ACCURACY, BIAS, AND NOTABLE OBSERVATIONS FOR EXPERIMENTAL ACTIGRAPHY DATA

Jürgen Symanzik*, Peter M. Sherick
Utah State University, Logan, UT, USA

William D. Shannon
Washington University, St. Louis, MO, USA

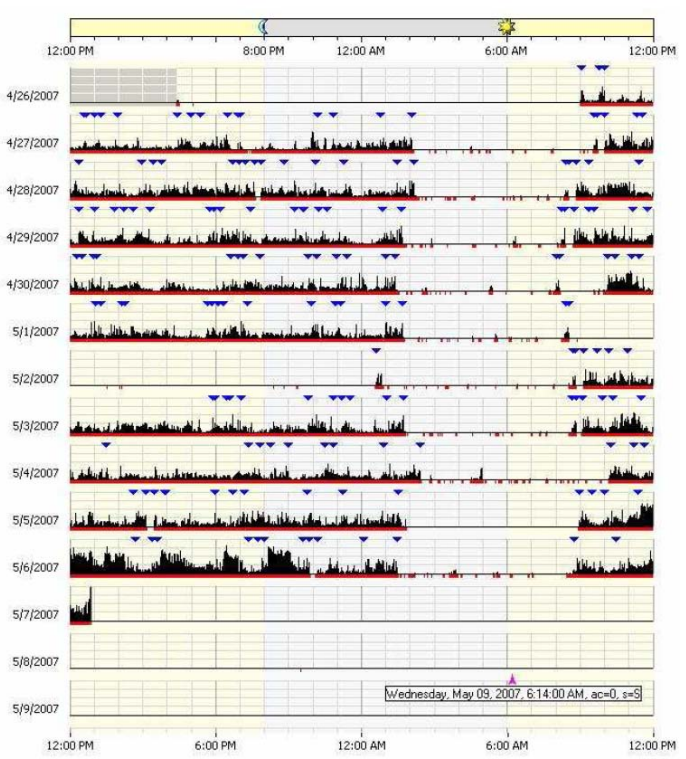
*e-mail: symanzik@math.usu.edu
Web: http://www.math.usu.edu/~symanzik
Funding Source: **NIH R01 HL092347**

Content

- Context & Motivation.
- Data.
- Methods.
- Results.
- Discussion & Conclusion.

Context

- Reliable knowledge of patient activity levels are desirable in many medical situations including:
  Serious Illness (e.g. Cancer, Heart Disease), Recovery, Depression, and Sleep Science.
- Activity can be measured in a laboratory setting, but this may be time-consuming, invasive or inconvenient.
- Self-reports are helpful but may be incomplete, inaccurate, and difficult to quantify.
- => Reliable activity measurements are needed.



ActigraphyDevices

- Actigraphy is a relatively new and promising approach to obtaining numerical activity data.
- An Actigraph device is essentially an accelerometer encased in a wrist-watch assembly. Movement intensity of the individual is recorded by epoch (15 or 60 seconds for instance) and stored in internal memory.
- These intensity measurements provide a summary of the individual's activity during the time period.

AW-2, AW-Spectrum, AW-Score

Motivation

- A measurement device is only as reliable as its measurements.
- **Accuracy** and **precision** are desirable features.
- **Accuracy** is the agreement of the measurement with the true (but unknown) value of the quantity measured. The difference between this quantity and the measurement is the device bias.
- **Precision** is the agreement of multiple measurements made on the same quantity with each other. Precision is assessed with the variance or standard deviation of the measurements.

Motivation (Cont.)

- Previous research (e.g., Esliger and Tremblay, 2006) demonstrated concerns with actigraphy devices.
- In a sample of 39 Actical actigraphy instruments, seven were found to be too biased for use.
- Misleading information could have negative effects, especially in medical studies.
- Frequent calibration and testing of devices has been recommended.

