Biometrics 61, 1115–1122
December 2005

On Analyzing Circadian Rhythms Data Using Nonlinear Mixed Models with Harmonic Terms

Paul S. Albert* **and Sally Hunsberger**

Biometric Research Branch, National Cancer Institute,
Executive Plaza North, Room 8136, Bethesda, Maryland 20892-7434, U.S.A.
**email:* albertp@ctep.nci.nih.gov

Summary. Wang, Ke, and Brown (2003, *Biometrics* **59,** 804–812) developed a smoothing-based approach for modeling circadian rhythms with random effects. Their approach is flexible in that fixed and random covariates can affect both the amplitude and phase shift of a nonparametrically smoothed periodic function. In motivating their approach, Wang et al. stated that a simple sinusoidal function is too restrictive. In addition, they stated that "although adding harmonics can improve the fit, it is difficult to decide how many harmonics to include in the model, and the results are difficult to interpret." We disagree with the notion that harmonic models cannot be a useful tool in modeling longitudinal circadian rhythm data. In this note, we show how nonlinear mixed models with harmonic terms allow for a simple and flexible alternative to Wang et al.'s approach. We show how to choose the number of harmonics using penalized likelihood to flexibly model circadian rhythms and to estimate the effect of covariates on the rhythms. We fit harmonic models to the cortisol circadian rhythm data presented by Wang et al. to illustrate our approach. Furthermore, we evaluate the properties of our procedure with a small simulation study. The proposed parametric approach provides an alternative to Wang et al.'s semiparametric approach and has the added advantage of being easy to implement in most statistical software packages.

Key words: Harmonic models; Nonlinear models; Periodic data; Seasonal data; Smoothing.

1. Introduction

The development of flexible models for characterizing circadian rhythms in longitudinal data is an important problem in biology and medicine. We commend Wang, Ke, and Brown (2003) for the development of a flexible semiparametric model for analyzing longitudinal circadian rhythm data. Specifically, they proposed that the effect of covariates and individual effects (random effects) alter the circadian rhythm by changing the amplitude and phase shift but not the shape of the pattern. They proposed a two-stage random effects model whereby in the first stage,

$$y_{ij} = \phi_{1i} + \exp(\phi_{2i}) f(t_{ij} - \text{alogit}(\phi_{3i})) + \epsilon_{ij}, \qquad (1)$$

where $i = 1, \ldots, m, j = 1, \ldots, n_i$, m is the number of individuals, n_i is the number of observations for the ith individual, $\text{alogit}(x) = \exp(x)/(1 + \exp(x))$, and t_{ij} is the time of the jth observation on the ith subject. In addition, the error terms ϵ_{ij} are assumed Gaussian with mean zero and variance σ^2. Wang et al. (2003) proposed that in the second stage,

$$\boldsymbol{\phi}_i = \boldsymbol{A}_i \boldsymbol{\beta} + \boldsymbol{B}_i \boldsymbol{b}_i, \quad \text{where} \quad \boldsymbol{b}_i \text{ i.i.d.} \sim N(\boldsymbol{0}, \boldsymbol{D}), \qquad (2)$$

where $\boldsymbol{\phi}_i = (\phi_{i1}, \phi_{i2}, \phi_{i3})'$, $\boldsymbol{\beta}$ is a p-dimensional vector of fixed-effect parameters, $\boldsymbol{b}_i$ is a q-dimensional vector of random effects associated with individual i which has mean zero and variance $\boldsymbol{D}$, and $\boldsymbol{A}_i$ and $\boldsymbol{B}_i$ are $3 \times p\}$ and $3 \times q$ design matrices. Wang et al. (2003) illustrated their methodology with the comparison of circadian rhythms in cortisol levels between small groups of depressed ($n = 11$) and normal ($n = 9$) patients with 12 equally spaced follow-up measurements over a 24-hour period. When modeling each of the two groups separately, the design matrix $\boldsymbol{A}_i$ can be specified as $\boldsymbol{A}_i = (1, 0, 0)'$. For fitting a model in which the overall mean, amplitude, and phase shift can vary by group,

$$\boldsymbol{A}_i = \begin{pmatrix} 1 & G & 0 & 0 \\ 0 & 0 & G & 0 \\ 0 & 0 & 0 & G \end{pmatrix}, \qquad (3)$$

where G is an indicator variable for group. For both models, $\boldsymbol{B}_i$ is a 3×3 identity matrix corresponding to a single random intercept for the overall mean, amplitude, and phase shift.

This two-stage approach generalizes the earlier, more simplified fixed-effects approach by Wang and Brown (1996) in which separate models were fit for each individual's data and these estimates were subsequently summarized across individuals. Wang et al. (2003) developed a framework for a formal two-stage approach for modeling longitudinal circadian rhythm data. Their methodology uses nonparametric estimation of the underlying function f, which involves maximizing a penalized likelihood to determine the degree of smoothing required for a particular data set. The authors use conditional t-tests for testing group differences in circadian rhythms, which condition on the choice of f as well as the variance estimates of the random effects and residual variance (Wang,

