.3 Cramer-von Mises Control Chart

Figure 12 shows the results obtained with the Cramer-von Mises Control Chart. Twenty change points are estimated and the precursor coincidence rate is $r_p(\Delta T, \tau) = 0.55$ with *p-value*=0.03 (Poisson) and *p-value*=0.061 (Shuffle).

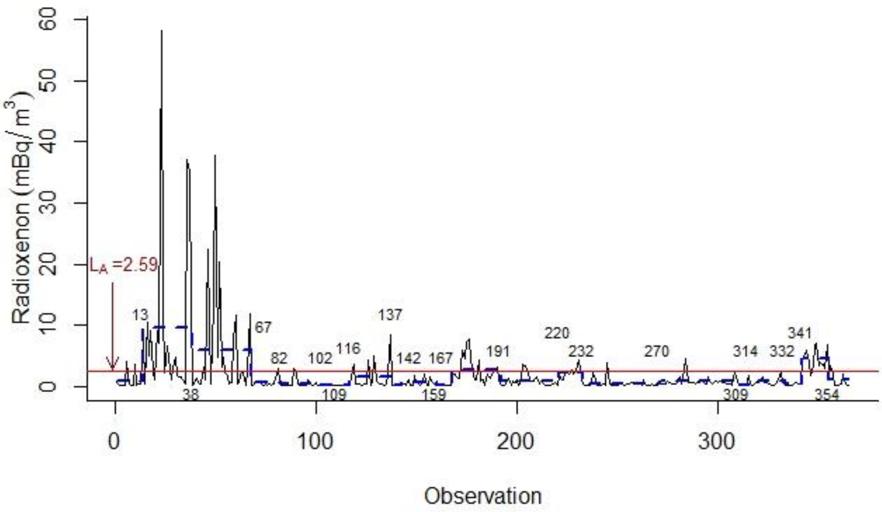

Figure 12. Station SEX63: results of CvM control chart

4.4 Kolmogorov Smirnov

Figure 13 shows the results obtained with the Kolmogorov-Smirnov Control Chart. Sixteen change points are estimated and the precursor coincidence rate is $r_p(\Delta T, \tau) = 0.51$ with *p-value*=0.01 (Poisson) and *p-value*=0.031 (Shuffle).

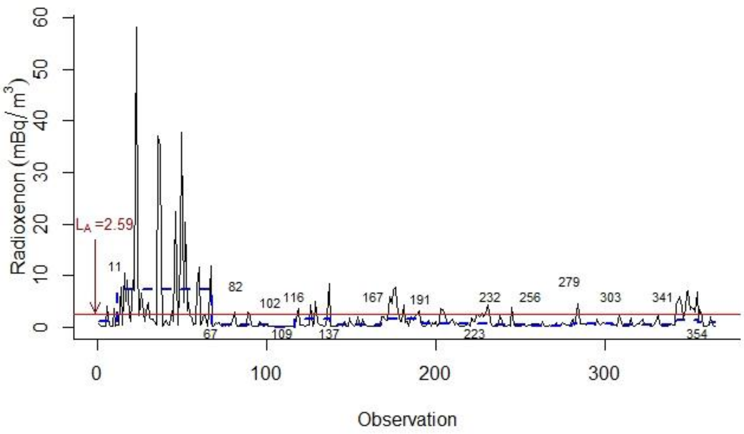

Figure 13. Station SEX63: results of KS control chart

5. Results for the monitoring station MNX45

For the monitoring station MNX45 we considered the 365 valid observations from March 4, 2013 to September 14, 2017. Figure 14 shows the ^{133}Xe activity concentrations along with the abnormal limit computed by the IDC (the red continuous line at 0.938 mBq/m^3).

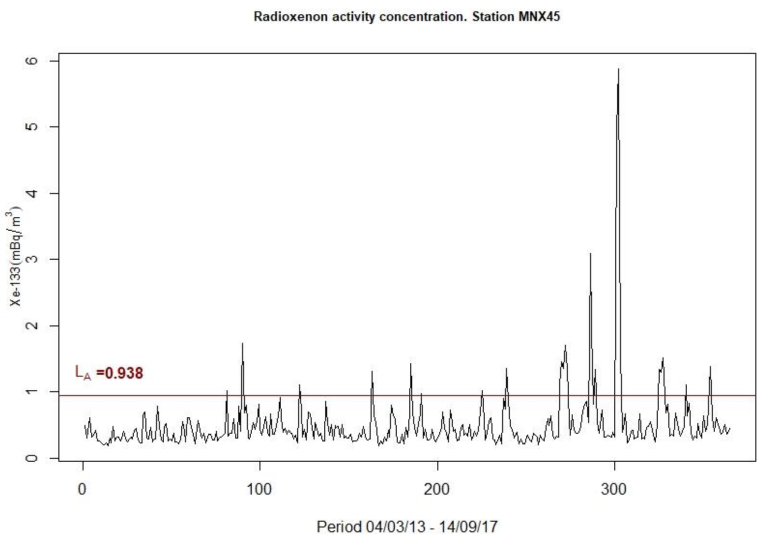

Figure 14. Station MNX45: time sequence plot of the available ^{133}Xe activity concentrations and L_A value for the period 04/03/2013-14/09/2017

In this case, there are twenty-five exceedances of L_A threshold (observations 81, 90, 122, 163, 185, 191, 225, 239, 269-273, 286-287, 289, 301-303, 325-328, 340, 354). The estimated global and partial autocorrelations are reported in Figure 15 and 16, respectively.

