Anomaly Detection in Activities of Daily Living with Linear Drift

Óscar Belmonte-Fernández · Antonio Caballer-Miedes · Eris Chinellato · Raúl Montoliu · Emilio Sansano-Sansano · Rubén García-Vidal

Received: 31 October 2018 / Accepted: 10 June 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

Anomaly detection in Activities of Daily Living (ADL) plays an important role in e-health applications. An abrupt change in the ADL performed by a subject might indicate that she/he needs some help. Another important issue related with e-health applications is the case where the change in ADL undergoes a linear drift, which occurs in cognitive decline, Alzheimer’s disease or dementia. This work presents a novel method for detecting a linear drift in ADL modelled as circular normal distributions. The method is based on techniques commonly used in Statistical Process Control and, through the selection of a convenient threshold, is able to detect and estimate the change point in time when a linear drift started. Public datasets have been used to assess whether ADL can be modelled by a mixture of circular normal distributions. Exhaustive experimentation was performed on simulated data to assess the validity of the change detection algorithm, the results showing that the difference between the real change point and the estimated change point was 4.90 ± 3.17 days on average.

ADL can be modelled using a mixture of circular normal distributions. A new method to detect anomalies following a linear drift is presented. Exhaustive experiments showed that this method is able to estimate the change point in time for processes following a linear drift.

Keywords Anomaly detection · Activities of Daily Living · Abrupt change · Linear drift · Circular normal distribution

Introduction

Anomaly detection in ADL is a particularly important issue in e-health applications such as monitoring the older adults [1–3], monitoring chronic diseases [4] and detection of depression [5], to cite but a few. In this kind of application, the subject is continuously monitored, commonly by means of a sensor network. Each sensor in the network generates a stream of data, which in turn is used to detect anomalies. The monitoring system, in these cases, is able to detect an abrupt change in the data stream; for example, an anomalous reading from the accelerometer might signal a fall [6]. Moreover, anomaly detection [7] plays a central role in areas such as intrusion/fraud detection [8, 9], quality control in industry [10], public health outbreaks [11], and climate change [12].

Statistical Process Control (SPC) [13], applied in industry for quality control, has been faced with the issue of abrupt change detection for a long time. The statistical approach is well suited when monitoring temporal data which follow a known probability distribution function (pdf) and the parameters of the pdf are also known before and after the change, which is the case of the cumulative sum algorithm, also known as the CUSUM algorithm [14, 15].

In contrast to abrupt changes, progressive changes, or changes with some trend in the data stream, are more difficult to detect. In the case of ADL, the time when some activities start or end, for example the time of going to sleep and leaving or entering home, could be considered a random variable, and hence this random variable could be suitably modelled following a circular normal distribution (see “Activity Modelling Using the Circular Normal Distribution”). Most illnesses that are more specific to the older adults, like cognitive decline, Alzheimer’s disease (AD), dementia or functional impairment in general, could follow a progressive change at some stage of the illness or during the whole illness process [16–19].
Due to the importance of early detection of cognitive decline in the older adults [20–22] and to the rate at which the world’s population is ageing [23, 24], developing techniques for the early detection of changes in ADL and estimating the point in time when the change took place are of great importance. This, in turn, can be integrated in existing e-health systems to detect changes in continuously monitored older adults [25]. In this kind of system the older adults are monitored on a continuous basis by means of devices attached to the user’s body and the data thus collected are sent using wireless communication to a server for later analysis. Analysed data can provide insights into physical aspects of health, and also psychological aspects of health such as loneliness, social isolation, and satisfaction with quality of life [26, 27].

Our first hypothesis is that periodic ADL can be modelled using circular normal distribution functions. Our second hypothesis is that abrupt changes in periodic ADL modelled by circular normal distribution can be detected and the change point in time can be estimated by means of the methods used in SPC. Our third hypothesis is that the methods applied for abrupt change detection can be extended to detect changes in a circular normal distribution which at some point in time starts to follow a linear drift, and this result can be applied to detect a linear drift in ADL.

Both cases are quite important when monitoring ADL performed by the older adults. Abrupt change detection, when applied to ADL, might indicate that a change in the medication taken by an older adult has had some effect on his/her behaviour [28]. The detection of a linear trend in ADL might indicate the onset of some cognitive decline, for example Alzheimer’s disease [29, 30].

Detection of changes in ADL is of special interest in the early detection of cognitive decline in older adults [31]. There is evidence that circadian rhythm disruption is a symptom of neurodegeneration, which in turn changes ADL patterns [32–34]. Recently, it has been suggested that there is a relationship between sleep characteristics and cognitive decline in the older adults, emphasizing the fact that sleep and cognition are closely related [35–41]. In [42] the authors present a 1-year study conducted with 77 older adults (21 mild AD, 27 moderate to severe AD, and 29 control subjects). The results show that patients with moderate to severe AD have earlier bedtimes (9:45, $\sigma = 1 \text{ h 31 min}$) compared with patients with mild AD (10:45, $\sigma = 1 \text{ h 36 min}$) and controls (11:30, $\sigma = 58 \text{ min}$). No differences regarding the wake-up times were reported among moderate to severe AD (7:15, $\sigma = 55 \text{ min}$), mild AD (6:55, $\sigma = 54 \text{ min}$), and control subjects (7:00, $\sigma = 47 \text{ min}$).

Furthermore, some works find a positive correlation between the severity of the disorder and the severity of the cognitive impairment [29–31, 43]. An early study by Stem, R. G. et al. [44] on 183 patients (111 with AD and 72 non-AD) over a period of 90 months showed a positive linear correlation between the scores of the AD patients on the Cognitive Subscale of the Alzheimer’s Disease Assessment Scale and the assessment time within 24 months. In addition, this study also showed that the rating score followed a quadratic correlation with the score at the time of assessment.

A 1-year study involving 53 elderly people living alone, and using infrared sensors to document their outdoors and indoors activities, showed that subjects with a decline in their cognitive function presented fewer number of outings and spent less hours outside home than subjects that showed no cognitive decline [45]. Another study involving 85 independent older adults who live alone indicated that a greater number of hours spent outside the home was associated with better cognitive functions, as measured by the Clinical Dementia Rating [46]. Similar results were reported in [47].

Cognitive-inspired techniques play an important role in Ambient Assisted Living to develop new systems which improve the quality of life of citizens in general and the older adults in particular [48–50]. Recognition of Activities of Daily Living is a key aspect in such systems and is mainly performed by applying machine learning techniques [51] to data provided by an in-home sensor network [25]. Other novel techniques are based on analysing the speech and emotional temperature [52] or performance in handwriting/drawing tasks on digitizing tablets [53].

The contributions of our work are the following:

1. We model Activities of Daily Living (ADL) by means of circular normal probability distributions, and we assess its validity by performing experiments on two public datasets.
2. We extend the CUSUM algorithm in order to detect a linear trend in a process following a circular normal distribution, and we use MLE to estimate the change point.
3. We perform extensive simulations in order to assess the validity of the presented approach in two cases: abrupt change and linear drift change.

The rest of the paper is organized as follows: Section “Background” presents the background in the fields of SPC, Human Activity Recognition (HAR), Circular Statistics, and Anomaly Detection in ADL. Section “Methodology” shows how to model ADL by means of a circular normal distribution; this section also presents the CUSUM algorithm and how it is applied to detect abrupt changes in a process following a circular normal distribution. A contribution extending the CUSUM algorithm to detect linear trends is then presented and applied to the case of a process following a circular normal distribution. Section “Results” describes an extensive set of experiments.
performed to assess the validity of the presented method. Finally, “Conclusion” presents the main conclusions.

**Background**

In this section, first SPC is introduced, as the method presented in this paper relies on SPC techniques. Then, HAR is presented as the base problem of identifying ADL. Finally, some related works in the field of anomaly detection are reviewed.

**Statistical Process Control**

SPC is an important field for quality control in industrial processes. Since its introduction by Shewhart [54], SPC control charts have become one of the most used tools for change detection in industrial processes. Other extensively and successfully used control charts are Cumulative Sum (CUSUM), introduced by Page [14, 15], and the Exponential Weighted Moving Average (EWMA), introduced by Roberts [55]. In general, the main objective of control charts is to detect, as soon as possible, abrupt changes in monitored processes [56].

After detecting an abrupt change, the next step is to estimate the time when the change occurred. Different approaches have been presented for change point estimation, such as neural networks [57–59], fuzzy sets [51, 60–63], and the bayesian approach [64, 65]. Nevertheless, most existing approaches are based on MLE [66–68].

In most cases, the detection algorithm and estimation method assume an abrupt change in the process. However, it may be important to consider a more gradual change, such as a linear trend disturbance [66–68]. The results in these last cases show that the performance and accuracy of the abrupt change methods are outperformed by those methods specifically designed for linear drift anomaly detection [66].

