

The paper provides a very valuable alternative modeling for the phase in cycles modeling that readers will find appealing. It also gives a number of other valuable insights on this topic. The paper is thoughtful and well written.

I disagree with the paragraph on page 15 rows 267-278. There is some sort of a notion here that you can look at individual data to evaluate if cycles exist and then if the majority don't show cycles take this as evidence that cycles do not exist. No. This is contrary to the principles of multilevel modeling and using the population as a whole to make inferences by accumulating information and borrowing information. What if you have only 3 observations per person only? Does that mean that you can't make inference on cycles (and anything else really) because in the majority nothing is significant? The way this is phrased it sounds like you are making a new statistical principle here but this totally contradicts the foundation of multilevel modeling. While I may agree that in your two examples one has stronger evidence of cycles than the other, I doubt your inference is correct as phrased.

Let me try to rephrase this. If you have a multilevel regression of two variables with a random coefficient, we do not look at individual posterior distribution and try to make inference for the population. Instead we estimate the mean and variance for the random effect and then decide if these are significant - there are only two parameters in the final verdict and they have a clear interpretation - if the mean of the random effect is significant we claim that the covariate has predictive power, and if the variance of the random effect is significant we claim that the predictive effect likely vary across individuals. The same principles apply to cycles as well. The existence of cycles is determined by first the significance of parameter A in formula (14). If A is significant, cycles exist. If A is not significant but the variance of (a_i) is significant then cycles still exist. When we say that cycles exist for the population and we model these with random effects - that doesn't imply that everyone has cycles. Because the coefficients are random many people may indeed have no cycles or the amplitude may be too small for any practical purpose. Nevertheless, on average, cycles would still be an essential part of modeling the variable for the entire population. Statistical significance on a particular individual may still be out of reach for most individuals.

Ultimately, statistical significance is secondary. What we really care about is the cycle's R^2 . If it is less than .01 - I think we wouldn't bother. If it is more than .10 I will insist that modeling the data must include the cycles, if it is between 0.01 and .10 then I would want to know if the evidence is statistically significant. R^2 summarizes the effect across the population and it has not much to do with statistical significance of individual random effects which you are using to make a claim on cycles.

Inevitably ... in practical applications ... the idea of using SIN and COS to model within day consistent patterns must always be on the fringes as there are countless alternatives to model within day-cyclical patterns including "linear trend within day", "dummy variables for particular parts of the day", "quadratic trend within day", "log of time within day". If you have hourly records for 16 hours in the day, then 15 dummy variables would give you the most detailed picture.

Using SIN and COS to reduce these 15 parameters to 2 is a great idea if the 15 parameter curve can be fitted with SIN and COS and you can get a more parsimonious model and smaller SE. The reason I say all of these is because you should put this in perspective - cyclical day patterns can be modeled in many different ways and some trigonometric difficulties are nice to understand better but are not the citadel. Using SIN and COS is a tool in the toolbox. But even if you use SIN and COS, you can model daily cycles / patterns without ever having to do inference on the peak and amplitude. You do not have to. This is optional. Psychological data is messy. Most of the time the data is too variable to be able to estimate individually specific peaks well and you might have to just settle for a predictive power of the time of the day. Obviously if you have many observations for each individual it may become feasible to estimate the peak more precisely. Nevertheless, throughout the paper, the process of modeling circular data with SIN and COS is not separated from the process of making a secondary model for the peaking and amplitude. Again this secondary model is optional. And only this optional part requires circular statistical inference.

Figure 2 is beautiful and very illuminating. But also very impractical. Drawing reliable heat maps would require an extraordinary amount of sample draws. The little black line that you are using to make the inference will be quite volatile and ultimately your conclusion will be based on rare events in the tail: estimate of the density at (0,0). The way I look at this - purposely ignoring the black line - all 4 scenarios have marginal proof of significance - with somewhat of varying degree but hardly a robust difference I would make a claim on.

Your method boils down to estimating the density at 0,0 which is more difficult than estimating $P(C1 \leq 0, S1 \leq 0)$ - using that for the D1-76 or $P(C1 \geq 0, S1 \leq 0)$ - using that for the D2-11
Is better and more reliable.

But why argue over this when simulation study can be made? And why not include the rest of the natural alternatives? Below is a list of methods that could serve as alternatives. One can compare type I and type II errors in the scenarios of cycles and no cycles with varying sample size. Here is what I would say are the competing methods

- A. Your method: density at (0,0) in HPD area
- B. Your method based in bivariate normal: density at (0,0) for bivariate normal fit of draws
- C. Smallest quadrant method I mention above
- D. Test the hypothesis $C=0$ or $S=0$ with the univariate credibility interval
- E. Joint hypothesis for $C=0$ and $S=0$ with wald test
- F. Z-score for $\sqrt{C^2+S^2} > 1.96$

My intuition is that the differences between these methods would be marginal - but I could be wrong. I understand that asking you to show us these results is a lot of work ... but acknowledging the need for this work would also be good enough for me as well.

In Section 2.3 “ ... as this will never contain zero “

maybe mention here that this is a test on the boundary of admissible space and similar to testing variance=0

Figure 7, This is pretty confusing ... I am not sure what the goal is but it looks bad. You seem to be combining “linear” and “circular” for the same computation. Why would you do that? Pick a method and do the random effect computation with that method and then compute the mean with the same method

Line 398 - I already talked about that earlier so I won't repeat it here, but give a p-value, a test, something that is a statistical method. You are using an unorthodox argument for something we usually use a p-value for. Include here for both data sets R^2 of the cycles with and without random amplitude and phase, include mean and variance for S_i and C_i together with significance. For the non-random model give S and C and confidence limits - this is where you should have started. Again, the existence of cycles are determined from these quantities not circular statistics or heat maps.

Line 443 - all of these choices are odd/artificial - without looking at the data you want to choose as the start of day when folks are asleep so 3 am.

There is another odd/artificial choice in your analysis. I would consider using cycles for data that looks quadratic in nature and a peak or a bottom happens in the middle of the day. Presumably this is the case in your data. Then model the thing that is in the middle of the day - not the thing that is at the end of the day where you don't have data or the data runs out and you have problems with circular effects. What I am trying to say You should have modeled the bottom and not the peak. This looks like artificially creating a problem that doesn't exist.

Line 512 - as far as I can see circular correl (Z, Φ) = circular correl ($-Z, \Phi$). Is that true? Maybe clarify - essentially this is an absolute value of a correlation in a way

Is the circular correlation a part of the estimation process? Or is it estimated after the model estimation? If it is not a part of the estimation maybe remove the first appendix.

There are a lot of limitations of the circular correlation (not being a part of the estimation is one). The circular correlation boils down to using $\sin(\Phi_i)$ and $\cos(\Phi_i)$ as predictors for the other random effects. But this points to the other limitations. You can use Φ to predict A but not the other way around (line 577). It will be hard for the random Φ to be on the LHS of an equation with multiple predictors, or on the RHS with other predictors. Maybe mention these limitations. Ultimately if we have enough data, the random effects posteriors are narrow and do

not have “broken up” distribution then not using circular methodology would have an advantage as you can use it as predictor in multiple regression as well as to be predicted. Maybe include such a discussion in the paper to give a more balanced presentation.

Line 1136- “Since LResultant can never reach one”

It doesn’t make sense, please explain. In fact I think it is wrong since $TC > N$ if ranks of ϕ and Z agree