




Analyzing cyclic patterns in psychological data: A tutorial

Peter Verboon^a,  and Ruslan Leontjevas^a

^aDepartment of Statistics and Methodology, Faculty of Psychology and Educational Sciences, Open University, The Netherlands

Abstract ■ A growing number of studies on psychological phenomena employ the Ecological Momentary Assessment (EMA) method for obtaining intensive longitudinal data in daily life. Whereas cyclic processes may underlie different psychological and physiological outcomes, the cyclic model that describes such processes is rarely used for fitting EMA data. The aim of this paper is to introduce the cyclic model to researchers, and to demonstrate its use in an empirical data set. It is shown how the cyclic terms can be incorporated in multilevel models. Based on secondary analyses on an existing EMA data set, it can be concluded that adding cyclic terms in EMA analyses may improve model fit and may help understanding the dynamic processes.

Keywords ■ ESM, EMA, Cyclic models, multilevel analysis, intensive longitudinal data. **Tools** ■ R.

 Peter.Verboon@ou.nl

 [PV: 0000-0001-8656-0890](#); [RL: 0000-0002-9689-7611](#)

 [10.20982/tqmp.14.4.p218](https://doi.org/10.20982/tqmp.14.4.p218)

Acting Editor ■ [Denis Cousineau](#) (Université d'Ottawa)

Reviewers

■ One anonymous reviewer

Introduction

Intensive longitudinal data are becoming more widely used in psychological research (Conner & Mehl, 2015). Modern technology, e. g., applications for tablets and smartphones, facilitates the use of the *Ecological Momentary Assessment* or *Experience Sampling Method* (further addressed as EMA) for obtaining data from respondents on a daily basis. These methods have a density of assessments that varies from several minutes or hours between assessments over a period of several days to a less dense schedule (e. g., once a day) over a period of several days or months. The data on behavior, (bio)physiological processes or psychological states are collected in the natural environment of participants in which the outcomes of interest occur, while the focus is rather on subjects' short moment states than on the recall over a long period (Shiffman, Stone, & Hufford, 2009).

Because data are collected in the context in which they occur, the ecological validity of the EMA is thought to be better than in research with retrospective questioning. Compared to retrospective questionnaire research, the recall bias of the measurements is reduced, which improve validity of data (Trull & Ebner-Priemer, 2014; Shiffman et al., 2009).

The validity of the EMA has proved to be useful in research on a very wide range of outcomes including most clinical symptoms and syndromes, and a plethora of health-related behaviors and conditions (see e. g., Shiffman et al., 2009). For example, recently published systematic reviews emphasize the benefits of the EMA in the investigation of substance use and craving in daily life (Serre, Fatseas, Swendsen, & Auriacombe, 2015), the impact of psychotropic medication on patients' experiences (Bos, Schoevers, & aan het Rot, 2015), anxiety symptom dynamics (Walz, Nauta, & Aan het Rot, 2014), and the dynamics of everyday mood in patients with a major depressive disorder (aan het Rot, Hogenelst, & Schoevers, 2012, 6).

The growing number of published EMA studies makes researchers more familiar with the methods. However, the complexity of statistical techniques for analyzing intensive longitudinal data is challenging. For example, the temporal nature of longitudinal data with repeated measures within subjects require a thorough inspection of possible time trends and/or cycles (Beal & Weiss, 2003). By ignoring these time patterns, statistical models can be misspecified, which results in biased outcomes. On the contrary, the use of time related predictors may reduce the error in the criterion and may, therefore, increase the probability to find theoretically relevant effects.



While growth models are relatively easy to regress with additional linear or transformed terms that match a curvilinear pattern, the procedures for cycles are more complex. The aim of this article is to support applied researchers in their choice of appropriate models by demonstrating how cyclic models can be used to fit longitudinal data.

Various cyclic physiological processes may influence psychological outcomes. For example, there is a circadian rhythm in hormone production that influences psychological states and behavior (Chow, Hamaker, Fujita, & Boker, 2009). Another well-documented cyclic predictor of psychological outcomes is the menstrual cycle (Dimmock, Wyatt, Jones, & O'Brien, 2000). Cyclic patterns can be monitored in all kinds of time frames. For EMA data, patterns within a day will often be relevant. However, weekly cycles have been shown to exist in various biological variables (for an overview, see Larsen & Kasimatis, 1990). Researchers may be interested in underlying cyclic patterns of specific outcomes or they need to account for cycles to eliminate spurious relationships that are a by-product of the time pattern (Beal & Weiss, 2003).

The cyclic model is not entirely new in psychological research. Recently, Huh, Kaysen, and Atkins (2015) demonstrated its use in modelling daily alcohol consumption and concluded that the cyclic model offers a compromise between the simplicity of the use of a dummy variable for weekends versus workdays and a full set of dummies for all days. A model with dummies for each time point (each day for example) will more accurately estimate the means per time point. However, such a model is less efficient than a cyclic model because of the larger number of parameters to estimate. When time factors are modelled to interact with other covariates in the model with dummy variables, the number of parameters may result in unstable and unreliable estimates. Cyclic models with sine and cosine terms can be an elegant solution to capture cyclic patterns in such data.

In a sample of students, Larsen and Kasimatis (1990) showed that mood had a weekly cyclic pattern, which explained 40% of the variance of mood variation on the aggregated level. However, in a more recent study (Ram et al., 2005), the weekly cycles only explain a small amount of variance in mood fluctuations. Cyclic models can be used for the estimation of effects that show both cyclic patterns common to all participants and patterns that may vary between participants (so called fixed and random effects). By using multilevel analysis, Ram et al. (2005) showed large variability between individuals with respect to cyclic patterns of mood.

Cyclic models have also been used in several other studies on psychological outcomes (e. g., Bodenmann,

Atkins, Schär, & Poffet, 2010; Chow et al., 2009; Huh et al., 2015). To further improve EMA research, it is necessary to enhance the knowledge of EMA researchers about the cyclic models and to provide tools to analyze the models. In the present paper, which should serve as a tutorial, we firstly introduce the cyclic model for applied researchers with some basic knowledge of linear multilevel modelling. Secondly, we introduce an R-package, “cyclic”, which was designed to make the cyclic analyses easy for applied researchers, and we illustrate the use of this package for several models, with the focus on multilevel models. Finally, in accordance with our aim to demonstrate the additional value of the cyclic model in EMA research, we apply the cyclic model on three variables obtained from a much larger empirical EMA data set about smoking lapse (Bolman et al., 2018) and interpret the effects of the cyclic terms.