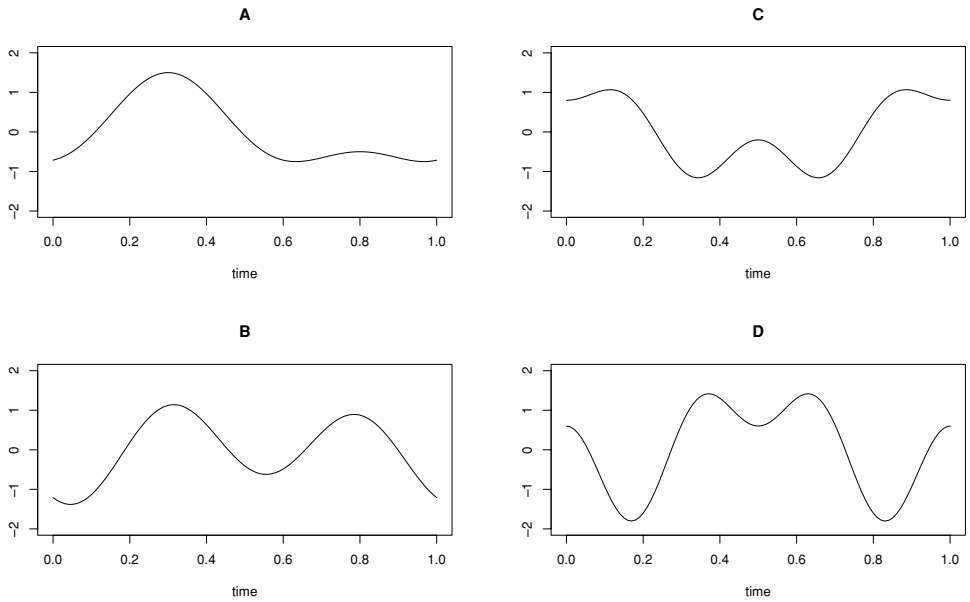

Figure 1. Circadian rhythms represented by f with harmonic models. (A) and (B) show two-harmonic models, while (C) and (D) show three-harmonic models. Patterns are generated as $f(t) = A \times \cos(2\pi(t + B)) + C \times \cos(4\pi(t + D)) + E \times \cos(6\pi(t + F))$, where for (A): $A = 1$, $B = 0.7$, $C = 0.5$, $D = 0.7$, $E = 0$, $F = 0$, for (B): $A = 0.4$, $B = 0.5$, $C = 1$, $D = 0.2$, $E = 0$, $F = 0$, for (C): $A = 1$, $B = 0$, $C = 0.3$, $D = -0.5$, $E = 0.5$, $F = -0.5$, and for (D): $A = 1$, $B = -0.5$, $C = 0.6$, $D = 0$, $E = 1$, $F = 0$.

1998). Wang and Ke (2002) developed the software package ASSIST for fitting these models.

Wang et al. (2003) stated that the sinusoidal curves are too restrictive for describing complex circadian patterns and stated that "although adding harmonics can improve the fit, it is difficult to decide how many harmonics to include in the model, and the results are difficult to interpret." In this article, we show how parametric models can be used for making appropriate inferences when analyzing circadian rhythm longitudinal data. In Section 2, we demonstrate how nonlinear mixed models with harmonic terms can be used to flexibly fit circadian rhythms to longitudinal data. We analyze the cortisol data presented by Wang et al. (2003) with the parametric modeling approach in Section 3. In Section 4, we show using simulations that valid inferences can be made in small samples with this parametric approach. In Section 5, we contrast the parametric and semiparametric approaches.

2. Nonlinear Mixed Models with Harmonic Terms

An important feature of Wang et al.'s approach is the shape-invariant nature of how covariates and random effects affect the underlying circadian rhythm. Model (1) can be developed, where instead of nonparametrically estimating f, as is done by Wang et al. (2003), we parametrically model f with a series of harmonic terms to flexibly characterize the circadian pattern. Specifically, the function f can be parameterized as

$$f(t) = \sum_{k=1}^{K} \beta_k \cos\{2k\pi t_{ij} + \theta_k - \text{alogit}(\phi_{3i})\}, \qquad (4)$$

where t_{ij} is in fractions of a day ranging from 0 to 1 starting at midnight, and where K is the number of harmonics, where a large K results in a very flexible circadian rhythm at the expense of estimating additional parameters. The parameters $\beta_1, \beta_2, \ldots, \beta_K, \theta_1, \theta_2, \ldots, \theta_K$ characterize a flexible circadian rhythm. Although we agree with Wang et al. (2003) that a single harmonic term is very restrictive, a model with three harmonics can very flexibly describe most circadian rhythms. Figure 1 shows a series of circadian patterns with $K = 2$ and $K = 3$. The figure demonstrates the wide range of circadian patterns which can be characterized with a few harmonics.

Model (1) with f parameterized by (4) can be fit with various nonlinear mixed modeling software routines. We use the `nlme` procedure in R (Venables et al., 2004) or S-plus (2001), which uses methodology developed by Lindstrom and Bates (1990) for parameter estimation. The Appendix provides an example of code for fitting these models in R.

Wang et al. (2003) proposed a penalized-likelihood approach for choosing the smoothing parameter in their semiparametric approach. Similarly, we propose choosing the number of harmonics using a penalized likelihood method such as either the Akaike information criterion (AIC) (Akaike, 1973), where $\text{AIC}(p) = -2 \log L + 2p$ and p is the number of model parameters, or the Bayesian information criterion (BIC) (Schwartz, 1978), where $\text{BIC}(p) = -2\log L + p \log(\sum_{i=1}^{m} n_i)$. Since f is very flexible with three harmonics, in most situations, we recommend choosing between one, two, or three harmonics.