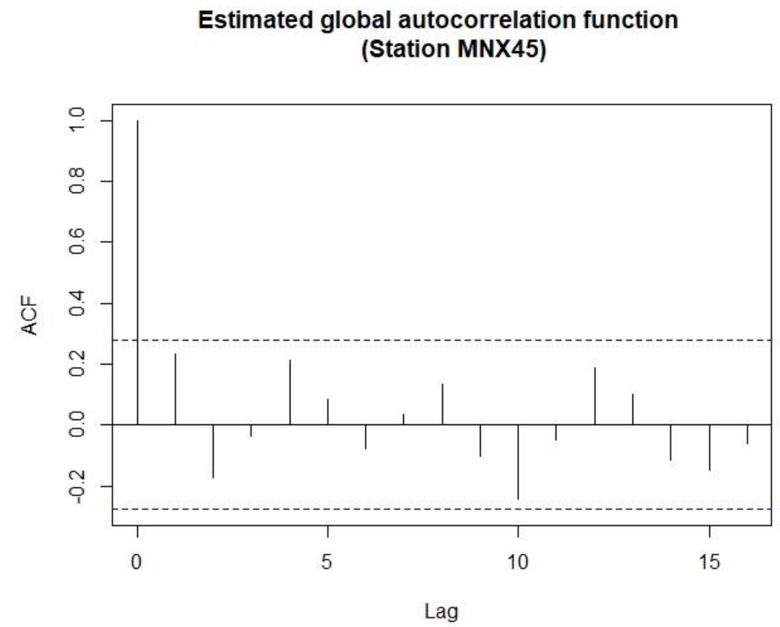

Figure 15. Station MNX45: estimated global autocorrelation function

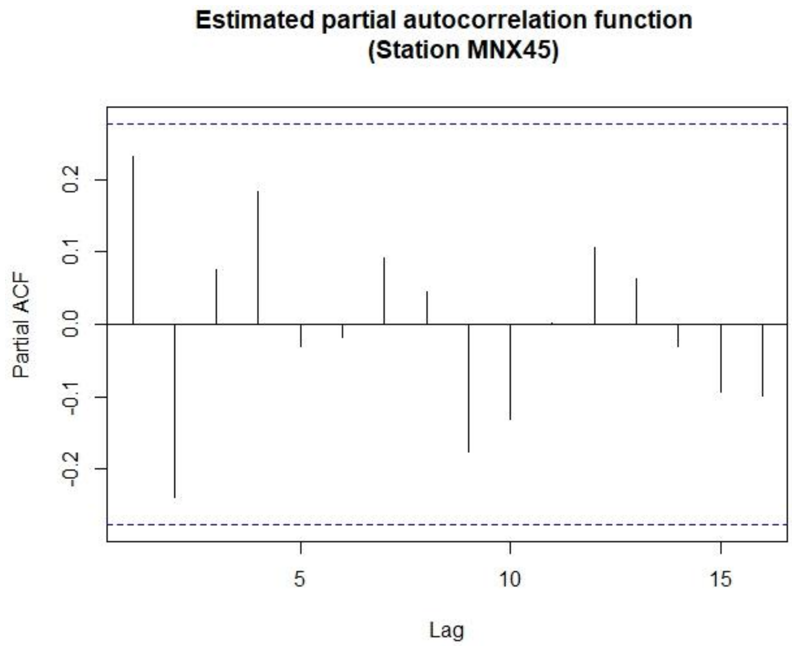

Figure 16. Station MNX45: estimated partial autocorrelation function

For station MNX45, the autocorrelation analysis also doesn't show any clear pattern in the estimated global and partial autocorrelation functions. Therefore, it is reasonable to assume the independence of the observations.

5.1. RS/P Control Chart

The RS/P control chart detects one change point in the location (Figure 17) and thirteen change points in the scale (Figure 18).

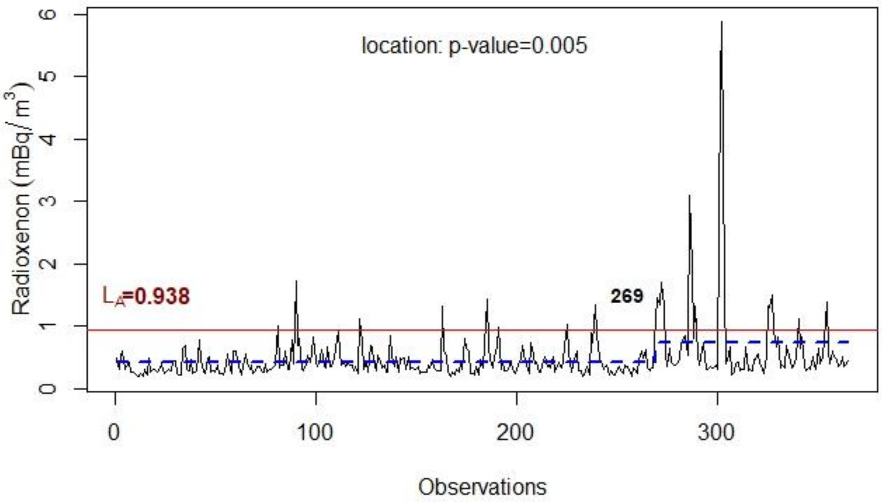

Figure 17. Station SEX63: results of RS/P control chart for detecting location shifts

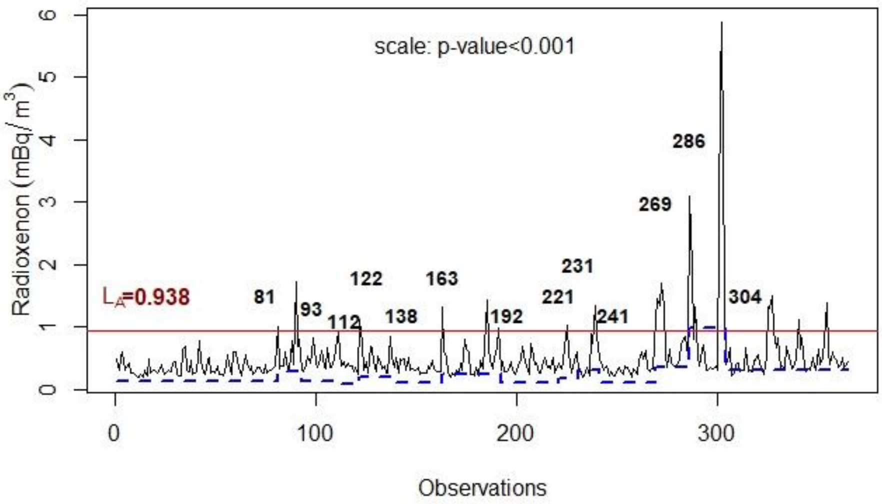

Figure 18. Station MNX45: results of RS/P control chart for detecting scale shifts

For Station MNX45, the estimated change point in the location coincides with one of the estimated change points in the scale. Thus, considering a total of thirteen change points, the precursor coincidence rate is $r_p(\Delta T, \tau) = 0.56$ with *p-value*=0.0066 (Poisson) and *p-value*=0.006 (Shuffle).

5.2 Mann-Whitney Control Chart

Figure 19 shows the results obtained with the Mann-Whitney CPM control chart.

