Practice in Growth Curve Modeling

for Women's Menstural Cycle
- Fundamental Overview -

Chel Hee Lee[1] Angela Baerwald[2]

[1]Clinical Research Supporting Unit
College of Medicine
University of Saskatchewan

[2]Department of Obstetrics, Gynecology and Reproductive Sciences
College of Medicine
University of Saskatchewan

Saskatoon SAS Users Group Meeting
September 16, 2015

Introduction

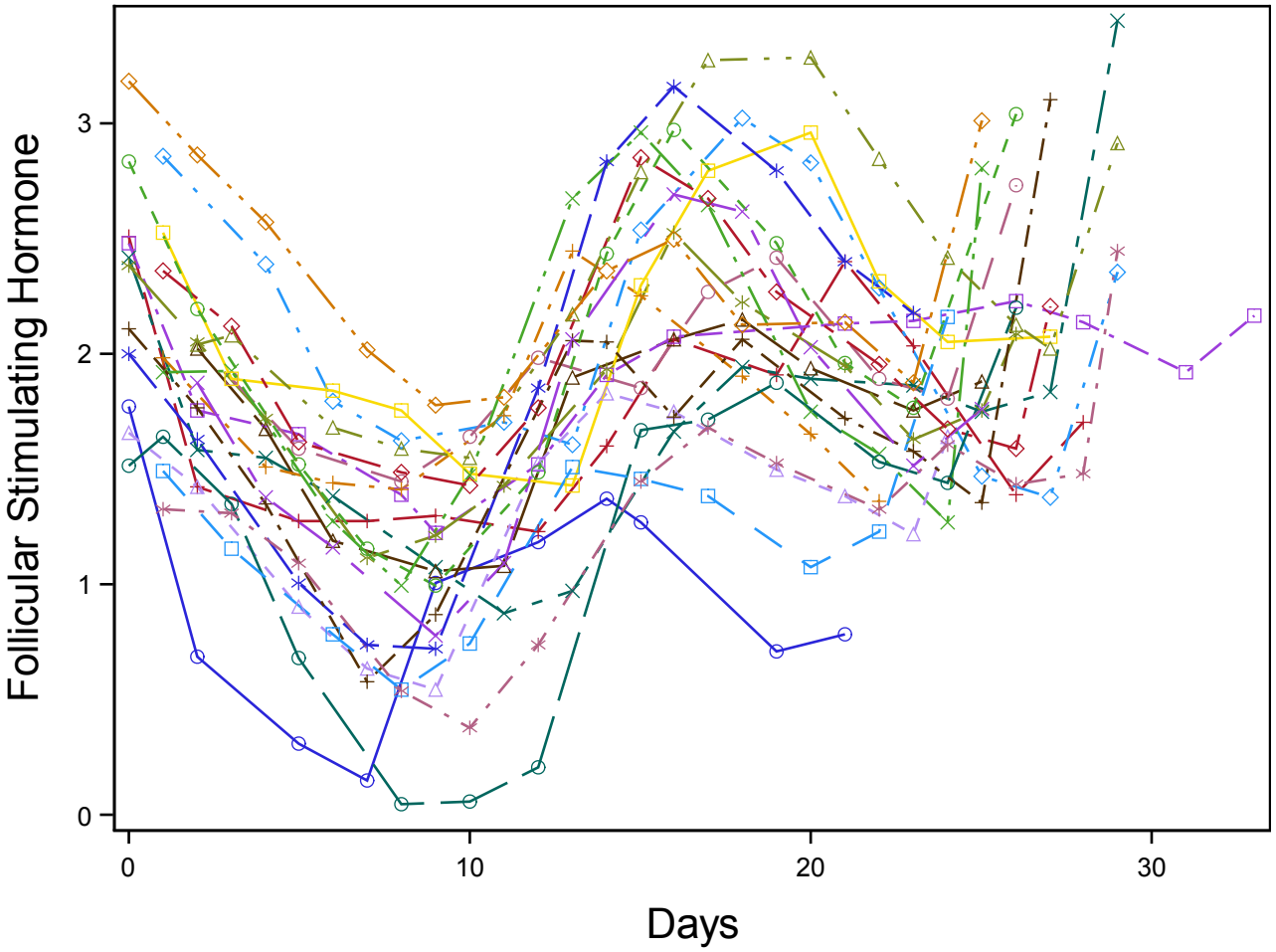
"Essentially, all models are wrong, but some are useful."

(Box and Draper, 1986, p. 424)

- It is a best guess.
- It is an approximation.
- It would be a translation of observation into a model.

