Fluctuation and Noise Letters

The Effects of Examination Methods on Describing the Scale-Free Nature of Human Locomotor Activity

Krisztina Panna Komáromi, Bálint Maczák and Gergely Vadai*

Department of Technical Informatics, University of Szeged, Szeged 6720, Hungary

**vadaig@inf.u-szeged.hu*

Received Day Month Year
Revised Day Month Year
Accepted Day Month Year
Published Day Month Year

Communicated by

Although human dynamics exhibit scale-independent properties, different results have emerged from the statistical analyses of various actigraphic or mobility datasets regarding the characterization of their heavy-tailed distributions. By applying approaches from significant studies of this nature of human locomotor activity on our multi-day long acceleration dataset, we examine how the non-standardised activity determination methods and the separation of the resulting activity data into active and passive periods affect the conclusions. We also show that the statistical or model fitting approaches commonly used to study these periods not only influence the results but are also highly sensitive to the actigraphic methodology employed. In contrast, the differently derived activity data – and even the underlying acceleration signals – measured on the wrist, follow universal $1/f$ noise-based spectral characteristics that have also been previously observed in the context of human spatial motion, thus offering the possibility of a more robust analysis of scale-independence.

Keywords: actigraphy; scale-free nature; $1/f$ noise; human dynamics.

1. Introduction

Human dynamics exhibit scale-free characteristics in both spatial movement data and locomotor activity measurements. However, the processing and statistical analysis of these datasets have been carried out in diverse ways; moreover, the examination of the heavy-tailed distributions of different datasets has led to various, sometimes conflicting models. Locomotor activity is often quantified in medical research using so-called actigraphs, devices that measure the acceleration of the wrist and compress it to an activity value for every epoch (non-overlapping, consecutive timeslots, typically of 1 minute). However, the methods used to derive activity values vary significantly across manufacturers and are not disclosed properly. Consequently, differently produced activity signals are often analysed without a sufficient description of the underlying methodology.

Additionally, statistical methods for analysing the device-dependent activity data are not standardised in the literature either. While all related studies separate activity signals into active and passive (or resting) periods and examine the distribution of their durations, they differ in key aspects, such as the threshold used for separating the periods, the type of distribution function (probability density function or complementary cumulative distribution function) used for examining the durations, the fitting method (least squares or maximum likelihood estimation) and its range, the selection of fitted distribution functions, and the approach used for determining the most suitable one [1,2]. In contrast, as we previously showed, the spectra of both acceleration signals measured on the wrist for several days, and the activity signals derived from them follow a universal, $1/f$ noise-based spectral characteristic – representing the spectral scale-free property –, independent of the actigraphic methodology [3]. It is important to note that we have identified very similar spectral characteristics in the case of location-data-based spatial displacement as well [4,5].

The aim of the present work is to systematically examine how different ways of signal and data processing, and subsequent model fitting affect the conclusions drawn from statistical analyses. To do this, we compared the statistical distributions of actigraphic datasets generated in various ways for the same acceleration data we measured and analysed previously [6]. In the examination of the effects of these various activity determination methods and the threshold rule separating activity signals into active and passive periods, we complement our previous work [7] by incorporating a broader range of methods and conducting a more detailed examination of the influence of each processing steps. Furthermore, our aim is to investigate the impact of different statistical approaches found in the relevant studies on model fitting and the sensitivity of these steps to actigraphic methodology by comparing it with the robustness of spectral analysis.

2. Different methods of processing and analysing actigraphic data

To examine the joint impact of the different actigraphic and statistical methodologies on the scale-free properties of human locomotor activity, firstly we need to highlight the key differences in the important works of the past 20 years in this field.

2.1. *Diversity of actigraphic methodology*

The collected studies used different actigraphs to record activity data: Actiwatch (Cambridge Neurotechnology Ltd., Cambridge, UK) [8–10], MicroMini Motionlogger (Ambulatory Monitoring Inc., Ardsley, NY, USA) [2,11–13], ActiGraph GT3X+ and wGT3X-BT (ActiGraph LLC., Pensacola, FL, USA) [1,14]. This alone leads to inconsistencies as manufacturers quantify locomotor activity differently. Previously, we established a generalised framework for these activity determination procedures that consists of two steps [6]: preprocessing the raw acceleration data, and applying an activity metric to determine epoch-wise activity values. Most of the related studies used 60-second epochs [1,2,9,11–13,15], consistent with standard practice of broader actigraphic literature, though a few of them diverged from this (e.g., 30 seconds [8,10], or 5-120 seconds [14,15]).

In the related works, three classical activity metrics were employed for the epoch-based data aggregation: Proportional Integration Method (PIM), Zero Crossing Method (ZCM), and Time Above Threshold (TAT). While PIM numerically integrates the acceleration values over every epoch, ZCM counts the times the acceleration signal crossed a threshold (whose level we optimally set to the standard deviation of the whole acceleration recording [6]), whereas TAT measures the duration spent above that same threshold. Beyond these, additional activity metrics are also commonly used in the broader literature: High-Pass Filtered Euclidean Norm (HFEN), Euclidean Norm Minus One (ENMO), Mean Amplitude Deviation (MAD), and Activity Index (AI), see [6] for details.

However, the specifics of how the acceleration data was preprocessed – prior to the activity-metric-based data aggregation – are mostly hidden from the scientists [6]; therefore, we could only approximate them based on the available information. Briefly, the raw acceleration data collected along the three axes (UFXYZ) is typically transformed into a single time series by calculating the magnitude data (UFM). The gravitational component (g) is commonly eliminated from it by subtracting 1 g and taking the absolute value (UFNM) [14]. Alternatively, the g can be attenuated by applying a digital filter (for details, see [3]) either on the already computed magnitude data (FMpost) [2,10,11,13] or on per-axis data (FXYZ) before the magnitude calculation (FMpre) [1].