An assumption for both the semiparametric and parametric shape-invariant models is that the underlying shape of the circadian rhythm does not vary across groups. Wang et al. (2003) proposed an informal test for comparing the shapes of the circadian rhythms between normal and depressed groups.

Furthermore, they stated that methodology based on a sinusoidal function cannot be used to test this assumption. Contrary to their statement, we show that a test comparing the shapes of the curves in different groups can be formulated for the parametric model. In model (1), we compare $f(t)$ given by (4) with $f(t)$ given by

$$f(t) = \sum_{k=1}^{K} [\beta_{1k} \cos\{2k\pi t_{ij} + \theta_{1k} - \text{alogit}(\phi_{3i})\} + \beta_{2k} G \cos\{2k\pi t_{ij} + \theta_{2k} G - \text{alogit}(\phi_{3i})\}], \quad (5)$$

where $\beta_{21} = \theta_{21} = 0$, and G is a group indicator. This can be tested with a likelihood-ratio test of whether $\theta_{2k} = 0$ and $\beta_{2k} = 0$ for $k = 2, 3, \ldots, K$ in (5). We show in Section 4 that this likelihood-ratio test has the correct type I error rate in small samples.

Assuming the scale-invariant model, for the two-group comparison in the cortisol example (normal versus depressed patients), we propose choosing the number of harmonics based on minimizing either the AIC or BIC for $K = 1$, 2, or 3 harmonics. Inferences on group differences can be made using the chosen model with conditional t-tests. For the two-group comparison, the conditional t-test, which is conditional on estimates of the random effects and residual error variances, takes the form of the estimated fixed-effect parameter divided by its standard error. A rule for determining the appropriate degrees of freedom for the conditional t-test is given by Pinheiro and Bates (2000). The t-test is provided in both the R and S-plus software package. We evaluate the properties of this approach with simulations in Section 4.

In the next section, we fit the harmonic models to the cortisol circadian rhythm data presented by Wang et al. (2003) and compare inferences with those made with the semiparametric approach.

3. Analysis

We fit the harmonic models to the cortisol data presented by Wang et al. (2003). Initially, we fit the model (1) with $f(t)$ given by (4) to the depressed and normal groups separately. We fit these models with one, two, and three harmonics. For the depressed group, AIC was 269.1, 242.3, and 244.9 for models with one, two, and three harmonics, respectively. For the normal group, AIC was 222.9, 175.9, and 179.9 for models with one to three harmonics. For a model with no group effects on the combined data, AIC was 481.4, 408.3, and 412.0 for models with one to three harmonics. Thus, AIC was minimized for models with two harmonics for the depressed, normal, and combined groups. The BIC, which tends to favor simpler models than AIC in data sets with many observations, was also minimized with two-harmonic models for each group separately as well as for the two groups combined (data not shown). Figure 2 shows estimates of f obtained using two harmonics in each group as well as the combined group. Ninety-five percent confidence intervals were obtained by the bootstrap. The forms of the circadian patterns are similar to those presented in Figure 4 of Wang et al. (2003), but the range of f appears to be attenuated relative to their figures. A comparison of AIC and BIC between the harmonic models and the semiparametric models proposed by Wang et al. (2003) favors the parametric models for the normal, depressed, and combined groups. For example, the AIC and BIC for the semiparametric model fit to the normal and depression data combined were 417.7 and 477.1, respectively. This can be compared with the parametric model with two harmonics, which resulted in an AIC and BIC of 408.3 and 450.0, respectively.

We tested whether the shape-invariant assumption between the normal and depressed groups is reasonable. We chose two harmonics for testing for an interaction, since this was the number of harmonics which minimized both the AIC and BIC for the normal, depressed, and combined groups. The likelihood-ratio test comparing $f(t)$ given by (4) and (5) (with $K = 2$, the likelihood-ratio test statistic has a chi-square distribution with 2 degrees of freedom) was significant ($P = 0.043$), suggesting that there is some evidence for a different shape between groups. Although statistically significant, the difference in the form of the pattern appears small, so we proceeded to fit the shape-invariant model.

We fit model (1) with $\boldsymbol{A}_i$ given by (3), $\boldsymbol{B}_i$ being a 3×3 identity matrix, and f given by (4) with $K = 1$, 2, and 3 harmonics.

The model can be written similarly to equation (11) in Wang et al. (2003),

$$y_{ijk} = \tau_k + b_{1i} + \exp\left(b_{2i} + d_1 \times I_{[g=2]}\right) \times f\left\{\left(t_{ij} - \text{alogit}\left(b_{3i} + d_2 \times I_{[g=2]}\right)\right)\right\} + \epsilon_{ijk}, \quad (6)$$

where $g = 1$ and 2 for the depressed and normal groups, respectively. In addition, $(b_{1i},\ b_{2i},\ b_{3i}) \sim N(\mathbf{0},\ \boldsymbol{D})$, where $\tau_2 - \tau_1$, d_1 and d_2 measure the difference of 24-hour mean, amplitude, and phase between the normal and depressed groups, respectively.