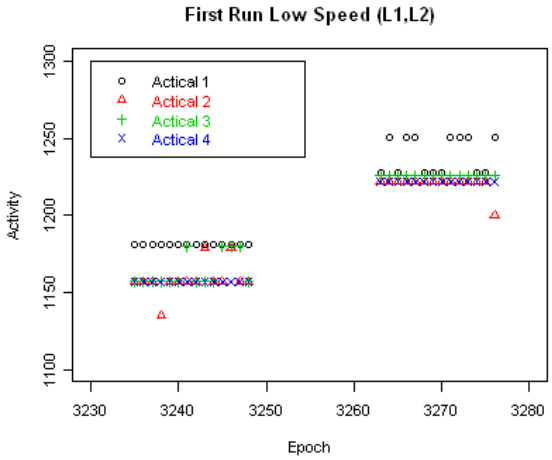
Main Experiment

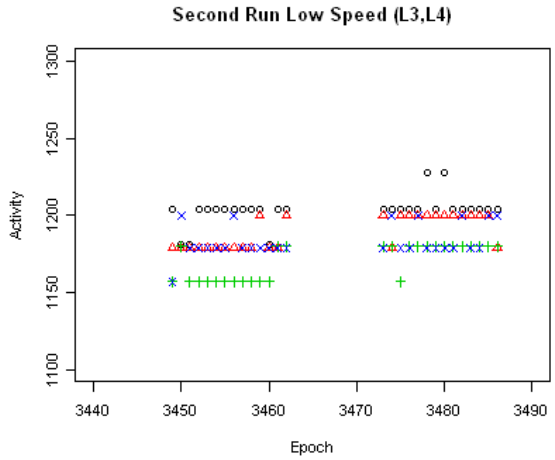
- Four Actical actigraphy devices were used in a controlled laboratory experiment.
- The devices were mounted to a shaker table. The table was turned to Speed 3 for four minutes then off for two minutes. Then, the table was turned back on at Speed 3 for four additional minutes. Due to the dial setup, it was not possible to achieve the exact same speed setting.
- This procedure was repeated for Speeds 6 and 8, with two four minute intervals apiece.
- The entire process was repeated twice.

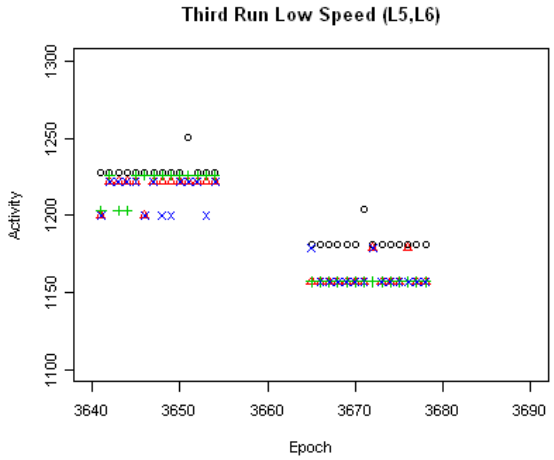
Data

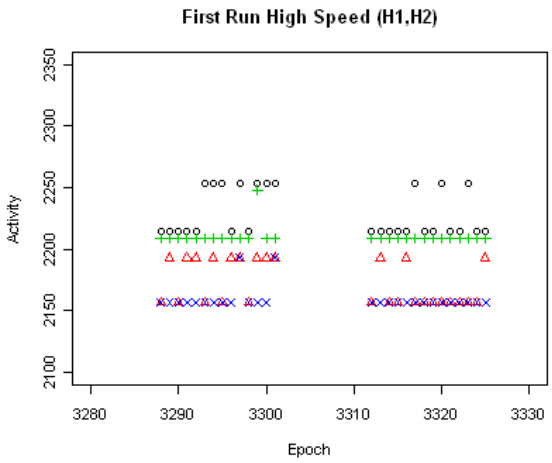
- All the data recorded at Speed 3 were zero. As such, these were excluded from the analysis.
- The Acticals were set to record at 15-second intervals over the four minutes. Only the inner 14 observations were used in order to eliminate measurements when the table was turned on or off.
- A total of 672 observations were used to compare the devices. That is:
  14 measurements x 6 series x 2 speeds x 4 Acticals