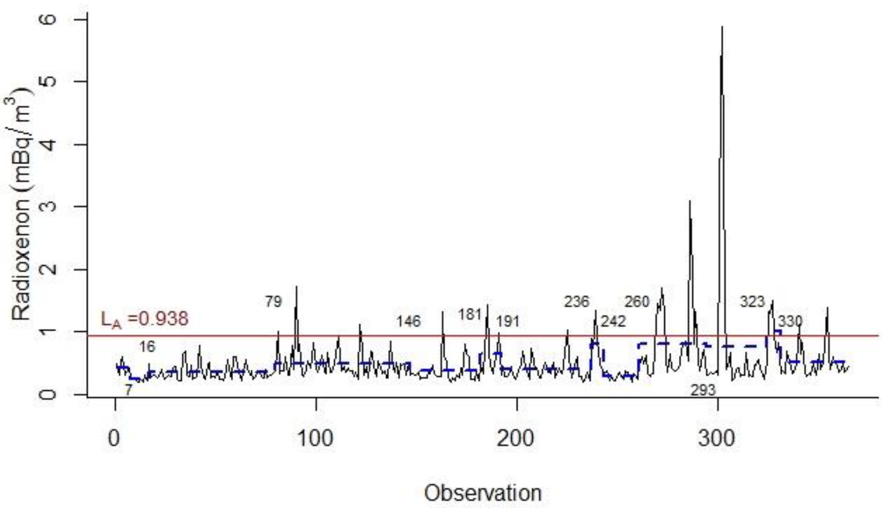

Figure 19. Station MNX45: results for Mann-Whitney control chart

In total, twelve change points are estimated and the precursor coincidence rate is $r_p(\Delta T, \tau) = 0.56$ with *p-value*=0.003 (Poisson) and *p-value*=0.002 (Shuffle).

5. 3. Cramer–von-Mises Control Chart

In Figure 20 the results obtained with the CvM Control Chart are reported. The estimated change points are also twelve in this case, with a precursor coincidence rate of $r_p(\Delta T, \tau) = 0.56$ and *p-value*=0.003 (Poisson) and *p-value*=0.005 (Shuffle).

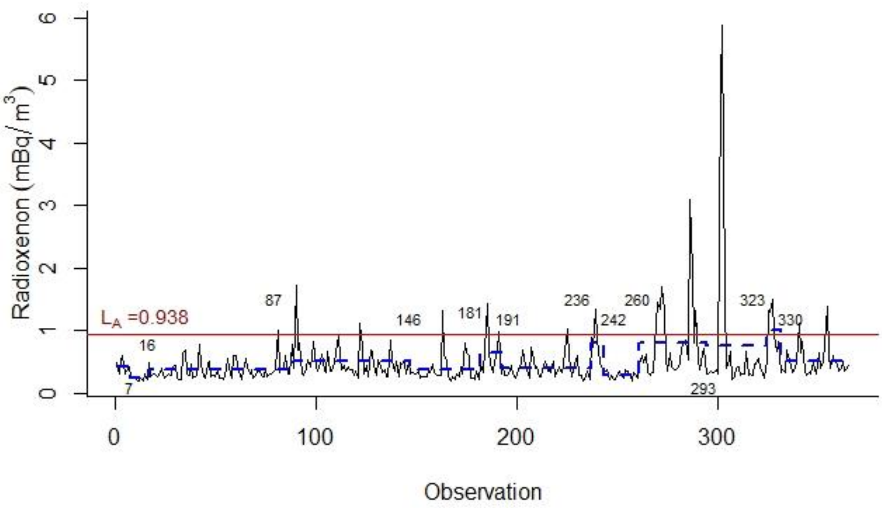

Figure 20. Station MNX45: results for CvM control chart

Please note that the results obtained in this case are very similar to those obtained with the MW Control Chart (Figure 19). There is only one difference: the CvM control chart estimates a change point at observation 87 while the MW control chart at instant 79.

5.4. Kolmogorov-Smirnov control chart

The results obtained with the KS control chart are reported in Figure 21. In total, nine change points are estimated with $r_p(\Delta T, \tau) = 0.52$ and *p-value*=0.001 (Poisson) and *p-value*=0.001 (Shuffle).

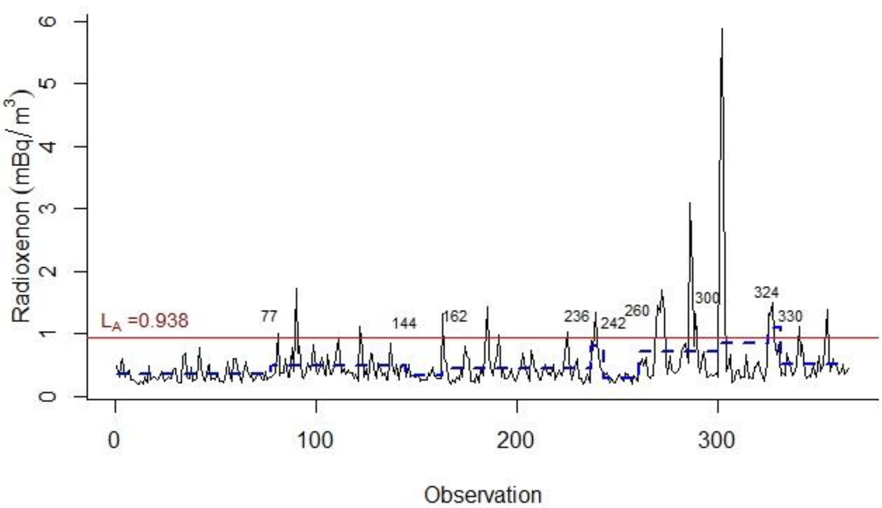

Figure 21. Station MNX45: results of KS Control chart

6. Simulation studies

In order to identify the most suitable monitoring algorithm to supplement the activity of the National Data Centres it is important to base our conclusions on objective criteria. With this aim, we are going to study and compare the performance of the considered control charts by means of two simulation studies.

To make the comparison fair, it is necessary to generate data similar to those observed. To this end, we focus our attention on the monitored data of the SEX63 station. By using the *R* packages *glamss* (Stasinopoulos and Rigby, 2008, Rigby et al., 2019) and *GeneralizedHyperbolic* (Scott, 2018) we looked for a distribution which fits to the data and the best result has been obtained with the generalized inverse Gaussian distribution.

The generalized inverse Gaussian distribution has density function given by

$$f(x)=\frac{(\psi/\chi)^{\frac{\lambda}{2}}}{2K_\lambda\left(\sqrt{\psi\chi}\right)}x^{\lambda-1}e^{-\frac{1}{2}\left(\chi x^{-1}+\psi x\right)} \tag{15}$$

for $x>0$, where $K_\lambda()$ is the modified Bessel function of the third kind with order λ (Jørgensen, 1982). The domain of variation of the parameters in (15) is given by

$$\lambda\in R,\ (\chi,\psi)\in\Theta_\lambda \tag{16}$$

where

$$\Theta_\lambda=\begin{cases}\{(\chi,\psi):\chi\geq 0,\psi>0\} & \text{if } \lambda>0\\ \{(\chi,\psi):\chi>0,\psi>0\} & \text{if } \lambda=0\\ \{(\chi,\psi):\chi>0,\psi\geq 0\} & \text{if } \lambda<0.\end{cases} \tag{17}$$

The expected value and the variance of (15) are given by

$$E(X)=\mu=\frac{\sqrt{\chi}K_{\lambda+1}\left(\sqrt{\chi\psi}\right)}{\sqrt{\psi}K_\lambda\left(\sqrt{\chi\psi}\right)} \tag{18}$$

and

$$V(X)=\sigma^2=\left(\frac{\chi}{\psi}\right)\left[\frac{K_{\lambda+2}\left(\sqrt{\chi\psi}\right)}{K_\lambda\left(\sqrt{\chi\psi}\right)}-\left(\frac{K_{\lambda+1}\left(\sqrt{\chi\psi}\right)}{K_\lambda\left(\sqrt{\chi\psi}\right)}\right)^2\right] \tag{19}$$

respectively.

