ARL-TR-9635 ● JAN 2023

Multiscale and Multimodal Characterization of Mobile Sensor Data

by Alexander F Danvers, Lidia S Obregon, Esther M Sternberg, Matthias R Mehl, and Evan C Carter

NOTICES

Disclaimers

ARL-TR-9635 ● JAN 2023

Multiscale and Multimodal Characterization of Mobile Sensor Data

Evan C Carter
DEVCOM Army Research Laboratory

Alexander F Danvers, Lidia S Obregon, Esther M Sternberg, and Matthias R Mehl
University of Arizona

REPORT DOCUMENTATION PAGE						*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED (From - To)
January 2023	Technical Report	10 January 2021–29 September 2022

4. TITLE AND SUBTITLE
Multiscale and Multimodal Characterization of Mobile Sensor Data

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Alexander F Danvers, Lidia S Obregon, Esther M Sternberg, Matthias R Mehl, and Evan C Carter

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
DEVCOM Army Research Laboratory
ATTN: FCDD-RLA-FA
Aberdeen Proving Ground, MD 21005

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-TR-9635

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release: distribution unlimited.

13. SUPPLEMENTARY NOTES
ORCID ID: Evan C Carter, 0000-0001-7471-8769

14. ABSTRACT
Physiological and behavioral processes unfold on multiple time scales. Traditional time-series analysis tools are designed to capture stationary, single-scale processes, which may miss important information. Several methods have been proposed in recent decades to capture multiscale properties of time series, such as detrended fluctuation analysis. This report examines the way that multiscale measurement of physiology and behavior fits passive sensing data from the Fitbit Charge 4 mobile sensor. Physical activity and heart rate (HR) data from a large, long-term study of office workers were analyzed using traditional time-series analyses and a newly developed multiscale method: multiscale regression analysis. These analyses were conducted at the day and month level. Results indicate that multiscale analyses lead to substantial improvements in model R^2 over single-scale analyses for autocorrelation analyses of HR and steps (13% to 108% increase) and for the cross-correlation or coherence between HR and steps (21% to 88% increase). Multiscale analyses that led to better fit statistics were most advantageous when considering physical activity as compared with HR. Overall results suggest that physiology and behavior in daily life are better captured by estimating multiscale rather than single-scale processes.

15. SUBJECT TERMS
multiscale regression analysis, time series, fractal analysis, passive sensing, heart rate, HR, physical activity, Humans in Complex Systems

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Evan C Carter
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER (Include area code)
Unclassified	Unclassified	Unclassified	UU	23	(240) 478-9295

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Contents

List of Figures

List of Tables

1. Introduction

Human behavior and physiology operate on multiple time scales with multiple interacting components including mental and physiological processes. Traditional methods of analyzing time-series data assume that data operate at a single time scale; that is, the data-generating process exists at a single time scale, so the patterns of interest are observable when measured at a given rate. These traditional methods will therefore often miss key properties of real-world human data, potentially obscuring important effects and leading to worse model fit and poor prediction performance. For example, heart rate (HR) data are often studied using multiscale methods (Peng et al. 1995), and results from numerous studies indicate that cardiac interbeat intervals are best characterized by multiscale rather than single-scale measures (Peng et al. 1995; Perkiömäki et al. 2000; Hu et al. 2010).

Several methods for characterizing multiscale properties of time-series data, including methods that capture the associations between multiple time series, have been developed in recent years. These methods have been applied in fields such as psychology, economics, geophysics, and urban planning (Yuan et al. 2015), but they have not yet been applied to large, long-term studies of people in their daily lives. Datasets of this scope are becoming more common as technological advancements make sophisticated data collection systems more accessible, and given the likely importance of multiscale processes reflected in such data, a marriage of multiscale modeling with data collected over a long time period and "in the wild" is an important step.

Here we aim to identify the gains in model fit from using these multiscale and multimodal methods on a large, long-term dataset of passively sensed human data. Ultimately, such methods may facilitate a low-cost, accurate assessment of the health, readiness, and psychological state of Soldiers, potentially improving decision making by commanders and forming the basis of adaptive technology such as artificial intelligence that can act as teammates rather than tools.

1.1 Multiscale Measurement of Heart Rate

HR fluctuates over time according to activity, health, and fitness. Researchers have found that unique insight into physiological and mental functioning can also be gained by analyzing patterns of variability in HR throughout a task, day, or even a baseline resting period. HR variability (HRV), a measure of the breathing-related variability around average HR, gives insight into the functioning of the sympathetic and parasympathetic nervous systems (van Ravenswaaij-Arts et al. 1993). Critically, the common measures used to study HRV examine the variability at one scale, such as the standard deviation (SD) of the interbeat interval of normal sinus beats or the root mean square of successive differences between normal heartbeats (Shaffer and Ginsberg 2017). In contrast, fractal measures are used to analyze HRV on multiple time scales, and studies of cardiac activity have shown that healthy hearts display fractal properties characteristic of complex systems, whereas certain pathological states show a lack of these properties (Peng et al. 1995).