Similar to models for individual and combined groups, the function f was best fit with two harmonics; the AIC for models with one to three harmonics was 483.1, 412.4, and 416.1. In addition, f with $K = 2$ minimized the BIC. Therefore, we made inference using (4) with $K = 2$. In contrast to the findings of Wang et al. (2003) using the semiparametric approach, we did not find significant differences between the normal and depressed groups using the harmonic model. The estimate of $\tau_2 - \tau_1$ was -0.210 (SE = 0.158, $t_{212} = -1.33$, $P = 0.18$), the estimate of d_1 was 0.116 (SE = 0.142, $t_{212} = 0.82$, $P = 0.42$), and the estimate of d_2 was 0.060 (SE = 0.139, $t_{212} = 0.43$, $P = 0.67$). These estimates were obtained with $K = 2$; however, very similar inferences were obtained when fitting models with $K = 3$ or 4. Additionally, we fit a two-harmonic model which included an indicator of group as a random effect in addition to including group as a fixed effect. This model allowed for a different variance structure in the random effects across groups. Inferences on group differences with this more complex model were nearly identical to the previously stated results (data not shown).

The approximately 12% ($\exp(\hat{d}_1) = 1.12$) increase in amplitude for the normal group compared with the depressed group is consistent with apparent differences between the two circadian rhythms observed in Figure 2. Our results differed from those by Wang et al. (2003), who showed highly significant group differences. However, there were some puzzling inconsistencies in their data analysis that we could not

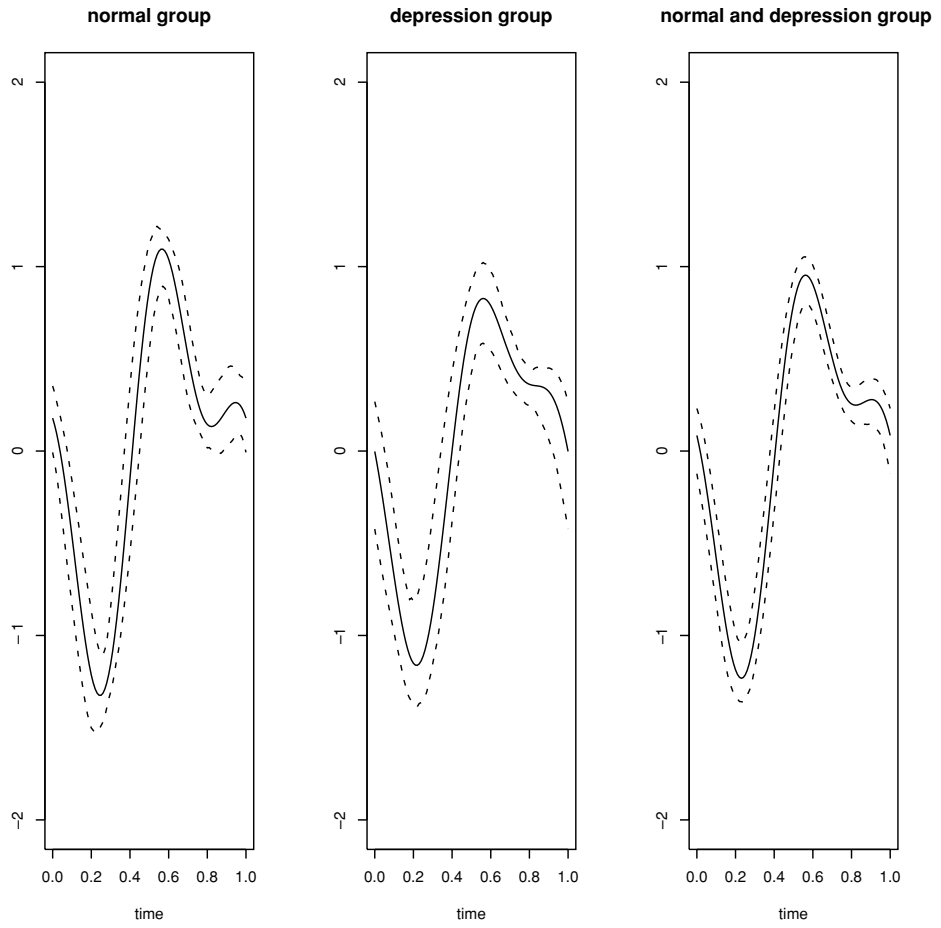

Figure 2. The solid lines are the estimates of f, and the dashed lines are the pointwise 95% confidence intervals for the normal ($n = 9$), depressed ($n = 11$), and combined ($n = 20$) groups. Confidence intervals are constructed using the bootstrap (520 samples) with the percentile method.

reconcile. Specifically, they estimated that the amplitude was 26% (exp(0.2350) = 1.26) larger in the normal group than in the depressed group, yet Figure 4 in Wang et al. (2003) appears to show that the depressed group has a larger amplitude than the normal group. Unfortunately, we were unable to reproduce Wang et al.'s results on group differences since, for this case, we had technical difficulties with ASSIST in the current version of either R or S-plus.

In Section 4, we evaluate the statistical properties of the harmonic models.

4. Simulation Results

In this section, we evaluate the statistical properties of our procedure for testing for interaction between groups (test of whether the scale-invariant model is appropriate) and for estimating group differences using the scale-invariant harmonic model. The simulations are based on the data structure (number of follow-up times and number of patients in each group) of the cortisol data set comparing the normal ($n = 9$) and depressed ($n = 11$) groups.