Hormones Involved in Women's Menstrual Cycle

Subject-specific Trajectory (in Logarithm Scale)



Linear Mixed-Effect Models (Singer, 1998)

Level 1 (within person)

$$Y_{ij} = \pi_{0j} + \pi_{1j}(TIME)_{ij} + r_{ij}, \quad \text{where } r_{ij} \sim N(0, \sigma^2)$$

Level 2 (between-person)

$$\begin{aligned} \pi_{0j} &= \beta_{00} + u_{0j}, \quad \text{where } \begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix}\right] \\ \pi_{1j} &= \beta_{10} + u_{1j}, \end{aligned}$$

Unconditional Linear Growth Curve

$$Y_{ij} = \underbrace{[\beta_{00} + \beta_{10} TIME_{ij}]}_{\text{FIXED}} + \underbrace{[u_{0j} + u_{1j} TIME_{ij} + r_{ij}]}_{\text{RANDOM}}$$

Note that a 3-level model can be made if individuals within groups are tracked over time.

Using PROC MIXED in SAS

Commands for a linear time trend model

```
PROC MIXED DATA = datafile;
  CLASS id time;
  MODEL y = time / SOLUTION CHISQ;
  REPEATED time / TYPE=UN SUBJECT=id;
RUN;
```

Commands for a subject-specific model

```
PROC MIXED DATA = datafile;
  CLASS id;
  MODEL y = time / SOLUTION CHISQ;
  RANDOM intercept time / TYPE=UN SUBJECT=id;
RUN;
```

Concerns Bothering Us

We are working with ...

- A small size of participants with many repeated measurements,

Concerns Bothering Us

We are working with ...

- A small size of participants with many repeated measurements,
- An unbalanced data set observed at unequally spaced times, and

Concerns Bothering Us

We are working with ...

- A small size of participants with many repeated measurements,
- An unbalanced data set observed at unequally spaced times, and
- An incomplete data due to missing values.

Concerns Bothering Us

We are faced with difficulties of ...

- Determining the level of a model of interest,

Concerns Bothering Us

We are faced with difficulties of ...

- Determining the level of a model of interest,
- Incorporating a person-level covariate,

Concerns Bothering Us

We are faced with difficulties of ...

- Determining the level of a model of interest,
- Incorporating a person-level covariate,
- Addressing a difference between groups,

Concerns Bothering Us

We are faced with difficulties of ...

- Determining the level of a model of interest,
- Incorporating a person-level covariate,
- Addressing a difference between groups,
- Dealing with missing data and dropout,

Concerns Bothering Us

We are faced with difficulties of ...

- Determining the level of a model of interest,
- Incorporating a person-level covariate,
- Addressing a difference between groups,
- Dealing with missing data and dropout,
- Binning data or using data as it is,

Concerns Bothering Us

We are faced with difficulties of ...

- Determining the level of a model of interest,
- Incorporating a person-level covariate,
- Addressing a difference between groups,
- Dealing with missing data and dropout,
- Binning data or using data as it is,
- Centralizing, normalizing, orthogonalizing data, and

Concerns Bothering Us

We are faced with difficulties of ...

- Determining the level of a model of interest,
- Incorporating a person-level covariate,
- Addressing a difference between groups,
- Dealing with missing data and dropout,
- Binning data or using data as it is,
- Centralizing, normalizing, orthogonalizing data, and
- Applying either or both models,

Concerns Bothering Us

We are faced with difficulties of ...

- Determining the level of a model of interest,
- Incorporating a person-level covariate,
- Addressing a difference between groups,
- Dealing with missing data and dropout,
- Binning data or using data as it is,
- Centralizing, normalizing, orthogonalizing data, and
- Applying either or both models,
- Choosing an appropriate pattern of covariance structure, etc.

Parameteric Approaches to Modeling Curves

Nonlinear Model

$$y_i = u(x_i) + \epsilon_i, \quad i = 1, 2, \ldots, n$$

- p-th Degree Polynomial Regession Model (Johnson et al., 2013)

$$u(x_i) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \cdots + \beta_p x_i^p$$

Parameteric Approaches to Modeling Curves

Nonlinear Model

$$y_i = u(x_i) + \epsilon_i, \quad i = 1, 2, \ldots, n$$

- p-th Degree Polynomial Regession Model (Johnson et al., 2013)

$$u(x_i) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \cdots + \beta_p x_i^p$$

- Exponential Growth Curve Model (Zwietering M H et al., 1990)

$$u(x_i) = \beta_0 e^{\beta_1 x_i}$$

Parameteric Approaches to Modeling Curves

Nonlinear Model

$$y_i = u(x_i) + \epsilon_i, \quad i = 1, 2, \ldots, n$$

- p-th Degree Polynomial Regession Model (Johnson et al., 2013)

$$u(x_i) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \cdots + \beta_p x_i^p$$