Measurements used to capture these properties, such as detrended fluctuation analysis (DFA) or multifractal DFA (MF-DFA), specifically attempt to quantify HRV occurring across the multiple time scales to capture regularities between degree of variability and scale of measurement and assess the fractal nature of the HR. Findings from the fractal HR literature have found that measures extracted from DFA are distinct from and can actually provide better predictive accuracy than traditional HRV measures in select tasks, especially as prognostic markers (Peng et al. 1995; Sen and McGill 2018). Furthermore, these measures are of particular interest because they naturally fit into theoretical conceptualizations that span multiple time scales and potentially have an effect on HR. For instance, we might expect that HR is influenced by overall patterns of health, such as sleep quality, nested in local patterns of mood, like an acute stressful event, and more immediate events like responding to a surprising stimulus. Therefore, in principle, multi-time-scale fractal measures should capture dynamic properties in HR that are missed by traditional, one-scale analysis.

1.2 Multiscale Regression Analysis

Early groundbreaking work on multiscale properties of HR used techniques like DFA and MF-DFA, which characterize the relationship between variability and scale of measurement. In DFA, a series of time scales are selected for analysis. Each corresponds to a number of data points to consider at a time. For example, if the scale is 10 points, the time series is divided into non-overlapping windows of 10 points each. In each window the time series is detrended using a linear trend, and then the variance of the points in that window is estimated. The variance of all windows at that scale is then summed, giving an overall estimate at that window size. The procedure is repeated for multiple scales/window sizes. The association between scale and the different variance estimates obtained is then estimated. This is the Hurst exponent, and it can be used in other analyses to predict outcomes of interest or differentiating between experimental conditions. Notably, this approach is limited to considering a single time series.

A more recent development in multiscale time-series analysis is multiscale regression analysis (MRA; Likens et al. 2019), which extends the logic of DFA from examinations of variance across time scales to examinations of regression coefficients across time scales. Multiscale regression proceeds in much the same way as the DFA: A set of time scales to analyze is chosen (Fig. 1) and the data are divided into windows of that size (e.g., 10 data points). The two time series—referred to as the predictor and the criterion—are detrended (simulation studies suggest a quadratic trend is best; Likens et al. 2019), and a regression coefficient predicting the criterion is estimated in each window. The average of these regression coefficients across all windows at a given time scale is taken as the association at that time scale (Fig. 2). This is repeated at all time scales of interest, and the relationship between time scale and regression coefficients can be estimated.

