

Motionwatch8 Wrist Actimetry Data Analysis: from ambulatory recording to real life monitoring.

Gianluigi G. Delucca

<https://orcid.org/0000-0002-7680-1256>

Parma, Italy

email: ggdelucca@medricerca.it

Abstract— On the National Sleep Research Resource (NSRR) platform, there's a year's worth in real-life wrist actigraphy data using a Motionwatch8. The data analysis suggested in this paper reveals a data structure that allows the design of parameters to be used in real life monitoring and may drive the developments of devices. The analysis is totally data driven and connects in a quantitative way the wrist movements (time base seconds) to the circadian rhythms (time base days). A simple example of parameters is described.

Keywords – Actimetry; Motionwatch8; Circadian .

I INTRODUCTION

The Author demonstrated the feasibility of the use of the wrist actigraph Motionwatch8 (CamNtech, Cambridge, UK) for real life, long term (one year), short epochs (1 sec) recordings [1]. The raw data are available on the National Sleep Research Resource (NSRR) platform [2]. This paper wish to explore that dataset looking for hints that could allow to develop monitors of that type of data to be used in real life. Section II Raw data - describes the raw data with the use of runs; Section III Data analysis – describes the creation and uses of Wrist Actigraphy Pulse (WAP); Section IV Discussion - points out some of the open issues and we conclude in Section V - Conclusion and Future work.

Unless otherwise noted, computations and graphs were made by programs written by the author in Octave [3]. In order to make the paper easier to read, the cultural environment of the model of the data is described in Appendix 1.

II RAW DATA.

Technical details related to the data acquisition of the one year recording are fully explained elsewhere [1]. In short, the MotionWatch8 samples at 50 Hz the output of an accelerometer filtered in the frequency range 3-11 Hz. As a measure of the movement, the device computes one value each second in a custom unit called “Counts”. Each stored sample brings the following information: Movement (in Counts), Time (year, month, day, hour, minute, second), Light (Lux). Out from the 31,576,501 samples of the year, there are 24,936,212 zeros (78.97%), 1,034,485 n/a (3.28%) and 5,605,804 (17.75%) non-zero values.

Visual inspection of the data

The time base is more than 30 millions seconds and that makes it impossible to visually evaluate the distribution of Counts values over time. As a consequence of the graphics limitations, it is needed to develop strategies to extract meaningful features from the dataset and display them.

Runs - Distribution in time

Since most of the data are zeros, we need to focus on them. A zero value brings at least three information:

- Value – The zero value means that the measured intensity of the movement did not reach the “noise removal threshold” decided by the manufacturer.
- Time – When the sample has been taken
- Position – The zero value sample has a position inside a series of zeros, from 1 to several thousands long.

As introduced in [1], one option to partially exploit the position information is to use the runs approach, i.e. to measure the length of the series of zeros and a non zero values. For instance, the series of Counts 1,1,1,0,1, 20, 3, 0, 0, 0, 3, 35, 128, 0, 0, 5, 78, 0, 0, 0. get runs 3,1,3,3,3,2,2,3.

In order to keep the yes/no information of the zero/non zero value, we use a negative sign for the runs of zeros and the series of runs becomes 3,-1,3,-3,3,-2,2,-3. We show in Figure 1 one day where Counts are displayed over time in the traditional way by using the manufacturer software Motionware [4] and the same time interval is in Figure 2 as Runs over time, each run positioned at its ending time.

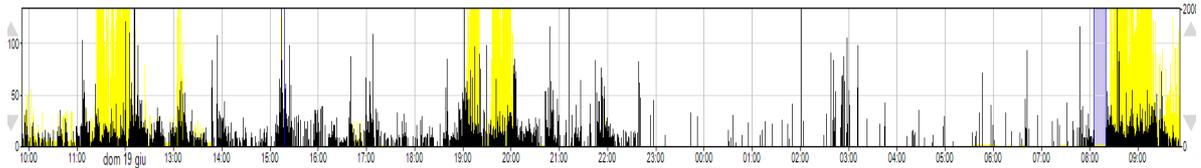


Fig.1 Activity of one day plotted in line. Start time 09:52:00. X axis: Seconds – Y-axis: Counts Yellow: Light in Lux – Grey: n/a data.