As all the aforementioned activity metrics are widely used in broader areas of actigraphic research, we applied all previously mentioned activity metrics to the differently preprocessed magnitude of acceleration data in proper combinations, generating 20 distinct activity signals for analysis. To identify how activity data was derived, we use the notation: activity metric as operator with preprocessing method as argument – e.g., ZCM(UFNM).

2.2. *Analysing scale-free nature*

Analysing the scale-free nature of human locomotor activity is most frequently done by examining the statistical distributions of active and passive periods in the data. The first step is to segment the already acquired activity data into such periods using a threshold. The method used to determine this threshold varies across studies but is typically defined relative to the mean activity level (commonly set at 100% [14], 70% [10], or 10% [9] of it) or as the mean of the non-zero activity values [1,2,15]. Among the related works, one study [8] adopted a significantly different approach and set the threshold value to be negligibly small (for example, 1% of the mean activity level), and as such, its results should not be considered in the same context as the others. Their approach serves as an analogy to the identification of time spent in a single location (i.e., waiting time), a similarly fundamental step commonly employed in spatial mobility analysis regarding bursty dynamics [16].

After identifying the active and passive periods, the related works have predominantly analysed their Complementary Cumulative Distribution Functions (CCDFs) [1,2,8–10,12–15] to assess their statistical properties, although Probability Density Functions (PDFs) can also be used [11], which is the more common approach in spatial mobility analysis. Although both functions are capable of capturing the scale-free properties implied by a power-law distribution (if the CCDF follows the form $x^{-\gamma}$, the PDF will also follow a power law, albeit with exponent of $\gamma + 1$), choosing between the two approaches is not straightforward, as each offers distinct advantages [7,17]. For the PDF, a trade-off must be made between resolution and noisiness as – in contrast to the CCDF – it requires binning the data; and even though logarithmic binning can mitigate the noise at the tail for heavy-tailed distributions, we found during our analysis that the number of bins can influence the estimated α exponent. Another influential factor is the finite length of the activity recordings, which limits the number of long active or passive segments that can form; as a result, the PDF becomes noisy, and the CCDF exhibits truncation at the tail of the distribution – a well-known limitation [18,19]. These advantages and limitations of the two approaches will also be graphically illustrated at the beginning of Section 3.1. Moreover, such artifacts may be further amplified if only one PDF/CCDF is estimated for the whole group by pooling (i.e., concatenating) active/passive periods of the individuals [10,15] instead of one PDF/CCDF for each subject [1,2,8,9,11,13,14]. Additionally, pooling can result in the emergence of power-law distributions, even if the distributions at the individual level are characterised by different models [20].

In related literature, there is no single dominant approach for parameterizing the distributions through model fitting; both Least Squares (LS) [2,9,10,12] and Maximum Likelihood Estimation (MLE) [1,2,13,14] are commonly used methods (although the latter is generally recommended when power-law distribution is assumed – indicative of scale-free properties [17]). However, the choice of fitting range – both for active and passive periods – varied considerably across studies. Moreover, a significant number of studies fitted multiple candidate models and selected the best-fitting one using a wide range of goodness-of-fit metrics. These included, for example, the Likelihood Ratio Test (LRT) [14], Sum of Squared Errors (SSE) [12], Reduced χ^2 [12,15], Akaike Information Criterion (AIC) [12,15], and Bayesian Information Criterion (BIC) [1,12,15], with some studies employing several of these metrics simultaneously [12,15].

The comparison of findings across studies about the models that explain the distributions is complicated by the differing objectives and methodological approaches described above. Without segmenting the data into intraday intervals (e.g., sleep and wakefulness, or their specific combinations), numerous studies have found power-law distribution for passive periods, with some reporting the γ exponent between 0.9 and 1.02 for healthy subjects [2,15,21,22] and between 0.7 and 0.8 for subjects with depression [15,21]. On the contrary, others have observed γ exponent in the latter range for healthy subjects [10,12], and although one study found the distribution to be heavy-tailed, they concluded it did not follow power law [13]. Another study found power-law behaviour only for periods shorter than 20 minutes, but the γ exponent they estimated contradicted past results: 0.7 for healthy and 0.93 for subjects with depression [9]. When analysing wake periods separately, a power-law distribution with γ exponent of 1.67 was reported [11], though the power-law nature of wake periods was also challenged by another study [13]. For specific health conditions, truncated power law was reported for passive periods for mental illness, whereas lognormal distribution was found for dementia [14]. The distribution of active periods was classically characterised by stretched exponentials [2,12,15,21,22], although some studies reported lognormal [11], truncated power law [10], or even power-law distributions for periods under 35 minutes [9]. In contrast, the study that used negligibly small threshold to separate active and passive periods according to their different theoretical approach found the exact opposite: the distribution of active periods followed power law, and the passive periods exhibited exponential distribution.

Beyond analysing the statistical distributions, the scale-free behaviour of human locomotor activity can also be assessed in the frequency domain, where the scale independence is indicated by the presence of $1/f$-like fluctuations, which has been carried out in the literature both in healthy individuals and in subjects with various health conditions for diagnostic purposes, mainly using Detrended Fluctuation Analysis (DFA) and separately for sleep and wakefulness [23–26]. Previously [3], focusing on the power spectral densities, we found that both activity data and the underlying acceleration signals generally follow the same spectral characteristics, regardless of the actigraphic methodology chosen. A key component of these universal characteristics is the presence of $1/f$ noise observed over at least 2 decades above the daily periodicity, see [3] for details. Although the presence of $1/f$ noise in activity signals had been noted by others before [12], we proved with our systematic investigation that even the raw acceleration of our wrist exhibits the same spectral properties, thus $1/f$ noise is an inherent part of human locomotor activity. Since then, we have confirmed the universality of the spectral characteristics we found in additional cohorts of healthy subjects [27].