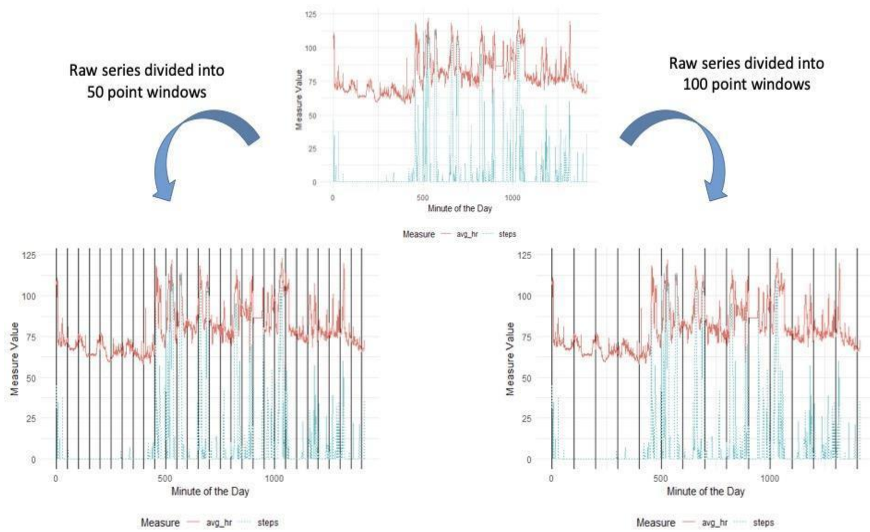


Fig. 1 **Illustration of multiscale approach**

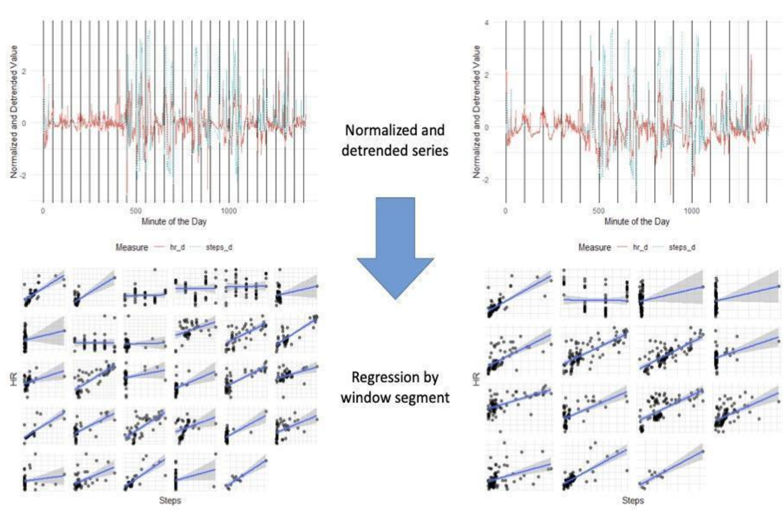


Fig. 2 **Illustration of multiscale regression**

Beyond offering a way to estimate the relationship between time scale and regression coefficients, multiscale regression also captures the strength of association at each time scale. The coefficient of determination (R^2) value for each window at a given time scale can be estimated and averaged to give an overall R^2 at that scale. These R^2 values provide an estimate of model fit at each time scale, allowing for a comparison to single-scale methods. For example, analyses of motor coordination found that, when people were balancing on a force plate, the association between ankle and hip motion was strongest at slower time scales (Likens et al. 2019). The association between ankle and shoulder motion was strongest at faster time scales.

The association or coherence between HR and physical activity has theoretical importance in psychophysiological research (Brouwer et al. 2018). Physical activity tends to increase HR, as during walking or exercising (Murray et al. 1985; Boulay et al. 1997; Nystoriak and Bhatnagar 2018). When a person is physically active, the cardiovascular system mobilizes to meet a systematic demand for more blood supply, which causes cardiac output, blood pressure, and HR to increase. Furthermore, as factors such as exercise intensity increase, HRV decreases (Michael et al. 2017).

In addition to physical activity, HR can also be elevated due to psychological factors, such as stress or engagement (Friedman and Thayer 1998; Thayer and Sternberg 2006). This observation has led numerous psychophysiologists to propose estimating the “excess” HR, or the residual of HR predicted by physical activity, as a potential measure of psychological influences on the heart (Brouwer et al. 2018; Brown et al. 2018). Several results have found that measures of excess HR are associated with psychological variables such as stress (Lambiase et al. 2012; Verkuil et al. 2016). Coherence between HR and activity—such as the activity of facial muscles—also plays an important role in theories of emotion (Mauss et al. 2005; Brown et al. 2020).

The coherence between HR and physical activity has typically been modeled using standard regression and time-series techniques that, as noted, assume stationarity and effects at a single time scale. Given that HR is influenced by processes operating at multiple time scales, however, a natural next step is to examine how the association between HR and physical activity changes across multiple time scales. Finding multiscale influences of physical activity and HR would suggest that the body self-organizes to adapt to extended periods of high (or low) activity as well as changes at shorter, more immediate scales in response to activity. Cross-correlation analysis was used to analyze two time series at a single scale, while multiscale regression analysis was used to analyze two time series at multiple time scales.

1.3 The Present Study

We conducted an analysis of HR and physical activity time series using a large-scale mobile sensor dataset. The Fitbit Charge 4 device was used as the mobile-sensing device for data collection in this study due to its widespread availability and inexpensiveness. The data used in this study were preprocessed by established Fitbit signal processing algorithms. We did not attempt to develop or validate any novel signal processing algorithms for the data analysis.

We focus on three key comparisons: 1) single modality versus multimodal analysis, 2) single time scale versus multiscale, and 3) short- versus long-time scale (i.e., day-level versus month-level). We focus on HR in the single modality analyses and the association between HR and physical activity in the multimodal analyses. We use traditional time-series methods—the autocorrelation function (ACF) and cross-correlation function (CCF)—for the single time-scale analyses and MRA for the multiscale analyses. For the comparison of short- and long-time scales, we repeat these analyses with data binned at either the day- or month-level.

2. Methods, Assumptions, and Procedures

2.1 Participants and Procedures

Participants were 206 office workers from an office in Silicon Valley who participated in a longitudinal study that lasted 60 days. After data cleaning (described in the following), the average number of days of data used for each person in the study was 56.

During the 60-day study, participants were instructed to wear mobile sensors while at work. Participants wore the sensors on different sets of days since enrollment in the study was rolling and not all participants were in the study at the same time. To preserve ecological validity and naturalistic sampling of participants' daily life in the office, there were no specific instructions given to participants regarding behavioral procedures or expectations.

2.2 Measures

Participants were given a Fitbit device (Fitbit Charge 4) to wear throughout the day. The Fitbit uses photoplethysmography (PPG) to non-invasively capture HR data by detecting blood volume changes from the skin. Fitbit technology estimates HR based on the raw PPG data and an internal processing algorithm. Data was measured at the minute-by-minute scale due to practical constraints with both Fitbit's activity output and the quality of the data. Mobile sensor data are noisy, so estimates of HR at the individual beat level can sometimes be inaccurate. Smoothing over multiple heartbeats occurring in a given minute creates more reliable and stable estimates and allows for more accurate analysis.