The cyclic model

The simplest way to incorporate time (t) in the prediction of the outcome variable (Y) is by using a linear model:

$$Y = b_0 + b_1t + e \quad (1)$$

This model assumes that the outcome is linearly dependent on time with intercept b_0 and a slope b_1 , where e is an error term indicating deviation from linearity. Assuming this model has properties of a regular regression model, the error term is normally distributed with a zero mean. For instance, Y increases during the day and the next day this linear process continues. For many variables, model in Eq. 1 is not very likely. A cyclic model with a cosine term can be more realistic to capture daily patterns (see also Flury & Levri, 1999):

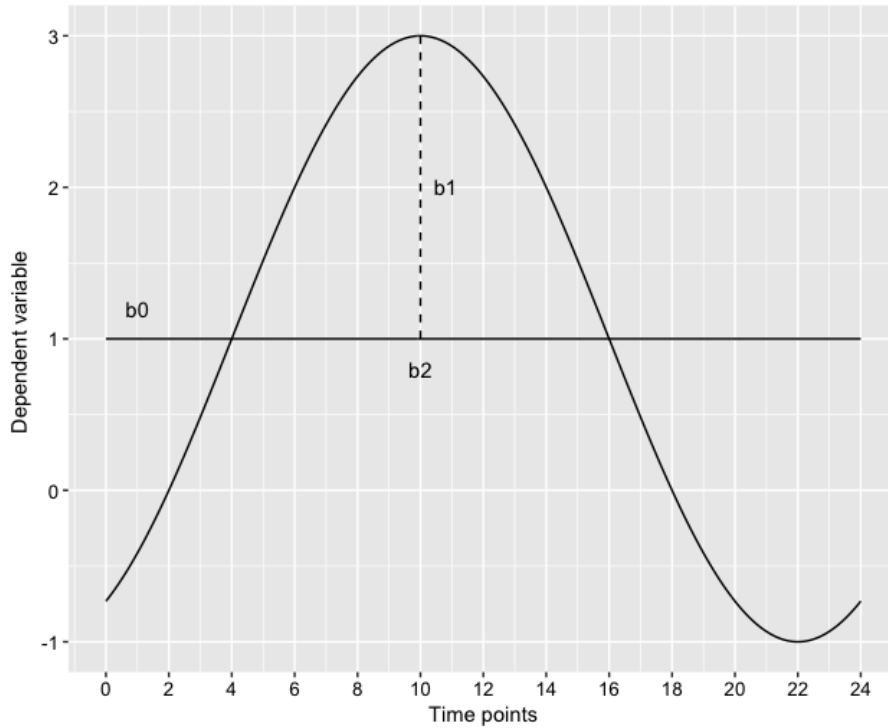
$$Y = b_0 + b_1 \cos\left(\frac{2\pi}{P}(t - b_2)\right) + e \quad (2)$$

Here, P is the period that represents the number of assessments within one cycle (e. g., 10 assessments within one day, or 7 days within one week) and t is the indicator of time at which measurement occurred; t runs from 1 to P . The cosine function ranges from -1 to +1 with the minimum value for $\cos(\pi)$, $\cos(3\pi)$, etc., and the maximum value for $\cos(0)$, $\cos(2\pi)$, $\cos(4\pi)$, etc. The parameter b_1 represents the amplitude (the height of the cycle relative to b_0), and b_2 represents the time point when this amplitude is established, i. e., the horizontal shift. The b_0 is the intercept that represents the mean value of the pattern (i. e. the vertical shift). Figure 1 depicts an example of a cyclic pattern for a 24 hours period.

In the appendix one can find how the data for this figure are generated and how this figure is constructed. The appendix also provides all R code for running the analyses,



Figure 1 ■ Example of a cyclic model for a 24 hours period, with $P = 24$, $b_0 = 1$, $b_1 = 2$ and $b_2 = 10$.



discussed in this paper, including explanations on how the figures are constructed.

How to estimate the model parameters. To estimate the cyclic model it can be rewritten using the trigonometric equality:

$$\cos(t - u) = \cos(t)\cos(u) + \sin(t)\sin(u).$$

Applied to Eq. 2 this gives:

$$\begin{aligned}
 Y &= b_0 \\
 &+ b_1 \cos\left(\frac{2\pi}{P} b_2\right) \cos\left(\frac{2\pi}{P} t\right) \\
 &+ b_1 \sin\left(\frac{2\pi}{P} b_2\right) \sin\left(\frac{2\pi}{P} t\right) \\
 &+ e
 \end{aligned}
 \tag{3}$$

Next, we define:

$$a_0 = b_0 \tag{4a}$$

$$a_1 = b_1 \cos\left(\frac{2\pi}{P} b_2\right) \tag{4b}$$

$$a_2 = b_1 \sin\left(\frac{2\pi}{P} b_2\right) \tag{4c}$$

and

$$C = \cos\left(\frac{2\pi}{P} t\right) \tag{5a}$$

$$S = \sin\left(\frac{2\pi}{P} t\right) \tag{5b}$$

which gives the following linear model:

$$Y = a_0 + a_1 C + a_2 S + e \tag{6}$$

The parameters in the model of Eq. 6 can be transformed to the original parameters b_1 and b_2 using another trigonometric rule and some elementary algebra:

$$a_1^2 + a_2^2 = b_1^2 \cos^2\left(\frac{2\pi}{P} b_2\right) + b_1^2 \sin^2\left(\frac{2\pi}{P} b_2\right) = b_1^2 \tag{7}$$

because the sin and cos form the perpendicular sides of a triangle with hypotenuse of length 1. Consequently,

$$b_1 = \sqrt{a_1^2 + a_2^2} \tag{7}$$

This shows that b_1 is always positive, representing the amplitude of the cyclic process. With b_1 known, b_2 can be derived from Eq. 4b or Eq. 4c. For example, the inverse function of the cosine (arccos) can be used to calculate b_2 from



Eq. 4b,

b2 = arccos(a1/b1) * (P/2pi) (8a)

or likewise the inverse function of the sine can be used to calculate b2 from Eq. 4c,

b2 = arcsin(a2/b1) * (P/2pi) (8b)

The R function cycpar() in the cyclic package calculates the cyclic parameters (i. e., the amplitude and the phase) from the parameter estimates obtained from the linear model (eq. 6). The amplitude (b1) is computed by Eq. 7 and the phase (b2) by Eq. 8a. For a meaningful interpretation, b2 is assured to be positive by accounting for the sign of the regression parameter a2. For example, when a1 and a2 are obtained from an analysis using Eq. 6, with values 0.5 and 1.0 respectively, and P equals 24, the function is called as: cycpar(a1 = 0.5, a2 = 1.0, P = 24), which yields a vector with two named elements:

amplitude phase
1.118 4.229.

In the present paper we assume that the period (P) is known and is not estimated from the data.

Analyzing multiple data sources

In EMA, data are obtained from a sample of persons, often measured over a number of days. This type of data thus contains time series from several sources that can be combined in a statistical analysis. Assessments are often nested within days, which are nested within subjects. Since the data from the same subject are not independent, a multi-level approach is often required to analyze this type of data (Beal & Weiss, 2003). Because Eq. 6 is transcribed as a regression model, the multilevel approach (MLA) can be applied like in other linear regression models. The intercept is usually taken as a random coefficient but, also, the cyclic parameters can be estimated as fixed or random factors. Random effect parameters in the model reflect a variation of effects across individuals or days (or both). For example, the amplitudes or phase shifts of a weekly cyclic pattern of positive affect may vary across individuals, which could imply that some persons have their peak on Saturday while others have their peak on Monday. Likewise, some persons may have a rather flat pattern (little variability in positive affect) while others have high peaks.

Another advantage of using an MLA is that, for outcome measures missing at random, missing data can be handled. EMA studies may have many missing data points. Participants may miss several individual assessments during a day, they may miss complete days, or they may drop

out from the study. Similar to the application of MLA for non-cyclic data, the researcher should be cautious while assuming that data are missing at random and should test this assumption.

In addition to a cyclic pattern, a linear trend within the time period can also be included in the MLA model. Furthermore, covariates can be added to the model to better understand the variation in the dependent variable or reduce the error variance. Finally, generalized linear models in the MLA can deal with dichotomous dependent variables.

Assuming a fixed and known periodicity (P), the cyclic multilevel model with an additional linear trend (b3) and all parameters assumed to be random effects, can be written as:

Yi = b0i + b1i*cos(2pi/P * (ti - b2i)) + b3i*ti + e (9)

Following the usual notation in the multilevel literature, the random effects are written as:

b0i = gamma00 + u0i,
b1i = gamma10 + u1i,
b2i = gamma20 + u2i,
b3i = gamma30 + u3i,

where the four gamma's are the sample means of the parameters in the model and the four u's are the individual deviations from those means. Like e in Eq. ??, the u's are assumed to be normally distributed. In most MLA software, the covariance structure of the u's can be specified.