In the data analysis, we tested for an interaction between groups using two harmonics ($K = 2$), since this was the number of harmonics that minimized AIC and BIC for both groups separately as well as the combined group. To evaluate the statistical properties of the interaction test, we simulated circadian rhythm data with f identical to Figure 1A and the variance matrix for the random effects $\boldsymbol{D}$ being a diagonal matrix with the diagonal elements given by 0.04, 0.01, and 0. In addition, we simulated data with $d_1 = d_2 = (\tau_1 - \tau_2) = 0$ and $\sigma^2 = 0.16$. We estimated the percentage of times that the likelihood-ratio test (chi-square with 2 degrees of freedom) resulted in a P-value less than or equal to 0.05. With 2000 simulated realizations, the empirical type I error rate was 0.059 (SE = 0.005) for a 0.05 level test. Thus, the interaction test appears to have a nearly correct type I error rate in small samples like the cortisol data set.

We examined the statistical properties of the parametric scale-invariant model, focusing on estimating d_1, with 2000 simulations. We evaluated the properties for two procedures for estimating d_1. First, we assumed that we knew that two harmonics was the correct number of harmonics. The average d_1 was −0.0004, which was very close to the simulated value of zero. The average asymptotic standard error was 0.077, which was close to the Monte Carlo standard deviation of 0.080. In addition, the estimated type I error rate for the conditional t-test (estimated parameter over SE following a t-distribution

with 212 degrees of freedom) of whether $d_1 = 0$ was 0.063 for a 0.05 level test. This slight inflation of the type I error rate was reduced to 0.048 when we doubled the sample size from 20 to 40 patients (18 and 22 normal and control patients, respectively). Thus, for a known fixed number of harmonics, estimation and testing on group differences have good properties. In the second approach, we estimated the number of harmonics by fitting a parametric scale-invariant model with either one, two, or three harmonics. We then chose the model with the smallest AIC and used this model for inference. As with the fixed known number of harmonics, the average d_1 was nearly zero (-0.0004) and the average model-based standard error and MC standard deviation were almost identical (0.077 and 0.080, respectively). In addition, the estimated type I error rate for the conditional t-test of whether $d_1 = 0$ was 0.064 for a 0.05 level test. This slight inflation of the type I error rate was reduced to 0.049 when we doubled the sample size. Thus, estimating the number of harmonics on combined data does not appear to influence the quality of our inferences with the parametric scale-invariant model.

5. Discussion

Others have proposed harmonic models for analyzing circadian rhythm time-series data (Greenhouse, Kass, and Tsay, 1987, among others). This note described how to use nonlinear mixed models with harmonic terms to compare the circadian rhythms between groups of longitudinal data. We show that contrary to statements by Wang et al. (2003) that harmonic models are inflexible and do not provide a rich enough framework for analyzing longitudinal circadian rhythm data, a nonlinear mixed model with harmonic terms provides a flexible framework for analysis. These nonlinear mixed models are simple to fit in statistical packages such as R and S-plus. Although the semiparametric method of Wang et al. (2003) is very attractive, we felt it important to emphasize that simpler methods could be used for this problem.

Wang et al. (2003) mentioned that choosing the number of harmonics is difficult and the models are difficult to interpret. We showed how choosing the number of harmonics can be done using penalized-likelihood techniques such as AIC and BIC. Further, appropriate inference on group differences can be made using standard nonlinear mixed modeling software with a penalized-likelihood approach. It is true that it is difficult to infer any physical meaning to the harmonic terms in any of our models. However, in fitting the scale-invariant harmonic models (1) and (6), the harmonics are used only as a method for flexibly modeling the underlying circadian pattern f. We do not interpret the harmonic terms. In this way, our harmonic approach is similar to Wang et al.'s semiparametric approach where the underlying circadian pattern f is modeled with cubic splines which themselves are difficult to interpret. Furthermore, we wish to emphasize the relationship between using penalized-likelihood approaches for choosing the number of harmonics in parametric models and for choosing the smoothing parameter in semiparametric models using either cross-validation or penalized likelihood. Both approaches attempt to minimize the inherent trade-off between bias and variance in estimating f.

Both the semiparametric and parametric models use nonlinear mixed modeling procedures which rely on approximate inference (Pinhero and Bates, 2000). Although limited, our simulations show that these approximations work well in our setting. Unfortunately, no simulations were provided for the semiparametric approach.

The data analysis and simulations suggested that harmonic modeling of f was a viable alternative to semiparametric modeling. However, as in many nonlinear regression problems, the algorithm may be sensitive to starting values. We suggest choosing different starting values to assure that a global maximum of the likelihood is achieved. Occasionally, the `nlme` routine in R did not converge. This usually occurred when fitting higher order harmonic models in which an estimated β_k coefficient was close to zero and there was little information for estimating the corresponding scale parameters θ_k. The likelihoods from iteration to iteration were nearly identical, yet the parameter estimates did not meet the rigid convergence criterion set as a default in `nlme`. This problem has been discussed in a nonlinear regression context by Gallant (1977). We recommend choosing less rigid convergence requirements or setting a maximum large number of iterations to solve this problem. There was little discussion by Wang et al. (2003) about the importance of choosing starting values and the potential for convergence problems. We suspect that the semiparametric models may also be sensitive to starting values and occasionally have convergence problems.