Physical activity in the present study was operationalized using step counts recorded from the Fitbit Charge 4 device. All minutes of physical activity analyzed were matched to those of the HR, using the same rules for cleaning (see data trimming section).

When conducting day level analyses, we wanted to capture only data where the participant was awake and active. To estimate the period of time at which the participant was awake, we used the first time an individual recorded five steps in a minute as the start of the day and the time an individual last recorded five steps in a minute as the end of the day. This simple operationalization removes the extended periods of sleep where an individual rarely moves at all, a period with very different dynamics from daily activity. For month-level analyses, all data—including data from overnight—was included.

2.3 Analyses

We compared single to multiscale, uni- to multimodality, and short to long time scales. Table 1 outlines the analysis method used to conduct each comparison. Unimodal analyses were conducted separately on HR and step-count time series. Multimodal analyses were conducted on the two together. All analyses are repeated at the day and month level. All comparisons were made using the model fit statistic, the R^2, often interpreted as the percent of variance explained.

Table 1 Analysis outline

Modality	Scale	
	Single	**Multi**
Uni	ACF	DFA
Multi	CCF	MRA

Single Scale, Single Modality. The ACF was used to analyze single modality data on a single time scale. The ACF is typically used at multiple lags. A range of lags 1 to 10 were estimated, and as is typical in this type of mobile sensor data, ACF for HR and steps were highest at lag-1 in almost all cases. This was therefore chosen as a common reference point (see Results section).

Multiscale, Single Modality. DFA was used to analyze data from a single variable across multiple time scales. The DFA method analyzes a single time series at multiple window sizes. These window sizes are typically evenly spaced along a log-base-2 scale, and enough data to reach 1024 points (2^{10}) are typically recommended for this analysis. (Note that a full day's data, recorded in minutes, yield 86,400 points—more than an adequate sample for DFA.)

DFA typically examines variance as a function of scale without consideration of time lags. However, to facilitate the comparison of this type of multiscale, single modal analysis with the rest of the analyses used, we modified the DFA procedure so that it involves estimation of the lag-1 autocorrelation as a function of scale. Also note that it is common for DFA to use log-2 scales for time points (e.g., 8, 16, 32 points). However, for a more accurate comparison with MRA (see the following), the same time scales employed in that analysis were also used for the multiscale, single modality analysis. The scales analyzed were therefore a range of 10 to 200 min in increments of 5 min.

Single Scale, Multiple Modalities. The CCF was used to analyze data on a single time scale from two modalities. The cross-correlation between HR and physical activity was estimated at lags from –10 to +10. Negative lags correspond to activity leading HR, and positive lags correspond to HR leading activity.

Multiscale, Multiple Modalities. MRA was used to analyze data from two modalities across multiple time scales. In this analysis, physical activity was used to predict HR, but our results, given as R^2, would be the same if HR was used to predict activity. Prediction of HR from activity was used here because it was deemed more physiologically plausible, and our primary interest was in fit rather than prediction. The window sizes to be used are set by the user. Unlike a DFA, which is typically performed with window sizes increasing on a log-base-two scale, MRA and related analyses have been performed over a linearly increasing range of time scales. The current analyses explored scales from 10 to 200 min separated by increments of 5 min. At the low end, enough points need to be included to make a stable estimate of the association. We deemed 10 points to be sufficient for this purpose. At the high end there should be enough data to create several independent segments that can be averaged. Additionally, physiological plausibility should be considered. We considered 200 min (3 and 1/3 h) to be long enough that effects of activity on HR were likely to have dissipated. As with DFA, we implement MRA with a lag of 1 to make it more directly comparable to the multiscale, multimodality CCF.

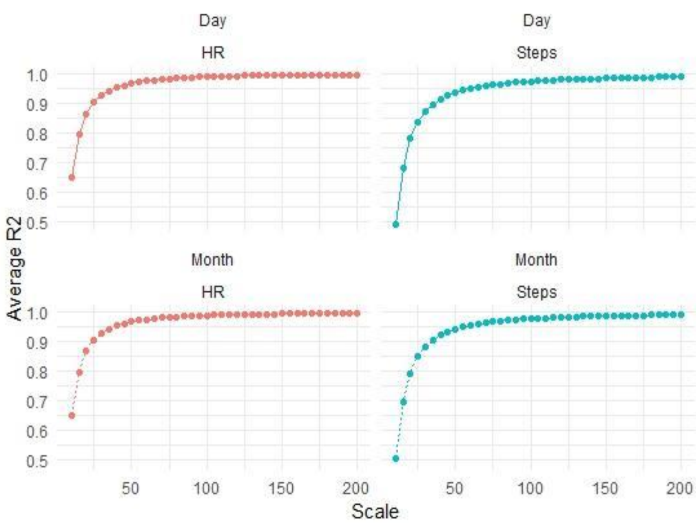
Day versus Month. All of the analyses described, which compare single versus multiscale and unimodal versus multimodal approaches, were conducted on two different versions of the data. One version represented the data by day, following the trimming rule described (to capture just the "active day"), and the other aggregated by the month.