Jørgensen (1982) proposed the symbol $N^{-1}(\lambda,\chi,\psi)$ for the distribution (15) and defined N_λ^{-1} as the class of distributions given by

$$N_{\lambda}^{-1} = \left\{ N^{-1}(\lambda, \chi, \psi) : (\chi, \psi) \in \Theta_{\lambda} \right\}. \tag{20}$$

In our case the estimated parameters are $\hat{\chi} = 0.596400$, $\hat{\psi} = 0.048177$ and $\hat{\lambda} = -0.738610$. The result of the Kolmogorov-Smirnov goodness of fit test is D=0.061162 (p-value = 0.1303) and in Figures (22-25) the histogram with the fitted density, the log-histogram, the Q-Q plot and P-P plot, respectively are reported.

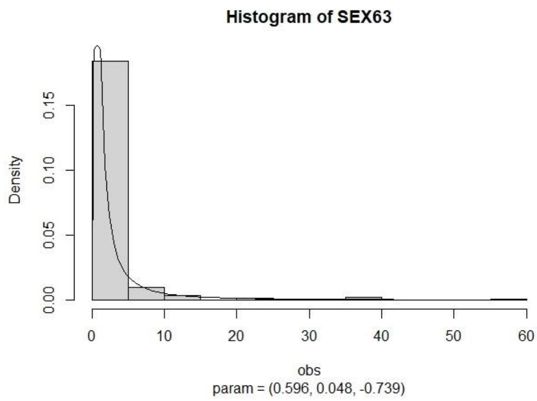

Figure 22. Histogram and fitted density of SEX63 data

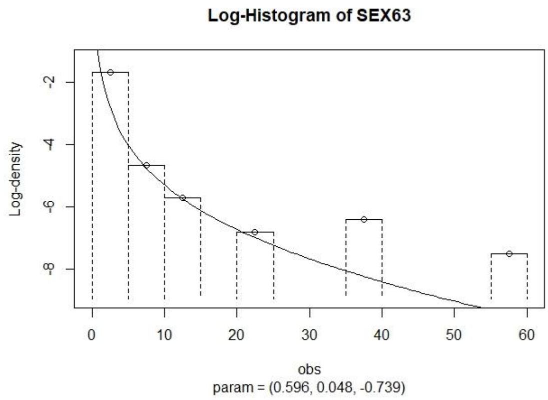

Figure 23. Log-histogram and fitted log-density of SEX63 data

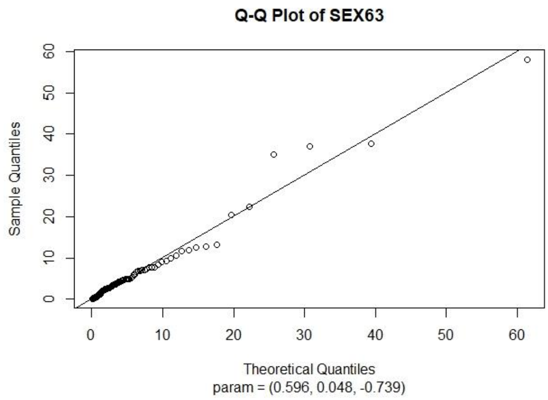

Figure 24. Q-Q plot SEX63 data.

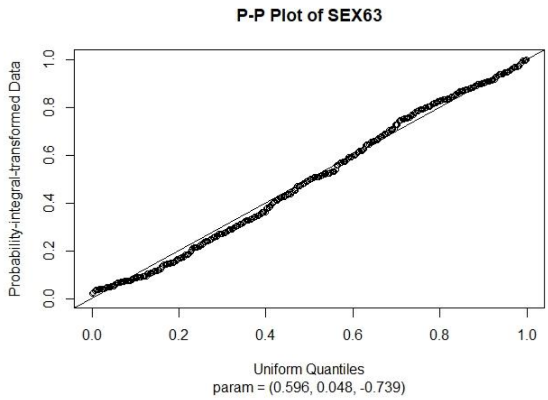

Figure 25. P-P plot of SEX63 data.

There are several measures for assessing the statistical properties of monitoring algorithms. The most commonly used measure is the average run length (ARL) (Montgomery, 2009). The run length of a control chart is a discrete random variable that is defined as the number of plotted statistics before an out-of-control point is observed on the chart. The ARL is the expected value of this random variable. Since the control charts considered are designed to achieve the same average run length in control (ARL_0) we set, as in Section 3, ARL_0=200 for all the control charts and we focus our attention on their out-of-control performance.

In detail, we consider as an initial benchmark the situation where the data come from a generalized inverse Gaussian distribution with parameters equal to the estimated ones: $\chi_0 =$0.596400, $\psi_0 =$ 0.048177 and $\lambda_0 =$-0.738610. Please note that $X_0 \sim N^{-1}\left(\lambda_0, \chi_0, \psi_0\right)$ has expected value (18) and variance (19) equal to μ_0=2.048641 and σ_0^2=30.413006 respectively.

In the first simulation study, to create scenarios as consistent as possible with those of the case under consideration for each run, of the k=10^4 replications, we generate n=365 observations: the first m

observations are generated from $X_0 \sim N^{-1}(\lambda_0, \chi_0, \psi_0)$ while the observations from m+1 to n are generated with specific changes in the parameters. In our case, without loss of any generality we set m=50. The number of observations taken from the point of change m till the alarm instant is the out-of-control run length (RL). The average of the k run lengths is our estimate of the out-of-control average run length (ARL_1).

In each run, if at the end of the n observations the control chart did not detect any alarm, we defined this event as a missed alarm. Therefore, we compute for each control chart the percentage of missed alarms obtained in the k replications. More specifically, as out-of-control cases, we considered the following scenarios.