Effect size measures in the cyclic model

To estimate the relevance of the cyclic process, a measure of effect size is necessary. The amplitude of the process can be used to specify effect size. In ordinary regression models, the standardized regression coefficients (beta's) are often used to compare effects of specific parameters. Likewise, the b1 parameter can also be standardized, which makes it possible to compare amplitudes for outcomes measured on different scales. In the context of cyclic models, it is sufficient to standardize the variable of interest before conducting the analysis, which results in an amplitude b1 measured in standard deviations.

An empirical example

The cyclic model was tested on empirical data obtained from an EMA design. The aim of these analyses was to investigate the presence of cyclic effects in EMA data. We also aimed to support applied researchers by showing, step-by-step, how this type of data can be analyzed in the software environment R (Team, 2018) with a rich graphical toolbox. We demonstrate how a multilevel analysis using cyclic models on EMA data can be performed in R.



Method

Data. Secondary analyses were conducted using the cyclic multilevel model with EMA data on smoking lapse (Bolman et al., 2018). In total, 49 individuals that had quit smoking were assessed during one week at 10 random time points per day distributed within fixed intervals. From the variables of this study, we used positive affect (PA) and intention to refrain from smoking, all measured at the level of the assessments within a day. In Bolman et al. (2018), these variables were studied as predictors of smoking lapse. PA was measured by a four-item scale originating from the PANAS (Watson, Clark, & Tellingan, 1988). An example item for PA was: ‘At this moment I feel happy’. The items used a 7-point scale (1=*not at all* to 7=*very*) and the higher total score, ranging from 4 to 28, means higher PA. Non-smoking intention was measured on a seven-point scale in line with a question used frequently in retrospective studies though adapted to measure the intention at that specific moment: ‘At this moment I do not intend to smoke’. Smoking lapse was measured by asking whether the participant had smoked a cigarette in the previous period (time from the previous assessment). There were 2935 valid records for these variables. Both variables were standardized to have zero mean and unit variance over the complete sample.

Subjects ($N = 8$) were removed from the analyses if they had less than 50 valid records (the maximum number is 70 records). In addition, three subjects were removed because their standard deviation of intention to refrain from smoking was less than 0.10. This is a very small value, which implies that their intention score was approximately constant during the whole research period. This yielded a sample of 38 subjects that were used in the analyses.

Analyses. Firstly, in analyses assuming only one level in data, the data patterns for the variable *intention to refrain from smoking* are shown for three subjects and cyclic models are fitted on each of them. Secondly, we illustrate the daily and weekly cyclic patterns for the variables stress and positive affect for one of the subjects. This is followed, by multilevel analyses in the whole sample. For the analyses of the variable *intention to refrain from smoking*, the lmer function in the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) in R (Team, 2018) was used. Six multilevel models were ran to obtain parameter estimates and measures of model fit.

The model fit was compared using the Likelihood ratio (LR) test of differences in deviances of the nested models. The information criteria AIC was additionally used to inspect the fit of the models. The final model with cyclic terms for daily and weekly cycles and the predictor stress

looks like:

$$\begin{aligned}
 Intention_i &= a_{0i} + a_{1i} C_d + a_{2i} S_d \\
 &+ a_{3i} C_w + a_{4i} S_w + a_{5i} Stress_i \\
 &+ e
 \end{aligned} \quad (10)$$

The C and S variables are defined in Eq. refoq.5a and Eq. 5b, respectively. The subscripts d and w refer to daily cycles and weekly cyclic terms, respectively. The index i refers to the random effect of these parameters, where i is the index of subjects. Table 1 gives an overview of all analyses.

Results

Cyclic patterns per participant. Three subjects (2, 15, and 18) and the average pattern across all subjects are used for the illustration of the data patterns. They were randomly selected from the top half of the participants with respect to their SD of intention. The SD of intention within a subject in the sample of 38 subjects varied between 0.22 and 1.14 (median = 0.69). The three subjects that were selected had a SD of respectively 0.70 (subject 2), 0.96 (subject 15) and 1.11 (subject 18). In Figure 2, the upper left panel shows the average pattern of the 38 subjects. The other three panels are the data from three subjects (2, 15, and 18), selected for illustration. The x-axis represents the 10 beeps within the 7 consecutive days (days have different colors) and the y-axis represents the (standardized) intention to refrain from smoking.

The four plots show unclear structures. In the aggregated plot (upper left), there seems to be slow rise in intention during the week. Based on this observation, the variable day number might be used as linear predictor in the model.

The next step was to fit the cyclic model exploring patterns within days to each of these four datasets shown in Figure 2. The results are depicted in Figure 3.

The figure shows the data points averaged over days and subjects (upper left plot) or over days only (other three plots). The cyclic lines connect predicted values of the cyclic model. The amplitude estimates for the these four analyses are respectively 0.04 ($R^2 = 0.04$ for the model) for the aggregated data, 0.14 for subject 2 ($R^2 = 0.02$), 0.32 for subject 15 ($R^2 = 0.06$), and 0.48 for subject 18 ($R^2 = 0.10$). The average amplitude is much smaller than those obtained in the individual analyses because it is based on data of all 38 subjects, including a substantial number of subjects with a very weak oscillation. Furthermore, the cycles do not have the same phase shift, thus the amplitudes are not at the same moment, which may cancel out part of the signal.

The phase shift estimates for these four analyses are respectively 7.5, 9.8, 8.2, and 3.4 for the averaged data, sub-

**Table 1** ■ Overview of the analyses.

Analysis	Dependent	Specifications	Figure
1. Single level	Intention	4 plots with raw data: for three randomly selected subjects and for the values, averaged over subjects.	Figure 2
2. Single level	Intention	4 cyclic analyses: for three randomly selected subjects and for the values, averaged over subjects.	Figure 3
3. Single level	Intention	1 cyclic analysis: for the values, averaged over subjects, with time as linear predictor added. -	
4. Single level	Intention, Stress, Positive affect	3 cyclic analyses: for three dependent variables, for one subject	Figure 4
5. Single level	Intention, Stress, Positive affect	3 cyclic analyses: for three dependent variables, for one subject. Cycle has a 7 days period, instead of a 10 beeps.	Figure 5
6. Multilevel	Intention	6 cyclic multilevel analyses with ascending model complexity Results of the fifth model are plotted	Table 2 Figure 6

ject 2, subject 15, and subject 18. The intercepts are slightly below zero, except for subject 18 who has an intercept of -0.5. Since the data were standardized, this implies that the intention to refrain from smoking of subject 18 is below the average of all the data points.

To model the apparent linear trend in intention to refrain from smoking as a function of day number, this variable was added to the model and this model was fitted on the aggregated data (i. e. the average person in Figure 3). The coefficient of this linear trend term is $b = 0.02$ ($SE = 0.01$, $p = 0.002$), while all other parameters are the same as in the previous analysis. For this model the $R^2 = 0.16$, which is substantially larger than without the linear term.

Daily cycles for stress, intention and positive affect. For illustration purposes, we analyzed the variables stress and positive affect (theoretical predictors of intention to refrain from smoking) for subject 15. In Figure 4 the data are shown for these two variables together with intention to refrain from smoking. The left panels of Figure 4 show the raw data (dots) and the predicted value connected with a line. The right panels show the means and predicted values per assessment number connected with a line.