3. Results and Discussion

3.1 HR Data

Day Level. We fit ACF to HR time series at the day level and the optimal lag was selected from among those calculated (1 to 10). As expected, in over 99.9% of cases (11,526 of 11,533), this was the lag-1 autocorrelation. The average R^2 for the single scale autocorrelation was $R^2 = 0.848$.

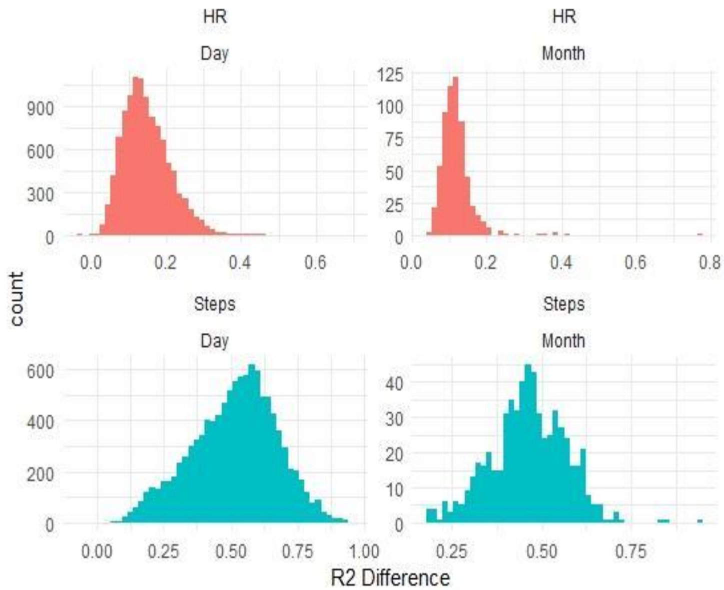
For the multiscale analysis, we fit a DFA, and the optimal time scale for the lag-1 autocorrelation was selected from among those calculated (10 to 200, in increments of 5). There was more variability in the optimal scale than the optimal lag. The most common optimal scale was 200 points (32.7%). Plotting the average R^2 value for each time scale (Fig. 3) illustrates that the R^2 values quickly approach and asymptote at a value close to $R^2 = 1.00$. The optimal scale for most cases was at $R^2 = 0.99$ with increasing variability explained out to the third and fourth decimal. This finding indicates that it is unlikely that we underestimated model fit given the range of scales we considered. The average optimal R^2 for the multiscale analysis was $R^2 = 0.997$. This is an increase of 17.5% in the average R^2 values when using multiscale analysis (DFA) rather than single-scale (ACF).



Note: Standard errors (SEs) are too small to be resolved clearly on this figure. Average SE of the mean is 0.0003.

Fig. 3 Average R^2 by scale for multiscale autoregressive analyses

The difference in model fit can also be assessed at the level of the individual day by subtracting the best R^2 for the single-scale model from the best R^2 for the multiscale model. The average R^2 *difference* for the single-scale versus the multiscale model is 0.149. This was an increase of 17.5% for the multiscale over the single-scale model. The distribution of these R^2 differences between single-scale and multiscale autocorrelation estimates is illustrated in Fig. 4.



Note: Positive values indicate advantage for multiscale model over single-scale model.

Fig. 4 Histograms of R^2 differences for single-scale vs. multiscale autoregressive analyses

Month Level. For 100% of cases the lag-1 autocorrelation was the optimal value for HR when data were represented at the month level and modeled with ACF. The average R^2 for the optimal HR autocorrelation was $R^2 = 0.878$.

The most common optimal scale for the month-level multiscale single-modal analysis was 200 points (52% of cases). As with the day level analyses, however, the optimal R^2 appeared to asymptote to approximately $R^2 = 1.00$ when scales greater than 100 were used. This suggests that using longer time scales would have yielded very little difference in the model comparison done here. The average R^2 for the optimal steps autocorrelation was $R^2 = 0.997$. This is an improvement of 13%. The average month-by-month R^2 difference was 0.119. This is also an improvement of 13%. The distribution of differences in R^2 between the single scale and multiscale analysis had a slightly more skewed distribution at the month level than the day level (Fig. 4, bottom panels).

3.2 Steps Data

Day Level. For 99.4% of cases, the lag-1 autocorrelation was the optimal value for steps. For 99.8% of cases, the optimal lag was 3 or less. The average R^2 value for the autocorrelation of steps data was $R^2 = 0.476$.

For the multiscale analysis, the optimal time scale displayed a similar pattern of variability as that shown in the HR autocorrelation. A time scale of 200 was selected in 31.6% of cases. In a plot of the average R^2 by scale, it was again revealed that as time scale increased, values tended to asymptote toward $R^2 = 1.00$. This indicates that a sufficient set of scales were considered to identify the optimal fit. The average R^2 value for the multiscale autocorrelation of steps was $R^2 = 0.991$. This is an improvement of 108%.