**Human Activity Recognition**

The goal of Human Activity Recognition (HAR) is to determine the set of activities performed by a person in a time interval, based on a number of measured attributes. Formally, given a set of activity labels $A = \{A_1, ..., A_t\}$, a time interval $T = [t_{start}, t_{end}]$ and a set of $j$ time series, each from a particular measured attribute $S = \{S_1, ..., S_j\}$, the goal is to determine the sequence of non overlapping activities $A = \{A_{tk1,tk2}, A_{tk2,tk3}, ..., A_{tkj,tkj+1}\}$ based on the data concerning $S$ [69]. Some typical activities are: getting out of bed, eating, watching TV, bathing and going to bed, to cite just a few.

The measured attributes can be image recordings [70–73], binary sensors attached to doors, drawers and windows in the case of external sensors, and accelerometers, gyroscopes, temperature or heart rate sensors worn by the human being [69, 70].

In this work, it is assumed that the set of non overlapping activities and the time interval for each of them $A = \{A_{tk1,tk2}, A_{tk2,tk3}, ..., A_{tkj,tkj+1}\}$ are already known. The goal of this paper is to detect changes in routine activities such as going to bed, based on their start or end times, since such changes in a person’s behaviour can indicate a decline in health.

**Circular Statistics**

Directional statistics deals with data that represent directions. In the case of two dimensions, directional statistics is called circular statistics. A particular case is periodic data, for example, the arrival times of patients at a hospital [74], where each time on the clock face represents a direction.

Circular statistics has been used successfully for modelling human activity recognition [75, 76], time patterns in crime analysis [77], and analysis of social interactions and mood [78]. In all the cited cases there is a clear periodicity of the studied phenomenon regarding time; this periodicity might be a day, a week, or even a year or longer periods of time.

Different circular pdfs have been used to model directional data. The circular Gumbel distribution was used in [79] to model ADL. Also, the Wrapped Normal distributions has been used in [74, 80] to model ADL. The circular normal distribution, also known as von Mises distribution, has been extensively used to model directional data [75–78].

The point of view of authors in [81] is that although circular data is being collected in different areas of cognitive and experimental psychology, analysis methods for circular for circular data are not common in psychology and the social sciences in general. The authors present two cases where the use of circular statistics is well suited: the human motor resonance [82], and the cognitive maps data [83].

**Anomaly Detection**

Following [84], we define anomalous activities as events with the following properties: a) they occur rarely, and b) they were not expected in advance.

These properties highlight two facts: (a) it is difficult to find HAR databases containing anomalous activities; (b) it is also difficult to know the parameters that define an anomalous activity.

Several techniques have been used to detect anomalous activities. A two-phase pipeline process is used in [85]. In the first phase, normal activities are filtered out using a one-class Support Vector Machine with the aim of reducing
Logic inference is used to identify anomalous behaviour. The second phase uses an iterative procedure in which a Hidden Markov Model (HMM) of the initial activities is first created, and each new activity instance is then checked against the current model. Every new activity is thus classified using the trained HMM, and if its likelihood is below a given threshold, it is considered anomalous and a new class is created for it.

A variant of the Hamming distance is used in [79] to measure the dissimilarity between two circular sequences of data. These sequences model the ADL of the older adults performed at home as circular Gumbel distributions. The data were acquired over a period of 6 months using a network of PIR motion sensors installed in an 80-year-old woman’s flat.

Probabilistic models including Naïve Bayes, Hidden Markov Model, Hidden Semi-Markov Models and Conditional Random Fields are compared in [73] using three datasets acquired in real scenarios. Data are provided by several sensors such as reed switches, pressure mats, mercury contacts, passive infrared and float sensors [86]. The results show that an increase in model complexity also improves the results.

The work in [87] also uses an HMM to detect anomalous activities, exponential smoothing on histories of time series is then used to predict future changes, and finally both outcomes are fused to detect anomalous activities. In the former two works, an activity that might be classified as anomalous could consist of two consecutive activities in the time series being exchanged, but any abrupt change in the start or end time of a single activity would not be detected as an anomalous behaviour. Conversely, the work presented in this paper is able to detect anomalous behaviours consisting in a delay in the start time of an activity.

A probabilistic anomalous detection method based on dynamic Bayesian networks is presented in [88]. Information provided by wearable sensors and the location context is used to detect anomalous behaviours in spatial elements, timing, duration, and sequences of ADL.

A two-layered extension of the hidden semi-Markov model is used in [89] to model ADL based on the data provided by a set of cameras. This model, in turn, is used to detect anomalies in the duration of activities.

Machine Learning techniques are commonly used to detect anomalous behaviours. In particular [90] and [91] use Support Vector Machines (SVM) to build models based on normal behaviour data to subsequently try to detect anomalous behaviours.

The authors in [92] use a semantic approximation to detect anomalous behaviours in mild cognitive impairment. With the help of cognitive neuroscience researchers, both normal and anomalous activities are represented using a Web Ontology Language version 2 (OWL 2) ontology. Logic inference is used to identify anomalous behaviour.

Infrared sensors and magnetic contacts in doorways are used in [93] to locate a user in order to measure her circadian activity. After registering long periods of activity, a probabilistic model is built. Activities with high entropy are thus considered to be anomalous behaviours.

Information provided by binary sensors and their activation times are used in [94] to create regularity and duration scores. The authors use density-based spatial clustering with noise on the scores to cluster activities. The cluster with the largest number of instances represents normal behaviour.

**Methodology**

Circular distribution is introduced first in this section to model periodic events. Then abrupt change detection using the CUSUM algorithm is presented. After that, the main contribution of this work is presented, that is, a novel method to detect a linear trend in a circular normal distribution.

**Activity Modelling Using the Circular Normal Distribution**

Circular normal distribution is a symmetric unimodal distribution describing circular data of a periodic nature. This kind of data arise in a number of fields such as biology, geography, geology, geophysics, medicine, meteorology and oceanography. Examples of such data include directions of wind and ocean currents, directional movement of animals in response to stimuli, and biorhythms [93, 95].

The probability density function of the circular normal distribution for a given angle $\phi$ is given by:

$$f(\phi | \mu, \kappa) = \frac{e^{\kappa \cos(\phi - \mu)}}{2\pi I_0(\kappa)}$$

where $\forall \phi \in \mathbb{R}, \mu \in [0, 2\pi)$ and $\kappa \geq 0$. The $\mu$ parameter is known as the mean direction parameter, the $\kappa$ parameter is known as a concentration parameter, and $I_0(\kappa)$ is the modified Bessel function of order 0 defined by:

$$I_0(\kappa) = \frac{1}{2\pi} \int_{0}^{2\pi} e^{\kappa \cos \varphi} d\varphi$$

The variance of the circular distribution is given by:

$$\sigma^2 = 1 - \frac{I_1(\kappa)}{I_0(\kappa)}$$

(1)

If the value of the mean direction $\mu \in [0, 2\pi)$ is taken as time units in the range $[0, 24)$ hours, the equivalence between the concentration parameter $\kappa$ and the square root of the variance $\sigma$ can be extracted using Eq. 1. Table 1 shows some equivalent values used in the experiments presented in “Results”. In the remainder of this work the parameter
Table 1 Equivalences between some concentration parameters $\kappa$ and the square root of the variance for a circular distribution

| $\kappa$ | $\sigma$ (in min) |
|----------|------------------|
| 33       | 40               |
| 58       | 30               |
| 84.5     | 25               |
| 131      | 20               |
| 233      | 15               |
| 517      | 10               |

In minutes, $\sigma$ will be used instead of $\kappa$ because it can be interpreted as the dispersion of the data, for example, the dispersion for a person’s wake-up time. Figure 1 shows some circular normal probability distribution function plots. Note that parameter $\kappa$ is related to the dispersion of the data: the larger the $\kappa$ is, the lower the dispersion of the data will be.

For a set of samples $\{\phi_i\}$ with $i \in [1, N]$, the MLE for $\mu$ and $\kappa$ are:

$\hat{\mu} = \tan^{-1} \left( \frac{\sum_{i=1}^{N} \sin \phi_i}{\sum_{i=1}^{N} \cos \phi_i} \right)$

$\hat{\kappa} = \frac{1}{n} \left[ \left( \sum_{i=1}^{N} \cos \phi_i \right)^2 + \left( \sum_{i=1}^{N} \sin \phi_i \right)^2 \right]$

These two expressions can be used to estimate the parameters $\mu$ and $\kappa$ from a set of data.

Two public datasets from the CASAS project were employed to assess the validity of using the circular normal probability distribution for modelling ADL. The first dataset is called Milan [96]. This dataset is annotated with ADL and was acquired between October 16 2009 and January 6 2010 by an older adult living alone in a flat. The second dataset is called Aruba [97]. It is annotated with ADL and was acquired between November 4 2010 and June 11 2011 by an older adult living alone in a flat.