3. Results

The differences in the conclusions drawn by the relevant studies may stem from the fact that different groups of subjects were examined and that, as discussed earlier, varying data processing and statistical analyses were applied. To understand the impact of the latter, we aimed to apply the main generalised methodological steps of these approaches to the same actigraphic dataset and compared the results obtained.

For this, we used an actigraphic acceleration dataset that we have previously examined in numerous works and that is publicly available [3,5–7,28]. The dataset consists of 10-day raw acceleration recordings from 42 healthy, free-living individuals, measured at a 10 Hz sampling rate within the ±8 g range using an actigraphic device placed on their non-dominant wrist, specifically developed to store raw triaxial acceleration data.

From the acceleration data, we determined 20 different types of activity signals for each subject with an epoch length of 1 minute (i.e., 1 activity value per minute), as described in Section 2.1. These activity signals were then divided into active and passive periods according to various threshold rules, after which the distributions of these periods were examined using different approaches and then we derived the best-fitting model from the activity data, as presented in Section 2.2. The parameters corresponding to these steps, which may vary across different methodologies, are presented in Table 1. In the following, to evaluate the impact of the individual parameters, we modified one or two of them at a time, while following the standard settings for the others. We assessed the effect of changing the parameters based on the shape of the active and passive period distributions, and the best-fitting model fitted to them. We used a widely adopted Python package *powerlaw* [29] to perform MLE-based fitting of power law, truncated power law, lognormal, exponential, and stretched exponential models.

Table 1. Standard settings of data process and analysis

Parameter	Value/Setting
Preprocessing	*FMpost*
Activity metric	*ZCM*
Threshold	*100% of mean*
Individual/pooled	*pooled*
Fitting range	*whole range*
Method of fit	*MLE*
Goodness of fit metric	*LRT*

3.1. *Effect of processing actigraphic data*

To investigate the impact of different activity determination methods, we modified the first two parameters (preprocessing technique and activity metric) of the processing pipeline described in Table 1., while keeping the parameters of all subsequent steps fixed according to standard settings. This allowed us to generate 20 distinct activity signals for each motion through combinations detailed in Section 2.1, and to compare the distributions of their active and passive periods, as seen in Fig. 1. It is evident that the CCDF and PDF of the passive periods differ significantly for several activity determination methods.

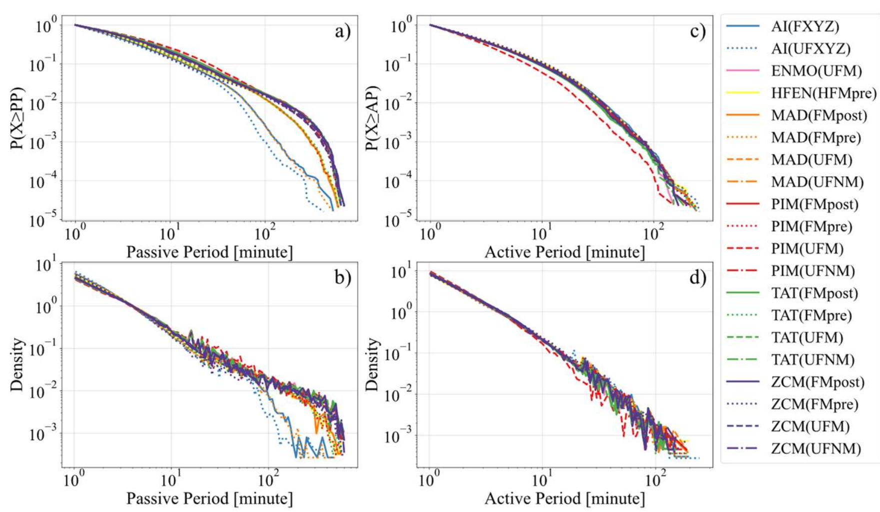

Fig. 1. The CCDF and PDF of pooled passive (subplots a) and b), respectively) and active periods (subplots c) and d), respectively) in the case of all the examined activity determination methods. The threshold separating active and passive periods was set according to the default settings (Table 1.), i.e., the mean activity level.

However, for the activity signals derived using the three classical metrics commonly applied in scale-invariance analyses (i.e., PIM, ZCM, and TAT), the distribution functions closely mirror each other regardless of how the acceleration signal was preprocessed, only the TAT(FMpre) shows a slight deviation. In the case of active periods, among all combinations, only the distribution of PIM(UFM) shows a slightly greater deviation. This can be attributed to the fact that Earth's gravity (i.e., the g) is not removed from the acceleration signals in this configuration, which has previously led to divergent behaviour in various types of our analyses [3,6].

In contrast, for the AI and MAD activity metrics, the resulting curves vary considerably depending on the type of preprocessing. This observation is further supported by the tables of Fig. 2. As shown in Fig. 2. a), when the threshold is set to 70% or 100% of the mean activity level, the best-fit model is consistently the same except for AI(UFXYZ), MAD(FMpre), and PIM(UFM): lognormal for $x_{min} = 1$ minute, and truncated power-law for $x_{min} = 10$ minutes. Fig. 2. b) confirms this and further highlights that, for a given preprocessing technique, changing the fitting range (a point we return to in Section 3.2) leads to transitions between best-fit models in a similar sequence across the classical metrics. This aligns with our previous work, which showed that under identical preprocessing conditions, most of these activity metrics yield activity signals that are strongly correlated [12].