- Exponential Growth Curve Model (Zwietering M H et al., 1990)

$$u(x_i) = \beta_0 e^{\beta_1 x_i}$$

- Trigonometric Model (Cornelissen Germaine, 2014)

$$u(x_i) = \mu + A\cos(wx_i + \phi)$$

Modeling Techniques

Piecewise Modeling Approach (Naumova et al., 2001)

$$u(x_i) = \beta_0 + \beta_1 x_i + \beta_2 (x_i - c)\delta,$$

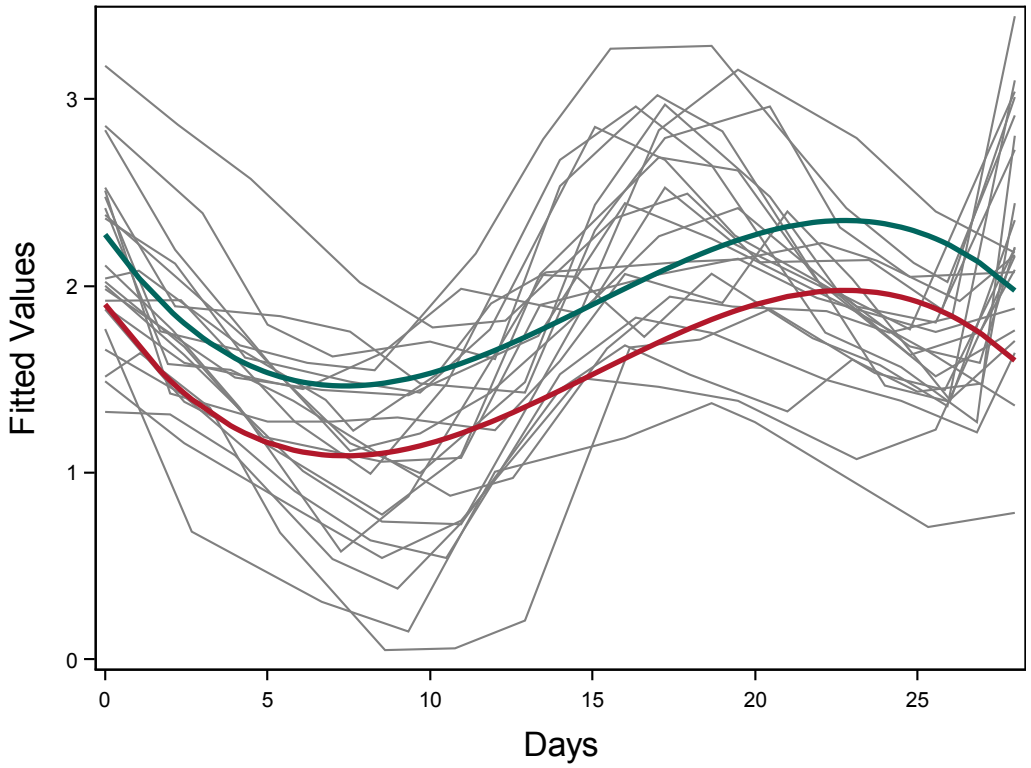
where $\delta = 1$ if $x_i > c$ and $\delta = 0$ if $x_i <= c$.

Cubic-Splines Modeling Approach (Gurrin et al., 2005)

$$u(x;\theta) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \sum_{i=1}^{m} \theta_i (x - c_i)_+^3$$

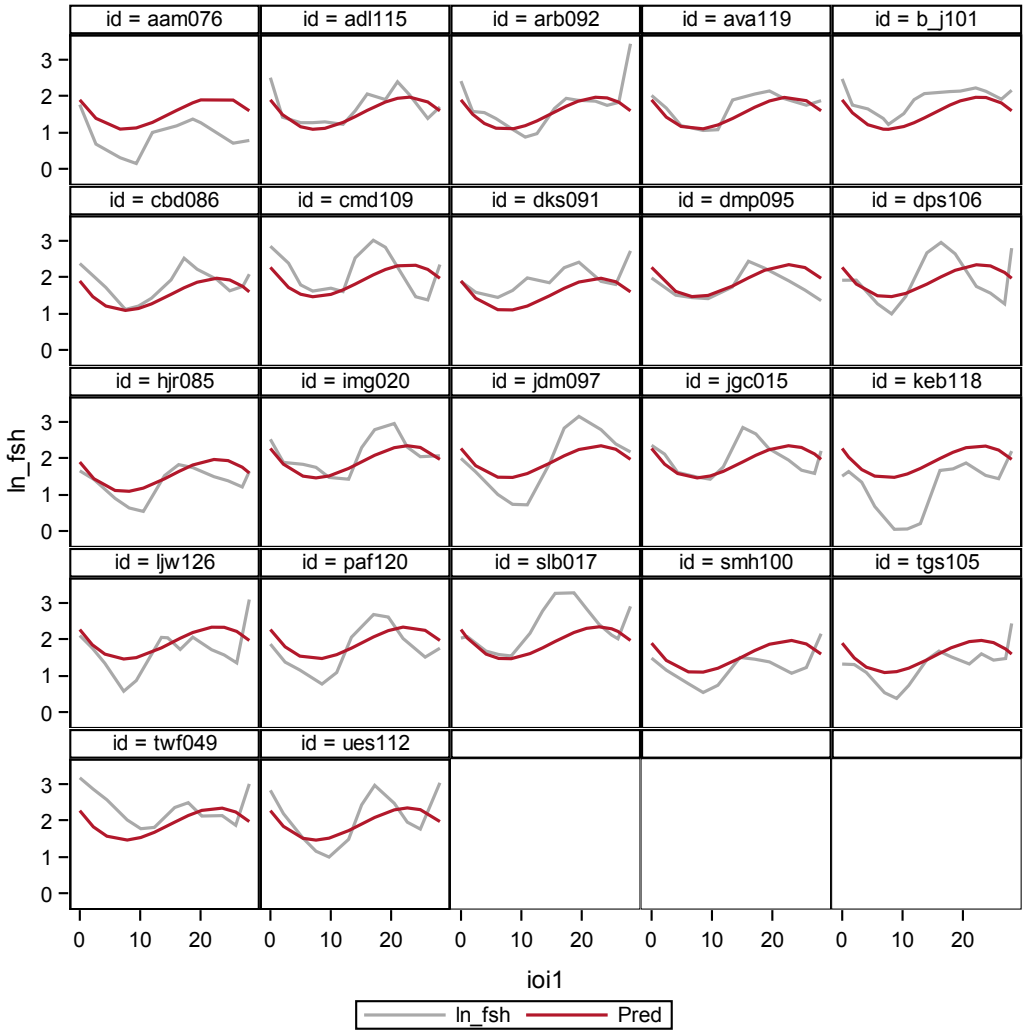
OH No! This is not what I wanted. (Problems?)