As far as the parameter χ is concerned, we considered cases where the "in-control" parameter χ_0 had been increased by 50%, 100% and 150%, while the other parameters remain unchanged. Therefore the data from observations m+1 to n are generated from $X_1 \sim N^{-1}(\lambda_0, \chi_1, \psi_0)$ with χ_1 equal to $1.5\chi_0$, $2\chi_0$ and $2.5\chi_0$. For the sake of completeness, it is worth reminding that in these cases the expected value and variance of X_1 are:

- for $\chi_1 = 1.5\chi_0$, μ_1=2.653687 and σ_1^2=40.323163 ($\delta = (\mu_1 - \mu_0)/\sigma_0 = 0.109713$ and $\varepsilon = \sigma_1/\sigma_0 = 1.151457$);
- for $\chi_1 = 2\chi_0$, μ_1=3.181181 and σ_1^2=49.161864 ($\delta = 0.205364$ and $\varepsilon = 1.271407$);
- for $\chi_1 = 2.5\chi_0$, μ_1=3.658311 and σ_1^2=57.262931 ($\delta = 0.291882$ and $\varepsilon = 1.372167$), respectively.

For the parameter ψ we considered cases where the "in-control" parameter $\psi_0 = 0.048177$ had been decreased by 25%, 50%, 75%, while the other parameters remain unchanged. Therefore, the data from observations m+1 to n are generated from $X_1 \sim N^{-1}(\lambda_0, \chi_0, \psi_1)$ with ψ_1 equal to $0.75\psi_0$, $0.5\psi_0$ and $0.25\psi_0$. In these cases, the expected value and variance of X_1 are:

- for $\psi_1 = 0.75\psi_0$, μ_1=2.268728 and σ_1^2=44.183769 ($\delta = 0.039908$ and $\varepsilon = 1.205318$);
- for $\psi_1 = 0.50\psi_0$, μ_1=2.612452 and σ_1^2=74.631170 ($\delta = 0.102236$ and $\varepsilon = 1.566500$);
- for $\psi_1 = 0.25\psi_0$, μ_1=3.301767 and σ_1^2=181.930436 ($\delta = 0.227223$ and $\varepsilon = 2.445812$), respectively.

For the parameter λ we considered cases where the "in-control" parameter $\lambda_0 = -0.738610$ had been increased by 25%, 50% and 75%, while the other parameters remain unchanged. Therefore, the data from observations m+1 to n are generated from $X_1 \sim N^{-1}(\lambda_1, \chi_0, \psi_0)$ with λ_1 equal to $0.75\lambda_0, 0.5\lambda_0$ and $0.25\lambda_0$. In these cases, the expected value and variance of X_1 are:

- for $\lambda_1 = 0.75\lambda_0$, μ_1=3.109226 and σ_1^2=60.285624 ($\delta = 0.192316$ and $\varepsilon = 1.407917$);
- for $\lambda_1 = 0.50\lambda_0$, μ_1=4.742899 and σ_1^2=114.065867 ($\delta = 0.488551$ and $\varepsilon = 1.936637$);
- for $\lambda_1 = 0.25\lambda_0$, μ_1=7.147671 and σ_1^2=203.226492 ($\delta = 0.924609$ and $\varepsilon = 2.585000$), respectively.

The simulation results are summarized in Tables 3-5 where for each control chart and each simulated scenario the following quantities are reported: %missAL, the percentage of missed alarms; ARL$_1$, the estimated average length; SDRL, the standard deviation of the RL.

	RS/P (location)	**RS/P (scale)**	**MW**	**CvM**	**KS**
$X_1 \sim N^{-1}(\lambda_0, \chi_1, \psi_0)$ with $\chi_1 = 1.5\chi_0$ ($\delta = 0.109713, \varepsilon = 1.151457$)					
%missAL	95.9	94.7	12.8	13.5	4.0
ARL$_1$	173.4	103.0	85.7	88.1	66.4
SD	110.6	85.6	78.3	79.7	67.8
$X_1 \sim N^{-1}(\lambda_0, \chi_1, \psi_0)$ with $\chi_1 = 2\chi_0$ ($\delta = 0.205364, \varepsilon = 1.271407$)					
%missAL	95.6	94.0	3.8	4.2	0.8
ARL$_1$	163.3	107.2	50.1	52.3	38.1
SD	116.3	87.1	57.6	59.2	46.6
$X_1 \sim N^{-1}(\lambda_0, \chi_1, \psi_0)$ with $\chi_1 = 2.5\chi_0$ ($\delta = 0.291882, \varepsilon = 1.37217$)					
%missAL	95.1	93.9	0.6	0.7	0.1
ARL$_1$	156.1	95.4	29.4	30.7	23.1
SD	120.6	85.2	36.0	37.6	28.4

Table 3. Simulation results for changes in the parameter χ

	RS/P (location)	RS/P (scale)	MW	CvM	KS
$X_1 \sim N^{-1}(\lambda_0, \chi_0, \psi_1)$ with $\psi_1 = 0.75\psi_0$ ($\delta = 0.039908$, $\varepsilon = 1.205318$)					
%missAL	95.5	94.2	19.2	18.8	8.9
ARL_1	180.2	100.6	114.4	115.0	100.9
SD	110.9	84.6	85.4	85.6	80.4
$X_1 \sim N^{-1}(\lambda_0, \chi_0, \psi_1)$ with $\psi_1 = 0.50\psi_0$ ($\delta = 0.102236$, $\varepsilon = 1.566500$)					
%missAL	95.5	94.1	18.6	18.8	8.7
ARL_1	182.1	111.6	113.3	112.4	98.0
SD	111.0	85.5	86.1	85.9	81.1
$X_1 \sim N^{-1}(\lambda_0, \chi_0, \psi_1)$ with $\psi_1 = 0.25\psi_0$ ($\delta = 0.227223$, $\varepsilon = 2.445812$)					
%missAL	95.1	93.5	17.1	17.9	7.9
ARL_1	177.8	101.7	110.0	111.3	92.8
SD	111.4	86.5	84.2	84.9	77.6

Table 4. Simulation results for changes in the parameter ψ

	RS/P (location)	RS/P (scale)	MW	CvM	KS
$X_1 \sim N^{-1}(\lambda_1, \chi_0, \psi_0)$ with $\lambda_1 = 0.75\lambda_0$ ($\delta = 0.192316$, $\varepsilon = 1.407917$)					
%missAL	95.1	93.8	11.8	12.2	3.6
ARL_1	167.9	104.2	87.9	89.5	66.6
SD	112.6	84.8	79.5	80.4	67.6
$X_1 \sim N^{-1}(\lambda_1, \chi_0, \psi_0)$ with $\lambda_1 = 0.50\lambda_0$ ($\delta = 0.488551$, $\varepsilon = 1.936637$)					
%missAL	94.3	92.0	1.9	2.5	0.5
ARL_1	131.6	76.8	41.0	44.4	30.8
SD	118.2	76.6	47.8	51.9	37.4
$X_1 \sim N^{-1}(\lambda_1, \chi_0, \psi_0)$ with $\lambda_1 = 0.25\lambda_0$ ($\delta = 0.924609$, $\varepsilon = 2.585000$)					
%missAL	90.1	88.0	0.01	0.05	0.0
ARL_1	86.1	54.8	16.4	17.6	14.0
SD	107.4	67.9	16.3	18.4	14.9

Table 5. Simulation results for changes in the parameter λ

In the second simulation study we consider cases where the data from observations m+1 to n are generated as $X_1 \sim N^{-1}(\lambda_0, \chi_0, \psi_0) + \delta$ and $X_1 \sim \varepsilon \cdot N^{-1}(\lambda_0, \chi_0, \psi_0)$ with δ=1, 1.5, 2.5 and ε=1.5, 2, 2.5 respectively. The results are summarised in Tables 6 and 7.