These plots show that at around assessment 8 (late in the afternoon) the reported stress is at the lowest level, the reported positive affect is at the highest level and the intention to refrain from smoking is also at the highest level. So, for person 15, stress and positive affect seem to correlate with the intention to refrain from smoking, which was confirmed by Pearson r 's (positive affect: $r = .47$; stress, $r = -.38$). The plots also show that there is substantial variation across days (e. g., a dip on the second day), especially for positive affect and intention mirrored by a positive peak in stress.

The daily cyclic terms explained 6.0% variance of stress, 5.6% of intention, and 1.3% of positive affect. These

are modest percentages, which are also illustrated by the left panels of Figure 3 in which the data points are not very close to the predicted line.

Weekly cycles for stress, intention and positive affect. Next, we analyzed the variables stress and positive affect again for subject 15, but this time a weekly cycle was tested instead of a daily cycle. The results are presented in Figure 5. The weekly cyclic terms explained 8.3% variance of stress, 4.8% of intention, and 6.4% of positive affect. This implies that more than 8% of the variance in stress is due to the day of the week, with the highest level at the third day. As can be expected, positive affect has a similar but opposite pattern to stress for this subject. The intention to refrain from smoking seems lowest at the first few days. These are again modest percentages, which are also illustrated by the left panels in the figure.

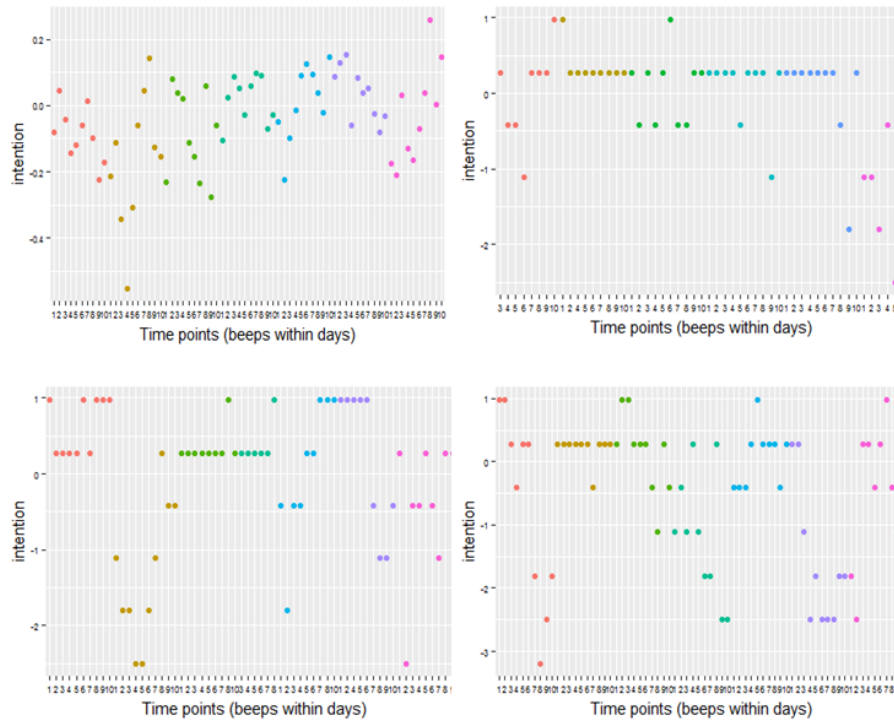
Until now, we have focused on individual data patterns. In the next section, we will simultaneously analyze the full sample of subjects using a series of models.

Multilevel analysis. Multilevel analysis using the sample ($N = 38$) was used to test whether cyclic patterns of the intention to refrain from smoking improve the fit of the model compared to a model without cyclic patterns.

Table 2 shows the results of the tested models. In the first column, the number of estimated parameters for each model is shown. Furthermore, the AIC and the deviance of the models are given (Spiegelhalter, Best, Carlin, & Van der Linde, 2002). In the column labelled " $\chi^2(df)$ ", the differences between the deviances and the difference in estimated parameters (degrees of freedom) are presented. Finally, R squared values are provided, using the method presented by (Xu, 2003). The cyclic models were tested against the null model. The null model has no predictors; it only has a fixed and random effect for the intercept. Together with the residual term this adds up to three parameters. The intraclass correlation, computed from this model, is



Figure 2 ■ Intention to refrain from smoking for all assessments (beeps within days). Upper left panel shows the results averaged over all participants (“average person”), the other three panels show data of subjects 2 (upper right), 15 (down left), 18 (down right). Different days are given separate colors.



38%, indicating that a substantial percentage of the variance of intention is due to variation between subjects (the highest level in the multilevel model).

When the model with fixed daily cyclic terms (model 2) was tested against the null model, the gain in deviance was only 3.7 with 2 degrees of freedom. As in the null model, a fixed and random intercept term and the residual term were also estimated in model 2. The likelihood ratio test indicates that adding daily cyclic terms yields no significant improvement over the null model, see Table 2. The parameters of daily cycles were, respectively, $b_1 = 0.04$ and $b_2 = 8.0$. Because the data are standardized, the intercept is approximately zero ($b_0 = -0.04$).

This result, indicating that there are no clear daily cycles for the whole sample, is not very surprising. Figures 2 and 3 already showed very different patterns in the data for different subjects.

In the third model, the cyclic terms were allowed to be random. This means that in this model, every subject can have its own amplitude and phase. Based on the significant likelihood ratio test, this model was preferred over the null model and the daily cyclic model with fixed effects. The

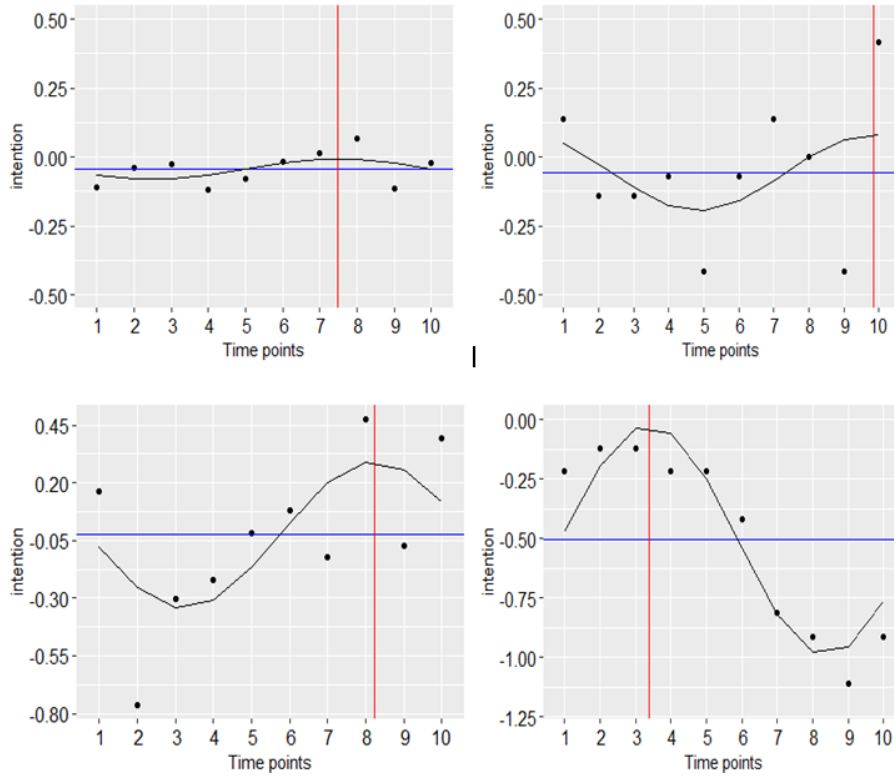
AIC (lower values indicate better fit) also indicated preference for this model over the other models. The estimated amplitude in standard deviations and phase shift were, respectively, $b_1 = 0.05$ and $b_2 = 7.9$.