3rd degree polynomial regression



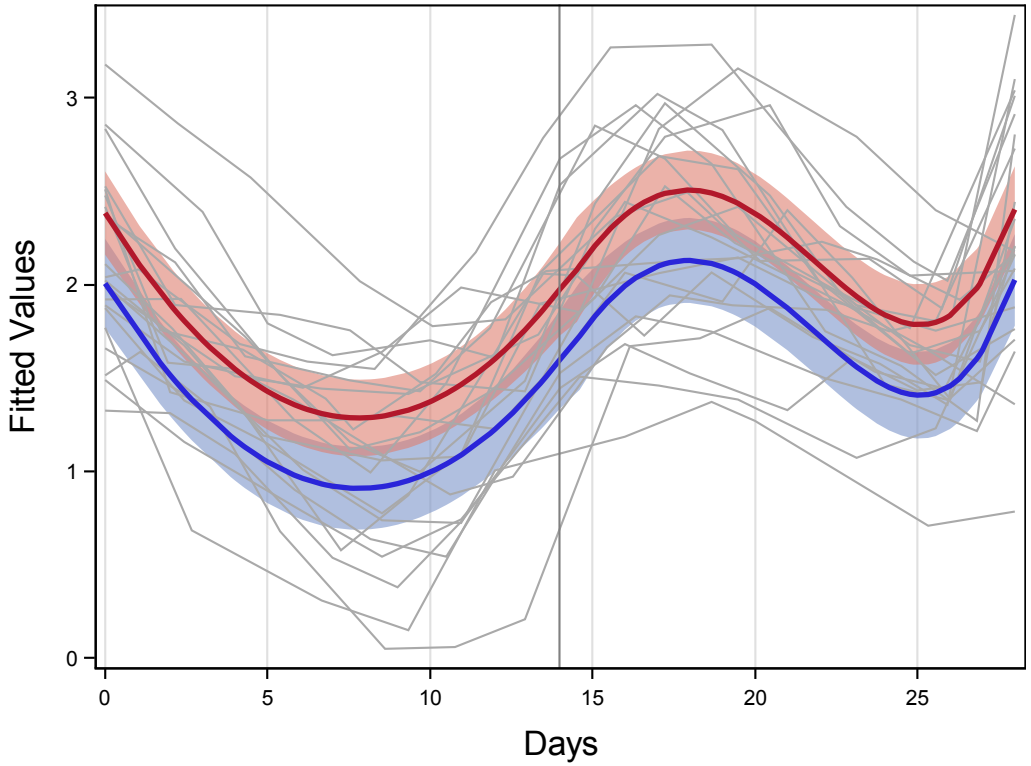
OH No! This is not what I wanted. (Problems?)