	RS/P (location)	**RS/P (scale)**	**MW**	**CvM**	**KS**
			$\delta = 1$		
%missAL	8.0	82.6	0.0	0.0	0.0
ARL1	1.9	83.1	4.2	3.9	3.0
SD	16.3	88.7	1.1	0.9	0.7
			$\delta = 1.5$		
%missAL	0.8	88.4	0.0	0.0	0.0
ARL1	1.0	107.7	3.9	3.6	2.9
SD	0.0	89.2	0.9	0.8	0.5
			$\delta = 2.0$		
%missAL	0.1	90.1	0.0	0.0	0.0
ARL1	1.0	108.3	3.8	3.5	2.8
SD	0.0	85.3	0.8	0.7	0.5

Table 6. Simulation results for $X_1 \sim N^{-1}(\lambda_0, \chi_0, \psi_0) + \delta$

	RS/P (location)	**RS/P (scale)**	**MW**	**CvM**	**KS**
			$\varepsilon = 1.5$		
%missAL	95.5	93.9	10.7	11.4	3.1
ARL1	165.3	101.7	80.0	83.8	61.8
SD	113.9	87.1	76.4	78.3	64.5
			$\varepsilon = 2.0$		
%missAL	95.1	93.2	2.1	2.5	0.4
ARL1	171.1	95.4	42.1	45.1	33.2
SD	117.1	86.0	50.0	53.0	41.4
			$\varepsilon = 2.5$		
%missAL	93.8	91.8	0.02	0.03	0.01
ARL1	135.0	80.3	15.6	16.8	13.6
SD	120.3	84.2	14.6	17.0	14.3

Table 7. Simulation results for $X_1 \sim \varepsilon \cdot N^{-1}(\lambda_0, \chi_0, \psi_0)$

Before commenting on the results, it is important to point out that we are aware that the scenarios examined in the simulations, although quite extensive, do not include all possible cases. The aim was to reproduce scenarios consistent and comparable with the case under study in order to compare the control charts using objective criteria. Recall that extensive simulation studies on the control charts examined in this article can be found in (Capizzi and Masarotto, 2013, Hawkins and Deng 2010, Ross and Adams 2012).

With these reservations in mind, the results of the first simulation study reveal that the RS/P control chart has a very high percentage of missed alarms and the worst performance in terms of ARL_1. While comparing MW, CvM and KS control charts the best performance in terms of percentage of missed alarms and ARL_1 is obtained with the KS control chart. In the second simulation study the RS/P control chart for monitoring the location has good performance only for $\delta \geq 1.5$, while for all the remaining cases the KS control chart has the best performance.

It is important to note that it would be appropriate to further investigate the reasons for such a poor performance of the RS/P control chart. For example, by extending the simulations to scenarios with multiple change-points where the RS/P control chart may have better performance. However, such a study is beyond the scope of the present research, but we reserve this issue for future investigations. For completeness, the *R* code for reproducing the simulation study for the case concerning the change in parameter χ is provided as supplementary material (SimSveziaGIGA.R and ElabResults1.R).

7. Discussion and conclusions

Before discussing the results further some issues need to be addressed.

An important question involves post-signal diagnostic of the CvM and KS control charts, because once detected a change it could be interesting determine whether a change represents a shift in location or a shift in scale. To carry out this post-signal analysis the authors (Ross et al. 2011 and Ross and Adams 2012) proposed the following procedure: first perform a two-sample non-parametric test on the subsets of data (before and after the change) to test whether they have the same location; second, perform a different two-sample non-parametric test for testing whether the two samples have equal scale parameters. If the location test gives a lower *p*-value, then it can be concluded that a change in location has occurred, while if the scale test gives a lower *p*-value, then it can be concluded that a scale shift has occurred. Ross and Adams (2012) proposed the Mann-Whitney test for the location parameter and the Ansari-Bradley test (Gibbons, 1985) for the scale parameter. The authors assessed the procedure using an extensive simulation study and concluded that it gives good performance since it correctly diagnoses the changes.

Secondly, it is worth remembering that the CvM and KS control charts are also able detect arbitrary changes to the process distribution. In this case, post-signal diagnostics may be difficult. However, any change in the distribution is likely to cause a shift in the first or second moment and so can be detected. Therefore, in situations where post-signal diagnosis is considered important, once a change has been signalled the above-described procedure can be applied in order to figure out if the shift has mostly involved the level or the variability (Ross et al., 2011 and Ross and Adams, 2012).
As an example, in Table 8 results are reported, obtained using the procedure described above, for the Station MNX45 and the change points estimated by the KS control chart (Figure 21). More precisely, in Table 8 the compared periods and the *p*-values obtained with the Mann-Whitney test for the location and Ansari-Bradley test for the scale are reported, respectively.

Compared periods (observations)	***p*-value (location)**	***p*-value (scale)**
1-76 *vs* 77-143	$1.845 \cdot 10^{-6}$	$9.944 \cdot 10^{-2}$
77-143 *vs* 144-161	$5.771 \cdot 10^{-4}$	$3.657 \cdot 10^{-1}$
144-161 *vs* 162-235	$1.013 \cdot 10^{-1}$	$5.170 \cdot 10^{-2}$
162-235 *vs* 236-241	$3.648 \cdot 10^{-2}$	$1.200 \cdot 10^{-2}$
236-241 *vs* 242-259	$1.136 \cdot 10^{-2}$	$4.019 \cdot 10^{-3}$
242-259 *vs* 260-299	$1.105 \cdot 10^{-5}$	$6.139 \cdot 10^{-1}$
260-299 *vs* 300-323	$3.503 \cdot 10^{-2}$	$3.526 \cdot 10^{-1}$
300-323 *vs* 324-330	$8.829 \cdot 10^{-3}$	$5.330 \cdot 10^{-1}$

Table 8. Post-signal analysis for Station MNX45 and change points estimated by KS control chart

Therefore, with the results reported in Table 8 an appropriate post-signal analysis can be performed. Having addressed these important issues, we now turn to discussing the results.
The aim of this work is to evaluate whether Change Point Model Control Charts could be useful for supplementing the interquartile filter method and for supporting the activity of the National Data Centres with an inferential approach. It should be remembered that the important activity of NDCs is to identify anomalous concentrations and determine whether they are attributable to nuclear tests or civil sources.
Fortunately, nuclear tests are rare and are usually framed in geopolitical contexts in which it is relatively easy to know if they have been carried out. The most frequent case is that of emissions from civil sources and these are more difficult to classify.