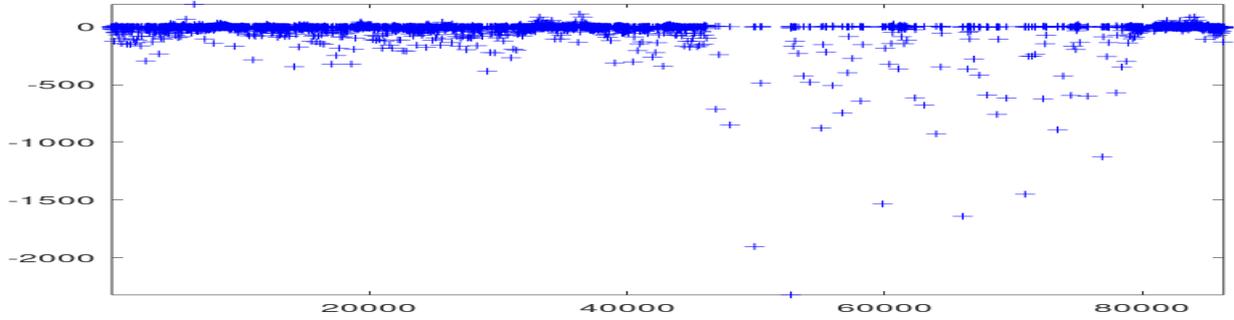


Fig.2 Start time 09:52:00. X axis: Seconds Y-axis: length of the series in seconds: non-zero series up, zero series down.

Runs - Distribution of lengths

In the recorded year there is a sequence of 1.429.113 zero series and, of course, the same amount of non-zeros ones; the zero series can be long up to 7400 seconds, the non-zero series up to 623 seconds. The distribution of lengths of runs in the year is shown in Figure 3 left side, where we split the diagram with non-zero series on the right and zero series on the left with a negative sign. There is a strong peak of several hundred thousand series around few short lengths. If we zoom in (Figure 3 right side), we see that the number of zero and non-zero series longer than 10 seconds is a small percent of the total.

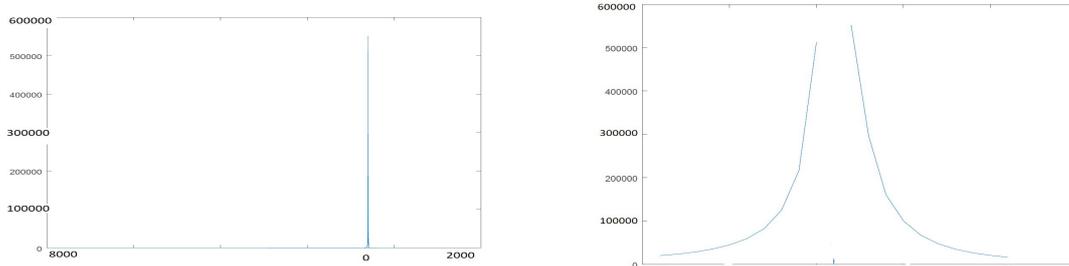


Figure 3. Full year. X axis: length of runs. Zeros series on the left, non-zero series on the right. Y axis: number of runs.

Runs - Sequences

From the exploratory analysis in [1], we also know that the sequence of runs is not random. Figure 4 left side shows that only short non-zero series can move to long zero ones; Figure 4 right side shows that only short non-zero series can move to long zero ones.

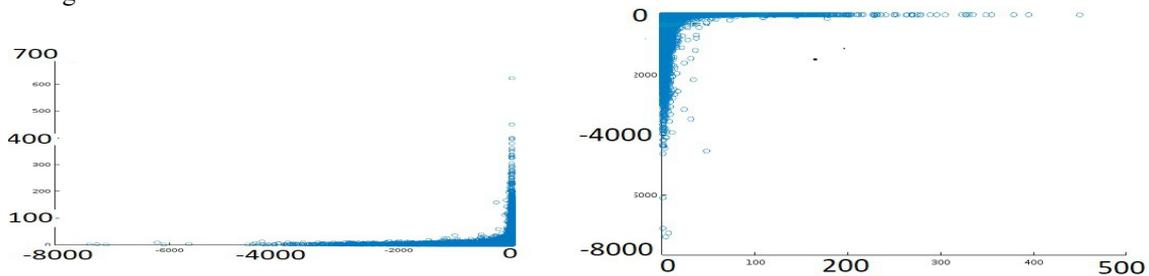


Figure 4. Full year. X axis: length of the run in seconds. Y axis :length of the following run in seconds .

Runs - Data compression

We may look at runs as a way to compress data. If we compute the length of the zero runs there is no information loss because it is possible to reconstruct the original data flow. For the non zero runs the loss of information depends on the

way we store the Counts values. For instance, we may store the sum or the mean of the Counts of each run as shown in Figure 5, but the important point is that we can evaluate for each parameter the loss of information from the compression strategy.