Our analysis of the cortisol data led to different inferences than those reported for the semiparametric approach by Wang et al. (2003). In contrast to the semiparametric approach, we found no significant differences in the circadian rhythms between normal and depressed groups using the parametric shape-invariant model. For a number of reasons, we favor conclusions with the parametric approach over those using the semiparametric approach. First, the reported parameter estimates for the shape-invariant model in Wang et al. (2003) were inconsistent with the estimated circadian rhythms presented in their Figure 4. The model suggested that the normal group has a larger amplitude than the depressed group, yet the figure indicated the reverse. This suggests that there may be a problem in Wang et al.'s analysis of the depression data or in the ASSIST code. An investigation of these inconsistencies in the data analysis by Wang et al. (2003) would be very useful. Second, as reported in Section 3, the parametric models had smaller AIC and BIC than the semiparametric models. Although one needs to cautiously interpret these results since there may be a problem with the ASSIST software, these findings suggest that the parametric models may describe the data better than the semiparametric models. Third, our simulations showed that the harmonic models have good statistical properties for small data sets such as the cortisol example. No simulations were reported by Wang et al. (2003) for their semiparametric approach.

We compliment Wang et al. (2003) for providing a comprehensive package (ASSIST) for fitting their proposed semiparametric models (Wang and Ke, 2002). However, we found numerous technical difficulties in reproducing the results in their paper using the implementation on the current versions of S-plus and R. For example, we were unable to reproduce the fit of their scale-invariant model in R because pdStrat is unavailable in R. An advantage of the parametric approach

is that it is easy to implement in most statistical software packages.

Acknowledgements

We thank Yuedong Wang and Chunlei Ke for helping us with their ASSIST package. We also thank the associate editor for helpful comments which substantially improved this article. We thank Sara Joslyn for editing the final version of the article.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Second International Symposium on Information Theory*, B. N. Petrov and F. Csaki (eds), 267–281. Budapest: Akademia Kaido.

Gallant, A. R. (1977). Testing a nonlinear regression specification: A nonregular case. *Journal of the American Statistical Association* **72,** 523–529.

Greenhouse, J. B., Kass, R. E., and Tsay, R. S. (1987). Fitting nonlinear models with ARMA errors to biological rhythm data. *Statistics in Medicine* **6,** 167–183.

Lindstrom, M. J. and Bates, D. M. (1990). Nonlinear mixed effects models for repeated measures data. *Biometrics* **46,** 673–687.

Pinheiro, J. and Bates, D. M. (2000). *Mixed-Effects Models in S and S-Plus*. New York: Springer.

Schwartz, G. (1978). Estimating the dimension of a model. *Annals of Statistics* **6,** 461–464.

S-Plus 6 for Windows: Guide to Statistics, Volume 1. (2001). Seattle: Insightful Corporation.

Venables, W. N., Smith, D. M., and the R Development Core Team. (2004). *Notes on R: A Programming Environment for Data Analysis and Graphics*, Version 1.7.1 Available at `http://cran.r-project.org/doc/manuals/R-intro.pdf`.

Wang, Y. (1998). Smoothing spline models with correlated random errors. *Journal of the American Statistical Association* **93,** 341–348.

Wang, Y. and Brown, M. B. (1996). A flexible model for human circadian rhythms. *Biometrics* **52,** 588–596.

Wang, Y. and Ke, C. (2002). ASSIST: A suite of S-plus functions implementing spline smoothing techniques. Available at `http://www.pstat.ucsb.edu/faculty/yuedong/research`.

Wang, Y., Ke, C., and Brown, M. B. (2003). Shape-invariant modeling of circadian rhythms with random effects and smoothing spline ANOVA decompositions. *Biometrics* **59,** 804–812.

Received May 2004. *Revised December* 2004. *Accepted December* 2004.

Appendix

R Code to Fit Nonlinear Mixed Model with Harmonic Terms

The following code was used to fit the parametric scale-invariant model with group effects to the cortisol data set where $f(t)$ is characterized with two harmonic terms

$$\begin{aligned}\text{Out} < -\text{nlme}(&\text{conc} \sim (A + \exp(B) * \\ &(C * \cos(2 * pi * (time + D + \exp(E)/(1 + \exp(E)))) \\ &+ F * \cos(4 * pi * (time + G + \exp(E)/(1 + \exp(E)))))), \\ &\text{data} = \text{dataset}, \text{fixed} = \text{list}(A \sim Group, B + E \sim -1 \\ &+ Group, C + D + F + G \sim 1), \\ &\text{random} = A + B + E \sim Group|ID, \\ &\text{start} = \text{c}(1.8, .2, .1, .2, .5, .1, .2, .1), \\ &\text{control} = \text{list}(\text{maxIter} = 100, \text{returnObject} = \text{T})).\end{aligned} \tag{A.1}$$

The authors replied as follows:

We agree with Drs Albert and Hunsberger that harmonic models can be a useful tool in modeling circadian rhythm data. The following discussion provides more comparisons between the parametric and nonparametric models.