The case-by-case comparison of R^2 values yielded similar results. The average difference in R^2 for a single-scale versus multiscale autoregressive analysis was 0.515. This is an improvement of 108%. In Fig. 4 a histogram of the R^2 differences for multiscale versus single-scale analyses shows substantial variability in the degree of improvement.

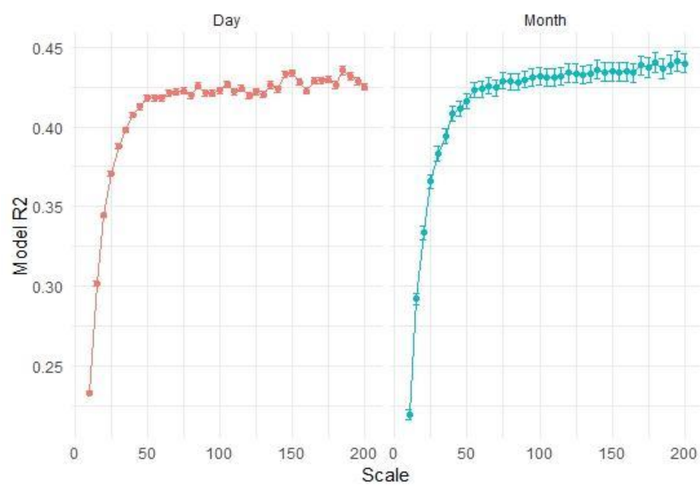
Month Level. The optimal lag for the single-scale-steps autocorrelation analysis was 1 for all cases but one (99.8% of cases). For this case a lag of 4 was optimal. The average R^2 value for the month-level, single-scale, autoregressive analysis of steps was $R^2 = 0.528$.

As with the month-level HR data, the optimal scale for approximately half (49.9%) of all autoregressive steps analyses was 200 points. This was, again, because the accuracy of these models asymptotes to approximately $R^2 = 1.00$ after 100 points are considered. The average R^2 value for the month-level, multiscale, autoregressive analysis of steps was $R^2 = 0.992$. This is an improvement of 88%. The average case-by-case R^2 difference for the month level was 0.464. This is also an improvement of 88%. The distribution of the R^2 difference scores is plotted in Fig. 4. Note that this is approximately normally distributed.

3.3 Multimodal Data

Day Level. As a single-scale analysis at the day level we fit a CCF to the HR and step data, considering 10 lags in either direction. The optimal lag was selected for each person on each date and for 95% of person days; this was a lag of –1, indicating that steps predicted HR 1 time point later. For approximately 4% of person days the optimal lag was –2, indicating steps predicted HR 2 time points later. However, in less than 1% of instances was the optimal lag outside of the range from –3 to 0. These findings suggest that the global optimal lag for the CCF was accurately identified through this analysis. The average best R^2 for the single-scale analysis was $R^2 = 0.34$.

For a multiscale analysis at the day level we fit an MRA on the HR and step data. As with lags in the single-scale case, the optimal scale was selected for each person on each day. The optimal time scales were much more evenly distributed across the scales considered (10 to 200 points, increasing in increments of 5). The most common optimal scales for person days were 185 points (6% of cases), 200 points (5% of cases), and 165 points (5% of cases). Overall, the optimal scale was typically longer (9 of the 10 top scales were 145 points or longer). This suggests that better fits might have been achieved if longer scales had been considered. However, plotting the average R^2 by time scale suggested an asymptote, so that the optimal R^2 was approaching $R^2 = 0.41$ (Fig. 5). If this asymptotic relationship were to hold, it would indicate that the approximately optimal time scale had been identified. The average best R^2 for the multiscale analysis was $R^2 = 0.41$, which is 21% higher than the average R^2 for the single-scale analysis.



Note: Error bars represent ±1 SE.

Fig. 5 Model R^2 by scale for day vs. month multiscale regressions

The difference in model fit can also be assessed at the level of the individual person day by subtracting the best R^2 for the single-scale model from the best R^2 for the multiscale model. The average R^2 difference for the single-scale versus the multiscale model is 0.30. This suggests that, considered at the level of an individual day, the multiscale model is an even larger improvement: the difference between the averages was only 0.07. This is an 88% increase over the average best single-scale model.

Investigating case-by-case differences (as opposed to average differences) also allows for the exploration of variability in the multiscale advantage. The SD of the R^2 difference was SD = 0.10. A histogram of these differences is provided in Fig. 6. This distribution has minimal skewness (0.34) but notable kurtosis (3.63). This increased "peakiness" of the distribution indicates that the middle of the distribution is a good representation of the typical difference between single and multiscale analyses.

