Modelling menstrual cycle length using state space models

Thiago de Paula Oliveira

thiago.oliveira@ed.ac.uk

https://prof-thiagooliveira.netlify.com

The Roslin Institute

13th May 2021

@HighlanderLab

Contents

Background

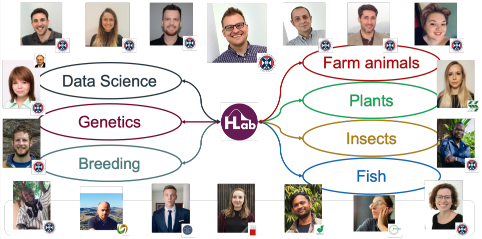
- BSc (2007-2012) in **Agriculture**, and MSc and PhD (2012-2018) in **Statistics**
  - University of São Paulo (ESALQ/USP)
  - Mixed-Effects Models; Concordance Analysis (applied to post-harvest)
  - Development of R packages
- Lecturer and Postdoc (2018-2019) in **Statistics**
  - ESALQ/USP
  - Development of R package called `lcc`
  - Lecture in Calculus and Experimental Statistics
- Postdoc (2019-2020) in **Biostatistics** - Athlete Performance and COVID-19
  - Extreme Values Theory applied to the prediction of new COVID-19 waves
  - Multilevel models, Longitudinal Data, State-Space Models
  - Evaluation of NBA athletes performance (Basketball)
  - One-step-ahead forecast interval to predict menstrual cycle length

Collaborations

- **Brazil**:
  - University of São Paulo - GLMM applied to entomology and biological control
  - The Center for Nuclear Energy in Agriculture (CENA) - models for gas emissions and soil carbon storage
  - Federal University of Paraná - development of statistical model applied to biological control
- **Ireland**:
  - NUI Galway - development of statistical models applied to sports and COVID-19
  - Maynooth University - development of R packages and statistical models
  - NUIG + ORRECO - athlete performance index for basketball NBA players
- **Qatar**: Aspire Academy - development of predictive models to evaluate athlete performance
- **France**: Limagrain - breeding company

University of Edinburgh - The Roslin Institute

- Researcher in Statistics Applied to Genetics
- Quantifying the Drives of Genetic Change in Plant Breeding (& Animal Breeding)
- Opportunities:
  - Twitter - @HighlanderLab
  - University of Edinburgh - Vacancy



Projects - 2019-2020

NUI Galway
OÉ Gaillimh

Insight

ORRECO
Game Changing Analytics

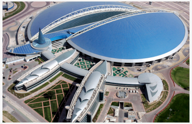
THE USE OF BIOSTATISTICS FOR OPTIMIZING ATHLETES PERFORMANCE

Uniquely blending Data Science and Sports Science to generate customized strategies by athlete

1. Modelling menstrual cycle length in athletes using state-space models
2. Oliveira-Newell Index - Athlete performance measure
    - Basketball
    - Football

ORRECO
Game Changing Analytics

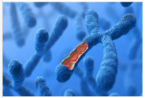
3. Development of predictive models
    - Mixed-effects models
    - Hierarchical Models
    - Shiny app



ASPIRE
ACADEMY

Projects - 2021

ROSLIN



HLab

STATISTICAL MODELS APPLIED TO QUANTITATIVE GENETICS AND GENOMICS OF PLANT BREEDING

Quantitative genetics is a field of biology studying the effect of genetic and environmental factors on quantitative traits.

AlphaPart

development active | CRAN 0.8.1 | build passing | codecov 65% | downloads 512/month | downloads 13K

A software that implements a method for **partitioning genetic trends** to quantify the sources of genetic gain in breeding programmes. The partitioning method is described in Garcia-Cortes et al. (2008). The package includes the main function `AlphaPart` for partitioning breeding values and auxiliary functions for manipulating data and summarizing, visualizing, and saving outputs.