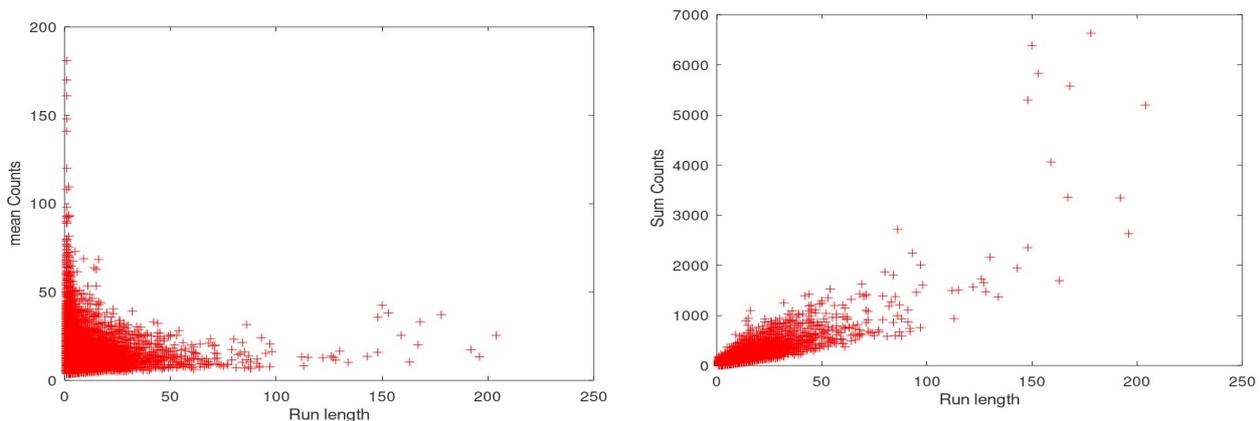


Figure 5 Distribution of Counts. x axis = length of non zero runs in seconds; Right side, y axis =mean Counts of the run ; Left side, y axis =sum of Counts of the run.

Summing up, shifting the focus of the description of raw data from Counts to Runs demonstrated that:

- there is a daily bimodal pattern (Figure 2) ;
- the lengths of runs, both of zeros and non zero, in real life, are limited: roughly 10,000 seconds for zeros and 1.000 seconds for non zero (Figure 3) ;
- the number of those runs is strictly related to the length of the run: shorter the length, more frequent the run (Figure 3);
- the sequence of runs is in part deterministic (Figure 4);
- runs lengths are partially related to the Counts (Figure 5).

III Data analysis

Let's restart from Figure 1, where there is the graph of Counts values on a time axis.

1. We focus on the evaluation of zeros and rewrite the data giving to the samples with a Counts value the only meaning of "non zero", i.e. we transform the series of Counts from 1, 1, 1, 0, 1, 20, 3, 0, 0, 0, 3, 35, 128, 0, 0, 5, 78, 0, 0, 0. to 1, 1, 1, 0, 1, 1, 1, 0, 0, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0.
2. The theory of runs approach helps to underscore that 0 and 1 are not a measure any more and now are symbols that carry only the meaning of over/below the threshold. Selecting +/-1 as symbols is a way to extremely simplify computation. With that selection, the toy series becomes 1, 1, 1, -1, 1, 1, 1, -1, -1, -1, 1, 1, 1, -1, -1, 1, 1, -1, -1, -1, -1, and a simple sum provides the sequence of the lengths of the runs with a sign $L(i) = 3, -1, 3, -3, 3, -2, 2, -3$.

Creode

Recordings in real life require an overall evaluation of the model we use to analyze the data (See Appendix 1). The creode analogy helps to graphically shows the relationship between the part of the raw data described with runs (lengths of symbols sequences) and time. We put the line of the time as the vertical axis, pointing downward, and we put the runs, still using negative and positive values for the zero and non zero runs, on the horizontal axis.

Starting from a 0,0 point, for each second added on the time vertical axis, the runs add one second on the horizontal one. So, on the graph those points follow the diagonals of +/- 45 degrees. If the sum of those runs $R(i) = \text{sum}(L(i))$ is a sinusoidal type, there will be a fluctuation around the vertical time line.

Instead, we get a progressive negative line as in Figure 6. What we see in Figure 6 is that the path, i.e. the sum of the runs lengths with alternate sign, has a well defined trend.