To improve the model further, we added weekly cyclic terms, implying that we allowed for both daily cycles and weekly cycles. The weekly terms were added as fixed effects in this model, which means that the weekly cycles are assumed to be similar for all subjects. This model appeared to show a significant improvement upon the previous models. So, fixed weekly cycles seem to be present for intention to refrain from smoking in these data. The estimated amplitude in standard deviations and phase shift were, respectively, $b_3 = 0.10$ and $b_4 = 5.5$, while the daily parameters in this model were approximately the same as in the previous model, respectively, $b_1 = 0.05$ and $b_2 = 7.9$.

The next model was similar to the fourth model but with the weekly cyclic terms added as random effects, implying that the weekly cycles are not assumed to be equal for everybody, but are allowed to vary across subjects. This model fitted the data a lot better than a model without the random effect for weekly cycles. Table 2 shows that the



Figure 3 ■ Intention to refrain from smoking for assessments (beeps averaged across days). Upper left panel shows average person, the other three are subjects 2 (upper left), 15 (down left), 18 (down right). Predicted values from cyclic model are also plotted, and red vertical line indicates phase shift.



AIC and the deviance are much lower than for model 4. The estimated amplitude and phase shift for the daily cycles were, respectively, $b_1 = 0.04$ and $b_2 = 7.9$, while the weekly parameters for this model respectively, $b_3 = 0.09$ and $b_4 = 5.5$. Although this model fits the data better than model 4, the fixed parameter estimates are almost the same.

Figure 6 shows how the latter model fits the data. The two cyclic patterns are visible, intention is highest at the fifth and sixth day of the week and during the days there is a peak late in the afternoon. The ratio of the residual variance relative to the total variance of intention is 48% and the predicted scores from this model correlate highly with the observed values ($r = .72$).

To illustrate the use of covariates, we added the predictor stress to the model. Stress is known to be an important predictor of smoking. This model improves model 5, as indicated by the likelihood ratio test, $\chi^2(7) = 26.1$, $p < 0.001$. The corresponding fixed effect coefficient is $b_5 = -0.05$. So, an increase in reported stress is weakly associated with less intention to refrain from smoking, see

also Bolman et al. (2018). All cyclic terms appear to be similar to those in model 5. This model also fits better than the model with only stress as a (random effect) predictor without cyclic terms (deviance = 5598, $df = 6$).

Discussion

This paper has illustrated that cyclic models may be useful for modelling dynamic processes in intensive longitudinal designs such as EMA. First, the cyclic model was introduced and it was shown how cyclic terms could be incorporated in linear models. Second, a step-by-step example was presented to show how analyses using a cyclic model can be performed.

When comparing models with and without cyclic terms for the data used in the example, the model improvements were quite small. In general, small effects of cyclic terms would probably not change the estimates of other regression parameters. In the smoking data the effects of daily stress and craving to smoke changed only slightly after adding the cyclic terms in the model predicting the non-smoking intention as in the original model presented by



Table 2 ■ Model comparisons for intention to refrain from smoking.

	# pars	R^2	AIC	Deviance	$\chi^2(df)$	p
Model						
1. Null	3	.374	5640	5634		
2. DF	5	.375	5640	5630	3.7 (2)	.151
3. DR	10	.407	5620	5600	29.9 (5)	.000
4. DR_WF	12	.413	5603	5579	21.1 (2)	.000
5. DR_WR	21	.515	5369	5327	252.4 (9)	.000
6. DR_WR_SR	28	.527	5357	5301	26.1 (7)	.000

Note. Note. DF = model with fixed effects for daily cyclic terms, DR = model with random effects for daily cyclic terms. DR_WF = model with daily and weekly cyclic terms, daily terms are random, weekly terms are fixed. DR_WR = model with daily and weekly cyclic terms, all terms are random. DR_WR_SR = as model DR_WR with an additional random effect for the covariate stress.

Bolman et al. (2018).

However, when cyclic effects are stronger, or when both predictor and criterion follow a similar cyclic pattern, it is likely that the cyclic process may influence a relationship between some predictors and criterion. Theoretical arguments for the causal process underlying the model should then be used to establish whether the cyclic process provides valuable information to the model or whether it is merely masking the interesting theoretical relationships.

The secondary analyses on the smoking data showed a cyclic daily pattern for stress and positive affect. This is in accordance with Larsen and Kasimatis (1990) who studied mood fluctuations and Chow et al. (2009) who discussed cyclic models in state-space modelling and presented an example with daily diary data. In the example of Chow et al. with daily measurements of affect during 52 days, weekly cyclic patterns of affect were found.

In this tutorial, we used the beep numbers as proxies for the time points (t), because the exact time of the measurements were not available. This is actually not correct because the beeps are most likely not equally spaced in time. It is important to realize that the time term, t , should be measured at intervals with equal distances between all measurements (an interval-contingent EMA design with a fixed interval). However, some authors consider this a “true” EMA design when reports are conducted at randomly generated signals within a number of scheduled time blocks (a signal-contingent design). This method was originally intended to assess the events and experiences in a representative way during the individual’s daily life (Fisher & To, 2012). If the time blocks are not of the same length, the distance between the baseline point and the center of a time block can be used to code the variable t ’s (e. g., 2, 4, 7 hours, etc.), while the total length of all time blocks describes the period (P) in the formula of the cyclic model. If the exact time of responding is recorded, it is best to use it as time variable in the cyclic model.

Once the basic ideas behind the cyclic model are understood, it is easy to build models with additional predictors (covariates). Also, moderation effects can be taken into account by adding product terms to the model (Hayes, 2018). For instance, interaction effects between person’s characteristics and the time factor could be of interest. The interpretation of such a moderation effect is essentially the same as in linear regression models; namely, the strength of an effect of a predictor on Y varies with time (of the day or of the week) or the time pattern varies over personal characteristics. However, with cyclic terms, this variation is not a linear function of time but a cyclic one. Another situation that involves interaction effects with time is when the trends differ between groups, for instance between men and women, where the cycle of women differs from the men’s cycle. When a researcher has learned how to analyze and interpret cyclic models, such extensions are rather straightforward and may provide insight into the causal processes underlying the data.

A limitation of this paper is the fact that we assumed a fixed period (7 days, 10 beeps) in all our analyses. An important question in cyclic modelling is how to establish the period of a cyclic process. Usually the number of days or the number of hours within a day is taken as a convenient period but other periods are possible. To explore which period suits best with the data, spectral analyses can be used. Spectral analysis decomposes the variance in the data by applying a series of sine and cosine waves to account for the variance. See for example Larsen and Kasimatis (1990) to learn more about this approach.

We recommend researchers with intensive longitudinal data, such as EMA data, to explore cyclic terms like any other trend that may be present in the data. Adding cyclic terms may not only improve the model fit, but could also yield better insight in the dynamic processes underlying the variable.