Maize breeding programme

Aim:

Utilise genomic information to identify which genome regions drive genetic change and which sources contribute to favourable alleles in these genome regions

LG

Suggestions of Training

1. Courses:
   - Advanced - GitHub: Advance your journey
   - Beginner - Bash/Shell
   - Fundamental/Intermediate - C++
   - Beginner - SFTP, SSH server
   - Additional:
     - Interactive Web Applications (Shiny)
     - LaTex - Overleaf: online real-time collaborative LaTeX editor

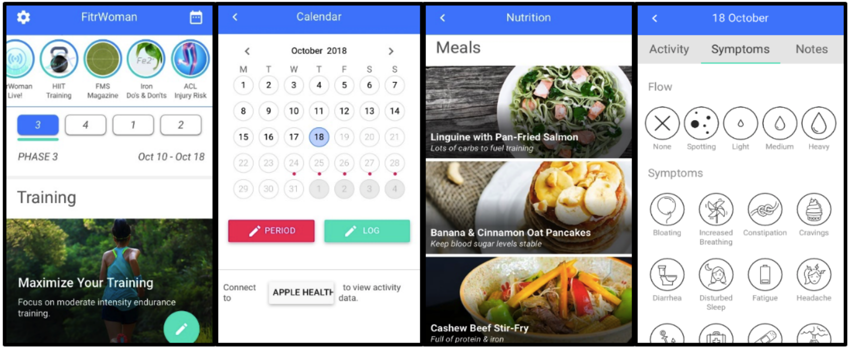
2. Everyday
   - Organise yourself - Evernote and Joplin are a free, open source note taking and to-do application
   - Build powerful Slack workflows + GitHub integration
   - Zoom and Microsoft Teams

Modelling menstrual cycle length using state space models

Introduction & Motivation

Introduction and motivation

- FitrWoman: New opportunity to study the menstrual cycle in athletes
- Record detailed menstrual cycle data
- Investigate associations between menstrual cycle symptoms and cycle length



Introduction and motivation

1. Target population: Female Athletes who are using the FitrWoman
2. Data came from FitrWoman app (Rights of Individuals under the GDPR[a])
3. Statistical analyses are required to improve predictions of menstrual cycle length in athletes

[a]GDPR - General Data Protection Regulation

Objective

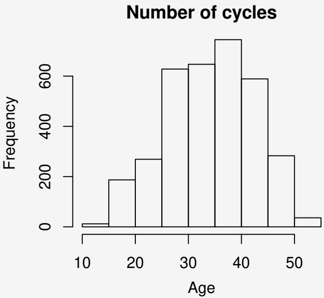
- Derived a predictive distribution based on individual repeated measurements using a state-space model formulation
- One-step ahead forecasting of menstrual cycle length based on individual past information, considering:
  1. Standard menstrual cycle
  2. Non-standard: menstrual dysfunction - greater than 35 days

Descriptive Analysis & Methods

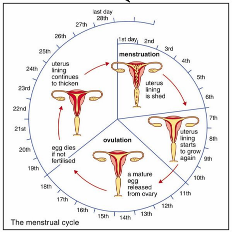
Data characteristics

2178 Female Athletes

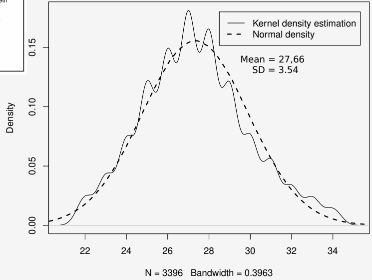
- Age: 13.18 - 53.34
- Mean age: 35.16 years old