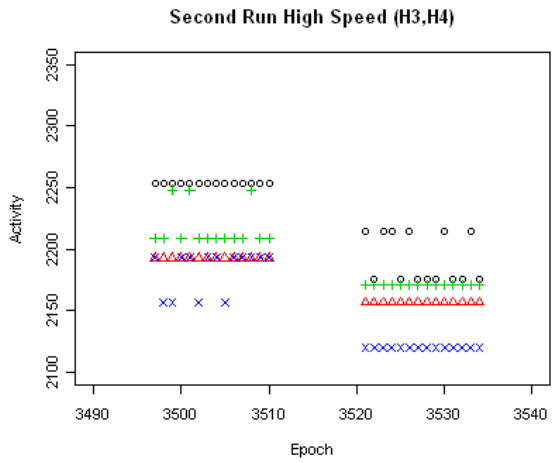
Data Used in Analyses

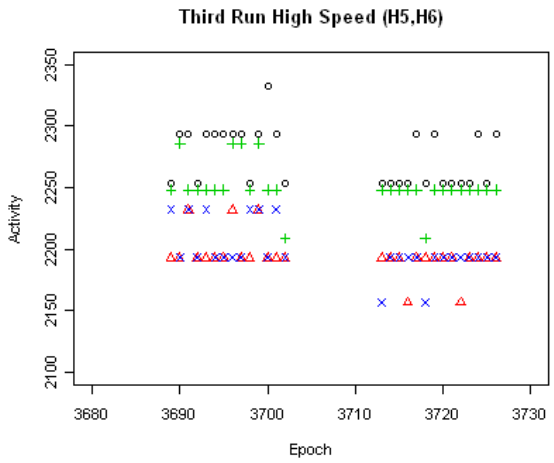












Methods – Precision Model

- Precision is assessed using the Multivariate Regression framework of Christensen and Blackwood (1993): *Tests for Precision and Accuracy of Multiple Measuring Devices*.

Methods – Precision Model

- A multivariate regression (Eq. 1) is performed on each phase where the deviations of three Acticals (Eq. 2) are regressed on the mean of all four Acticals (Eq. 3). Only three regression are performed to avoid dependence with the mean.

(1) $$\tilde{y}_{ijkl} = \delta_{jkl} + \gamma_{jkl}\bar{y}_{i.kl} + \varepsilon_{ijkl}$$

(2) $$\tilde{y}_{ijkl} = y_{ijkl} - \bar{y}_{i.kl}$$

(3) $$\bar{y}_{i.kl} = \frac{1}{4}\sum_{j=1}^{4} y_{ijkl}$$

Measurement i =1,…,14 ;
Actical j = 1,2,3 ;
Speed k = 1,2 ;
Series l = 1,…,6
j = 1,…,4

- Where the intercept is represented by δ_{jkl} and the slope is accounted for by γ_{jkl}.

Methods – Precision Test

- To test the precision, or whether the variance of the measurements are equivalent, the full model (Eq. 1) is tested against the reduced (Eq. 4).

  (4) $\tilde{y}_{ijkl} = \delta_{jkl} + \varepsilon_{ijkl} \qquad j = 1,2,3$

- This is a test that all the regression slopes are zero.
- The test is determined by Pillai's Trace Statistic which is approximately F-distributed.
- The twelve tests (over six series and two speeds) were conducted at the $\alpha = 0.05/12 = 0.00416$ level, applying a Bonferroni correction to account for multiple comparisons.