3rd degree polynomial regression



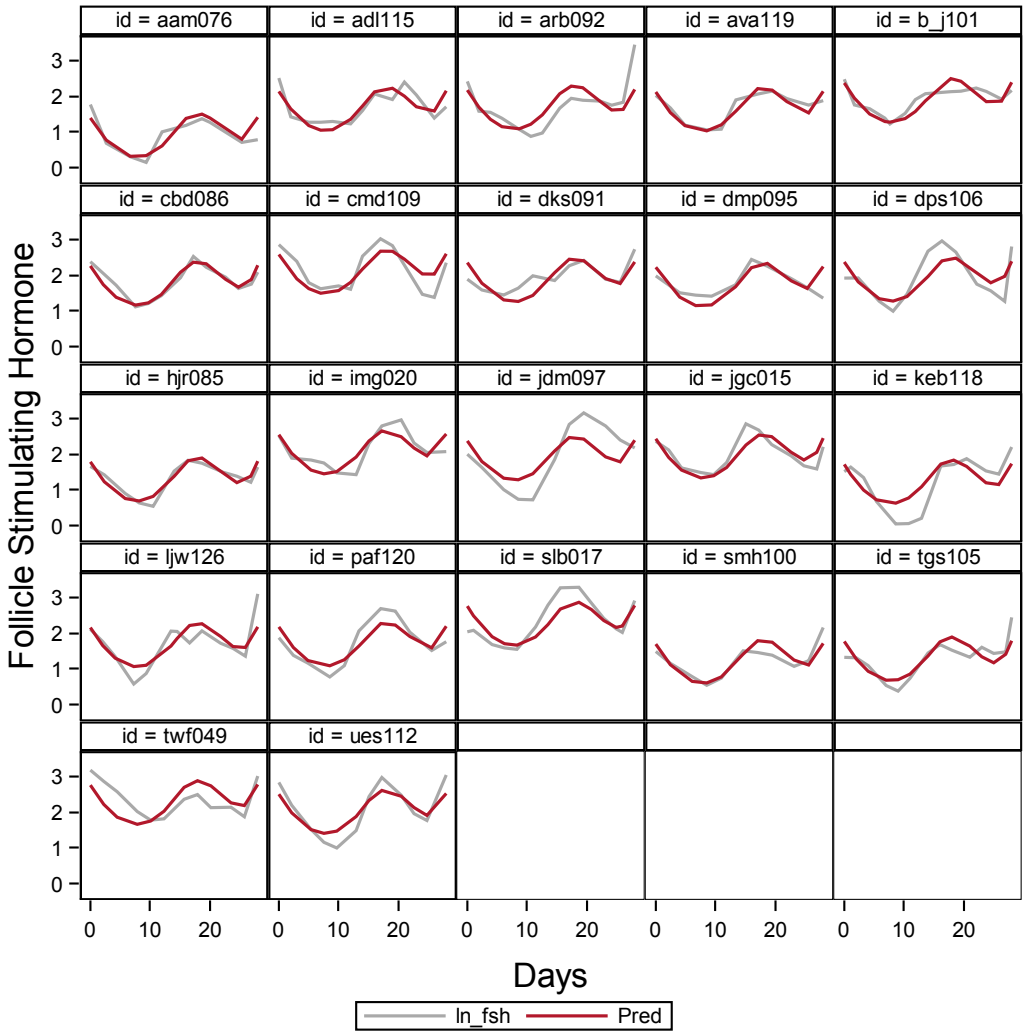
Oh! It's better! However, ... Disadvantages???

Piecewise-4th degree polynomial regression



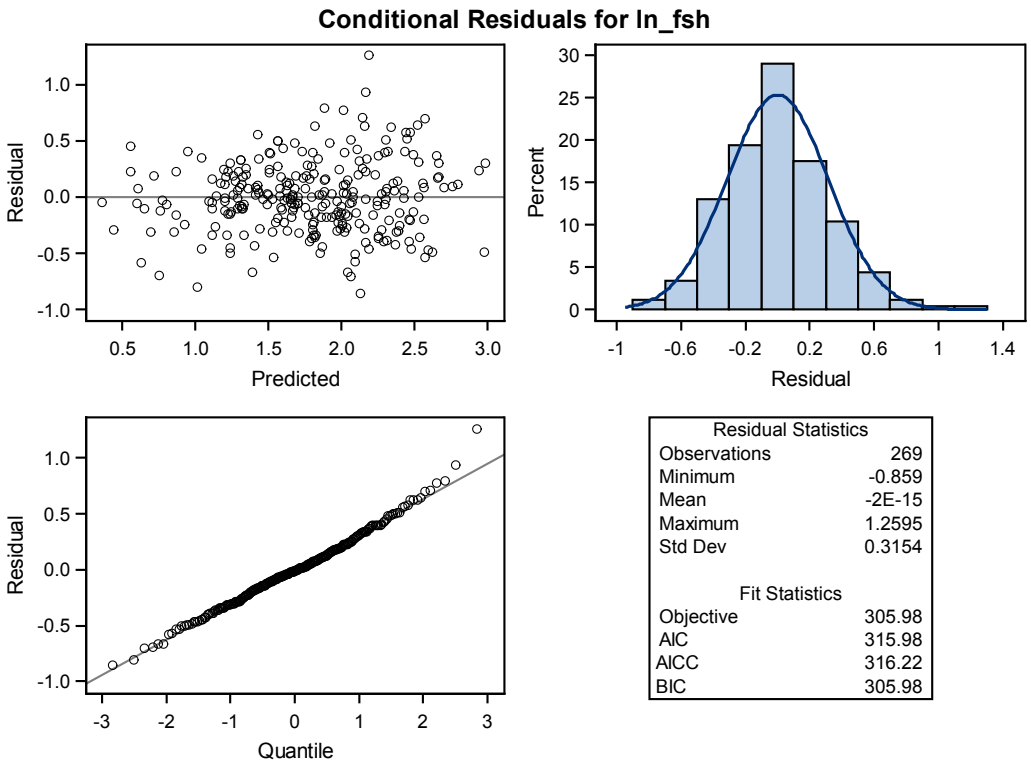
Oh! It's better! However, ... Disadvantages???