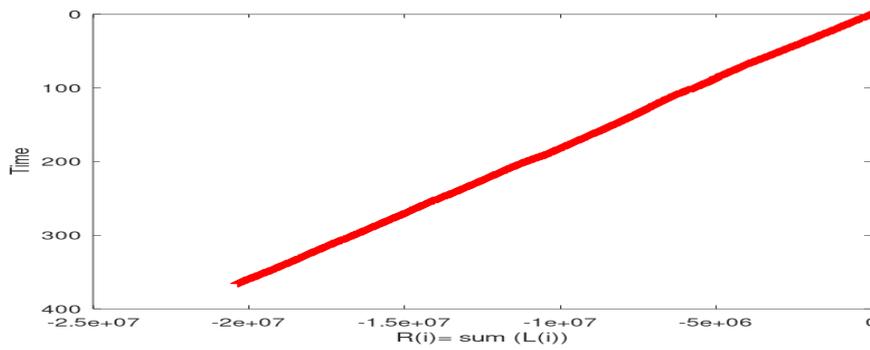


Figure 6. One year X axis =R(i) sum of runs - Y axis = Time in days

We knew (Section II Raw data) that around 80% of the samples of the year is zero. Therefore, after one hundred samples we have 80 negative and 20 positive values that provide a sum of -60. What was not documented until now, as far as we know, is that that general rule applies on a daily basis.

“Components” of R(i)

We can start the analysis with a simple linear interpolation of R(i) that provides results around 0.6, as expected.

When we subtract that linear interpolation from the path, the visual inspection shows (Figure 7) one periodic signal with a sawtooth shape.

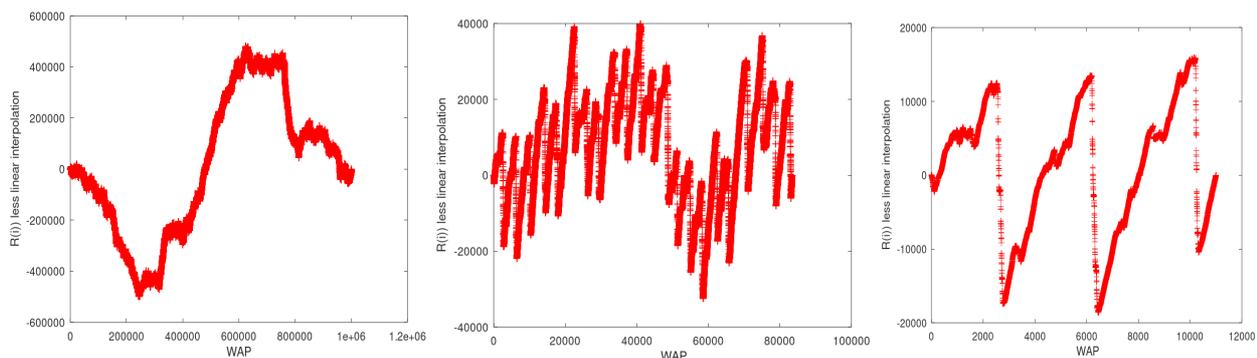


Figure 7 . X axis = Runs, scale 1/20. Y axis = R(i) less the linear interpolation. Left side one year, center one month, right side 3 days.

The “constant”.

Let's consider the “constant” component from the linear interpolation first.

Runs are by definition alternatively positive and negative. If we want to create a constant we have to use a sequence, with a minimum of 2 components like (3,-4). We suggest the term Wrist Actigraphy Pulse (WAP) for such a couple. $WAP(1) > 0 ; WAP(i) = L(i) + L(i+1)$. What we measure with a WAP, is when and how much the length of one burst of activity (non zero run) is locally “compensated” by the length of the following rest (zeros run). In our example, we move from the runs sequence $L(i) = 3, -1, 3, -3, 3, -2, 2, -3$. to the WAP sequence $W(j) = 2, 0, 1, -1$. With the addition operation, we loose information and are not able to reconstruct the original signal. In order to avoid it, we need to store also the length of the WAP $Length\ WAP(i) = L(i) - L(i+1)$ and the sequence becomes $W(j) = (2, 4), (0, 6), (1, 5), (-1, 5)$. We can normalize each WAP against its own length, $nWAP = L(i) + L(i+1) / L(i) - L(i+1)$ and show in that way the contribution of the single WAP to the “constant”.

It may help to understand the new perspective, if we compare the normalized WAPs with the Immobility parameters often used in clinical actimetry , i.e. the percent of zero values of a time interval [1]. The movement interval time m and the following immobility time i provide the immobility ratio $i/m+i$ while the nWAP is the ratio $m-i/m+i$.