References

- aan het Rot, M., Hogenelst, K., & Schoevers, R. A. (2012). Mood disorders in everyday life: A systematic review of experience sampling and ecological momentary assessment studies. *Clinical Psychology Review, 32*, 510–523. doi:[10.1016/j.cpr.2012.05.007](https://doi.org/10.1016/j.cpr.2012.05.007)
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. doi:[10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01)
- Beal, D. J., & Weiss, H. M. (2003). Methods of ecological momentary assessment in organizational research. *Organizational Research Methods, 6*(4), 440–464. doi:[10.1177/1094428103257361](https://doi.org/10.1177/1094428103257361)
- Bodenmann, G., Atkins, D. C., Schär, M., & Poffet, V. (2010). The association between daily stress and sexual activity. *Journal of Family Psychology, 24*(3), 271–279. doi:[10.1037/a0019365](https://doi.org/10.1037/a0019365)
- Bolman, C., Verboon, P., Jacobs, N., Thewissen, V., Boonen, V., & Soons, K. (2018). Predicting smoking lapses in the first week of quitting: An ecological momentary assessment study. *Journal of Addiction Medicine, 12*(1), 65–71. doi:[10.1097/ADM.0000000000000365](https://doi.org/10.1097/ADM.0000000000000365)
- Bos, F. M., Schoevers, R. A., & aan het Rot, M. (2015). Experience sampling and ecological momentary assessment studies in psychopharmacology: A systematic review. *European Neuropsychopharmacology, 25*(11), 1853–1864. doi:[10.1016/j.euroneuro.2015.08.008](https://doi.org/10.1016/j.euroneuro.2015.08.008)
- Chow, S.-M., Hamaker, E. L., Fujita, F., & Boker, S. M. (2009). Representing time-varying cyclic dynamics using multiple-subject state-space models. *The British Journal of Mathematical and Statistical Psychology, 62*(3), 683–716. doi:[10.1348/000711008X384080](https://doi.org/10.1348/000711008X384080)
- Conner, T. S., & Mehl, M. R. (2015). Ambulatory assessment: Methods for studying everyday life. In R. Scott & S. Kosslyn (Eds.), *Emerging trends in the social and behavioral sciences: An interdisciplinary, searchable, and linkable resource* (pp. 1–15). doi:[10.1002/9781118900772.etrds0010](https://doi.org/10.1002/9781118900772.etrds0010)
- Dimmock, P. W., Wyatt, K. M., Jones, P. W., & O'Brien, P. M. (2000). Efficacy of selective serotonin-reuptake inhibitors in premenstrual syndrome: A systematic review. *Lancet, 356*(9236), 1131–1136. doi:[10.1016/S0140-6736\(00\)02754-9](https://doi.org/10.1016/S0140-6736(00)02754-9)
- Fisher, C. D., & To, M. L. (2012). Using experience sampling methodology in organizational behavior. *Journal of Organizational Behavior, 33*, 865–877. doi:[10.1002/job.1803](https://doi.org/10.1002/job.1803)
- Flury, B. D., & Levri, E. P. (1999). Periodic logistic regression. *Ecology, 80*(7), 2254–2260. doi:[10.1890/0012-9658](https://doi.org/10.1890/0012-9658)
- Hayes, A. F. (2018). *Introduction to moderation, mediation, and conditional process analysis*. A regression based approach. New York: Guilford Press.
- Huh, D., Kaysen, D. L., & Atkins, D. C. (2015). Modeling cyclical patterns in daily college drinking data with many zeroes. *Multivariate Behavioral Research, 50*(2), 184–196. doi:[10.1080/00273171.2014.977433](https://doi.org/10.1080/00273171.2014.977433)
- Larsen, R. J., & Kasimatis, M. (1990). Individual differences in entrainment of mood to the weekly calendar. *Journal of Personality and Social Psychology, 58*(1), 164–171. doi:[10.1037/0022-3514.58.1.164](https://doi.org/10.1037/0022-3514.58.1.164)
- Ram, N., Chow, S. Y. M., Bowles, R. P., Wang, L., Grimm, K., Fujita, F., & Nesselroade, J. R. (2005). Examining interindividual differences in cyclicity of pleasant and unpleasant affects using spectral analysis and item response modeling. *Psychometrika, 70*(4), 773–790. doi:[10.1007/s11336-001-1270-5](https://doi.org/10.1007/s11336-001-1270-5)
- Serre, F., Fatseas, M., Swendsen, J., & Auriacombe, M. (2015). Ecological momentary assessment in the investigation of craving and substance use in daily life: A systematic review. *Drug and Alcohol Dependency, 148*, 1–20. doi:[10.1016/j.drugalcdep.2014.12.024](https://doi.org/10.1016/j.drugalcdep.2014.12.024)
- Shiffman, S., Stone, A. A., & Hufford, M. R. (2009). Ecological momentary assessment. *Annual Review of Clinical Psychology, 4*, 1–32. doi:[10.1146/annurev.clinpsy.3.022806.091415](https://doi.org/10.1146/annurev.clinpsy.3.022806.091415)
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society, Series B, 64*(4), 583–616. doi:[10.1111/1467-9868.00353](https://doi.org/10.1111/1467-9868.00353)
- Team, R. C. (2018). *R: A language and environment for statistical computing*. Vienna, Austria. URL: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Trull, T. J., & Ebner-Priemer, U. (2014). The role of ambulatory assessment in psychological science. *Current Directions in Psychological Science, 23*(6), 466–470. doi:[10.1177/0963721414550706](https://doi.org/10.1177/0963721414550706)
- Walz, L. C., Nauta, M. H., & Aan het Rot, M. (2014). Experience sampling and ecological momentary assessment for studying the daily lives of patients with anxiety disorders: A systematic review. *Journal of Anxiety Disorders, 28*(8), 925–937. doi:[10.1016/j.janxdis.2014.09.022](https://doi.org/10.1016/j.janxdis.2014.09.022)
- Watson, D., Clark, L., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The panas scales. *Journal of Personality & Social Psychology, 54*, 1063–1070. doi:[10.1037/0022-3514.54.6.1063](https://doi.org/10.1037/0022-3514.54.6.1063)
- Xu, R. (2003). Measuring explained variation in linear mixed effects models. *Statistics in Medicine, 22*(22), 3527–3541. doi:[10.1002/sim.1572](https://doi.org/10.1002/sim.1572)



Appendix: A tutorial for analyzing cyclic models

Preliminary Steps

There are three functions used in this appendix, which are in the R package `cyclic`. This package can be installed from Github by running the following command in R:

```
devtools::install_github("PeterVerboon/cyclicpkg/cyclic")
```

The package is then activated by running:

```
require(cyclic)
```

The package consists of three functions:

- The function `fitCyclic()` fits a linear model with `lm()` to obtain cyclic parameters.
- The function `fitCyclicMLA()` fits a MLA model with `lmer()` from the `lme4` package to obtain cyclic parameters.
- The function `cycpar()` is an auxiliary function, which computes the cyclic parameters from the linear model parameters.

These functions should be helpful for applying a cyclic model to your data. The functions have methods for printing the results and plotting the fitted cyclic patterns.

As an illustration of a cyclic process we generate perfect cyclic data and construct Figure 1 with these data using the following code. The package `ggplot2` is used for plotting.

```
require(ggplot2)

a0 <- 1; b1 <- 2; b2 <- 10; P <- 24; T <- seq(0, 24, by = 0.1);
y <- a0 + b1*cos(2*(pi/P)*(T - b2))
g <- ggplot() + geom_line(aes(x=T, y=y))
g <- g + labs(x = "Time_points", y = "Dependent_variable_")
g <- g + theme(axis.text = element_text(size = 10, colour="black"))
g <- g + scale_x_continuous(breaks = seq(0,24,2))
g <- g + geom_line(aes(x=b2, y=seq(1,3,0.1)), lty="dashed")
g <- g + geom_line(aes(x=seq(0,24,0.1), y=a0))
g <- g + annotate("text", x = 1, y = 1.2, label = "b0")
g <- g + annotate("text", x = 10.8, y = 2, label = "b1")
g <- g + annotate("text", x = 10, y = 0.8, label = "b2")
g
```