Piecewise-4th degree polynomial regression



Oh! It's better! However, ... Disadvantages???

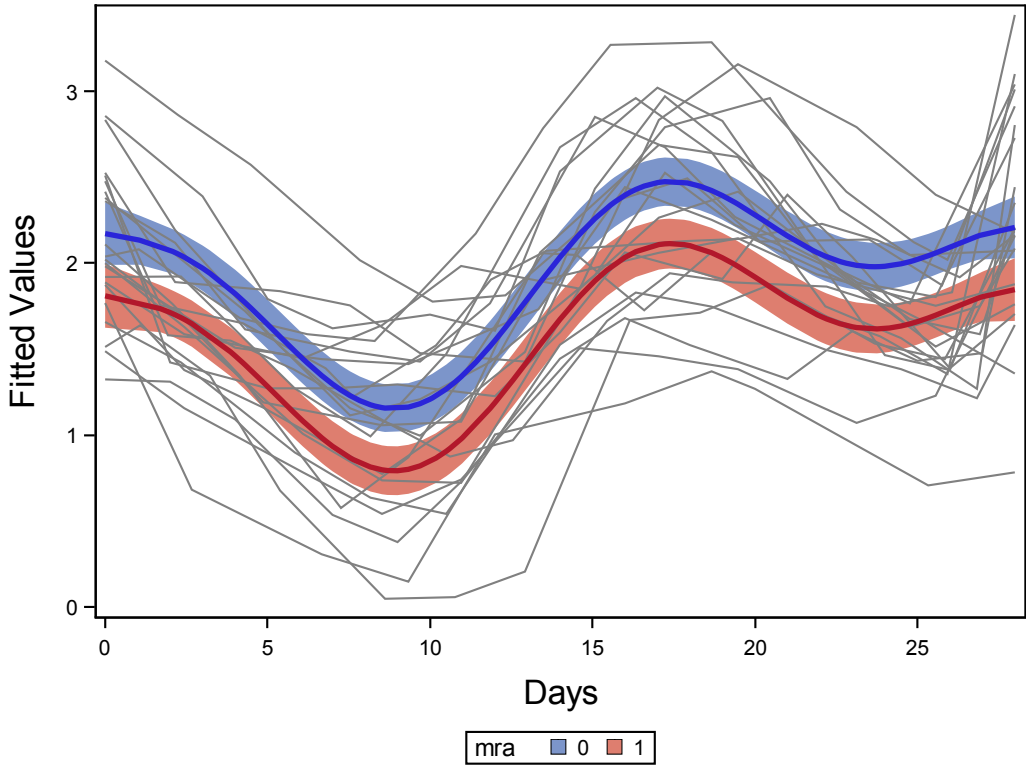
Piecewise-4th degree polynomial regression



Residual Statistics	
Observations	269
Minimum	-0.859
Mean	-2E-15
Maximum	1.2595
Std Dev	0.3154
Fit Statistics	
Objective	305.98
AIC	315.98
AICC	316.22
BIC	305.98