Parametric versus nonparametric models. It is reasonable to assume that the common shape function f in model (1) of Albert and Hunsberger (2005) is periodic and smooth. Naturally, f can be modeled parametrically using harmonics (Albert and Hunsberger, 2005)

$$f(t) = \sum_{k=1}^{K} \beta_k \cos(2\pi k t + \theta_k). \tag{1}$$

The order K is unknown in practice. Thus one needs to select K using a model selection procedure. An alternative nonparametric approach is to assume that f belongs to the following infinite dimensional periodic spline model space (Wahba, 1990; Gu, 2002; Wang, Ke, and Brown, 2003)

$$\begin{aligned}W_2^0(per) = \Big\{ & f : f \text{ and } f' \text{ are absolutely continuous}, \\ & f(0) = f(1), f'(0) = f'(1), \\ & \int_0^1 f(t)\,dt = 0, \int_0^1 (f''(t))^2\,dt < \infty \Big\}.\end{aligned} \tag{2}$$

For equally spaced designs, the periodic spline estimate is essentially a low-pass filter: components at frequency k are down-weighted by a factor of $1 + \lambda(2\pi k)^4$, where λ is a smoothing parameter. Consequently, selecting an order K for the harmonic model (1) may be viewed as hard thresholding and selecting the smoothing parameter for the periodic spline may be viewed as soft thresholding. See Wahba (1990) and Wang (2004) for details.

Selection of the tuning parameters K and λ are the key to success of the parametric and nonparametric models. Our original statement, "Although adding harmonics can improve the fit, it is difficult to decide how many harmonics to include in the model," was prompted by the lack of research for selecting K. We commend Albert and Hunsberger for the development of penalized likelihood model selection methods. We agree with the authors that a parametric model should be used when it fits data appropriately. In fact, there are several other parametric models proposed in the literature for modeling different forms of circadian rhythms (Batschelet, 1981; Ruf, 1996). Rather than developing methods for a

specific form of circadian rhythm, our nonparametric procedure in Wang et al. (2003) can adopt various forms of circadian rhythms. The philosophy of the nonparametric procedure is to let the data speak for themselves. The family of harmonic parametric model (1) with $K = 1$, 2, and 3 as suggested in Albert and Hunsberger (2005) may still be restrictive for certain forms of circadian rhythms. Even if one decides to use model (1), the nonparametric methods can be used as diagnostic tools to check the appropriateness of these parametric models. Specifically, using different reproducing kernels as in Gu (2002, p. 122), one can test the hypothesis that f has the form (1) for any K using methods in Wang et al. (2003). This can be easily implemented using our R function `snm`. The nonparametric methods can also be used to suggest possible families of parametric models when it is unclear which one to use.

The parametric model is simple. However, when $K \geq 2$, it is not necessarily easier to fit a parametric model using the `nlme` function because more terms will have to be included in the formula, and more initial values will have to be determined and may lead to numerical instability. As noted by the authors the estimate of the phase parameter θ_k is unstable when the coefficient β_k is close to zero. Strictly speaking, θ_k is unidentifiable when $\beta_k = 0$. The chance of having at least one very small coefficient increases as K increases. Combining all harmonics together as a nonparametric function f, our function `snm` is just as easy to implement as the `nlme` function (see the R code in the Appendix). The common shape function f is often treated as a nuisance parameter. We are interested in the whole function rather than coefficients β_k and θ_k. Important parameters, 24-hour mean, amplitude, and phase, are the same for both parametric and nonparametric models with the same interpretations.

To summarize, both parametric and nonparametric models have their advantages and limitations. They complement each other.

ASSIST package and data analysis. Conclusions in Wang et al. (2003) were based on the following model

$$\begin{aligned} y_{ijk} = {} & \beta_k + b_{1i(k)} + \exp\left(b_{2i(k)} + d_1 \times I_{[k=2]}\right) \\ & \times f\left(t_{ijk} - \text{alogit}\left(b_{3i(k)} + d_2 \times I_{[k=2]}\right)\right) + \epsilon_{ijk}, \\ & i = 1, \ldots, m, \quad j = 1, \ldots, n_{ik}, \quad k = 1, 2, \end{aligned} \tag{3}$$

where k represents the group with $k = 1$ and $k = 2$ corresponding to depression and normal groups, respectively, $\beta_2 - \beta_1$, d_1, and d_2 measure the differences of 24-hour mean, amplitude, and phase between the normal group and the depression group, and $f \in W_2^0(per)$. Random effects $b_{1i(k)}$, $b_{2i(k)}$, and $b_{3i(k)}$ represent the ith subject's deviation of 24-hour mean, amplitude, and phase. Note that subjects are nested within group, which is reflected in our notation. Since covariance structures for normal group and depressed group were different, we assumed that $\boldsymbol{b}_{i(k)} = (b_{1i(k)}, b_{2i(k)}, b_{3i(k)})^T \overset{\text{iid}}{\sim} N(0, \sigma^2 \boldsymbol{D}_k)$, where $\boldsymbol{D}_k$'s are unstructured positive-definite matrices. Model (3) requires 12 parameters for the covariance matrices $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$. Model (7) in Albert and Hunsberger (2005) has a different covariance structure for random effects: it assumes six random effects for some subjects with an unstructured covariance matrix. Therefore, model (7) in Albert and Hunsberger (2005) requires 21 parameters for the covariance matrix of random effects.

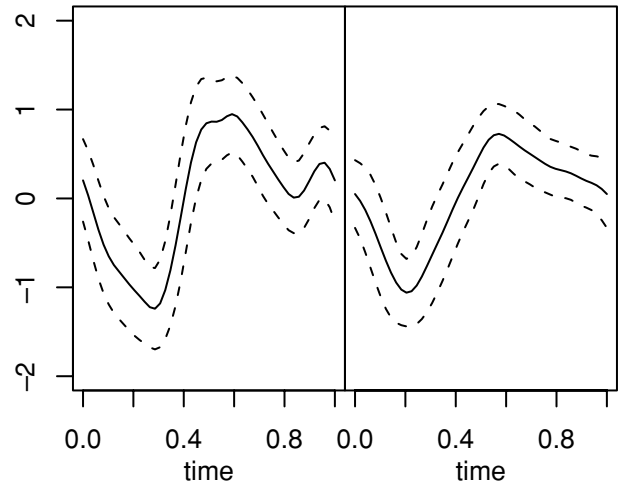

Figure 1. Solid lines are estimates of the common functions and dotted lines are 95% Bayesian confidence intervals. The left and right panels are estimated common curves for the normal and depression groups, respectively.