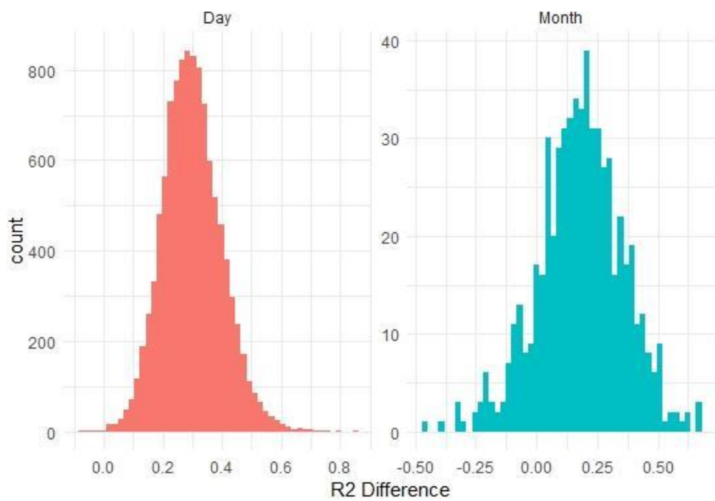


Fig. 6 Case-by-case R^2 difference for day vs. month multiscale regression

Month Level. The average R^2 of the best fitting CCF model at the month level was $R^2 = 0.33$, while for the best scale MRA model the average was $R^2 = 0.51$. This is an improvement of 52%, a substantially greater improvement than was gained in the comparable day-level analysis (21%). When examining the differences calculated on a case-by-case basis (e.g., subtracting each month's best CCF R^2 from that month's best MRA R^2), the difference was 0.17. This was a 52% improvement over the average best CCF, which was less than the improvement seen in the day-level analysis (88%). Note that the distribution of R^2 difference scores observed at the month level is similar to that of the day level but with the center of the distribution shifted slightly toward zero (see Fig. 6). The average R^2 by scale observed at the month level was also very similar to that observed at the day level. For both analyses, the overall R^2 appears to be approaching an asymptote of around $R^2 = 0.41$. The full set of results described here are presented in Table 2.

Table 2 Comparison of average R^2 values for the models investigated

Aggregation	Modality	Scale		Change	
		Single	Multi	Difference	Percent
Day	HR	0.848	0.997	0.149	17.57
	Steps	0.476	0.991	0.515	108.19
	Multimodal	0.34	0.41	0.070	20.59
Month	HR	0.878	0.992	0.114	12.98
	Steps	0.528	0.992	0.464	87.88
	Multimodal	0.33	0.51	0.180	54.55

Note. "Difference" is the average R^2 for the multiscale method minus the average R^2 for the single-scale method. "Percent" is "Difference" divided by the average R^2 for the single-scale method multiplied by 100.

4. Conclusions

Fitbit data collected during the daily life of a large group of office workers were analyzed using multiscale versus single-scale and multimodal versus unimodal analyses. Minute-by-minute HR and step counts were found to have multiscale properties both in their autocorrelations and their associations with each other. Multiscale analyses were able to identify better-fitting models than single-scale analyses in all the cases examined: HR autocorrelation, steps autocorrelation, and steps–HR coherence. These results demonstrate the importance of considering the multiscale properties of time-series data of human behavior.

One practical application of these analyses is identifying an optimal time scale for predictive or forecasting accuracy. For autoregressive analyses of both HR and steps the optimal lag reached approximately $R^2 = 0.99$ after considering 100 or more points. This was true at either the day or month level. This value was substantially higher than the value obtained when using a traditional single-scale autocorrelation function. This suggests that local autocorrelation analyses that account for about 2 h are better at forecasting than a single forecast value that uses all data.

Also of note, single-scale analyses of minute-by-minute HR fit much better than single-scale analyses of minute-by-minute steps. This suggests that multiscale analyses are particularly important for appropriately characterizing physical activity in daily life. Step counts may have this sort of multiscale patterning because of the way activities are nested in time. Periods of sitting, walking around, gesturing, or presenting during meetings, exercising after work hours, and other daily life activities are likely to produce characteristic autocorrelations that last for the specific period of the activity—from minutes to hours. HR, on the other hand, may be more similar across several different daily activities, including heads-down work and meetings.

For multimodal analyses, results indicated that optimal scale differed widely from person to person and day to day. In general, longer time scales were selected—more than 145 min considered. Yet the overall optimum could be closer to 2 to 3 h depending on the person and day. This suggests that the optimal scale of measuring HR and activity coherence may be an individual difference, which depends on characteristics of the individual and of the patterns of their daily life.

5. References

Boulay MR, Simoneau JA, Lortie G, Bouchard C. Monitoring high-intensity endurance exercise with heart rate and thresholds. Medicine & Amp Science in Sports & Amp Exercise. 1997;29(1):125–132. doi: 10.1097/00005768-199701000-00018.

Brouwer AM, Van Dam E, Van Erp JB, Spangler DP, Brooks JR. Improving real-life estimates of emotion based on heart rate: a perspective on taking metabolic heart rate into account. Frontiers in Human Neuroscience. 2018;12:284.