In order to test whether ADL can be modelled by circular normal distributions, six ADL were selected from each dataset and were adjusted to a mixture of circular normal distributions using the R library movMF Hornik2014. The Kullback-Leibler divergence [98] was used to calculate the divergence between the empirical probability density function and the fitted probability density function. Results are shown in Tables 2 and 3. Figures 2 and 3 show the data for the Milan and Aruba datasets using a rose diagram; the solid lines show the empirical density from the data, and the dashed lines show the density for the fitted mixture of circular normal distributions. The radial segments in these figures correspond to the example times shown in Tables 2 and 3. The KL-divergence of all analysed data is less than 0.03, and the maximum of the lobes for the activities are close to the density of the real data. The analysis of these
distributed samples, the
by
\( p_\theta(y_i) \)

The names of the ADL are the same as those used in the public dataset. #pfs is the number of circular normal distributions used to fit the data. KL is the KL-divergence between the empirical density and the density of the fitted data. Ex: Time (\( \sigma \)) stands for an example of the time of the activity and its fitted \( \sigma \) in minutes.

| ADL                  | Samples | #pfs | KL       | Ex: Time (\( \sigma \)) |
|----------------------|---------|------|----------|-------------------------|
| Sleep begin          | 156     | 8    | 0.0252   | 22:17 (38.05)           |
| Sleep end            | 156     | 8    | 0.0120   | 08:14 (17.28)           |
| Bed to toilet begin  | 88      | 8    | 0.0072   | 02:11 (13.52)           |
| Bed to toilet end    | 88      | 8    | 0.0045   | 02:13 (13.40)           |
| Leave home begin     | 214     | 8    | 0.0145   | 17:16 (30.95)           |
| Leave home end       | 214     | 8    | 0.0059   | 17:22 (31.00)           |

The names of the ADL are the same as those used in the public dataset.

Abrupt Change Detection and Change Point Estimation

The CUSUM algorithm is a quality control method which was first proposed in [15] as a continuous inspection scheme to detect changes in a sequence of independent random variables following some known probability function.

If it is assumed that a signal can be modelled as a discrete random signal \( \mathcal{Y} \), with independent and identically distributed samples, the pdf of each sample is given by \( p_\theta(y_i) \), where \( \theta \) is a deterministic parameter. The occurrence of an event is modelled by an instantaneous change in \( \theta \), so that \( \theta = \theta_1 \) before the event at \( i = \tau \), and \( \theta = \theta_2 \) when \( i \geq \tau \). Thus, the two possible hypotheses are:

- \( H_1 : \theta = \theta_1 \), No change has occurred.
- \( H_2 : \theta = \theta_2 \), A change has occurred.

The pdf of the signal \( \mathcal{Y} \) observed between the initial sample \( y_1 \) and the current sample \( y_N \) can take two forms depending on the above hypotheses. Under the ‘no change’ hypothesis, also known as the in-control process, \( H_1 \), the pdf is:

\[
p_{\theta_1|H_1} = \prod_{i=1}^{N} p_{\theta_1}(y_i).
\]

On the other hand, under the ‘change’ hypothesis, also known as the out-of-control process, \( H_2 \), the pdf is:

\[
p_{\theta_2|H_2} = \prod_{i=1}^{\tau-1} p_{\theta_1}(y_i) \prod_{i=\tau}^{N} p_{\theta_2}(y_i).
\]

With these assumptions, the CUSUM algorithm defines two quantities, the cumulative sum:

\[
S_N^\tau = \sum_{i=\tau}^{N} \ln \frac{p_{\theta_2}(y_i)}{p_{\theta_1}(y_i)} = \sum_{i=\tau}^{N} s_i
\]

where \( s_i = \ln \frac{p_{\theta_2}(y_i)}{p_{\theta_1}(y_i)} \) is defined as the instantaneous log-likelihood, and the decision function:

\[
g_N = \max_{1 \leq \tau \leq N} S_N^\tau
\]

Given a user-defined threshold \( h \), if the decision function \( g_N < h \), the hypothesis \( H_1 \) is true and so no change has occurred. In contrast, if the decision function \( g_N > h \), the hypothesis \( H_2 \) becomes true and so it is assumed that a change has taken place. In the latter case, Maximum Likelihood can be used to find an estimate of the change point considering (3):

\[
\hat{\tau} = \arg \max_{1 \leq \tau \leq N} S_N^\tau
\]

Equations 2 and 3 can be rewritten in a recursive form:

\[
S_N^\tau = S_N^{\tau-1} + s_N \quad \text{and} \quad g_N = (g_{N-1} + s_N)^+,
\]

where \( (z)^+ = \sup(0, z) \). These recursive equations are convenient to evaluate the cumulative sum and decision function as data are produced, as in the case of ADL monitoring, and to raise an alarm in the case of an anomaly being detected. For example, an abrupt change in a monitored person’s wake-up time over several consecutive days would trigger an alarm.
Activities in the Milan dataset. Rose diagrams show the data at intervals of 10 min. The solid line shows the density of the data. The dashed line shows the density of the fitted data. The solid segment in each activity is the time for the fitted circular normal distribution shown in Table 2.

\[
s_i = \ln \frac{f(\phi_i | \mu_2, \kappa)}{f(\phi_i | \mu_1, \kappa)} = \kappa [\cos \phi_i (\cos \mu_2 - \cos \mu_1) + \sin \phi_i (\sin \mu_2 - \sin \mu_1)]
\]

An MLE for \( s_i \) before the change point \( (i < \tau) \) is proportional to:

\[
\hat{s}_i(\hat{\phi}_i = \mu_1) \propto \cos(\mu_1 - \mu_2) - 1 \leq 0 \tag{4}
\]

In the same way, an MLE for \( s_i \) after the change point \( (i \geq \tau) \) is proportional to:

\[
\hat{s}_i(\hat{\phi}_i = \mu_2) \propto 1 - \cos(\mu_1 - \mu_2) \geq 0 \tag{5}
\]

Regarding \( S_i \), it is a sum of the negative values before the change point (4) and of the positive values after the change point (5), so \( S_i \) will have a local minimum in the interval \([1, N]\).

In sum, the most likely value for the instantaneous log-likelihood before the change point is less than zero. Conversely, the most likely value for the instantaneous log-likelihood after the change point is greater than zero, and so the cumulative sum will decrease before the change point and will increase after the change point. Figure 4 shows a set of plots with an example for the cumulative sum \( S_i \) and decision function \( g_i \) for processes following a circular normal distribution, when an abrupt change was introduced at time \( \tau = 50 \).

**Linear Trend Detection and Change Point Estimation**

Let us assume an older adult living alone and continuously monitored in her own home [35, 48]. The monitoring system detects a drift in her daily routines, for example, bedtime, wake-up time and time of leaving home. Health practitioners decide to base their answer on the Mini-Mental State Examination, and the results suggest that she is starting cognitive decline. Some time later, she is diagnosed with AD and starts to take medication under medical prescription. Again, some time later the monitoring...
Fig. 3  Activities in the Aruba dataset. Rose diagrams show the data at intervals of 10 min. The \textit{solid line} shows the density of the data. The \textit{dashed line} shows the density of the fitted data. The \textit{solid segment} in each activity is the time for the fitted circular normal distribution shown in Table 3

system detects another drift in her daily routine which might be related to the medication provided by the health practitioners \cite{28}. This scenario describes a hypothetical case for the main contribution of our work, i.e. to detect such a linear drift in the data for ADL.

If a process starts to follow a linear drift at event $i_0 = \tau$ at rate $\beta$, it can be modelled as:

$$
\theta(t \leq \tau) = \theta_1 \\
\theta(t > \tau) = \theta_2 = \theta_1 + (i - \tau)\beta
$$

Note that the drift rate $\beta$ is unknown, so a heuristic should be used to detect that the process has started to follow a linear drift.

To perform this detection, the following pair of hypotheses is used:

$H_1 : \theta = \theta_1$, \textit{No change has occurred.} \\
$H_2 : \theta = \theta_1 + \delta$, \textit{A change has occurred.}$

where $H_1$ is the no change hypothesis, and $H_2$ is the change hypothesis. If $\delta > 0$, the change occurred after time $\theta$ whereas, on the contrary, if $\delta < 0$, the change occurred before time $\theta$. Without loss of generality, it will be assumed that $\delta > 0$ in the remainder of this work. Hereafter, the parameter $\delta$ will be called \textit{the detector time}.