The fast sawtooth shaped “component”.

That is obviously a circadian rhythm. If we trace the local maximum and minimum of the W(i) (Figure 8), we can use them to compute some parameters. We know the transformation from W(i) to W(t) because each WAP length is the time interval itself: $t(WAP(i)) = t(WAP(i-1)) + Length\ WAP(i)$.

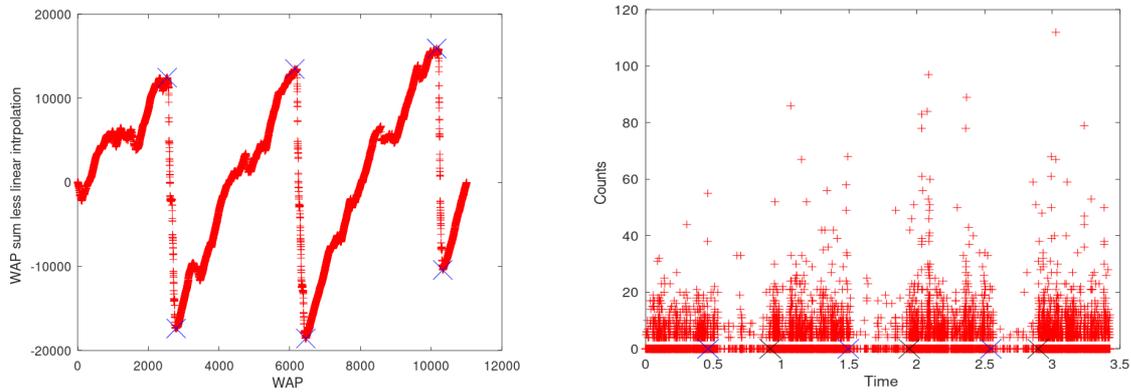


Figure 8 . Left side X axis = WAP. Y axis = sum of W(i) less the linear interpolation. Right side X axis = Time in days. Y axis = Counts.

Table 1 shows some measured parameters of the first 3 days in Figure 8:

- Circadian Max Rhythm = Time interval between two local maximums = (Max2-Max1)
- StepTime Max 1 = Time interval of the first part of the Circadian Max Rhythm = (Min1-Max1)
- StepTime Max 2 = Time interval of the second part of the Circadian Max Rhythm = (Max2-Min1)
- Circadian Min Rhythm = Time interval between two local minimums = (Min2-Min1)
- StepTime Min 1 = Time interval of first part of the Circadian Min Rhythm = StepTime Max 2
- StepTime Min 2 = Time interval of second part of the Circadian Min Rhythm = (Min2-Max2)

Circadian Max Rhythm starts at the beginning of the negative (passive, night) slope, Circadian Min Rhythm starts at the beginning of the positive (active, day) slope.

Table 1

	Day 1 /2	Day 2 /3	Day 3/4
Circadian Max Rhythm	(24:48:53)	(25:14:12)	
StepTime Max 1	(11:00:03)	(10:48:08)	(08:23:28)
StepTime Max 2	(13:48:50)	(14:26:04)	
Circadian Min Rhythm	(24:36:58)	(22:49:32)	
StepTime Min 1	(13:48:50)	(14:26:04)	
StepTime Min 2	(10:48:08)	(08:23:28)	

Counts.

A Counts value brings at least four information. Three reflect the information carried by zeros:

- Value – A Counts value means that the measured intensity of the movement is over the “noise removal threshold” decided by the manufacturer.
- Time – When the sample has been taken
- Position – The Counts value sample has a position inside a series of Counts, from 1 to several hundred long.

One information is new.

- Value – A Counts value compare samples against a scale, decided by the manufacturer

Until now we used only the time information. We know from Figure 5 that there is a rough linear relationship between the length of the non zero runs and the sum of its Counts and therefore, somehow, part of that information is already included in the WAPs evaluation. If adding a weight to the non zero runs will allow to use the remaining information, will depends on the design of the monitor. For instance, in Figure 9 on the upper line there are again the data of Figure

1 (Counts) in line and in polar coordinates. In the lower line there is the sum of Counts of the WAPs, also in line and in polar coordinates. While the two plots in line look similar, in the polar plot the daily distribution of activity is easier to understand. There is also a difference in eidology complexity because there are 13343 samples of Counts and 3294 sums of Counts of WAPs.