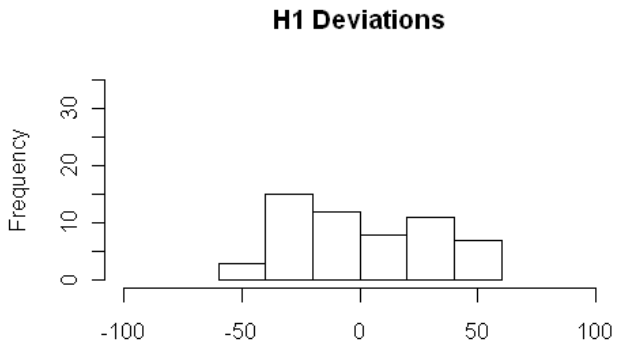
Methods – Precision Groups

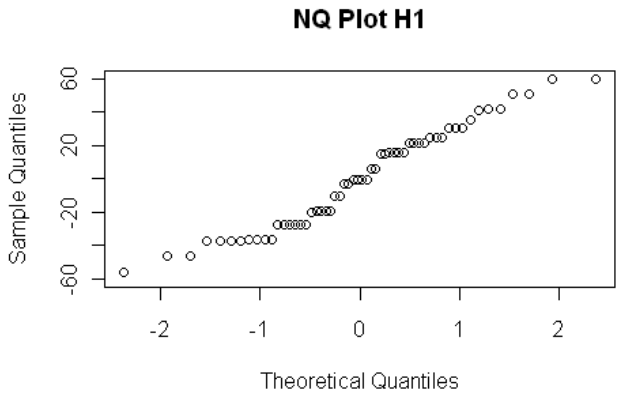
- Acticals can be grouped by their precision by performing the same analysis on pairs of devices.
- To compare Actical 1 and 2 for instance, a univariate regression of the deviations of Actical 1 on the mean of the two devices is performed.
- Computing Pillai's Trace Statistic allows one to test the hypothesis that two Acticals have equal variances in the series.
- Since some series had identical values for all 14 measurements, the regression could not be performed due to zero variability.
- Acticals were grouped separately if only one value was recorded in a series from Acticals with multiple values recorded.

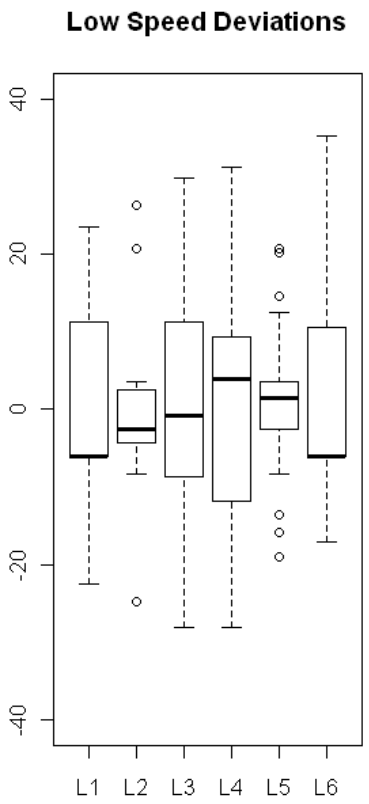
Methods – Accuracy Model

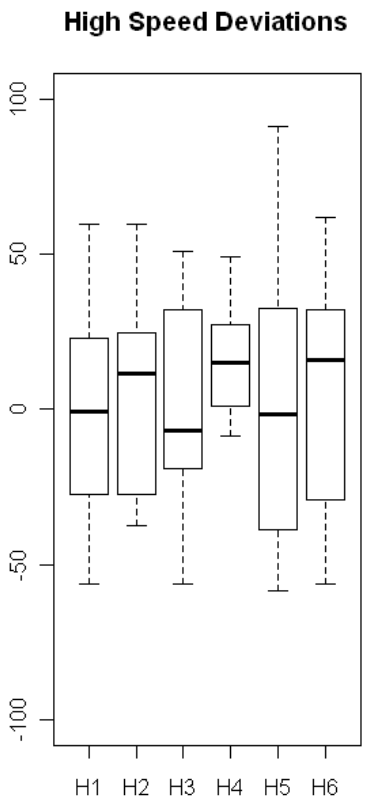
- To assess the accuracy of the devices, a similar Analysis of Variance (ANOVA) Approach was utilized.
- Similar tests were conducted as for the Precision Model.