Particularizing for the circular normal distribution, without loss of generality, it can be assumed that $\mu_1 = 0$ and, given that $\mu_2 = \delta$, the instantaneous log-likelihood in Eq. 4 becomes:

$$
\hat{s}_i = \kappa[\cos(\phi_i)(\cos(\delta) - 1) + \sin(\phi_i) \sin(\delta)]
$$

An MLE for $s_i$ before the change is proportional to:

$$
\hat{s}_i \propto \cos(\delta) - 1
$$

and so the MLE of the instantaneous log-likelihood is negative before the change, and after the change it is proportional to:

$$
\hat{s}_i \propto \cos(i\beta)(\cos(\delta) - 1) + \sin(i\beta) \sin(\delta)
$$
Fig. 4 Temporal series following a circular normal distribution (left), cumulative sum (middle), and decision function (right) for two sets of temporal data following a circular normal distribution. The change point was set to $\tau = 50$ in both experiments. The plots in the first row were generated with $\mu_1 = 0, \mu_2 = 5$ min, $\sigma = 25$ min and $h = 2.9$ (horizontal line). The plots in the second row were generated with $\mu_1 = 0, \mu_2 = 10$ min, $\sigma = 10$ min and $h = 6.2$.

The cumulative sum algorithm can be applied to each pair of consecutive hypotheses, $H_w$ and $H_{w+1}$ with $w \in [1, W]$. The change point for each pair of hypotheses can be estimated by:

$$S_{t+1, w} = \sum_{i=t}^{N} \ln \frac{p_{\theta_1+w\delta}(X_i)}{p_{\theta_1+(w-1)\delta}(X_i)}$$

$$\hat{\tau}_w = S_{1, w}^{N} - \arg \min_{1 \leq \tau \leq N} S_{\tau}^{N}_{1, w}$$  \hspace{1cm} (7)

Any time the cumulative sum $S_{t+1, w}^{N}$ reaches a minimum it means that hypothesis $H_{w+1}$ becomes valid and $H_w$ is no longer valid. This minimum is reached, for the pair of hypotheses $H_w$ and $H_{w+1}$, at $\hat{\tau}_w$, as stated by Eq. 7. As the out-of-control process evolves at rate $\beta$ over time, the minima $\hat{\tau}_w$ should evolve at the same rate, so we can write:

$$\frac{\partial \hat{\tau}_w}{\partial w} = \frac{\partial \hat{\theta}_{1, \tau > \tau}}{\partial i} = \hat{\beta}$$ \hspace{1cm} (8)

Equation 8 is the key point of our proposed algorithm, and if we are able to estimate the linear trend rate of the minima for a set of pairs of hypotheses, we will be able to measure the linear trend of the process after the change.

Simulated Data Generation

Random datasets were generated before the change point ($\tau$) using a circular normal distribution with constant mean.
Fig. 5 Temporal series following a circular normal distribution (left), cumulative sum (middle), and decision function (right) for two sets of temporal data following a circular normal distribution. The change point was set to $\tau = 50$ in both experiments. The plots in the first row were generated with $\sigma = 25$ min, $\beta = 0.25$ min per day and $h = 2.9$ (horizontal line). The plots in the second row were generated with $\sigma = 10$ min, $\beta = 0.25$ min per day and $h = 6.2$.

The simulated data were generated following the expression presented in “Simulated Data Generation”.

This section presents the results of applying our proposal to use the CUSUM algorithm for a process following a circular normal distribution, in accordance with the procedures described above. First, the performance in abrupt change detection is presented and discussed, and this is then followed by the performance in linear trend detection.

The $\sigma$ values for experimentation were chosen so as to agree with those data extracted from the experimentation with the real datasets presented in “Activity Modelling Using the Circular Normal Distribution” (Tables 2 and 3), which range between $13.40 \leq \sigma \leq 38.05$ min for the Milan dataset, and between $8.80 \leq \sigma \leq 36.69$ min for the Aruba dataset.

Abrupt Change Detection and Change Point Estimation in a Process Following a Circular Normal Distribution

The CUSUM algorithm was developed to estimate the change point in time when a process following a known pdf defined by a set of parameters suddenly changes the value of one of its parameters. In the case of circular normal pdf such a parameter is $\mu$, and it is assumed that $\kappa$ is unchanged. Taking into account the periodicity of the circular normal direction $\mu_1$ and concentration parameter $\kappa$, and with white noise added to it $N(0, \epsilon)$:

$$y_i(i \leq \tau) \rightarrow f(\phi|\mu_1, \kappa) + N(0, \sigma); \quad \mu_1 = \text{cte}; \sigma \ll \mu_1 \quad (9)$$

After the change point, the mean direction $\mu_2$ parameter of the circular normal distribution starts to follow a linear drift with rate $\beta$, and the concentration parameter $\kappa$ does not change:

$$y_i(i > \tau) \rightarrow f(\phi|\mu_2, \kappa) + N(0, \sigma); \quad \mu_2 = \mu_1 + (i - \tau)\beta; \sigma \ll \mu_1 \quad (10)$$

Figure 5 shows, on the left, some realizations of simulated data.

Results

As stated in “Anomaly Detection”, anomalous activities occur rarely, and they are not expected in advance. This means that it is difficult to find public ADL databases containing anomalous behaviours, and consequently it is also difficult to figure out the rate of anomalous behaviours in ADL. For this reason, in order to assess the validity of the presented method, anomalous behaviour data were simulated, as carried out in related approaches (e.g. [93] and [87]).
distribution, without loss of generality it can be assumed that $\mu = 0$, in the case of the in-control process.

The first set of experiments were performed to assess the validity of the CUSUM algorithm to detect a change in a time series when a sudden change happens. As mentioned before, an example is a sudden change in a person’s wake-up time due to, for instance, the onset of a cold.

For each experiment 50 samples following a circular normal distribution were generated with parameter $\mu_{before} = 0$ and given that, as the basis, 24 h are $2\pi$ radians. The change point in time was set at the 50th sample in all cases. After the change, 100 new samples following a circular normal distribution with parameter $\mu_{after}$ ranging from 5 min (0.022 radians) to 30 min (0.131 radians) were generated. As previously stated, parameter $\kappa$, and equivalently $\sigma$, was left unchanged. For each combination of parameters, the experiment was repeated 10,000 times. The value of the threshold $h$ must be set carefully: if it is too low, a great number of false alarms could be detected; in contrast, if the value of the threshold is too high, the decision function could never reach it. This behaviour can be used as a strength rather than a weakness of the algorithm in the following way: one could choose more than one threshold at the same time. For example, three different threshold levels could be used, the lowest threshold could be associated with a low risk in an older adult mental health, the middle threshold could be associated with a moderate risk in an older adult mental health, finally the highest threshold could be associated with high risk in an older adult mental health. In this way a health practitioner could associate the risk in an older adult mental health with a different threshold. In addition, adaptive techniques can be explored in order to change the threshold dynamically according to the working context and in relation to experience in similar situations or related data. Figure 6 shows how the percentage of successful experiments and the MRL depend on the value of the decision function threshold ($h$). An experiment is considered as successful if the CUSUM algorithm is able to provide an estimate of the change point $\hat{\tau}$. Table 4 shows the experimental results for various combinations of $\sigma$ (note the equivalence between $\sigma$ and $\kappa$ in Table 1) and $\mu_{after}$, each row presenting the results when the maximum number of successful experiments was reached (column labelled Success). The MRL is the interval between the start of the experiment and the change detection. $\hat{\tau}$ is the mean of the estimated change point. A False Alarm occurs when a change is detected before the real change has taken place, namely when $\hat{\tau} < \tau$. Change Fails is the number of changes that were not detected by the algorithm. $\hat{\tau}$ Fail happens when a change is correctly detected but the algorithm is unable to estimate the change point $\hat{\tau}$. Finally, Success experiments is the percentage of experiments which were able to estimate the change point $\hat{\tau}$.

In all the experiments, the mean estimated change point $\hat{\tau}$ is quite close to the real value $\tau = 50$. The differences between the real value of the change point $\tau$ and its

![Percentage of successful experiments](image1)

![Run Length](image2)