The “smoking” data

The “smoking” data are in the object `smokingdat.rda`. This is a data frame, which contains a selection of variables from a larger data file (for more information about the data, see Bolman et al., 2018). Check the names in the data frame and then put the data in `dat1`, the first data set used in the analysis.

```
data("smokedat")
names(smokedat)
## [1] "subjnr"      "beepnr"      "daynr"      " positiveAffect "
## [5] "intention"   "stress"
dat1 <- smokedat
nall <- length(unique(dat1$subjnr))
```

Count the number of records per subject and remove subjects with less than 50 records. The resulting data frame is called `dat3`.

```
dat1$count <- 1
dat2 <- aggregate(dat1[,c("count")], by=list(dat1$subjnr), FUN=sum, na.rm=T)
dat2$subjnr <- dat2$Group.1
```



```
dat1$count <- NULL
dat2$count <- dat2[,2]
dat3 <- merge(dat1, dat2[,c("subjnr", "count")], by.x = "subjnr")
rm(dat2)
dat3 <- subset(dat3, dat3$count >= 50)
nfew <- nall - length(unique(dat3$subjnr))
```

We also compute variability of intention per subject and remove subjects with very small variation in intention. The result is also in dat3.

```
dat2 <- aggregate(dat3$intention, by=list(dat3$subjnr), FUN=sd, na.rm=F)
dat2$subjnr <- dat2$Group.1
dat2$Group.1 <- NULL
dat2 <- merge(dat3, dat2, by.x = "subjnr")
dat3 <- subset(dat2, dat2$V1 > .10)
rm(dat2)
nunique <- length(unique(dat3$subjnr))
nsmall <- (nall - nfew) - nunique
```

The result of these steps are summarized below.

```
## Number of subjects in data set (nall):           49
## Number of subjects with less than 50 records (nfew): 8
## Number of subjects with SD smaller than .10, nsmall: 3
## Number of subjects used in analysis (nunique):    38
```

The next step is to construct a dataset that is aggregated over subjects, which is called dat4. This data frame will be used in subsequent analyses.

```
dat4 <- aggregate(dat3[,c("positiveAffect", "stress", "intention")],
  by=list(dat3$beepnr, dat3$daynr), FUN=mean, na.rm=F);
dat4$daynr <- dat4$Group.2
dat4$beepnr <- dat4$Group.1
dat4$Group.1 <- NULL
dat4$Group.2 <- NULL
```

Single-level Analyses

Step 1

The raw data for three subjects and average are plotted. We have to repeat step 1 for the aggregated data in dat4, and for subjects 2, 15 and 18. This makes Figure 2 from the tutorial.

```
pdat <- dat4 # averaged over all subjects
pdat <- subset(dat3, dat3$subjnr == 2) # ppn 2, 15, 18 selected for Figure |ref{fig:1}
npoints <- dim(pdat)[1]
x <- c(1:npoints)
pdat$day <- as.factor(pdat$daynr)
g <- ggplot(pdat, aes(x=x, y=pdat$intention, colour=pdat$day)) +
  geom_point() +
  scale_x_discrete(name = "Time_points_(beeps_within_days)",
    labels=pdat$beepnr, limits=c(1:npoints)) +
  theme(axis.text = element_text(size = 6, colour="black"),
    legend.position="none")
```

g



Step 2

Now we analyze the cyclic model and plot the result, again for the average and for subjects 2, 15, and 18. This is Figure 3 from the tutorial.

```
model_a <- fitCyclic(pdat, yvar = "intention", xvar = "beepnr",
  grp = "daynr", ymin = -2.5, ymax = 1.5, step = .30)
print(model_a)
plot(model_a, type = c("raw", "means"))
```

Step 3

Fit an extra term in the model with day as covariate in aggregated data. Show the parameters and the fit of the model.

```
model_b <- fitCyclic(dat4, yvar = "intention", xvar = "beepnr",
  grp = "daynr", cov = "daynr", ymin = -0.5, ymax = 0.5)
print(model_b)
```

Step 4

Apply the cyclic model with daily period for the variables stress and positive affect for subject 15 to obtain a plot of the raw data with the fitted cycles and a plot of the aggregated data. Together with intention (model_a) this makes Figure 4.

```
pdat <- subset(dat3, dat3$subjnr == 15)
model_c <- fitCyclic(pdat, yvar = "stress", xvar = "beepnr",
  grp = "daynr", ymin = -1.0, ymax = 0.5, step = 0.25)
plot(model_c, type = c("raw", "means"))
model_c <- fitCyclic(pdat, yvar = "positiveAffect", xvar = "beepnr",
  grp = "daynr", ymin = -1.0, ymax = 0.5, step = 0.25)
plot(model_c, type = c("raw", "means"))
```

Step 5

Apply, only for subject 15, the cyclic model with a weekly period for intention, stress and positive affect, instead of a daily period. Below the code is given which produces the six separate plots of Figure 5.

```
pdat <- subset(dat3, dat3$subjnr == 15)
model_d <- fitCyclic(pdat, yvar = "intention", xvar = "daynr",
  grp = "daynr", ymin = -2.0, ymax = 1.0, step = 0.25)
plot(model_d, type = c("raw", "means"))

model_d <- fitCyclic(pdat, yvar = "stress", xvar = "daynr",
  grp = "daynr", ymin = -2.0, ymax = 1.0, step = 0.25)
plot(model_d, type = c("raw", "means"))

model_d <- fitCyclic(pdat, yvar = "positiveAffect", xvar = "daynr",
  grp = "daynr", ymin = -2.0, ymax = 1.0, step = 0.25)
plot(model_d, type = c("raw", "means"))
```

Multilevel Analyses

Step 6

Analyze the data with the cyclic model using MLA and model comparison. Each consecutive model is more complex. The intraclass correlation is computed with the null model (model1). This is the first model that is fitted on the data.

```
model1 <- fitCyclicMLA(dat=dat3, yvar="intention", id = "subjnr")
```




```
print(model1)
```

The next two models fit a daily cycle to the data. Model2 with only the intercept as random term and model3 includes the cyclic terms as random effects.

```
model2 <- fitCyclicMLA(dat=dat3, yvar="intention", xvar1="beepnr",  
  xvar2="daynr", id = "subjnr", ncycle = 1, random = "intercept")  
print(model2)  
model3 <- fitCyclicMLA(dat=dat3, yvar="intention", xvar1="beepnr",  
  xvar2="daynr", id = "subjnr", ncycle = 1, random = "first")  
print(model3)  
plot(model3)
```

Model 4 and 5 are fitting two cyclic processes: patterns within days and patterns within weeks. In model 5 every effect is assumed to be random, whereas in model 4 only the within days cyclic terms are assumed random. Plotting model 5 gives a plot as is shown in Figure 6.

```
model4 <- fitCyclicMLA(dat=dat3, yvar="intention", xvar1="beepnr",  
  xvar2="daynr", id = "subjnr", ncycle = 2, random = "first")  
print(model4)  
model5 <- fitCyclicMLA(dat=dat3, yvar="intention", xvar1="beepnr",  
  xvar2="daynr", id = "subjnr", ncycle = 2, random = "all",  
  ymin = -0.5, ymax = 0.5, step=0.10 )  
print(model5)  
plot(model5)
```

Model 6 also contains a covariate. Finally all models are compared with each other, this is shown in Table \ref{tab:2}.

```
model6 <- fitCyclicMLA(dat=dat3, yvar="intention", xvar1="beepnr",  
  xvar2="daynr", id = "subjnr", cov = "stress", ncycle = 2,  
  random = "all" )  
print(model6)
```

```
anova(model6$fit, model5$fit, model4$fit, model3$fit,  
  model2$fit, model1$fit)
```

Citation

Verboon, P., & Leontjevas, R. (2018). Analyzing cyclic patterns in psychological data: A tutorial. *The Quantitative Methods for Psychology*, 14(4), 218–234. doi:10.20982/tqmp.14.4.p218

Copyright © 2018, Verboon and Leontjevas. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Received: 16/07/2018 ~ Accepted: 15/10/2018

Figures 4 to 6 follows.