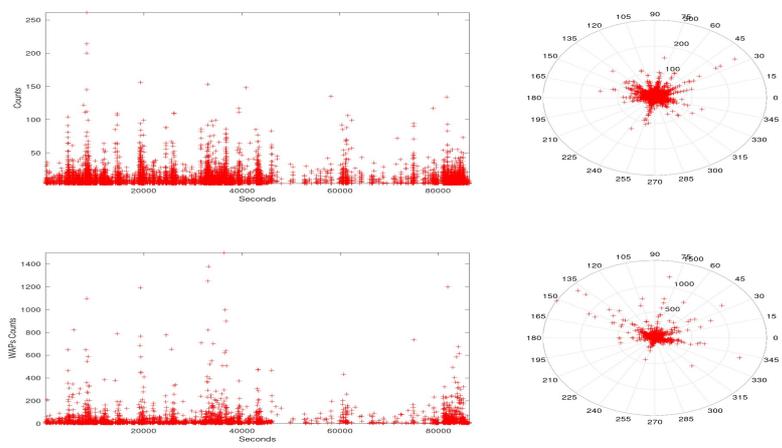


Figure 9 . 24 hours. Upper line: Counts over time in line and in polar coordinates. Lower line: sum of Counts of the WAPs, also in line and in polar coordinates. Time starting point 0 = 09:52 AM

Table 2 shows some measured parameters of the first 3 days in Figure 8:

- Circadian Max Counts = Sum of the Counts of the Circadian Max Rhythm
- StepCounts Max 1 = Sum of the Counts of the first part of the Circadian Max Rhythm
- StepCounts Max 2 = Sum of the Counts of the second part of the Circadian Max Rhythm
- Circadian Min Counts = Sum of the Counts of the Circadian Min Rhythm
- StepCounts Min 1 = Sum of the Counts of the first part of the Circadian Min Rhythm Step = StepCounts Max 2
- StepCounts Min 2 = Sum of the Counts of the second part of the Circadian Min Rhythm

	Table 2		
	Day 1 /2	Day 2 /3	Day 3 /4
Circadian Max Counts	(185067)	(189799)	
StepCounts Max 1	(19438)	(14533)	(8954)
StepCounts Max 2	(165642)	(175286)	
Circadian Min Counts	(180164)	(184205)	
StepCounts Min 1	(165642)	(175286)	
StepCounts Min 2	(14533)	(8954)	

Summing up, shifting the focus from Runs to WAPs allows to understand in a quantitative way the connection between the alternate of positive and negative Runs (time scale in seconds) with the building up of the Sum of Runs (time scale in days) and that offers the base for the development of monitoring parameters. Since WAPs can be calculated in real time, it is possible to construct parameters with any granularity.

Dynamic Time Warping

We can also look at WAPs as the results of a Dynamic Time Warping method based on the events: “Wrist starts moving” and “Wrist stop moving”. That underscore the data driven characteristic of the analysis. The simplified analysis procedure coded in Octave is listed in Appendix 2.

IV. DISCUSSION

Threshold

When we use only the lengths of the runs, we remove the information related of the intensity of the movement and keep the information that the intensity crossed the “noise removal” threshold decided by the manufacturer [1]. That brings the issue of the impact of that threshold on the raw data.

The 80% of zeros means that there is a ratio roughly of 4 zero samples for each non zero sample. We can modify it in several ways. For instance, we may decide that the threshold is too low and that all samples below the new threshold (1,2,...Counts) are zeros. That will increase the lengths of the zero series and lower the ratio. Another way is to change the epochs length (2,3,...sec), as it is traditionally done in clinical actigraphy, and that will increase the length of non zero series and higher the ratio.

In a way, it is a good new. Now that there are expected results, we may find strategies to compare acquisitions with different parameters of the same device and maybe from different devices and hopefully be able to re-use the large amount of actigraphy data available for new parameters development.

N/A samples

If we consider the length of the run itself as a symbol, we can label it. In our toy example, the series of Counts 1, 1, 1, 0, 1, 20, 3, 0, 0, 0, 3, 35, 128, 0, 0, 5, 78, 0, 0, 0. get runs with a sign 3, -1, 3, -3, 3, -2, 2, -3. that will be labeled 1,2,1,3,1,4,5,3. In Figure 10 left side we display approximately one week of the labels flow.