Fig. 6 Percentage of successful experiments as a function of the decision function threshold (left). MRL as a function of the decision function threshold (right). For this particular experiment $\sigma = 25$ min For the first row $\mu_{after} = 5$ min For the second row $\mu_{after} = 30$ min
Table 4 Results for abrupt changes experiments

| $\sigma$ | $\mu_{after}$ | $h$ | MRL | $\hat{\tau}$ | False Alarms | Change Fails | $\hat{\tau}$ Fails | Success |
|---|---|---|---|---|---|---|---|---|
| 40 | 5 | 1.2 | 90.76±0.31 | 51.82±0.35 | 1426 | 1760 | 134 | 66.80 |
| 40 | 10 | 2.5 | 90.01±0.26 | 51.40±0.27 | 681 | 969 | 88 | 82.62 |
| 40 | 15 | 3.7 | 85.84±0.22 | 51.22±0.21 | 351 | 361 | 34 | 92.54 |
| 40 | 20 | 5.2 | 83.25±0.19 | 50.44±0.16 | 106 | 107 | 20 | 97.67 |
| 40 | 25 | 6.7 | 80.01±0.16 | 50.09±0.12 | 31 | 32 | 5 | 99.32 |
| 40 | 30 | 9.1 | 79.71±0.14 | 50.04±0.09 | 0 | 5 | 2 | 99.93 |
| 30 | 5 | 1.5 | 88.78±0.29 | 52.90±0.32 | 1565 | 1116 | 90 | 72.29 |
| 30 | 10 | 3.4 | 88.79±0.24 | 51.27±0.23 | 392 | 593 | 44 | 89.71 |
| 30 | 15 | 5.2 | 83.08±0.19 | 50.61±0.16 | 93 | 125 | 24 | 97.58 |
| 30 | 20 | 7.7 | 80.84±0.15 | 50.14±0.11 | 14 | 20 | 6 | 99.60 |
| 30 | 25 | 9.4 | 75.01±0.11 | 50.00±0.07 | 2 | 0 | 0 | 99.98 |
| 30 | 30 | 10.4 | 69.71±0.09 | 49.95±0.05 | 0 | 0 | 0 | 100.0 |
| 25 | 5 | 1.9 | 88.85±0.28 | 51.76±0.30 | 1165 | 1093 | 83 | 76.59 |
| 25 | 10 | 4.1 | 85.49±0.21 | 50.94±0.19 | 232 | 365 | 34 | 93.78 |
| 25 | 15 | 6.8 | 82.35±0.17 | 50.18±0.12 | 22 | 57 | 6 | 99.15 |
| 25 | 20 | 9.8 | 77.97±0.12 | 50.04±0.08 | 0 | 4 | 1 | 99.95 |
| 25 | 25 | 10.4 | 69.44±0.09 | 49.98±0.05 | 0 | 0 | 0 | 100.0 |
| 25 | 30 | 10.7 | 63.95±0.06 | 50.01±0.04 | 0 | 0 | 0 | 100.0 |
| 20 | 5 | 2.4 | 88.13±0.26 | 51.63±0.27 | 774 | 891 | 83 | 82.52 |
| 20 | 10 | 5.1 | 82.35±0.18 | 50.62±0.15 | 96 | 120 | 26 | 97.58 |
| 20 | 15 | 8.2 | 75.97±0.13 | 49.94±0.09 | 5 | 5 | 0 | 99.99 |
| 20 | 20 | 11.4 | 71.60±0.09 | 50.01±0.05 | 0 | 0 | 0 | 100.0 |
| 20 | 25 | 10.4 | 62.52±0.06 | 50.00±0.03 | 0 | 0 | 0 | 100.0 |
| 20 | 30 | 10.5 | 58.78±0.04 | 50.02±0.02 | 0 | 0 | 0 | 100.0 |
| 15 | 5 | 3.4 | 87.42±0.22 | 50.57±0.21 | 331 | 582 | 74 | 90.13 |
| 15 | 10 | 7.6 | 80.15±0.15 | 50.12±0.10 | 10 | 20 | 2 | 99.68 |
| 15 | 15 | 10.1 | 68.81±0.09 | 49.98±0.05 | 0 | 0 | 0 | 100.0 |
| 15 | 20 | 10.3 | 60.94±0.05 | 50.01±0.03 | 0 | 0 | 0 | 100.0 |
| 15 | 25 | 10.9 | 57.38±0.03 | 49.99±0.02 | 0 | 0 | 0 | 100.0 |
| 15 | 30 | 10.5 | 54.92±0.02 | 50.00±0.01 | 0 | 0 | 0 | 100.0 |
| 10 | 5 | 5.4 | 84.67±0.19 | 50.61±0.16 | 90 | 152 | 23 | 97.35 |
| 10 | 10 | 10.2 | 69.25±0.09 | 50.06±0.05 | 0 | 0 | 0 | 100.0 |
| 10 | 15 | 10.3 | 58.80±0.04 | 49.98±0.02 | 0 | 0 | 0 | 100.0 |
| 10 | 20 | 11.5 | 55.48±0.02 | 50.00±0.01 | 0 | 0 | 0 | 100.0 |
| 10 | 25 | 11.7 | 53.45±0.02 | 50.00±0.01 | 0 | 0 | 0 | 100.0 |
| 10 | 30 | 10.5 | 52.02±0.01 | 50.00±0.01 | 0 | 0 | 0 | 100.0 |

$\mu_{after}$ and $\sigma$ are set in minutes. The Mean Run Lengths (MRL) and $\hat{\tau}$ are in days. False Alarms, Change Fails, and $\hat{\tau}$ are in units as the number of experiments. Success is the percentage of successful experiments. The estimated value $\hat{\tau}$ is within the range $0 \leq \|\tau - \hat{\tau}\| \leq 1.82 \pm 0.35$ for all the experiments performed. The biggest difference corresponds to $\sigma = 40$ min and $\mu_{after} = 5$ min, which is the most challenging case tested, since there is a high degree of variability in the data ($\sigma = 40$ min), and the abrupt change in the mean $\mu$ is only 5 min. In the example case presented above, if a person’s wake-up time is 7:00 a.m. with a dispersion of 40 minutes, and she suddenly changes her wake-up time by 5 minutes, namely 7:05 a.m., the CUSUM algorithm might detect the change after an average of 41 days, and estimate the day of the change with an error of only 2 days. For each pair of values ($\mu, \sigma$) the value of the threshold $h$ can be set in such a way that the maximum percentage of successful experiments increases when $\mu_{after}$ also increases. Conversely, the MRL, the number of False alarms, the
Fig. 7 Percentage of successful experiments vs. threshold for different detector times $\delta$ in minutes. $\sigma = 25$ min for all experiments. The dashed line is for 90% of successful experiments.

Table 5 Mean Run Length (MRL) and mean change point estimate $\hat{\tau}$ for the same percentage of successful experiments as in the case of an abrupt change.

| $\beta$ | $\delta$ | $h$ | MRL    | $\hat{\tau}$ | False Alarms | $\hat{\tau}$ Fails | Success |
|---------|----------|-----|--------|-------------|--------------|---------------------|---------|
| 0.25    | 2.5      | 1.2 | 90.52±0.17 | 65.17±0.10 | 749 | 160 | 90.91 |
| 0.25    | 5.0      | 2.0 | 90.43±0.17 | 65.08±0.10 | 919 | 173 | 89.08 |
| 0.25    | 7.5      | 2.6 | 91.25±0.17 | 65.26±0.10 | 959 | 114 | 89.27 |
| 0.25    | 10.0     | 3.2 | 93.95±0.17 | 66.07±0.10 | 798 | 118 | 90.84 |
| 0.5     | 2.5      | 1.2 | 78.59±0.11 | 55.65±0.16 | 705 | 188 | 91.07 |
| 0.5     | 5.0      | 2.0 | 78.11±0.11 | 55.79±0.16 | 917 | 154 | 89.29 |
| 0.5     | 7.5      | 2.6 | 78.37±0.11 | 55.92±0.16 | 902 | 151 | 89.47 |
| 0.5     | 10.0     | 3.2 | 79.73±0.11 | 56.05±0.16 | 795 | 130 | 90.75 |
| 0.75    | 2.5      | 1.2 | 73.07±0.08 | 51.38±0.14 | 752 | 210 | 90.38 |
| 0.75    | 5.0      | 2.0 | 72.63±0.08 | 51.94±0.14 | 871 | 155 | 89.74 |
| 0.75    | 7.5      | 2.6 | 72.71±0.08 | 51.93±0.14 | 906 | 160 | 89.34 |
| 0.75    | 10.0     | 3.2 | 73.58±0.08 | 52.11±0.14 | 763 | 127 | 91.10 |
| 1.0     | 2.5      | 1.2 | 69.94±0.07 | 49.67±0.13 | 699 | 198 | 91.03 |
| 1.0     | 5.0      | 2.0 | 69.44±0.07 | 49.88±0.13 | 890 | 180 | 89.30 |
| 1.0     | 7.5      | 2.6 | 69.41±0.07 | 50.18±0.13 | 965 | 143 | 88.92 |
| 1.0     | 10.0     | 3.2 | 69.90±0.07 | 50.08±0.13 | 785 | 152 | 90.63 |
| 1.5     | 2.5      | 1.2 | 66.06±0.05 | 47.37±0.12 | 767 | 184 | 90.49 |
| 1.5     | 5.0      | 2.0 | 65.61±0.05 | 47.66±0.12 | 899 | 175 | 89.26 |
| 1.5     | 7.5      | 2.6 | 65.51±0.05 | 47.90±0.12 | 995 | 149 | 88.56 |
| 1.5     | 10.0     | 3.2 | 65.71±0.05 | 47.87±0.12 | 800 | 147 | 90.53 |
| 2.0     | 2.5      | 1.2 | 63.76±0.04 | 46.48±0.11 | 735 | 182 | 90.83 |
| 2.0     | 5.0      | 2.0 | 63.31±0.04 | 46.58±0.11 | 882 | 172 | 89.46 |
| 2.0     | 7.5      | 2.8 | 63.75±0.04 | 46.50±0.11 | 739 | 191 | 90.70 |
| 2.0     | 10.0     | 3.2 | 63.39±0.04 | 46.80±0.11 | 776 | 168 | 90.56 |

Note that both variables are approximately the same regardless of the detector time $\delta$. The values were set to $\mu = 0$ and $\sigma = 25$ min for all experiments. Each experiment was performed 10,000 times with same parameters. Light grey rows highlight experiments with the same pair of time detector $\delta$ and threshold $h$. There were no change fails in any experiment.
number of Change Fails, the number of \( \hat{\tau} \) Fails and the number of successful experiments decrease when \( \mu_{\text{after}} \) increases.