For performing their tasks, the National Data Centres should cross-check the anomalous radioactivity measurements with information on the activities of the civil sources operating in the area of the monitoring station in order to make a proper classification: i.e., to establish whether the anomalous values can be related to a specific civil source or not.

This cross-checking activity should be carried out within a short time frame and is very demanding in terms of resources employed. Excluding the RS/P control chart, due to its poor performance, from the results summarised in Table 9 it can be noted that at least 50% of the values above L_A are preceded by a change point estimated by the CPM Control Charts.

	Station SEX63	**Station MNX45**
CPM Control Chart	**Precursor Coincidence Rate**	**Precursor Coincidence Rate**
MW	$r_p(\Delta T, \tau) = 0.61$	$r_p(\Delta T, \tau) = 0.56$
CvM	$r_p(\Delta T, \tau) = 0.55$	$r_p(\Delta T, \tau) = 0.56$
KS	$r_p(\Delta T, \tau) = 0.51$	$r_p(\Delta T, \tau) = 0.52$

Table 9: Summary of the results for the two monitoring stations.

Therefore, CPM Control Charts can give a valuable contribution to the activities of the NDCs by strengthening the discrimination efforts because, on an inferential basis, they allow priorities to be established. Based on the outcomes of CPM control charts, the NDCs can give precedence to anomalous values preceded by an estimated change point.

All things considered, we believe that the Kolmogorov-Smirnov control chart may provide the better support to the NDCs, since it has the best results in terms of percentage of missed alarm and ARL_1. Furthermore, using the post-signal procedure described in this Section detailed results for location and scale can be obtained.

Commenting on the results, it is also noteworthy to consider the relevant issue introduced in Section 3. The two monitoring stations represent very different scenarios. The measurement technologies are different. The geo-economical contexts in which the monitoring stations are located are different. The average magnitude of the monitored phenomenon, the variability (Tables 1 and 2) and the shape of the distributions show evident differences (Figures 3 and 4). As an example, SEX63 has an estimated skewness equal to 9.9 while for MNX45 the skewness is 5.2. It should also be noted that SEX63 had a reliable operation providing 73.34% of valid data, while for the MNX45 only 23.39% of data have been validated and are available for the analyses.

In our opinion, carrying out these analyses for such heterogenous monitoring stations, not limited to "ideal" situations, can guarantee an appreciable level of robustness in the results obtained - in the sense that these methodologies might provide useful results for the activities of NDCs, even if used on other monitoring stations.

The current research can be extended in several directions. As an example, as far as the issue of missing data is concerned, it is worth mentioning that in the procedure of the IDC the missing data are ignored and the abnormal limit is computed on the last 365 available data. To allow fair comparisons with the inter-quartile filter method the CPM control charts have been applied in the same way: using only the available data. We are aware that this aspect is very important and that it should be adequately investigated. For example, by studying both the effect of missing data on the autocorrelation structure of observations, and how often missing data occur in proximity to anomalous concentrations. This is because some anomalous readings may be caused by instrument malfunctions and not by shifts in the phenomenon being monitored. Furthermore, spatial models that consider observations not only at the single monitoring station, but also at nearby stations, may provide insights into the interconnectedness of the stations, in such a way as to improve the monitoring ability of the control charts. In this study we have limited ourselves to considering non-parametric control charts. However, other methodologies could also be of interest. As an example, spike detection in time series, which has attracted the increasing interest of researchers in many fields (Rey et al., 2015, Rafajłowicz and Rafajłowicz 2018), should be adequately investigated both in connection to our problem and for possible detection of malfunctions in the measurement devices. These aspects remain the subject of further research.

Acknowledgments

This work is the result of a collaboration between the Department of Statistical Sciences, University of Bologna and the National Data Centre - Radionuclides (NDC-RN) of ENEA, Bologna Research Centre. The data used in this work were obtained within the framework of a two-year contract known as “virtual Data Exploitation Centre” (vDEC), signed with the Preparatory Commission of the CTBTO. The views expressed herein are those of the authors and do not necessarily reflect the views of the CTBTO Preparatory Commission. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors are grateful to the Editors and the Reviewers for very useful comments and suggestions.

References

Capizzi, G. & Masarotto, G. (2013). Phase I Distribution-Free Analysis of Univariate Data. *Journal of Quality Technology*, 45(3), 273-284. https://doi.org/10.1080/00224065.2013.11917938.

Capizzi, G. (2015). Recent Advances in Process Monitoring: Nonparametric and Variable-Selection Methods for Phase I and Phase II. *Quality Engineering*, 27(1), 44-67. https://doi.org/10.1080/08982112.2015.968046.

Capizzi, G. & Masarotto, G. (2017). dfphase1: Phase I Control Charts (with Emphasis on Distribution-Free Methods). R package version 1.1.1. https://CRAN.R-project.org/package=dfphase1.

Chakraborti, S. (2011). *Nonparametric (distribution free) quality control charts*. In F. Ruggeri, R.S. Kennet, W.F. Frederick (Eds.), *Encyclopedia of Statistics in Quality and Reliability* (pp. 1–27). New York: Wiley. https://doi.org/10.1002/9781118445112.stat02699.

Chakraborti, S. & Graham, M.A. (2019). Nonparametric (distribution-free) control charts: An updated overview and some results. *Quality Engineering*, 31(4), 523-544. https://doi.org/10.1080/08982112.2018.1549330.

CTBT (1996). Comprehensive Nuclear Test-ban Treaty. Comprehensive Nuclear-Test-Ban Treaty Organization New York. Retrieved from (https://www.ctbto.org/fileadmin/content/treaty/treaty_text.pdf).

Donges, J.F. Schleussner C.-F. Siegmund J.F & R.V. Donner (2016). Event coincidence analysis for quantifying statistical interrelationships between event time series. *The European Physical Journal Special Topics*, 225(3), 471-487. https://doi.org/10.1140/epjst/e2015-50233-y.