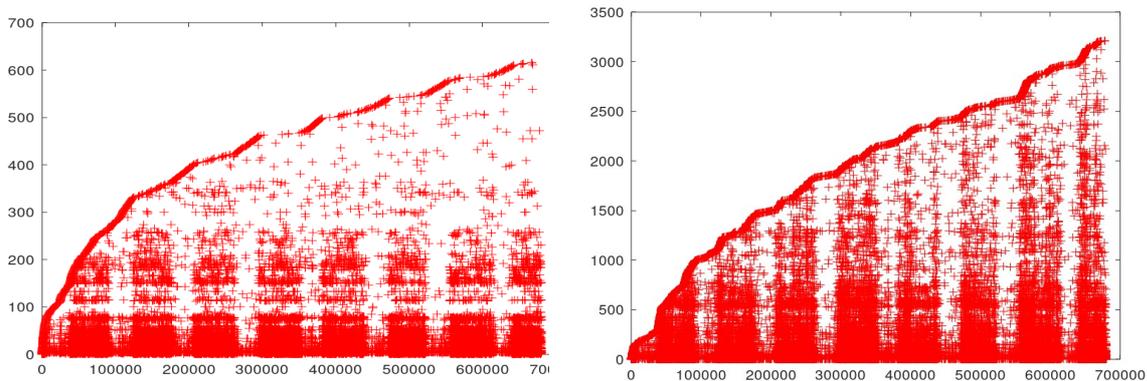


Figure 10. One week. X axis time in seconds, Y axis labels of lengths of runs. Left side labels of runs without Counts weight. Right side labels of runs with Counts weight

It is possible to note “columns” that are easily interpreted as days. “Days” are full of the same labels (few hundreds) while “nights” have a larger range to fill (few thousands), and provide the majority of the new labels. The “columns” also confirm that a limited number of labels is always present and create the daily structure, while many more are added to complete the picture.

If we couple one value related to Counts to the non zero runs, we differentiate among non zero series of the same length and the number of labels will be higher. In Figure 10 right side we see the same week with the new labeling. We may note that the increment of the labels number reverses between right and left. That shows the variability of the efforts, and therefore of the wrist Counts, among days.

From the labels flow we understand that in about a couple of days (roughly 10 thousands runs) the most of the information about the structure is recorded, but that creates one issue. It does not seem possible, at least with the device used, to provide such a recording in real life because there are various situations that will cut the recording into shorter or longer pieces. In our recording, most pieces are long 12 to 24 hours and in such a situation we will never get a stream of data long enough to compute some parameters.

“Components”.

The “components” of $R(i)$ do not exist, of course. The WAP is a way to describe the alternate of activity and rest of the wrist. Using the proposed structure, it is easy to see how variations of the length of the runs will provide both the “constant” $(20-80), (19-79), \dots = (-60, -60,)$ and the other “periodic components” $(21-79), (19-81), (20-78), (18-80), \dots = (-58, -62)$ of the WAP sum. With our acquisition setting, the constant value is about $20/80 = 1/4$, but it would be different for other settings and instruments.

There could be many “components”, both infra and ultradian. They could be tied to the lifestyle, the sleep phenotype, solar monthly cycles, moon monthly cycles, seasonal cycles, ... The year length has been enough to evaluate the daily variation (365 sets) and it is probably enough for ultradian parameters. Infradian “components” may require longer recordings and maybe one year is not enough as a dataset. Since that year is part of a 5 years recording [5], there is the possibility to evaluate other “components”.

V. CONCLUSION AND FUTURE WORK

Even if in this paper the data analysis is made in a more formal way than in the exploratory analysis and some basic results are obtained, it is still a ground breaking type of work that creates questions more than answers.

WAPs are easy to compute and their relationship with the measures from the device is intuitive, but the sequence of runs is open to a variety of analysis methodologies, from Machine Learning to Chaos, that may extract additional information. The suitable methodology is an open issue. The relationship between WAP's length and Counts is unexplored.

In this article only the movement data are evaluated. Two more papers are planned with the analysis of the co-recorded light data from Motionwatch8 and of the correlation between actimetry and heart rate. After those two papers, hopefully, the area of studies of the research project of the wrist actigraphy described in [5] will be better framed and it will be possible to move out from ground breaking type of work to more constructive work and possibly some applications.

REFERENCES

- [1] G.G. Delucca, Wrist Actigraphy Analysis From Motionwatch8 Data. International Journal on Advances in Life Sciences, vol 11 no 1 & 2, year 2019, 75-84 http://www.thinkmind.org/index.php?view=article&articleid=lifsci_v11_n12_2019_8
- [2] Data set available from: <https://sleepdata.org/datasets/oya> [accessed May 2021].
- [3] Octave.org - <https://www.gnu.org/software/octave/about.html> [accessed May 2021].
- [4] <https://www.camntech.com/motionware/> [accessed May 2021].
- [5] Research project: <https://www.actigraphy.eu/actigraphy-analysis-project.html> [accessed May 2021].