From the results presented in Table 4 it can be stated that the CUSUM algorithm for abrupt change detection can be successfully applied to data following a circular normal distribution within the parameter values tested.

**Linear Trend Detection and Change Point Estimation in a Process Following a Circular Normal Distribution**

The first step for detecting a change in a process following a circular normal distribution with a linear drift is to use the log-likelihood (6), and so a value for the detector time \( \delta \) should be chosen. To do so, first we present some experimental insights.

Figure 7 shows the number of successful experiments with regard to the threshold \( h \) set, for the values \( \delta \in \{2.5, 5.0, 7.5, 10.0\} \) in minutes. Note that the same percentage of successful experiments was achieved for different values of the detector time \( \delta \). Hence, the first insight is that, once the detector time \( \delta \) has been set, the threshold \( h \) can be chosen to achieve approximately the same percentage of successful experiments.

Table 5 shows the values for the MRL and the mean estimated change point \( \hat{\tau} \) for values of the linear drift rate \( \beta \in \{0.25, 0.5, 0.75, 1.0, 1.5, 2.0\} \) min per day. Once the detector time \( \delta \) had been set, the value of the threshold \( h \) was chosen to achieve approximately the same percentage of successful experiments in all cases (90 %). Note that, for the same linear drift rate \( \beta \), and once the time detector \( \delta \) and threshold \( h \) values have been set to achieve approximately the same percentage of successful experiments, the differences in the MRL and the mean of the estimated change point \( \hat{\tau} \) are small. For example, in the case of \( \beta = 0.25 \) min per day, the maximum of the MRL is 93.95 ± 0.17 (\( \delta = 10, h = 3.2 \)), and the minimum of the MRL is 90.43 ± 0.17 (\( \delta = 5.0, h = 2.0 \)), a difference of 3.53 ± 0.34(3.89 %). For the estimated change point \( \hat{\tau} \), the maximum is 66.07±0.10(\( \delta = 10.0, h = 3.2 \)), the minimum is 65.08±0.10(\( \delta = 5.0, h = 2.0 \)), a difference of 0.99 ± 0.20(1.52 %). In the case of \( \beta = 2.0 \) minutes per day, the maximum of the MRL is 63.76±0.04(\( \delta = 2.5, h = 1.2 \)), and the minimum of the MRL is 63.31±0.04(\( \delta = 5.0, h = 2.0 \)), a difference of 0.45±0.08(0.71 %). For the estimated change point \( \hat{\tau} \) the maximum is 46.80±0.11(\( \delta = 10.0, h = 3.2 \)), the minimum is 46.48±0.11(\( \delta = 2.5, h = 1.2 \)), a difference of 0.32±0.22(0.69 %). There were no change fails in any experiment. This is the second insight, if the values of \( \delta \) and \( h \) are set to achieve a certain percentage of successful experiments, the MRL and the mean of the estimated change point \( \hat{\tau} \) are approximately the same for any linear drift rate \( \beta \).

As a conclusion, and bearing in mind the two previous insights, it can be stated that for any given time detector \( \delta \), the threshold \( h \) can be set to achieve a percentage of successful experiments, which in turn gives the same result for the MRL and the mean estimated change point \( \hat{\tau} \).

A third important insight can be extracted from Table 5 if one looks at the rows highlighted in light grey: once the time detector \( \delta \) has been set, and independently of the linear drift rate \( \beta \), the threshold \( h \) can be chosen to obtain the same percentage of successful experiments. Figure 8 shows this behaviour in detail: for the six values

---

**Fig. 8** Percentage of successful experiments vs. threshold for different drift rates in minutes per day. \( \sigma = 25 \) min and \( t_{\text{detector}} = 5.0 \) min for all experiments
of $\beta \in \{0.25, 0.50, 0.75, 1.00, 1.50, 2.00\}$ min per day, the percentage of successful experiments as a function of the threshold $h$ are quite similar (the profile of the six plots behave in the same way). Therefore, it can be concluded that once the detector time $\delta$ has been set, the threshold $h$ can be chosen to achieve the same percentage of successful experiments regardless of the linear drift rate $\beta$ of the process. This conclusion is important because it allows us to compare the MRL and the mean estimated change point $\hat{\tau}$ for different linear drift rates $\beta$ by just setting the threshold $h$, which in turn sets the percentage of successful experiments.

Table 6 shows the MRL and the mean estimated change point $\hat{\tau}$ for different values of $\sigma$ in minutes, and the linear drift rate $\beta$ in minutes per day. For all experiments the change point was set to $\tau = 50$, the detector time was set

| $\sigma$ | $\beta$ | $h$ | MRL | $\hat{\tau}$ | False Alarms | $\hat{\tau}$ Fails | Success |
|---|---|---|---|---|---|---|---|
| 40 | 0.25 | 0.8 | 99.80±0.21 | 73.11±0.25 | 670 | 188 | 90.68 |
| 40 | 0.50 | 0.8 | 86.28±0.14 | 61.45±0.19 | 730 | 187 | 90.83 |
| 40 | 0.75 | 0.8 | 79.54±0.11 | 55.83±0.16 | 668 | 223 | 91.09 |
| 40 | 1.00 | 0.8 | 75.75±0.09 | 53.28±0.15 | 660 | 236 | 91.04 |
| 40 | 1.50 | 0.8 | 70.82±0.07 | 49.69±0.13 | 648 | 234 | 91.18 |
| 40 | 2.00 | 0.8 | 70.82±0.07 | 49.69±0.13 | 648 | 234 | 91.18 |
| 30 | 0.25 | 1.0 | 93.33±0.18 | 67.97±0.22 | 829 | 186 | 89.83 |
| 30 | 0.50 | 1.0 | 80.80±0.12 | 57.83±0.17 | 745 | 207 | 90.48 |
| 30 | 0.75 | 1.0 | 80.80±0.12 | 57.83±0.17 | 745 | 207 | 90.48 |
| 30 | 1.00 | 1.0 | 71.72±0.07 | 50.99±0.13 | 775 | 168 | 90.57 |
| 30 | 1.50 | 1.0 | 71.72±0.07 | 50.99±0.13 | 775 | 168 | 90.57 |
| 30 | 2.00 | 1.0 | 71.72±0.07 | 50.99±0.13 | 775 | 168 | 90.57 |
| 25 | 0.25 | 1.2 | 90.27±0.17 | 64.94±0.22 | 720 | 200 | 90.80 |
| 25 | 0.50 | 1.2 | 78.61±0.11 | 55.56±0.17 | 761 | 180 | 90.59 |
| 25 | 0.75 | 1.2 | 73.16±0.08 | 51.72±0.14 | 731 | 185 | 90.84 |
| 25 | 1.00 | 1.2 | 69.92±0.07 | 49.66±0.13 | 758 | 202 | 90.40 |
| 25 | 1.50 | 1.2 | 66.19±0.05 | 47.71±0.12 | 751 | 191 | 90.58 |
| 25 | 2.00 | 1.2 | 63.80±0.04 | 46.42±0.11 | 735 | 183 | 90.82 |
| 20 | 0.25 | 1.4 | 85.64±0.15 | 61.37±0.20 | 874 | 181 | 89.45 |
| 20 | 0.50 | 1.4 | 74.99±0.09 | 53.43±0.15 | 831 | 181 | 89.88 |
| 20 | 0.75 | 1.4 | 70.33±0.07 | 50.40±0.13 | 823 | 194 | 89.83 |
| 20 | 1.00 | 1.4 | 67.45±0.06 | 48.44±0.13 | 836 | 194 | 89.70 |
| 20 | 1.50 | 1.4 | 64.04±0.05 | 46.88±0.11 | 859 | 178 | 89.63 |
| 20 | 2.00 | 1.4 | 62.01±0.04 | 46.09±0.11 | 822 | 165 | 90.13 |
| 15 | 0.25 | 1.8 | 81.38±0.12 | 57.60±0.17 | 762 | 184 | 90.54 |
| 15 | 0.50 | 1.8 | 73.19±0.07 | 50.95±0.14 | 778 | 195 | 90.27 |
| 15 | 0.75 | 1.8 | 67.64±0.06 | 48.35±0.13 | 745 | 181 | 90.74 |
| 15 | 1.00 | 1.8 | 65.09±0.05 | 47.33±0.12 | 754 | 183 | 90.63 |
| 15 | 1.50 | 1.8 | 62.06±0.04 | 45.89±0.11 | 783 | 158 | 90.59 |
| 15 | 2.00 | 1.8 | 60.30±0.03 | 45.55±0.10 | 802 | 140 | 90.58 |
| 10 | 0.25 | 2.4 | 75.84±0.09 | 53.61±0.15 | 783 | 146 | 90.71 |
| 10 | 0.50 | 2.4 | 67.68±0.06 | 48.56±0.13 | 817 | 175 | 90.08 |
| 10 | 0.75 | 2.4 | 64.18±0.04 | 46.73±0.12 | 792 | 182 | 90.26 |
| 10 | 1.00 | 2.4 | 62.00±0.04 | 46.02±0.11 | 775 | 152 | 90.73 |
| 10 | 1.50 | 2.4 | 59.51±0.03 | 45.36±0.10 | 756 | 131 | 91.13 |
| 10 | 2.00 | 2.4 | 58.04±0.02 | 45.19±0.09 | 805 | 126 | 90.69 |