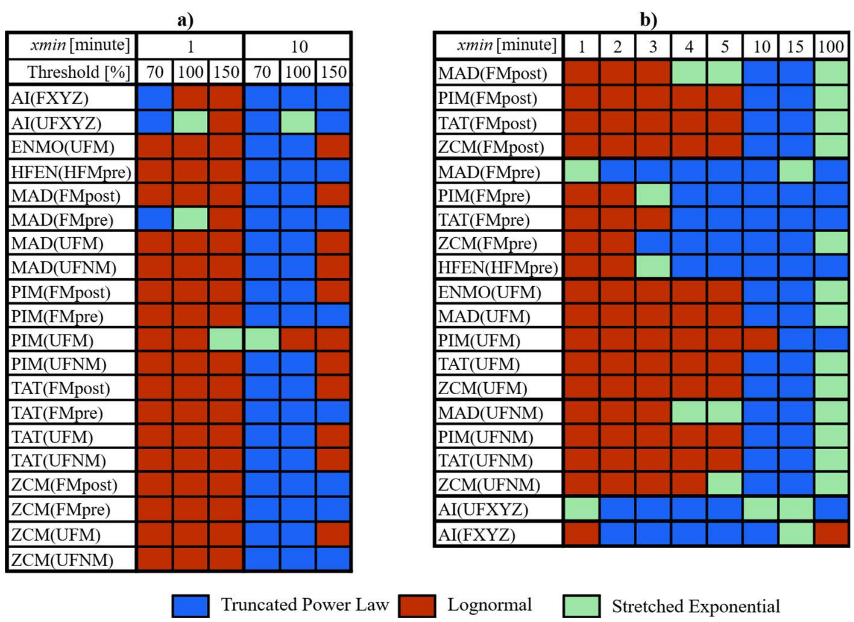

Fig. 2. a) Best-fitting models (indicated by colours) on pooled passive periods, determined by MLE using the LRT goodness-of-fit metric, as a function of x_{min} (1 or 10 minutes) and threshold rule (70%, 100%, or 150% of mean activity level), shown for all the examined activity determination methods (coloured rows). b) Same as in a), but x_{min} is incremented in finer steps, and the threshold is fixed at 100% of the mean activity level.

Although these results may suggest that changing the activity metric to any of the three classical ones commonly found in the related literature does not lead to differences in the outcomes, this may depend on the parameters of the further steps as well. This is clearly illustrated in Fig. 2 a): when the threshold rule is changed from 100% to 150% of the mean activity level and $x_{min} = 10$ minutes, the best-fit model for activity signals derived using these classical metrics is lognormal in half of the cases and truncated power-law in the other half. Notably, 150% of the mean activity level approximately corresponded to the mean of non-zero activity values in our datasets, and as it is a very common threshold rule in the related works [1,2,15], it underscores that the conclusions of such analyses can be influenced by the specific activity determination method employed.

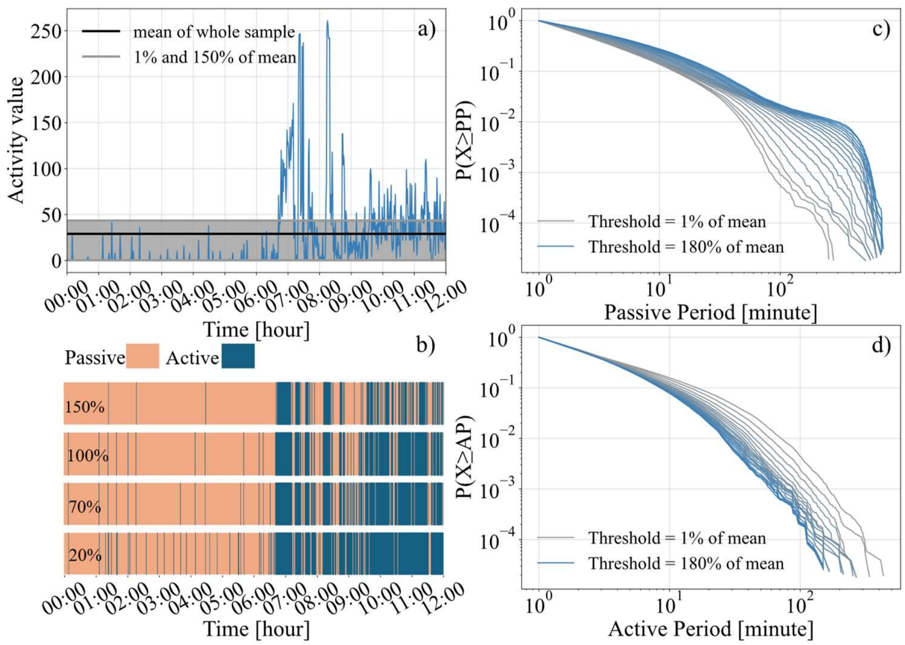

Fig. 3. a) The first 12 hours of an individual's ZCM(FMpost) activity signal. The black horizontal line indicates the mean activity level, while the grey shaded area highlights the interval from 1% to 150% of it. b) Formation of passive and active periods under different threshold rules. The CCDF of passive (c) and active (d) periods as the threshold increases gradually from 1% (grey) to 180% (blue) of the mean activity level.