Results – Plots of Deviations

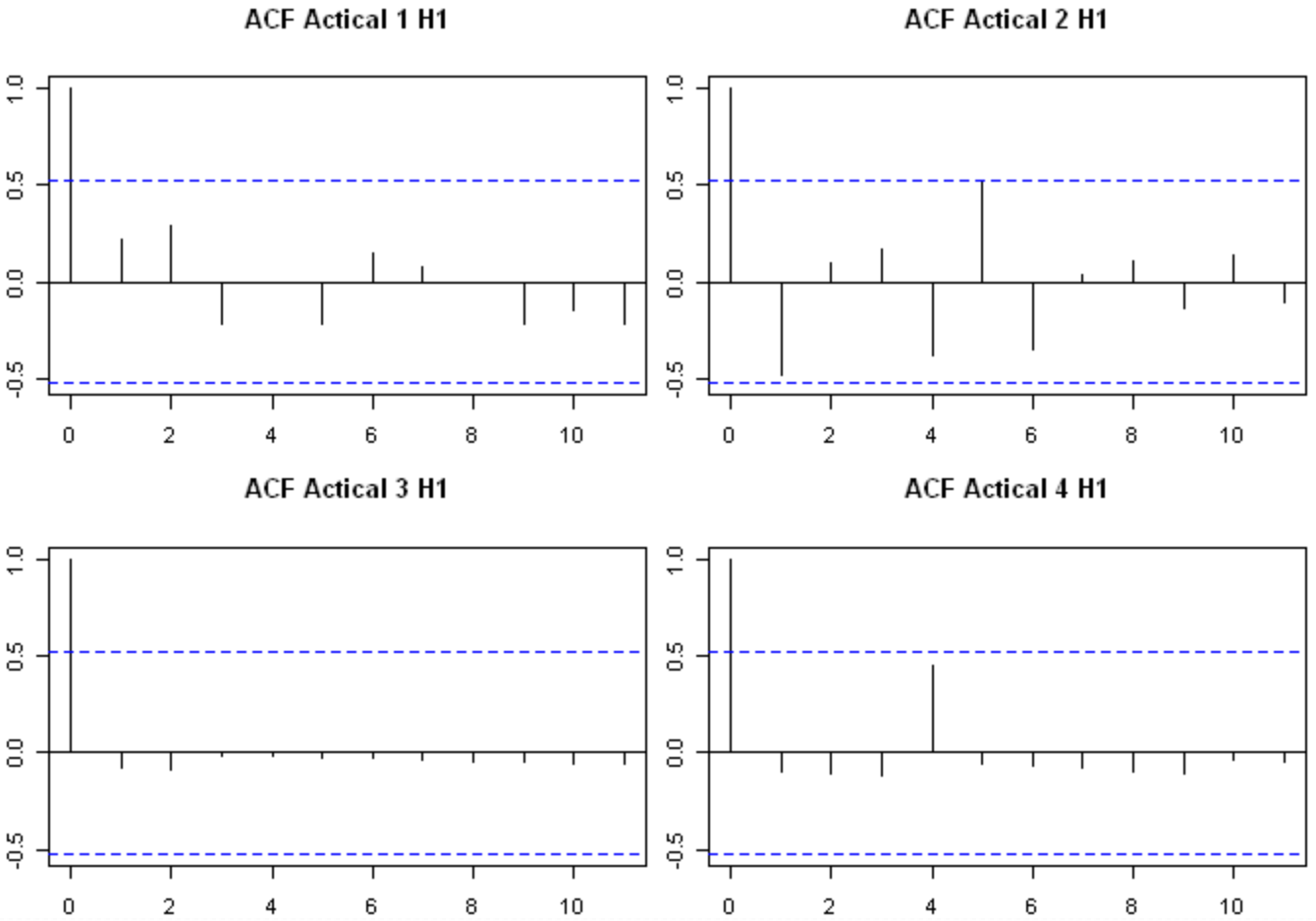








Results – ACF Plots



Results – SD Table/Groups

Low Speed	SD L1		SD L2		SD L3	
Actical 1	0.00(1)	B	**11.93(2)**	A	9.79(2)	A
Actical 2	10.44(3)	A	5.88(2)	A	7.63(2)	A
Actical 3	**10.78(2)**	A	0.00(1)	B	9.79(2)	A
Actical 4	0.00(1)	B	0.00(1)	B	**10.14(3)**	A
Low Speed	SD L4		SD L5		SD L6	
Actical 1	8.72(2)	A	6.15(2)	A	6.15(2)	A
Actical 2	7.63(2)	A	7.99(2)	A	**7.99(2)**	A
Actical 3	6.15(2)	A	9.79(2)	A	0.00(1)	B
Actical 4	**10.44(2)**	A	**10.94(2)**	A	**7.99(2)**	A
High Speed	SD H1		SD H2		SD H3	
Actical 1	**20.24(2)**	A	**16.61(2)**	A	0.00(1)	B
Actical 2	17.90(2)	A	15.33(2)	A	0.00(1)	B
Actical 3	10.42(2)	A	0.00(1)	B	16.61(2)	A
Actical 4	13.07(2)	A	0.00(1)	B	**16.88(2)**	A
High Speed	SD H4		SD H5		SD H6	
Actical 1	**20.03(2)**	A	**23.00(3)**	A	**18.75(2)**	A
Actical 2	0.00(1)	B	16.61(2)	A	13.07(2)	A
Actical 3	0.00(1)	B	22.16(3)	A	10.42(2)	A
Actical 4	0.00(1)	B	20.03(2)	A	13.07(2)	A

Results – Mean Table (Phase)

Low Speed	Mean L1		Mean L2		Mean L3	
Actical 1	**1181.0**	A	**1239.5**	A	**1199.1**	A
Actical 2	1158.6	B	1220.4	B	1182.0	B
Actical 3	1163.6	B	1226.0	B	1161.9	C
Actical 4	1157.0	B	1222.0	B	1180.4	B
Low Speed	Mean L4		Mean L5		Mean L6	
Actical 1	**1207.4**	A	**1229.6**	A	**1182.6**	A
Actical 2	1197.0	AB	1218.9	AB	1160.1	B
Actical 3	1178.4	C	1221.1	AB	1157.0	B
Actical 4	1186.5	BC	1214.1	B	1160.1	B
High Speed	Mean H1		Mean H2		Mean H3	
Actical 1	**2234.5**	A	**2223.4**	A	**2254.0**	A
Actical 2	2180.1	C	2164.7	B	2193.0	C
Actical 3	2211.8	B	2209.0	A	2217.4	B
Actical 4	2162.1	C	2157.0	B	2182.7	C
High Speed	Mean H4		Mean H5		Mean H6	
Actical 1	**2192.7**	A	**2285.4**	A	**2265.4**	A
Actical 2	2157.0	C	2201.4	C	2187.9	C
Actical 3	2171.0	B	2256.1	B	2245.2	B
Actical 4	2120.0	D	2209.7	C	2187.9	C

Discussion