Figure 4 ■ The variables intention to refrain from smoking, stress and positive affect for subject 15. Left panels show raw data with estimated daily cycle, right panels show scores averaged over days. Predicted values from cyclic model are also plotted, and red vertical line indicates phase shift.

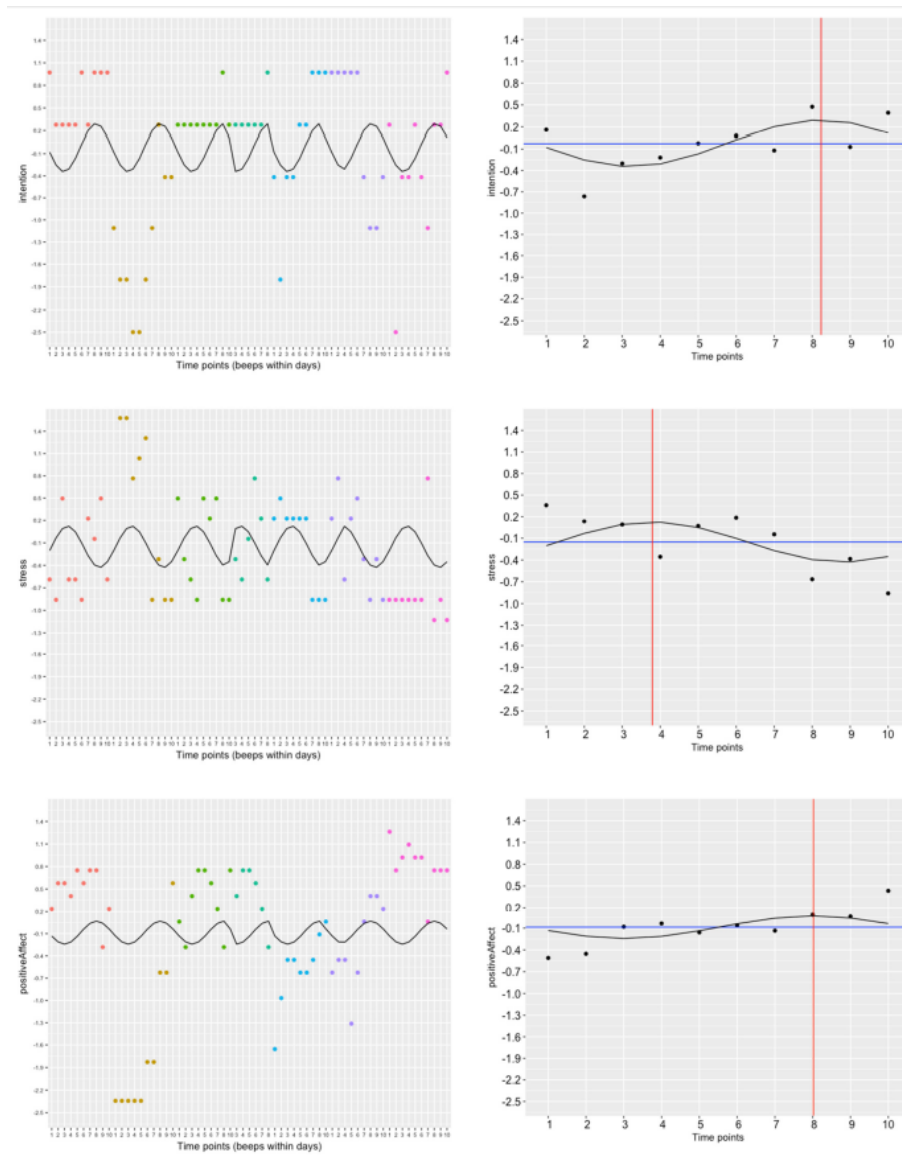




Figure 5 ■ The variables intention to refrain from smoking, stress and positive affect for subject 15. Left panels show raw data with estimated weekly cycle, right panels show scores averaged over beeps. Predicted values from cyclic model are also plotted, and red vertical line indicates phase shift.

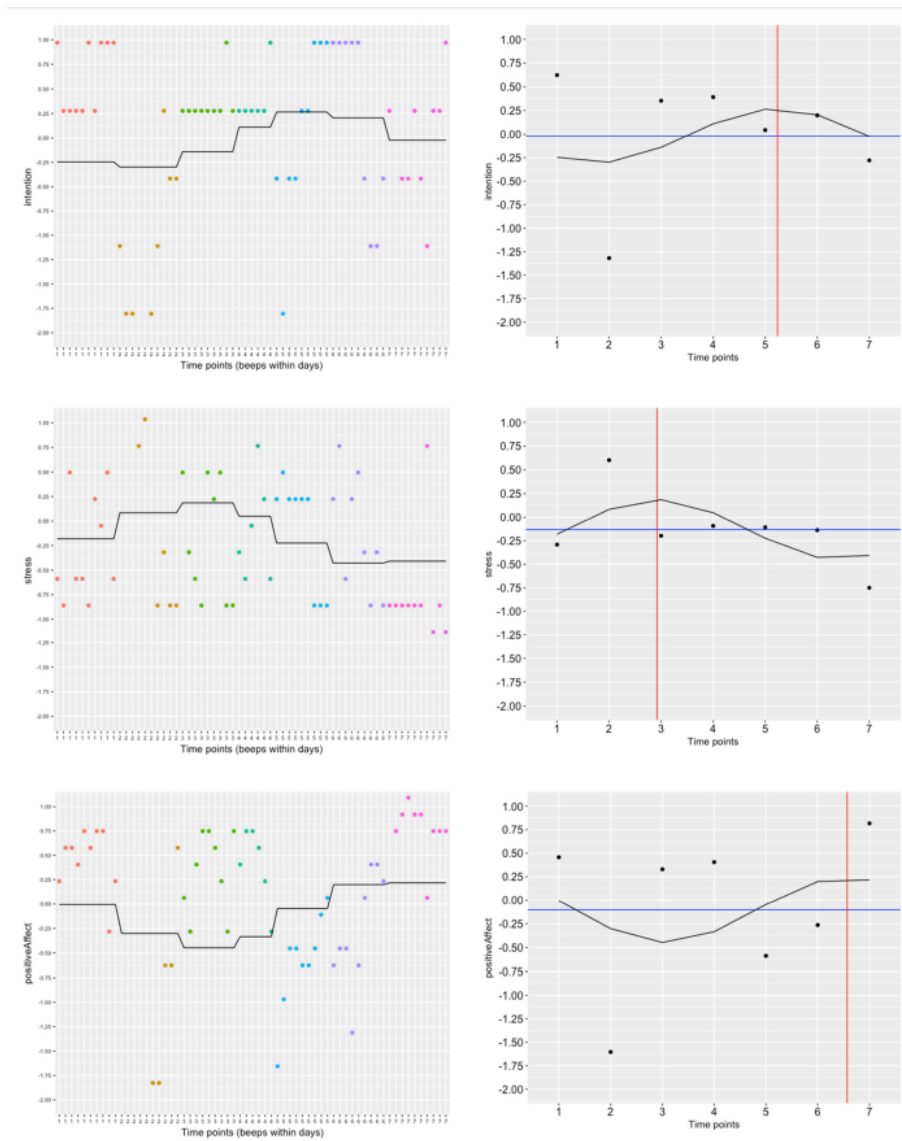




Figure 6 ■ Intention to refrain from smoking aggregated over subjects and predicted values from model 5 shown by line.