This effect of choosing different thresholds is well illustrated in Fig. 3, where it can be observed that the threshold used to determine active and passive periods can also influence the shape of the PDF and CCDF. In Fig. 3 a), the activity signal (ZCM(FMpost)) from an individual's first 12 hours of the day is shown. In this case, when the threshold is defined as the mean of the entire 10-day recording, several short active periods appear between midnight and 6:30 AM – likely reflecting brief movements during sleep. As shown in Fig. 3 b), lowering the threshold results in long passive periods (e.g., sleep) being fragmented into shorter segments due to the emergence of short intervening active periods. Consequently, the number of short active episodes rises, which in turn elevates the CCDF of active periods (Fig. 3 d)), while the number and duration of long passive periods decrease, altering the shape of their CCDF (Fig. 3 c)). Conversely, when the threshold is increased to 150% of the overall mean activity level, such brief nighttime movements are no longer classified as active periods, resulting in passive periods that can extend to over 600 minutes. Due to these phenomena, some studies analyse sleep and wake phases separately, applying different thresholds for each [1,13].

Altogether, the choice of threshold clearly influences the distribution of passive periods, as well as the selection of the best-fitting model. At very low threshold values, under certain settings, the stretched exponential can emerge as the best fit, while increasing the threshold tends to favour the truncated power law as the dominant model. This result is consistent with findings in the literature: studies that identified a best-fitting model other than the power law or truncated power law (or that exclusively fitted these models) typically used a threshold less than or equal to the mean [9,13,14]. However, as shown in Fig. 2, the relationship between the threshold level and the best-fitting model strongly depends on the value of x_{min}. Moreover, when the threshold is set at 150% of the mean, the lognormal distribution can also emerge as the best fit for specific values of x_{min}. These findings highlight that such analyses cannot be decoupled from the choice of fitting range.

3.2. *Effect of statistical approaches*

Once the active and passive periods are obtained, their distribution can be examined in various ways, as described in Section 2.2 in relation to the closely related studies. However, when examining the distributions, and especially when fitting models to them, the individual analytical steps can strongly influence the conclusions drawn, which is a well-known concern across scientific disciplines and calls for careful consideration by researchers. A substantial portion of the related studies also investigated the impact of modifying certain parameters (e.g., threshold [1,15,21,22], or epoch length [14,15], which were not found to have a substantial effect) and, alongside their conclusions, acknowledges the associated limitations and potential sources of variability [1,15]. Therefore, beyond highlighting the possible impact of key steps in statistical approach, we place particular emphasis on their sensitivity within the context of the methodology of actigraphy. In line with this objective, we examined either the distribution curves or the best-fitting models of datasets generated in different ways in the previous section.

The impact of processing steps on the distributions can not only be visualised, but also evaluated by fitting several models, from which the best-fitting one can be algorithmically selected using a goodness-of-fit metric. Model fitting is typically performed using either LS or MLE, as discussed in Section 2.2. In the case of LS, a function is fitted to a distribution curve, meaning that the choice of fitting to the PDF or the CCDF – as well as the methodological trade-offs associated with each – can influence the outcome. In our analysis, we identified several settings where different distributions emerged as the best fit depending on whether MLE was applied to the data or LS to either the PDF or the CCDF.

As also demonstrated by the above results, the selection of the model fitting range (i.e., the appropriate interval of active and passive periods considered) has a substantial influence on the outcome. In the case of LS, the effect of the aforementioned distortions can be mitigated by appropriately selecting the fitting range (x_{min} and x_{max}), thereby excluding the extremes and fitting only to the central portion of the distributions, which is a common practice in the related literature [2,12,15]. In contrast, when using MLE, most studies determine only the x_{min} value, since excluding the upper end of the data (i.e., applying x_{max}) is typically avoided unless supported by theoretical justification, particularly in the case of heavy-tailed distributions. Based on these considerations – and for the sake of simplicity, given that the sensitivity of the model fitting to actigraphic methodology is already apparent – we focus in the following sections on assessing the effect of x_{min} only.

The tables in Fig. 2 clearly demonstrate how substantially the best-fitting model (which, in all cases, was found to be either the truncated power law, lognormal, or stretched exponential) varies with the value of x_{min} for each combination of activity determination method and threshold rule. The results from Fig. 2 b) are summarised in Fig. 4, which depicts how the best-fitting models for passive periods – derived from 20 different activity signals using the default settings – change as a function of x_{min}. While the lognormal distribution is selected most often when fitting across the entire data range, increasing x_{min} tends to favor the truncated power law, and fitting only to the upper end of the data (i.e., tail of the distribution) frequently yields the stretched exponential as the best-fitting model. Although these shifts in the best-fitting models naturally follow from the shape of the empirical distributions, it is important to emphasise that the specific way in which the best-fitting model changes with x_{min} substantially varies depending on the activity determination method and threshold rule used.

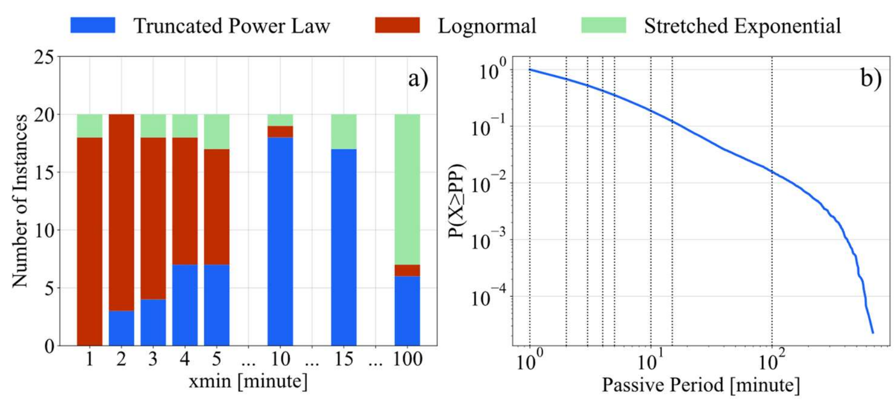

Fig. 4. a) The distribution of the best-fitted models (indicated by colours), determined by MLE using the LRT goodness-of-fit metric, for the pooled passive periods of activity data generated using the 20 different activity calculation methods as a function of changing x_{min}. b) An example for the CCDF of pooled passive periods for ZCM(FMpost) activity data, and the fitting regions determined by x_{min}.