At least 4 cycles

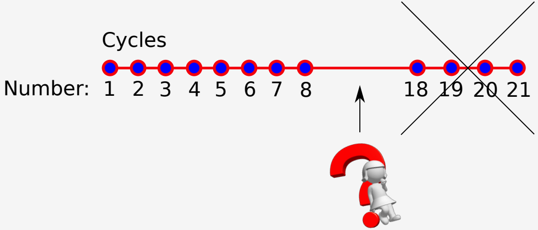


- 30 covariables
  - 28 True or False
  - 2 Continuous
- N. of Cycles: 16,990
- Longest recorded sequence of consecutive cycles
  - 53 measurements

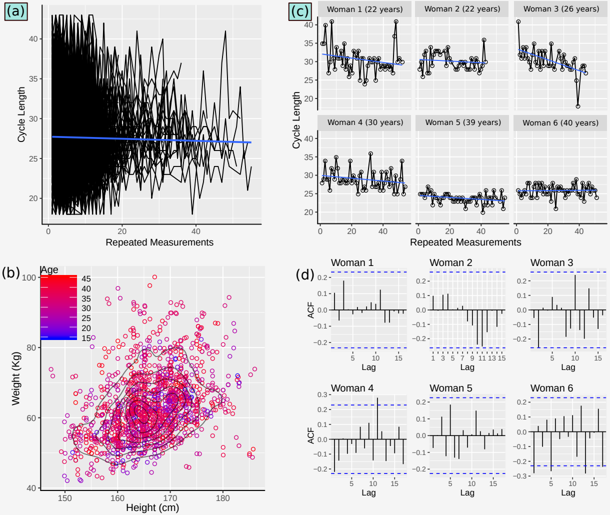


Data characteristics

1. Women can contributed with more than one sequence to the dataset
   1. no information on the reasons that caused this temporary dropout;
   2. we consider only the first sequence available;
   3. inclusion of the following sequences might bias the analysis



Descriptive Analysis & Methods



(e)

	Symptoms	No	Yes
1	Bloating	0.82	0.18
2	Constipation	0.94	0.06
3	Cravings	0.87	0.13
4	Diarrhea	0.91	0.09
5	Disturbed Sleep	0.82	0.18
6	Fatigue	0.81	0.19
7	Headache	0.87	0.13
8	Heavy legs	0.91	0.09
9	High temperature	0.92	0.08
10	Illness	0.96	0.04
11	Increased breathing	0.96	0.04
12	Injury	0.98	0.02
13	Irritability	0.86	0.14
14	Muscle ache	0.89	0.11
15	Nausea	0.94	0.06
16	Poor concentration	0.92	0.08
17	Stomach cramps	0.78	0.22
18	Stressed	0.89	0.11
19	Tender breasts	0.88	0.12
20	Weak	0.91	0.09
	Levels of flow amount:		
21	Heavy	0.78	0.22
22	Medium	0.17	0.83
23	light	0.70	0.30
24	Spotting	0.81	0.19
25	None	0.89	0.11
	Levels of training Intensity:		
26	Low	0.93	0.07
27	Medium	0.92	0.08
28	High	0.93	0.07

Methods

Statistical Model

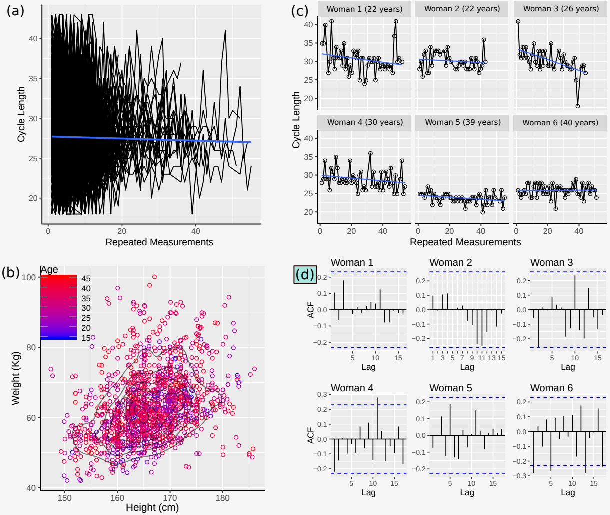
$$
\begin{aligned}
y_{ij} &= m_{ij}, \quad i = 1, 2, \ldots, I,\ j = 1, 2, \ldots, \mathcal{J} \\
m_{ij} &= m_{i,j-1} + \eta_{ij},\ \text{with } m_{i1} \sim N(\beta_0, \sigma^2_\eta) \text{ and } \eta_{ij} \sim N(0, \sigma^2_\eta),
\end{aligned}
\tag{1}
$$