The time of the detector was set to $\delta = 2.5$ min in all experiments. Threshold $h$ was set to obtain approximately 90% of successful experiments. Each experiment with the same parameters was performed 10,000 times.
to $\delta = 2.5$ mins and the threshold $h$ was chosen to achieve approximately 90% of successful experiments. Note that the value for the chosen threshold was the same for all linear drift rates $\beta$ for each value of $\sigma$, as was concluded in the previous paragraph.

The biggest MRL in Table 6 is 99.80 ± 0.21, which corresponds to a value of $\sigma = 40$ min, and the linear drift rate $\beta = 0.25$ min per day. In this case, the change was detected 49.80 days after it happened. After these 49.80 days, the mean of the circular normal process is $\mu = 49.80 \times 0.25 = 12.45$ min. This might seem a big delay for detecting an anomalous behaviour, but it is not so big when compared with the value of $\sigma = 40$ min. In this case, in fact, this is the most challenging of the series of experiments that were performed. Moreover, all the processes were random and the drift rate was linear, so the difference between the original mean and the final mean increases by only 0.25 min per day. In the case of $\sigma = 10$ min and $\beta = 0.25$ min per day, the MRL was 75.84 ± 0.09, so the change in the process was detected 25.84 days after it happened, and the mean of the process at this time was $\mu = 6.25$ min. Note the same behaviour for other values of $\sigma$ and $\beta$ from Table 6, that is, the smaller the value of $\sigma$ is, the quicker the change was detected. Additionally, for the same value of $\sigma$, when the linear drift increases, the MRL decreases.

The biggest difference between the real value of the change point $\tau$ (50) and the mean estimated change point $\tilde{\tau}$ is 23.11 ± 0.25 days corresponding to $\sigma = 40$ min and $\beta = 0.25$ min per day, which is the most challenging case tested as there is a high degree of variability in the data $\sigma$ and a low drift rate $\beta$. The second biggest difference is 17.97, which corresponds to $\sigma = 30$ min and $\beta = 0.25$ min per day. In the other cases, the differences range between 0.40 ≤ $\|\tau - \tilde{\tau}\|$ ≤ 14.94 days. The average difference between the real value of the change point and the mean estimated change point is 4.90 ± 1.98 days for the whole set of experiments performed.

The results from the experiments show that the proposed method can be used first to detect and then to estimate a linear change in data series following a circular normal distribution.

## Conclusion

A new method to detect linear drift behaviour in processes following a circular normal distribution is presented in this work. This method is based on the use of multiple hypotheses to estimate the linear drift rate of a process following a circular normal distribution. The method was applied on public ADL datasets, which were modelled as mixtures of circular normal distribution. The method is able to detect and then estimate a change point in a time series. The circular normal distribution is used to model periodic data, which is the case of Activities of Daily Living. Experimental results have shown the validity of the method in detecting and estimating the change point in time.

Some examples of application of the proposed method are: to detect and estimate the onset of cognitive or physical decline in a monitored older adult, or to detect and estimate the change in ADL as a result of changing the medication a person takes. These may be of great interest in the health domain, as it can help health practitioners to better understand behavioural changes in patients.

A key point in the method is to select a threshold $h$ to detect the change, since the lower the threshold is, the more false positives are detected. Given this insight, a manifold strategy can be developed to supervise the behaviour of a monitored person. First, a low threshold can be used to check for changes in ADL. If the threshold is reached, a first level of alarm can be set and a specific supervision of the person could be established. Then the threshold is increased, and if the new threshold is reached, a higher level alarm could be set and special supervision or even some kind of intervention could be implemented.

Although the MRL could seem quite long in the challenging case that was tested (41 days for $\beta = 0.25$ and $\sigma = 40$ min), this delay is not so long when compared with the time between the reported onset of symptoms and diagnosis, which is said to be 2.8 years [99].

In order to assess the validity of the algorithm using real data, we have planned to perform a set of experiments with volunteers. The data will be obtained, in part, with an improved version of the acquisition system developed by the members of our research group [100].

### Funding Information
This work has been partially funded by the Spanish Ministry of Science, Innovation and Universities through the “Retos investigación” programme (RTC2018-095168-B-C53) and by the Universitat Jaume I “Pla de promoció de la investigació 2017” programme (UJI-B2017-45). Óscar Belmonte-Fernández had a grant from the Spanish Ministry of Science, Innovation and Universities (PRX18/00123) for developing part of this work.

### Compliance with Ethical Standards

#### Conflict of Interest
The authors declare that they have no conflict of interest.

#### Ethical Approval
This article does not contain any studies with human participants or animals performed by any of the authors.

### References

1. Fozard JL, Rietsema J, Bouma H, Graafmans JAM. Gerontechnology: creating enabling environments for the challenges and opportunities of aging. Educ Gerontol. 2000;26(4):331–344.
42. Liguori C, Romigi A, Nuccetelli M, Zannino S, Sancesario G, Martorana A, et al. Orexinergic system dysregulation, sleep impairment, and cognitive decline in Alzheimer disease. JAMA Neurology. 2014;71(12):1498–1505.

43. Mohs RC, Schmeidler J, Aryan M. Longitudinal studies of cognitive, functional and behavioural change in patients with Alzheimer’s disease. Stat Med. 2000;19(11–12):1401–1409.

44. Stem RG, Mohs RC, Davidson M, Silverman J, Kramer-Ginsberg E, et al. A Longitudinal Study of Alzheimer’s Disease. Measurement, Rate, and Predictors of Cognitive Deterioration. 1994. 390–396.

45. Suzuki T, Murase S. Influence of outdoor activity and indoor activity on cognition decline: use of an infrared sensor to measure activity. Telemed. e-Health. 2010;16(6):686–690.

46. Petersen J, Austin D, Mattick N, Kaye J. Time out-of-home and cognitive, physical, and emotional wellbeing of older adults: a longitudinal mixed effects model. PLoS ONE. 2015;10(10):e0139643.

47. Harada K, Lee S, Lee S, Bae S, Harada K, Suzuki T, et al. Objectively-measured outdoor time and physical and psychological function among older adults. Geriatr Gerontol Int. 2016;17(10):1455–1462.

48. Dawadi PN, Cook DJ, Schmitter-Edgecombe M, Parsey C. Automated assessment of cognitive health using smart home technologies. Technol Health Care. 2013;21(4):323–343.

49. Austin J, Dodge HH, Riley T, Jacobs PG, Thielke S, Kaye J. A smart-home system to unobtrusively and continuously assess loneliness in older adults. IEEE J. Transl. Eng. Health Med.-JTEHM. 2016;4:1–11.

50. Kötteritzsch A, Weyers B. Assistive technologies for older adults. Technol Health Care. 2013;21(4):323–343.

51. Impedovo D, Pirlo G, Vessio G, Angelillo MT. A Handwriting-Based protocol for assessing neurodegenerative dementia. Cogn. Comput.; May; 2019.

52. Zarandi MHF, Aalaedini A. A general fuzzy-statistical clustering approach for estimating the time of change in variable sampling control charts. Inf Sci. 2010;180(16):3033–3044.

53. López-de Ipiña K, Alonso JB, Solé-Casals J, Barroso N, Hentriquez F, Faundez-Zanuy M, et al. On automatic diagnosis of alzheimer’s disease based on spontaneous speech analysis and emotional temperature. Cogn. Comput. 2015;7(1):44–55.

54. Köttertzsch A, Weyers B. Assistive technologies for older adults in urban areas: a literature review. Cogn. Comput. 2016;8(2):299–317.