Brown CL, Van Doren N, Ford BQ, Mauss IB, Sze JW, Levenson RW. Coherence between subjective experience and physiology in emotion: individual differences and implications for well-being. Emotion. 2020;20(5):818.

Brown SB, Brosschot JF, Versluis A, Thayer JF, Verkuil B. New methods to optimally detect episodes of non-metabolic heart rate variability reduction as an indicator of psychological stress in everyday life. International Journal of Psychophysiology. 2018;131:30–36.

Friedman BH, Thayer JF. Autonomic balance revisited: panic anxiety and heart rate variability. Journal of Psychosomatic Research. 1998;44(1):133–151.

Hu J, Gao J, Tung WW, Cao Y. Multiscale analysis of heart rate variability: a comparison of different complexity measures. Annals of Biomedical Engineering. 2010;38(3):854–864. doi: 10.1007/s10439-009-9863-2.

Lambiase MJ, Dorn J, Chernega NJ, McCarthy TF, Roemmich JN. Excess heart rate and systolic blood pressure during psychological stress in relation to metabolic demand in adolescents. Biological Psychology. 2012;91(1):42–47.

Lickens AD, Amazeen PG, West SG, Gibbons CT. Statistical properties of multiscale regression analysis: simulation and application to human postural control. Phsica A: Statistical Mechanics and its Applications. 2019;532. Article no.: 121580.

Mauss IB, Levenson RW, McCarter L, Wilhelm FH, Gross JJ. The tie that binds? Coherence among emotion experience, behavior, and physiology. Emotion. 2005;5(2):175.

Michael S, Graham KS, Davis GM, Oam GMD. Cardiac autonomic responses during exercise and post-exercise recovery using heart rate variability and systolic time intervals-A review. Frontiers in Physiology. 2017;8:301. doi: 10.3389/fphys.2017.00301.

Murray MP, Spurr GB, Sepic SB, Gardner GM, Mollinger LA. Treadmill vs. floor walking: kinematics, electromyogram, and heart rate. Journal of Applied Physiology. 1985;59(1):87–91. doi: 10.1152/jappl.1985.59.1.87.

Nystoriak MA, Bhatnagar A. Cardiovascular effects and benefits of exercise. Frontiers in Cardiovascular Medicine. 2018;5:135. doi: 103389/fcvm.2018 .00135

Peng C, Havlin S, Stanley HE, Goldberger AL. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. Chaos: An Interdisciplinary Journal of Nonlinear Science. 1995;5(1):82–87. doi: 10.1063/1.166141.

Perkiömäki JS, Mäkikallio TH, Huikuri HV. Nonlinear analysis of heart rate variability: fractal and complexity measures of heart rate behavior. Annals of Noninvasive Electrocardiology. 2000;5(2):179–187. doi: 10.1111/j.1542-474x.2000.tb00384.x.

Sen J, McGill D. Fractal analysis of heart rate variability as a predictor of mortality: a systematic review and meta-analysis. Chaos: An Interdisciplinary Journal of Nonlinear Science. 2018;28(7):072101. doi: 10.1063/1.5038818.

Shaffer F, Ginsberg JP. An overview of heart rate variability metrics and norms. Frontiers in Public Health. 2017;5. doi: 10.3389/fpubh.2017.00258.

Thayer JF, Sternberg E. Beyond heart rate variability: vagal regulation of allostatic systems. Annals of the New York Academy of Sciences. 2006;1088(1):361–372.

van Ravenswaaij-Arts CM, Kollee LA, Hopman JC, Stoelinga GB, van Geijn HP. Heart rate variability. Annals of Internal Medicine. 1993;118(6):436–447. doi: 10.7326/0003-4819-118-6-199303150-00008.

Verkuil B, Brosschot JF, Tollenaar MS, Lane RD, Thayer JF. Prolonged non-metabolic heart rate variability reduction as a physiological marker of psychological stress in daily life. Annals of Behavioral Medicine. 2016;50(5):704–714.

Yuan N, Fu Z, Zhang H, Piao L, Xoplaki E, Luterbacher J. Detrended partial-cross-correlation analysis: a new method for analyzing correlations in complex system. Sci Rep. 2015;5:8143. doi: 10.1038/srep08143.

List of Symbols, Abbreviations, and Acronyms

ACF	autocorrelation function
CCF	cross-correlation function
DFA	detrended fluctuation analysis
HR	heart rate
HRV	heart rate variability
MF-DFA	multifractal DFA
MRA	multiscale regression analysis
PPG	photoplethysmography
R^2	coefficient of determination
SD	standard deviation
SE	standard error

1	DEFENSE TECHNICAL
(PDF)	INFORMATION CTR
	DTIC OCA
1	DEVCOM ARL
(PDF)	FCDD RLB CI
	TECH LIB
1	DEVCOM ARL
(PDF)	FCDD RLA FA
	E CARTER
4	UNIVERSITY OF ARIZONA
(PDF)	A DANVERS
	L OBREGON
	E STERNBERG
	M MEHL