- The results of the precision analysis suggest the devices are very precise. There was no consistent pattern of imprecision by device.
- Actually, in twelve of the 48 series the same value was recorded for all 14 measurements. This is variability of zero. At most, three distinct values were recorded in 14 measurements.
- The results of the accuracy analysis suggest a possible issue with Actical 1.
- This Actical recorded a mean activity that was greater than the other three at all phases of the experiment.
- This mean activity was statistically different from the other Actical means indicating the device is biased.

Further Research

- An unanticipated result was that no activity was recorded at Speed 3.
- This seems to indicate there is a threshold below which the Actical is unable to measure activity.
- Whether this threshold is relevant to corresponding human movements warrants further investigation. There may also be a level of activity the Actical maxes out at (this has been demonstrated for other actigraphs).
- Other issues also need to be investigated including human testing, device placement, and device orientation in future research.

Conclusions

- One Actical of four was found to be biased.
- Intra-device comparisons are fairly stable, but inter-device comparisons and magnitude dependent analysis may be problematic.
- Equipment should be checked and recalibrated frequently.
- This is not easily accomplished in the field, no standard apparatus is available.
- Actigraphy is a useful supplement to current methods but should not serve as a full replacement for those methods. Data should be carefully reviewed.

Further Reading:

Sherick, P. (2010) “Assessing the Precision and Accuracy in a Small Sample of Actical Devices”, MS Project Report, Utah State University, Department of Mathematics and Statistics.

Questions ?