The goodness-of-fit metric used to select the best-fitting model among the candidates can be defined in various ways, potentially influencing the outcome. However, across the 5 goodness-of-fit metrics we examined, we observed only minor differences in the selected best-fitting models. Notably, the LRT and Reduced χ^2 tended to favour the truncated power law, whereas SSE, AIC, and BIC often selected the pure power law instead. Since SSE cannot adequately capture the goodness of the fit in the heavy tail of the distribution on a log-log scale – and both AIC and BIC are frequently approximated using SSE –, we mainly relied on LRT to determine the best-fitting model throughout the analysis.

3.3. *Comparison of specific methodologies*

In the previous sections, we showed how the different analytical steps can influence the final results; however, the effects of these steps could not always be fully distinguished; moreover, we were especially interested in the joint effect of the steps and the actigraphic methodology. For this, we applied the data processing approaches of the most relevant studies [1,2,8–15] from the past 20 years investigating the scale-free nature of human motor activity on the same actigraphic dataset, and then compared the resulting distributions and best-fit models to assess their methodological dependence. It is important to emphasize that among these studies, two cases examined exclusively individuals suffering from dementia [14] and mental illness [1] (and in the latter case, the device was placed on the hip rather than the non-dominant wrist), but we considered the inclusion of these works valuable, as their analytical approaches, findings and conclusions were in line with the others from a methodological perspective. Moreover, because insufficient information was available about processing actigraphic data in several cases of the related works, simplifications and approximations were necessary for their comparison. Ultimately, we examined the CCDF of active and passive periods produced by 9 different combinations of activity determination methods and threshold rules, which we will henceforth refer to as data processing combinations.

Unravelling and breaking down the statistical approaches into basic steps is even more challenging, as the specific aims and theoretical approach of the related studies often differ. Some studies started from a theoretical assumption, premising a given distribution and seeking the parameter values that provided the best fit [2,9,12], while others examined the distributions in the light of changes in various parameters [1,15]. Moreover, as not all steps are known, or specific data processing techniques (e.g., k-means-based active-passive separation [13], rescaling [15], sleep-wake separation [1,13], or specific combination of such segments [11,12]) may have been applied, the exact reproduction of the complete analytical pipeline of each study is infeasible. Although this prevented us from fully reproducing the complete methodology of the related studies, this does not pose a limitation in examining the robustness against data processing, as even when excluding model fitting from the examined steps and focusing solely on the influence of the most basic ones (actigraphic methodology and the selection of $x_{\min}$), clear variations can already emerge.

As shown in Fig. 5, comparing the 9 different data processing combinations, substantial differences between the CCDFs of passive periods were primarily observed among the three classical metrics in relation to the applied threshold levels. This aligns with the observations in Section 3.1, and the CCDFs changed depending on the threshold in the same manner as observed in Fig. 3 c).

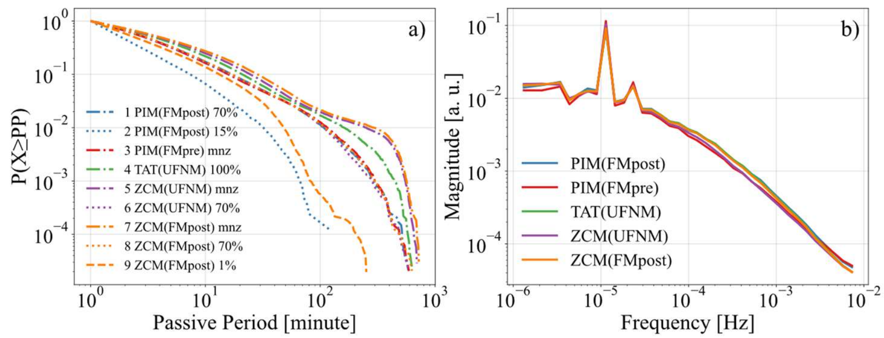

Fig. 5. a) The CCDF of pooled passive periods for the different – numbered in the legend – data processing combinations (i.e., activity determination methods and threshold rules) found in the literature. b) The PSD (whose variance was reduced using Daniel method – i.e., the magnitude of frequency components was averaged within logarithmically spaced bins) of the same type of activity signals, each of them ensemble-averaged over the subjects (for more details, see [3]).