55. Shewhart WA. Economic control of quality of manufactured product. ASQ Quality Press; vol. 509. 1931.

56. Basseville M, Nikiforov IV. Detection of abrupt changes: theory and application. Upper Saddle River: Prentice-Hall, Inc.; 1993.

57. Cheng CS. A multi-layer neural network model for detecting changes in the process mean. Comput Ind Eng. 1995;28(1):51–61.

58. Guh RS, Hsieh YC. A neural network based model for abnormal pattern recognition of control charts. Comput Ind Eng. 1999;36(1):97–108.

59. Cheng CS, Cheng SS. A neural network-based procedure for the monitoring of exponential mean. Comput Ind Eng. 2001;40(4):309–321.

60. Chang S, Aw C. A neural fuzzy control chart for detecting and classifying process mean shifts. Int J Prod Res. 1996;34(8):2265–2278.

61. El-Shal SM, Morris AS. A fuzzy expert system for fault detection in statistical process control of industrial processes. IEEE Trans Syst Man Cybern Part C Appl Rev. 2000;30(2):281–289.

62. Aalaedini A, Ghazanfari M, Nayeri MA. A hybrid fuzzy-statistical clustering approach for estimating the time of changes in fixed and variable sampling control charts. Inf Sci. 2009;179(11):1769–1784.

63. Lu KP, Chang ST, Yang MS. Change-point detection for shifts in control charts using fuzzy shift change-point algorithms. Comput Ind Eng. 2016;93:12–27.

64. Raftery A, Akman V. Bayesian analysis of a poisson process with a change-point. 1986.

65. Assareh H, Noorossana R, Mengersen KL. Bayesian change point estimation in Poisson-based control charts. J Ind Eng Int. 2013;9(1):32.

66. Perry MB, Pigatiello JJ, Simpson JR. Estimating the change point of a Poisson rate parameter with a linear trend disturbance. Qual Reliab Eng Int. 2006;22(4):371–384.

67. Noorossana R, Shadman A. Estimating the change point of a normal process mean with a monotonic change. Qual Reliab Eng Int. 2009;25(1):79–90.

68. Noorossana R, Heydari M. Change point estimation of a process variance with a linear trend disturbance. J Ind Eng Int. 2009;2:25–30.

69. Lara OD, Labrador MA. A survey on human activity recognition using wearable sensors. IEEE Commun. Surv. Tutor. 2013;15(3):1192–1209.

70. Guan D, Ma T, Yuan W, Lee YK, Jehad Sarkar A. Review of sensor-based activity recognition systems. IETE Tech Rev. 2011;28(5):418–433.

71. Ke SR, Thuc HLU, Lee YJ, Hwang JN, Yoo JH, Choi KH. A review on video-based human activity recognition. Computers. 2013(2):88–131.

72. Vrigkas M, Nikou C, Kakadiaris IA. A review of human activity recognition methods. Frontiers in Robotics and AI. 2015;2:28.

73. van Kasteren TL, Englebienne G, Kröse BJ. Human activity recognition from wireless sensor network data: benchmark and software. In: Activity recognition in pervasive intelligent environments. Springer; 2011. p. 165–186.

74. Mardia KV, Jupp PE. Directional statistics. New York: Wiley; 2009. vol. 494.

75. Diethe T, Twomey N, Flach P. Bayesian modelling of the temporal aspects of smart home activity with circular statistics. In: Joint european conference on machine learning and knowledge discovery in databases. Springer; 2015. p. 279–294.

76. Chinellato E, Mardia K, Hogg D, Cohn A. An Incremental von Mises Mixture Framework for Modelling Human Activity Streaming Data. In: International Work-Conference on Time Series Analysis (ITISE); 2017. p. 379–389.

77. Brunsdon C, Corcoran J. Using circular statistics to analyse time patterns in crime incidence. Comput Environ Urban Syst. 2006;30(3):300–319.

78. Kubiak T, Jonas C. Applying circular statistics to the analysis of monitoring data. Eur J Psychol Assess. 2007;23(4):227–237.

79. Franco C, Demongeot J, Villeneuve C, Vuillerme N. Behavioral telemonitoring of the elderly at home: detection of nycthemeral rhythms drifts from location data. In: 2010 IEEE 24th international conference on advanced information networking and applications workshops. IEEE; 2010. p. 759–766.

80. Jammalamadaka SR, Sengupta A. Topics in circular statistics. World Scientific, vol. 5. 2001.

81. Cremers J, Klugkist I. One direction? A tutorial for circular data workshop. IEEE; 2010. p. 759–766.
84. Yin X, Shen W, Samarabandu J, Wang X. Human activity detection based on multiple smart phone sensors and machine learning algorithms. In: 2015 IEEE 19th international conference on computer supported cooperative work in design (CSCWD); 2015. p. 582–587.
85. Yin J, Yang Q, Pan JJ. Sensor-based abnormal human-activity detection. IEEE Trans Knowl Data Eng. 2008;20(8):1082–1090.
86. van Kasteren T.L., Englebienne G., Kröse B.J. Human activity recognition from wireless sensor network data: Benchmark and software. In: Activity recognition in pervasive intelligent environments (pp. 165–186). Atlantis Press.; 2011.
87. Forkan ARM, Khalil I, Tari Z, Fofou S, Bouras A. A context-aware approach for long-term behavioural change detection and abnormality prediction in ambient assisted living. Pattern Recogn. 2015;48(3):628–641.
88. Zhu C, Sheng W, Liu M. Wearable sensor-based behavioral anomaly detection in smart assisted living systems. IEEE Trans Auton Sci Eng. 2015;12:1225–1234.
89. Duong TV, Bui H, Phung DQ, Venkatesh S. Activity recognition and abnormality detection with the switching hidden semi-Markov model. In: Proceedings of the 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR’05) - volume 1 - volume 01. CVPR ’05. Washington, DC, USA: IEEE Computer Society; 2005. p. 838–845.
90. Shin JH, Lee B, Park KS. Detection of abnormal living patterns for elderly living alone using support vector data description. IEEE Trans Inf Technol Biomed. 2011;15(3):438–448.
91. Palaniappan A, Bhargavi R, Vaidehi V. Abnormal human activity recognition using SVM based approach. In: 2012 international conference on recent trends in information technology; 2012. p. 97–102.
92. Riboni D, Bettini C, Civitarese G, Janjua ZH, Helaoui R. SmartFABER: recognizing fine-grained abnormal behaviors for early detection of mild cognitive impairment. Artif Intell Med. 2016;67(Supplement C):57–74.
93. Viron G, Noury N, Demongeot J. A system for automatic measurement of circadian activity deviations in telemedicine. IEEE Trans Biomed Eng. 2002;49(12):1463–1469.
94. Fahad LG, Rajarajan M. Anomalies detection in smart-home activities. In: IEEE 14th international conference on machine learning and applications (ICMLA), 2015. IEEE; 2015. p. 419–422.
95. Fisher N. I. Statistical analysis of circular data. Cambridge University Press. 1995.
96. Cook DJ, Schmitter-Edgecombe M. Assessing the quality of activities in a smart environment. Methods Inf Med. 2009;48(05):480–485.
97. Cook DJ. Learning setting-generalized activity models for smart spaces. IEEE Intel Syst. 2010;2010(99):1.
98. Kullback S, Leibler RA. On information and sufficiency. Ann Math Stat. 1951;22(1):79–86.
99. Brookmeyer R, Corrada MM, Curriero FC, Kawas C. Survival following a diagnosis of Alzheimer disease. Arch Neurol. 2002;59(11):1764–1767.
100. Belmonte-Fernández O, Puertas-Cabedo A, Torres-Sospedra J, Montoliu-Colás R, Trilles-Oliver S. An indoor positioning system based on wearables for ambient-assisted living. Sensors. 2017;17(1):36.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Óscar Belmonte-Fernández1 · Antonio Caballer-Miedes2 · Eris Chinellato3 · Raúl Montoliu4 · Emilio Sansano-Sansano5 · Rubén García-Vidal5

Antonio Caballer-Miedes
caballer@uji.es

Eris Chinellato
E.Chinellato@mdx.ac.uk

Raúl Montoliu
montoliu@uji.es

Emilio Sansano-Sansano
esansano@uji.es

Rubén García-Vidal
vidalr@uji.es

1 Institute of New Imaging Technologies and Department of Computer Languages and Systems, Universitat Jaume I, 12071 Castelló de la Plana, Castelló, Spain
2 Institute of New Imaging Technologies and Department of Developmental, Educational and Social Psychology and Methodology, Universitat Jaume I, 12071 Castelló de la Plana, Castelló, Spain
3 Department of Design Engineering & Maths, Middlesex University, London NW4 4BT, UK
4 Institute of New Imaging Technologies and Department of Computer Architecture and Technology, Universitat Jaume I, 12071 Castelló de la Plana, Castelló, Spain
5 Institute of New Imaging Technologies, Universitat Jaume I, 12071 Castelló de la Plana, Castelló, Spain