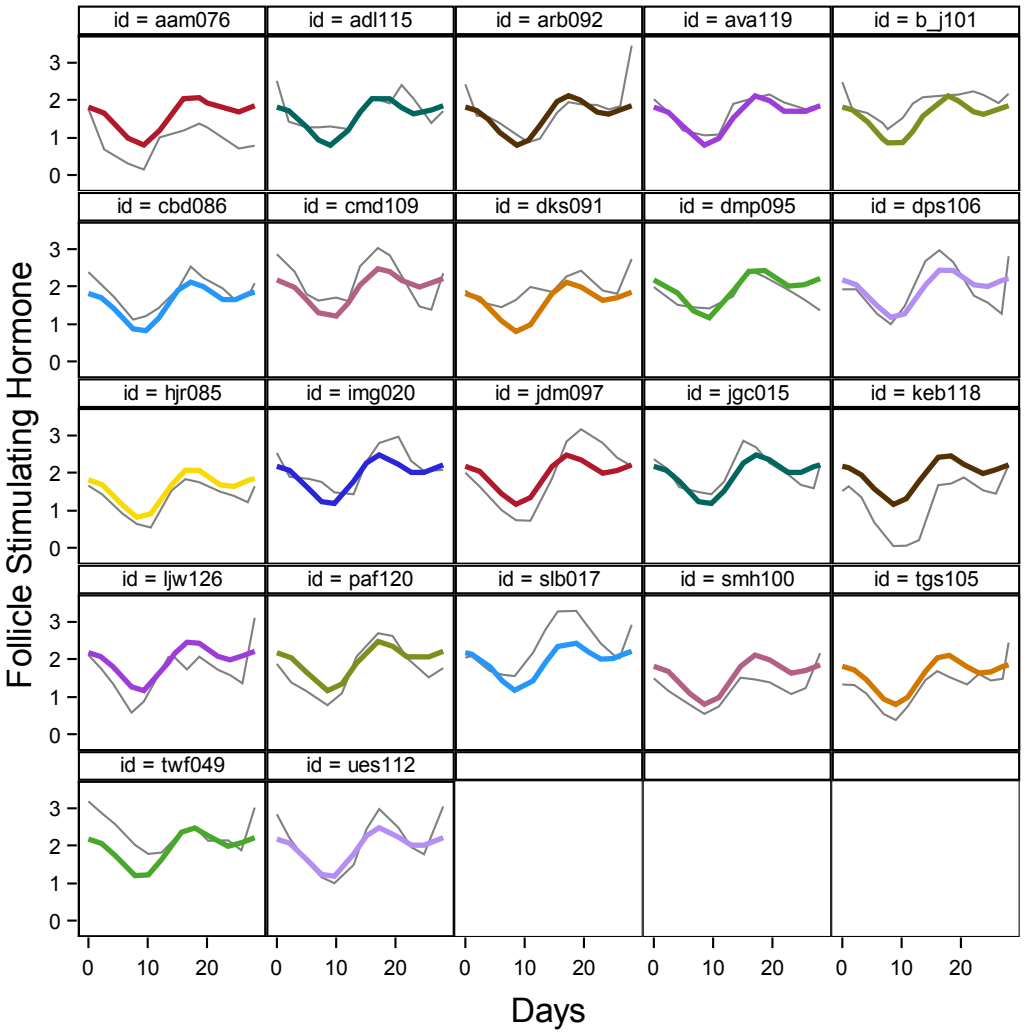
Let's do more work! Still, ... Disadvantages???

3rd degree polynomial regression with trigonometric functions



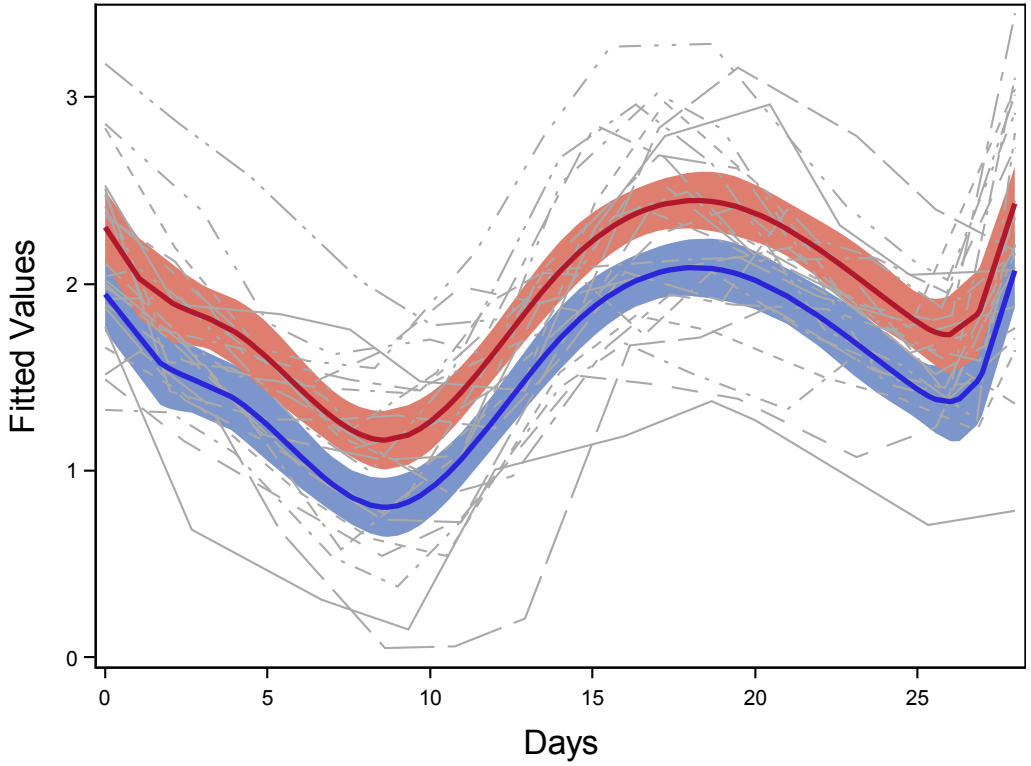
Let's do more work! Still, ... Disadvantages???