Nonefficient parameterization in model (7) of Albert and Hunsberger (2005) was due to a limitation of the current version of `nlme`. We first developed the ASSIST package under S-Plus 3.4 and subsequently switched to R (Wang and Ke, 2002). We fitted model (3) using the function `snm` in the ASSIST package under S-Plus 3.4. To allow different correlation structure for each strata (group), we used the `pdStrat` class in the `nlme` package. Unfortunately, this class of covariance structures is currently not available in the new R version of `nlme`. Hopefully, the `pdStrat` class will be available in the future (communication with Professor Bates).

For model (3), estimates of $\beta_2 - \beta_1$ and d_1 are -0.2724 (SE = 0.1311, p-value = 0.0389) and 0.2350 (SE = 0.0767, p-value = 0.0024). For comparison, we also fit a reduced model of (3) with the same covariance structure for both groups (i.e., $\boldsymbol{D}_1 = \boldsymbol{D}_2$). The R code for fitting this reduced model is shown in the Appendix. Estimates of $\beta_2 - \beta_1$ and d_1 based on the reduced model are -0.2593 (SE = 0.1440, p-value = 0.0733) and 0.2153 (SE = 0.1122, p-value = 0.0562), respectively. Comparing to the estimates in Albert and Hunsberger (2005), which were based on model (1) with $K = 2$ and the same covariance structure for both groups, it is clear that the resulting different conclusions are caused by the combination of a nonparametric model for f and covariance structure. More flexibility for the nonparametric f and covariance structure makes the absolute value of estimates a little bit larger and the SEs smaller. Note that conclusions based on our reduced model agree with those in Albert and Hunsberger (2005).

Which model to choose is, however, a more difficult question. We used the REML method (the default of `snm`) to fit model (3). The REML score is not comparable with the likelihood used in Albert and Hunsberger (2005). We note that the likelihood (or REML score) in both `nlme` and `snm` are approximated values. It is unclear how good these approximations

are. In addition, the degrees of freedom for nonparametric models may not be directly comparable with those for parametric models. Therefore, direct comparisons between parametric and nonparametric models using AIC and BIC need to be taken with a grain of salt. We also note that the sample size is relatively small, which makes it difficult to reach a definite conclusion.

We thank the authors for pointing out an apparent inconsistency between Figure 4 and conclusions in Wang et al. (2003). Plots in Figure 4 are indeed wrong due to a bug in the `intervals` function. The bug is now fixed and the corrected plots are shown in Figure 1. The conclusions based on model (3) remain the same.

References

Albert, P. S. and Hunsberger, S. (2005). On analyzing circadian rhythms data using nonlinear mixed models with harmonic terms. *Biometrics* **61,** 1115–1120.

Batschelet, E. (1981). *Circular Statistics in Biology.* London: Academic Press.

Gu, C. (2002). *Smoothing Spline ANOVA Models.* New York: Springer-Verlag.

Ruf, T. (1996). The baseline cosinus function: A periodic regression model for biological rhythms. *Biological Rhythms Research* **27,** 153–165.

Wahba, G. (1990). *Spline Models for Observational Data.* CBMS-NSF Regional Conference Series in Applied Mathematics, Volume 59. Philadelphia: SIAM.

Wang, Y. (2004). Model selection. In *Handbook of Computational Statistics*, Volume I, J. Gentle, W. Härdle, and Y. Mori (eds), 437–466. Berlin: Springer.

Wang, Y. and Ke, C. (2002). ASSIST: A suite of s-plus functions implementing spline smoothing techniques. Manual for the ASSIST package. Available at `http://www.pstat.ucsb.edu/faculty/yuedong/software`.

Wang, Y., Ke, C., and Brown, M. B. (2003). Shape-invariant modeling of circadian rhythms with random effects and smoothing spline ANOVA decompositions. *Biometrics* **59,** 804–812.

Appendix

R Code

The following code was used to fit model (3) with the same covariance structure for both groups.

```
snmfit <- snm(conc~b1+exp(b2)*f(time-alogit(b3)),
              func=f(u)~list(periodic(u)), data=cort.nordep,
              fixed=list(b1~Group, b2+b3~-1+Group),
              random=b1+b2+b3~1, groups=~ID, method="ML",
              spar="m", start=c(1.9, -0.3, 0, 0),
              control=list(prec.out=0.005, converg="PRSS"))
```


Yuedong Wang
Department of Statistics and Applied Probability
University of California
Santa Barbara, California 93106, U.S.A.
email: *yuedong@pstat.ucsb.edu*

Chunlei Ke
St. Jude Medical
Cardiac Rhythm Management Division
Sylmar, California 91342, U.S.A.
email: *cke@sjm.com*

and

Morton B. Brown
Department of Biostatistics
University of Michigan
Ann Arbor, Michigan 48109, U.S.A.
email: *mbb@umich.edu*