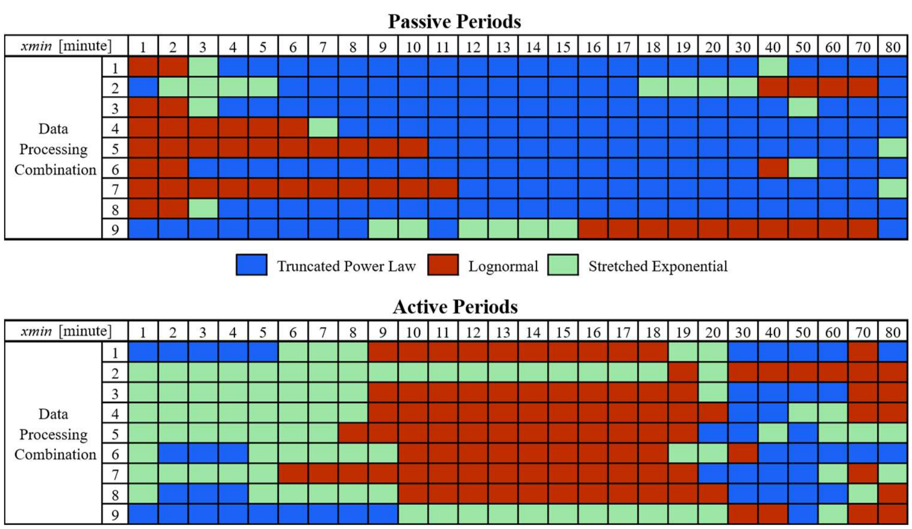

Fig. 6. The best-fitting models (indicated by colours), determined by MLE using the LRT goodness-of-fit metric, for the different data processing combinations (rows) found in the literature as a function of x_{min} (columns) for pooled passive (up) and active periods (bottom). The data processing combinations (i.e., activity determination methods and threshold rules) are numbered as for Fig. 5 to facilitate comparability.

Furthermore, Fig. 6 demonstrates the sensitivity of the statistical approaches by comparing the best-fit models for the passive and active periods from the 9 data processing combinations, with respect to the x_{min} parameter. In line with Section 3.2, for the full range of passive periods, the lognormal model fits in most cases, and as we narrow the range, the truncated power law model increasingly emerges, followed by the stretched exponential model at the end of the distribution. Contrary to this, in the case of the active periods, we observed a different order of best-fitting models. Both for passive and active periods, the exact pattern of transition between models differed in many cases of the 9 combinations. The only exceptions to the observed orders are the two settings (2 and 9), where the threshold is very small, as discussed in Section 3.1.

Overall, we observe that for the 9 data processing combinations derived from the related literature, even when examining the same movements there was no fitting range for which the best-fitting model would have been the same. Additionally, depending on the x_{min}, almost every method exhibited different patterns of how the best fit changed. In contrast, the spectra of both multi-day acceleration signals and the activity signals derived from them follow universal, $1/f$ noise-based characteristics, as shown in Fig. 5 b), regardless of the actigraphic methodology. Moreover, the spectral approach does not require segmenting the data into active and passive periods, and consequently, it does not demand the selection of a threshold level, which otherwise, as demonstrated above, can significantly influence the conclusions of statistical analyses. Although estimating the exponent of the power-law function is also essential in analysing $1/f$ noise, this does not seem to compromise the method's robustness to variations in data processing, as the spectra of activity signals determined in different ways are nearly identical.

We must remark that throughout this study, we focused solely on examining the best fits on the statistical distributions, as analysing the fitted model parameters was complicated by the fact that different studies applied specific steps and approaches due to their unidentical and specific objectives. Moreover, numerous factors could influence the results, making such comparisons highly complex, thus requiring further specialised, detailed analysis. On the other hand, for our aims, this was not necessary, as the lack of robustness of statistical analyses with respect to actigraphic methodology and data processing was already clearly apparent.

4. Conclusion

The statistical investigation of the scale-free nature of human locomotor activity naturally offers several advantages and is a self-evident approach when linked to theoretical models. However, our work specifically proves that even when examining the same dataset using the different actigraphic and statistical methodological approaches from the most relevant studies in the field, different conclusions can be reached. Particularly, we showed that even when only considering the different data processing steps (i.e., the method used to derive the activity data from the raw acceleration signal, and the threshold rule applied to divide the activity signal into active and passive periods) of the related studies, these steps have an influence on the statistical distribution of the same data. Furthermore, when we also took into account the different model fitting approaches (e.g., data pooling, fitting method and its range, and goodness-of-fit metric) of these studies, they caused additional differences in the evaluation of the distributions that could not even be fully separated from the variations introduced by the aforementioned data processing steps. This is well illustrated by the fact that the best-fit models – from which the related works drew their conclusions – changed differently as the fitting range increased, depending on the studies' specific data processing steps, both for active and passive periods.

In contrast, the spectra of the differently determined activity signals follow the same, universal $1/f$ noise-based characteristics, and in this approach, it is not necessary to divide the data into active and passive periods. Since the statistical approaches of studies in the field of spatial mobility are similarly diverse [16], yet human displacement also exhibits the same spectral characteristics [4,5], these results suggest that the scale-free nature of human dynamics can be examined more robustly in the frequency domain. The extent to which statistical and spectral approaches provide information about the characteristics of individual movements and how they can be applied for classification, such as in identifying markers of various health conditions, is an important area for further research, in which the spectral approaches are receiving increasing attention [24,26,30,31].

Acknowledgements

This research was supported by project TKP2021-NVA-09, which has been implemented with the support provided by the Ministry of Innovation and Technology of Hungary from the National Research, Development and Innovation Fund, financed under the TKP2021-NVA funding scheme.

Bálint Maczák was supported by EKÖP-24-3-II University Research Scholarship Programme of the Ministry for Culture and Innovation from the National Research, Development and Innovation Fund.

ORCID

Krisztina Panna Komáromi - https://orcid.org/0009-0008-8137-7032
Bálint Maczák - https://orcid.org/0000-0002-1867-5861
Gergely Vadai - https://orcid.org/0000-0002-2720-3380

References

[1] J. J. Chapman, J. A. Roberts, V. T. Nguyen and M. Breakspear, "Quantification of free-living activity patterns using accelerometry in adults with mental illness", *Sci Rep* 7 (2017), 43174.

[2] M. Hirose, et al., "Altered distribution of resting periods of daily locomotor activity in patients with delayed sleep phase disorder", *Front. Psychiatry* 13 (2022), 933690.