3rd degree polynomial regression with trigonometric functions



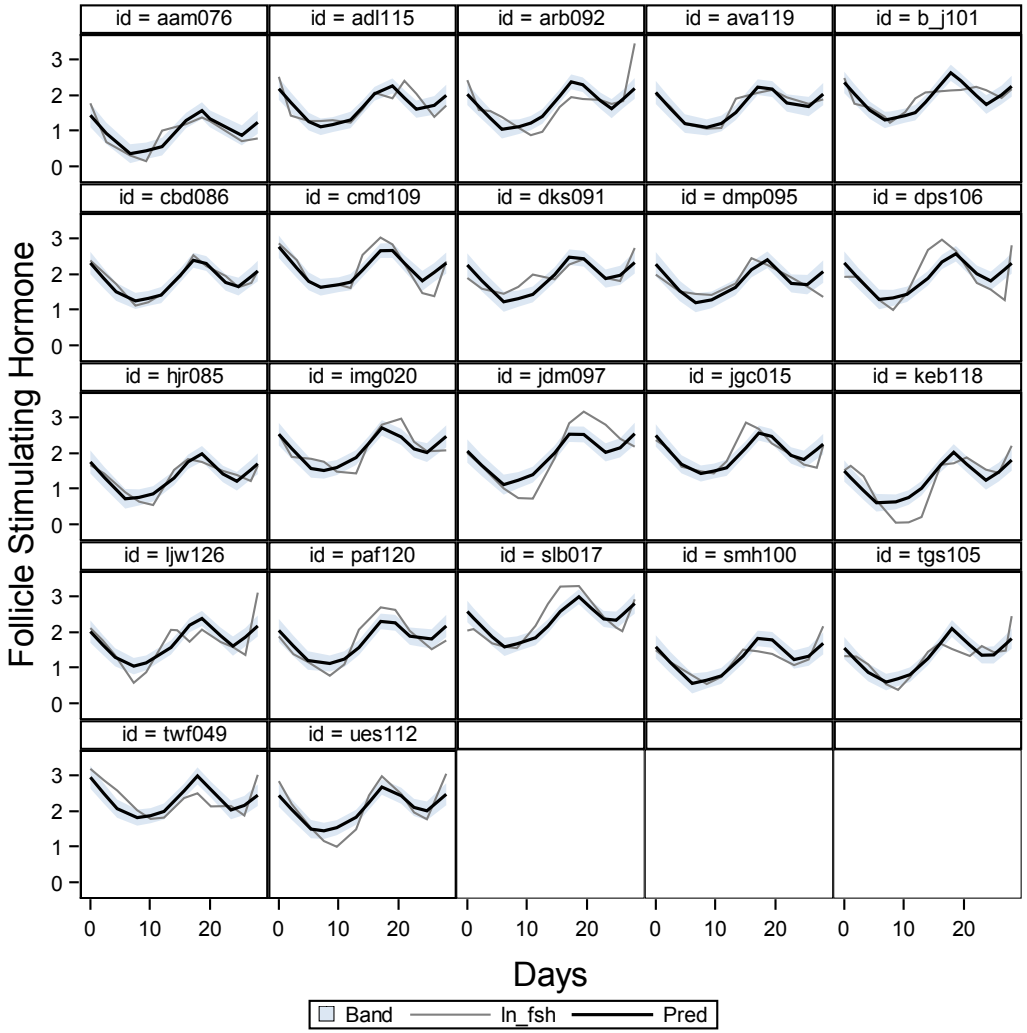
Much Better Now! Disadvantages???

Natural Cubic Splines



Much Better Now! Disadvantages???

Natural Cubic Splines



Questions

QUESTIONS?

References I

Box, G. E. P. and Draper, N. R. (1986). *Empirical Model-building and Response Surface*. John Wiley & Sons, Inc., New York, NY, USA.

Cornelissen Germaine (2014). Cosinor-based rhythmometry. *Theoretical Biology & Medical Modelling*, 11:16–16.

Fitzmaurice, G., Laird, N., and Ware, J. (2004). *Applied Longitudinal Analysis*. Wiley Series in Probability and Statistics - Applied Probability and Statistics Section Series. Wiley.

Gurrin, L. C., Scurrah, K. J., and Hazelton, M. L. (2005). Tutorial in biostatistics: spline smoothing with linear mixed models. *Statistics in Medicine*, 24(21):3361–3381.

Johnson, W., Balakrishna, N., and Griffiths, P. L. (2013). Modeling physical growth using mixed effects models. *American Journal of Physical Anthropology*, 150(1):58–67.

References II

Miles, J. and Shevlin, M. (2001). *Applying Regression and Correlation: A Guide for Students and Researchers*. SAGE Publications.

Naumova, E. N., Must, A., and Laird, N. M. (2001). Tutorial in Biostatistics: Evaluating the impact of 'critical periods' in longitudinal studies of growth using piecewise mixed effects models. *International Journal of Epidemiology*, 30(6):1332–1341.

Singer, J. D. (1998). Using SAS PROC MIXED to Fit Multilevel Models, Hierarchical Models, and Individual Growth Models. *Journal of Educational and Behavioral Statistics*, 23(4):323–355.

Zwietering M H, Jongenburger I, Rombouts F M, and van 't Riet, K. (1990). Modeling of the Bacterial Growth Curve. *Applied and Environmental Microbiology*, 56(6):1875–1881.