Erfanian, M. Sadeghpour Gildeh, B. & Reza Azarpazhooh, M. (2021). A new approach for monitoring healthcare performance using generalized additive profiles. *Journal of Statistical Computation and Simulation*, 91(1), 167-179. https://doi.org/10.1080/00949655.2020.1807981.

Fontaine J.P., Pointurier, F., Blanchard, X. & Taffary T. (2004) Atmospheric xenon radioactive isotope monitoring. *Journal of Environmental Radioactivity*, 72(1-2), 129–135. https://doi.org/10.1016/S0265-931X(03)00194-2

Gibbons, J. D. (1985). *Nonparametric Statistical Inference*. Boca Raton: McGraw-Hill.

Good, P.I. (2005). *Permutation, Parametric, and Bootstrap Tests of Hypotheses* (3rd ed.). New York, Springer. https://doi.org/10.1007/b138696.

Hawkins, D.M. & Deng Q (2010). *A Nonparametric Change-Point Control Chart,* Journal of Quality Technology, 42(2) 165-173. https://doi.org/10.1080/00224065.2010.11917814

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, 6(2), 65-70. https://doi.org/10.2307/4615733.

Jones-Farmer, L. A., Jordan, V., & Champ, C. W. (2009). Distribution-free Phase I control charts for subgroup location. *Journal of Quality Technology, 41(3)*, 304–316. https://doi.org/10.1080/00224065.2009.11917784

Jørgensen, B. (1982). *Statistical Properties of the Generalized Inverse Gaussian Distribution*. In Lecture Notes in Statistics. Vol. 9. New York, Berlin: Springer-Verlag. ISBN 0-387-90665-7, DOI: 10.1007/978-1-4612-5698-4.

Kalinowski, M, Axelsson, A., Bean, M., Blanchard, X, Bowyer, T., Brachet, G., Hebel, S. McIntyre, J., Peters, J., Pistner, C., Raith, M., Ringbom, A., Saey, P., Schlosser, C., Stocki, T., Taffary, T. & Ungiar, K. (2010). Discrimination of Nuclear Explosions against Civilian Sources Based on Atmospheric Xenon Isotopic Activity Ratios. *Pure Applied Geophysics*, 167(4-5), 517-539, https://doi.org/10.1007/s00024-009-0032-1.

Lehmann, E. L., & Romano, J. P. (2005). *Testing statistical hypotheses* (3rd ed.). New York: Springer. https://doi.org/10.1007/0-387-27605-X.

Montgomery, D. (2009) *Introduction to statistical quality control*, 6th edn. Hoboken: Wiley.

Niu, Y. S., Hao, N., & Zhang, H. (2016). Multiple change-point detection: a selective overview. *Statistical Science*, 31(4), 611-623. https://doi.org/10.1214/16-STS587.

Pesarin, F. (2001). Multivariate Permutation Tests: With Applications in Biostatistic. Wiley, New York.

Paroissin, C., Penalva, L. Pétrau, A & Verdiera, G. (2016). New control chart for monitoring and classification of environmental data. *Environmetrics*, 27, 182-193. https://doi.org/ 10.1002/env.2382.

Qiu, P & Li, Z. (2011). On Nonparametric Statistical Process Control of Univariate Processes. *Technometrics*, 53 (4), 390-405. https://doi.org/10.1198/TECH.2011.10005

Qiu, P. (2013). *Introduction to statistical process control*. Boca Raton: Chapman & Hall/CRC Press.

Rafajłowicz, W. & Rafajłowicz E. (2018). A rule-based method of spike detection and suppression and its application in a control system for additive manufacturing. *Applied Stochastic Models in Business and Industry*. 34 (5), 645-658. https://doi.org/10.1002/asmb.2298.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rey, G.H, Pedreira, C. & Quiroga, R.Q. (2015). Past, present and future of spike sorting techniques. Brain Research Bulletin, 119(Part B), 106-117. https://doi.org/10.1016/j.brainresbull.2015.04.007.

Rigby, R. A., Stasinopoulos, D. M., Heller, G. Z., & De Bastiani, F. (2019). *Distributions for modeling location, scale, and shape: Using GAMLSS in R*. Chapman and Hall/CRC: New York, https://doi.org/10.1201/9780429298547, ISBN: 9780429298547.

Ringbom, A., Larson, T. Axelsson, A., Elmgren, K. & Johansson, C. (2003). SAUNA-a system for automatic sampling, processing, and analysis of radioactive xenon. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 508(3), 542–553. https ://doi.org/10.1016/S0168 -9002(03)01657-7.

Rodbard, D., Rayford, P.L., Cooper, J.A. & Ross, G.T. (1968). Statistical quality control of radioimmunoassays. *The Journal of Clinical Endocrinology & Metabolism* 28(10), 1412-1418. https://doi.org/10.1210/jcem-28-10-1412.

Ross G.J. & Adams N.M. (2012). *Two Nonparametric Control Charts for Detecting Arbitrary Distribution Changes*. Journal of Quality Technology, 44(12), 102-116. https://doi.org/10.1080/00224065.2012.11917887.

Ross, G. J. (2015) Parametric and Nonparametric Sequential Change Detection in R: The cpm Package. Journal of Statistical Software 66(3), 1-20. https://doi.org/10.18637/jss.v066.i03.

Scott, D. (2018). GeneralizedHyperbolic: The Generalized Hyperbolic Distribution. *R* package version 0.8-4. https://CRAN.R-project.org/package=GeneralizedHyperbolic

Siegmund, J.F., Siegmund, N. & Donner R.V. (2017). CoinCalc-A new R package for quantifying simultaneities of event series. *Computers & Geosciences* 98, 64-72. http://dx.doi.org/10.1016/j.cageo.2016.10.004.

Stasinopoulos D.M. & Rigby R.A. (2008) Generalized additive models for location scale and shape (GAMLSS) in R. *Journal of Statistical Software*, 23(7), 1-46. https://doi.org/10.18637/jss.v023.i07

Suzuki, A., Murakami, H. & Mukherjee, A. (2021). Distribution-free Phase-I scheme for location, scale and skewness shifts with an application in monitoring customers'waiting time. Journal of Applied Statistics (Latest Articles), 1-21. https://doi.org/10.1080/02664763.2021.1994530

Zou, C., Yin, G., Feng, L. & Wang, Z. (2014). Nonparametric Maximum Likelihood Approach, to Multiple Change-Point Problems. *The Annals of Statistics*, 42(3), 970-1002. https://doi.org/10.1214/14-AOS1210

United Nations & United Nations. (1998). *CTBT, Comprehensive Nuclear Test-Ban Treaty*. New York, N.Y.: United Nations, Dept. for Disarmament Affairs and Dept. of Public Information.