In order to provide a direction of the analysis, we have to describe its cultural environment.

Goal of the analysis

The goal of our analysis is not to examine the acquired data, but to predict what is going to happen afterward. “Prediction” means to expect a result from a model of the system under study. That is what it is needed to implement a device that could be used in real life for unsupervised monitoring of patients. Without a prediction, the ongoing flow of data is meaningless. The above implies the hypothesis of a sort of “constant” in the parameters of the model.

Wrist-body connections

The position of the Motionwatch8 (MW8) on the wrist may be described as in Figure A1, where N is the movement as a direct reaction to the environment (e.g. the movement of the car we are traveling in) and S is the movement as requested from a body internal source (e.g. using a tool).

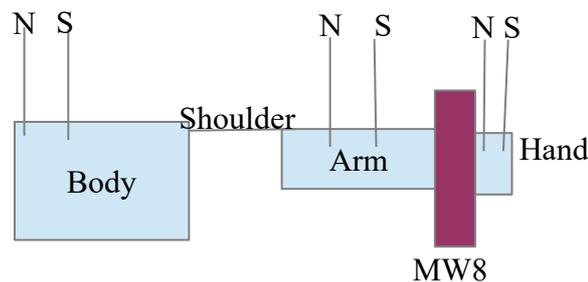


Figure A1 Wrist, arm, body schematics

From the wrist, we cannot distinguish the movement due to the hand (like when you use a tool) or to the arm and body (nearly all voluntary and involuntary movements). We may only say that from the body physical connections we get limits of the values of the measure.

We also do not know the feedback among events. For instance, we see a friend and we move the hand to call him, after a while he/her will arrive and we will shake hands. There is a relationship between the two events that probably influence also the wrist movements inside the time interval in between.

Furthermore, the quantity of movements due to direct impacts on the MW8 as well as movements due to the weight of the device itself, which is “carried by” and not “fixed to”, the wrist is not known.

Creode

As a model of the “environment” of the body we suggest the creode, i.e. the body follows an downhill path with many bumps and different speeds. The creode is unknown, but it does not mean that it has no structure. On the contrary, the life of the body itself is based on the existence of that structure.

While it is not possible to know the next external bump, there are expectations: night and day, the one way direction (time), many more small bumps than big ones, ... Those expectations, both from a long term past (e.g. genetics) and from the short term past, are somehow incorporated in the internal dynamics of the body, so that two equal external shocks/events may provide two different reactions due to the internal dynamics.

It is hard to differentiate the dynamics of the world from the dynamics of the body when the body itself is the result of an adaptation to the world.

In this approach, the “answer” of the body to a “stimulus” from the world is then the reaction suitable to “expected” settings of the environment. For instance, we may understand the jet-lag as one expression of that setting expectation: after the flight the body “search” for the circadian rhythm it is expecting and then try to adapt to what is “available”.

APPENDIX 2

```
Counts= [1, 1, 1, 0, 1, 20, 3, 0, 0, 0, 3, 35, 128, 0, 0, 5, 78, 0, 0, 0];
```

% stop moving

```
j=1
for i=1:size (Counts,2)-1
if ( Counts(i)>0 && Counts(i+1)==0 )
    Stop(j)=i
    j=j+1
endif
endfor
% Stop = 3 7 13 17
```

% start moving

```
j=1
Start(j)=1
j=j+1
for i=1:size (Counts,2)-1
if ( Counts(i)==0 && Counts(i+1)>0 )
    Start(j)=i
    j=j+1
endif
endfor
% Start = 1 4 10 15
```

%WAP length

```
WLength(1)=Start(2)
for i=2:size (Start,2) -1
WLength(i)=Start(i+1)-Start(i)
endfor
WLength(i+1)=size (Counts,2) - Start(i+1)
% WLength = 4 6 5 5
```

%WAP value

```
WValue(1)= Stop(1) - ( Start(2) - Stop(1) )
for i=2:size (Stop,2)-1
WValue(i)= ( Stop(i) - Start(i) ) - ( Start(i+1) - Stop(i) )
endfor
WValue(i+1)= ( Stop(i+1) - Start(i+1) ) - ( size (Counts,2) - Stop(i+1) )
% WValue = 2 0 1 -1
```

%WAP and nWAP

```
W= [WValue;WLength;WValue./WLength ]
%W = 2 4 0.5
% 0 6 0
% 1 5 0
% -1 5 -0.2
```