Consequently, in n further steps we have:

$$
m_{i,j+n} = m_{i,j-1} + \sum_{k=j}^{j+n} \eta_{ik} \text{ with } E\left[m_{i,j+n}\right] = \beta_0 \text{ and } \underbrace{Var\left[m_{i,j+n}\right] = n \times \sigma^2_\eta}_{\text{serie is not stationary}}
$$

- y_{ij} is the menstrual cycle length for the i-th woman at j-th cycle;
- m_{it} is a random walk model that allows a individual trend in the series
- η_{ij} assumed to be normally distributed with mean 0 and variance σ^2_η.

Descriptive Analysis



(e)

	Symptoms	No	Yes
1	Bloating	0.82	0.18
2	Constipation	0.94	0.06
3	Cravings	0.87	0.13
4	Diarrhea	0.91	0.09
5	Disturbed Sleep	0.82	0.18
6	Fatigue	0.81	0.19
7	Headache	0.87	0.13
8	Heavy legs	0.91	0.09
9	High temperature	0.92	0.08
10	Illness	0.96	0.04
11	Increased breathing	0.96	0.04
12	Injury	0.98	0.02
13	Irritability	0.86	0.14
14	Muscle ache	0.89	0.11
15	Nausea	0.94	0.06
16	Poor concentration	0.92	0.08
17	Stomach cramps	0.78	0.22
18	Stressed	0.89	0.11
19	Tender breasts	0.88	0.12
20	Weak	0.91	0.09
	Levels of flow amount:		
21	Heavy	0.78	0.22
22	Medium	0.17	0.83
23	light	0.70	0.30
24	Spotting	0.81	0.19
25	None	0.89	0.11
	Levels of training Intensity:		
26	Low	0.93	0.07
27	Medium	0.92	0.08
28	High	0.93	0.07

Methods

Statistical Model

$$y_{ij} = m_{ij} + \gamma_{ij}$$

$$m_{ij} = m_{i,j-1} + \eta_{ij}, \text{ with } \eta_{ij} \sim N(0, \sigma_\eta^2), \text{ with } m_{i1} \sim N(\beta_0, \sigma_\eta^2) \text{ and } \eta_{ij} \sim N(0, \sigma_\eta^2),$$

$$\gamma_{ij} = \phi\gamma_{i,j-1} + \theta\epsilon_{i,j-1} + \epsilon_{ij}, \text{ with } \epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$$

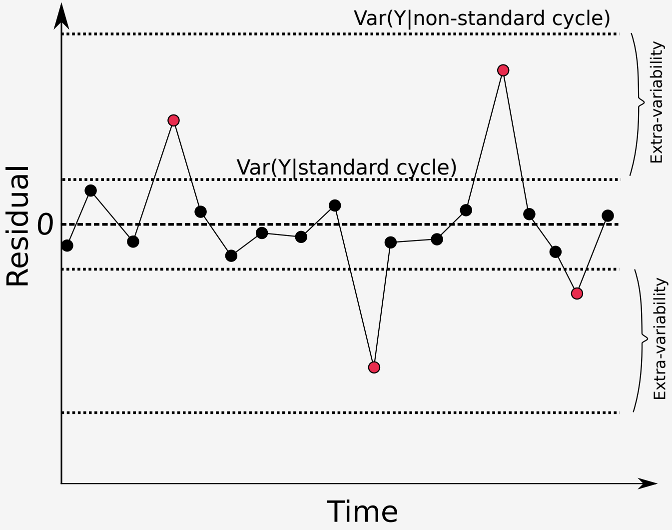
Consequently,

$$E\left[Y_{ij}\right] = E\left[m_{ij}\right] + E\left[\gamma_{ij}\right] = \beta_0 \tag{2}$$