[3] B. Maczák, Z. Gingl and G. Vadai, "General spectral characteristics of human activity and its inherent scale-free fluctuations" *Sci Rep* 14 (2024), 2604.

[4] G. Vadai, A. Antal and Z. Gingl, "Spectral Analysis of Fluctuations in Humans' Daily Motion Using Location Data", *Fluct. Noise Lett.* 18 (2019), 1940010.

[5] B. Maczák, A. Antal and G. Vadai, "The Noise of Our Daily Motion: General Spectral Characteristics of Human Mobility and Activity", *Fluct. Noise Lett.* 24 (2024), 2440061.

[6] B. Maczák, G. Vadai, A. Dér, I. Szendi and Z. Gingl, "Detailed analysis and comparison of different activity metrics", *PLOS ONE* 16 (2021), e0261718.

[7] B. Maczák, K. P. Komáromi and G. Vadai, "Different Aspects of the Scale-Free Nature of Human Activity – Examination of Its Spectral and Statistical Properties", in *2023 International Conference on Noise and Fluctuations (ICNF), IEEE Xplore* (2023), 1-4.

[8] D. Chialvo, et al., "How we move is universal: Scaling in the average shape of human activity", *Papers in Physics* 7 (2015), 070017–070017.

[9] O. B. Fasmer, E. Hauge, J. Ø. Berle, S. Dilsaver and K. J. Oedegaard, "Distribution of Active and Resting Periods in the Motor Activity of Patients with Depression and Schizophrenia" *Psychiatry Investig* 13 (2016), 112–120.

[10] J.-H. Lee, et al., "Statistical properties of human activity and criticality in active behavior" *EPL* 126 (2019), 68001.

[11] A. Búzás, et al., "Hierarchical organization of human physical activity", *Sci Rep* 14 (2024), 5981.

[12] J. K. Ochab, et al., "Scale-free fluctuations in behavioral performance: Delineating changes in spontaneous behavior of humans with induced sleep deficiency", *PLOS ONE* 9 (2014), e107542.

[13] M. Kawabata, et al., "Temporal organization of rest defined by actigraphy data in healthy and childhood chronic fatigue syndrome children", *BMC Psychiatry* 13 (2013), 281.

[14] S. E. Huber, et al., "Assessment of Fractal Characteristics of Locomotor Activity of Geriatric In-Patients With Alzheimer's Dementia" *Front Aging Neurosci.* 11 (2019), 272.

[15] T. Nakamura, et al., "Of Mice and Men — Universality and Breakdown of Behavioral Organization", *PLOS ONE* 3 (2008), e2050.

[16] M. Karsai, H.-H. Jo and K. Kaski, "Bursty Human Dynamics", *Springer International Publishing* (2018).

[17] A. Clauset, C. R. Shalizi and M. E. J. Newman, "Power-Law Distributions in Empirical Data", *SIAM Rev.* 51 (2009), 661–703.

[18] S. M. Burroughs and S. F. Tebbens, "Upper-truncated Power Laws in Natural Systems", *Pure appl. geophys*. 158 (2001), 741–757.

[19] N. Marshall, et al., "Analysis of Power Laws, Shape Collapses, and Neural Complexity: New Techniques and MATLAB Support via the NCC Toolbox", *Front. physiol.* 7 (2016), 250.

[20] S. Petrovskii, A. Mashanova and V. A. A. Jansen," Variation in individual walking behavior creates the impression of a Lévy flight" *PNAS* 108 (2011), 8704–8707.

[21] T. Nakamura, et al., "Universal Scaling Law in Human Behavioral Organization", *Phys. Rev. Lett.* 99 (2007), 138103.

[22] W. Sano, et al., "Enhanced Persistency of Resting and Active Periods of Locomotor Activity in Schizophrenia", *PLOS ONE* 7 (2012), e43539.

[23] K. Hu, et al., "Non-random fluctuations and multi-scale dynamics regulation of human activity", *Physica A: Stat. Mech. Appl.* 337 (2004), 307–318.

[24] P. M. Holloway, et al., "Complexity analysis of sleep and alterations with insomnia based on non-invasive techniques", *J. R. Soc.* 11, (2014), 20131112.

[25] E. Gudowska-Nowak, et al., "Seeking for a fingerprint: analysis of point processes in actigraphy recording", *JSTAT* 2016 (2016), 054034.

[26] C. Lin, et al., "Different Circadian Rest-Active Rhythms in Kleine-Levin Syndrome: A prospective and case-control study", *Sleep* 44 (2021), zsab096.

[27] B. Maczák, C. G. Horváth, R. Bódizs and G. Vadai, “Revealing the Generality of 1/f Noise Based Spectral Characteristics of Human Activity Across Different Datasets”, in 2023 *International Conference on Noise and Fluctuations (ICNF), IEEE Xplore* (2023), 1-4.

[28] B. Maczák, G. Vadai, A. Dér, I. Szendi and Z. Gingl, “Raw triaxial acceleration data of actigraphic measurements – supporting information of “Detailed analysis and comparison of different activity metrics””, doi:10.6084/M9.FIGSHARE.16437684 (2021).

[29] J. Alstott, E. Bullmore and D. Plenz, “powerlaw: A Python Package for Analysis of Heavy-Tailed Distributions”, *PLOS ONE* 9 (2014), e85777.

[30] L. Gao, et al., “Fractal motor activity regulation and sex differences in preclinical Alzheimer’s disease pathology”, *A&D: DADM* 13 (2021), e12211.

[31] N. Furutani, et al., “Complexity of Body Movements during Sleep in Children with Autism Spectrum Disorder”, *Entropy* 23 (2021), 418.