and

$$Var\left[Y_{ij}\right] = Var\left[m_{ij}\right] + Var\left[\gamma_{ij}\right] + 2Cov\left(m_{ij}, \gamma_{ij}\right) = j\sigma_\eta^2 + \frac{\left(1 + 2\phi\theta + \theta^2\right)\sigma_\epsilon^2}{1 + \phi^2} \tag{3}$$

Extra-Variability due to unusual menstrual cycles



Methods

Statistical Model

$$y_{ij} = m_{ij} + \gamma_{ij} + r_{ij}$$

$$m_{ij} = m_{i,j-1} + \eta_{ij}, \text{ with } \eta_{ij} \sim N(0, \sigma^2_\eta), \text{ with } m_{i1} \sim N(\beta_0, \sigma^2_\eta) \text{ and } \eta_{ij} \sim N(0, \sigma^2_\eta),$$

$$\gamma_{ij} = \phi\gamma_{i,j-1} + \theta\epsilon_{i,j-1} + \epsilon_{ij}, \text{ with } \epsilon_{ij} \sim N(0, \sigma^2_\epsilon)$$

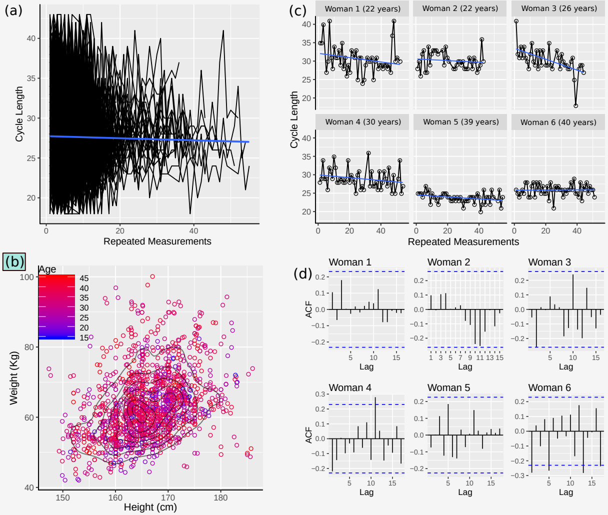
$$r_{ij} = \lambda_{ij} w_{ij}, \; w_{ij} \sim N\left(0, \sigma^2_w\right), \; \lambda_{ij} \sim \text{Bernoulli}\left(\pi\right), \; \pi \sim \text{Uniform}\left(0, 1\right)$$

Consequently,

$$Var\left[Y_{ij}|\lambda_{ij} = 0\right] = Var\left[m_{ij}\right] + Var\left[\gamma_{ij}\right] + 2Cov\left(m_{ij}, \gamma_{ij}\right) = j\sigma^2_\eta + \frac{\left(1 + 2\phi\theta + \theta^2\right)\sigma^2_\epsilon}{1 + \phi^2}$$

$$Var\left[Y_{ij}|\lambda_{ij} = 1\right] = Var\left[m_{ij}\right] + Var\left[\gamma_{ij}\right] + 2Cov\left(m_{ij}, \gamma_{ij}\right) = j\sigma^2_\eta + \frac{\left(1 + 2\phi\theta + \theta^2\right)\sigma^2_\epsilon}{1 + \phi^2} + \sigma^2_w$$

Descriptive Analysis



(e)

	Symptoms	No	Yes
1	Bloating	0.82	0.18
2	Constipation	0.94	0.06
3	Cravings	0.87	0.13
4	Diarrhea	0.91	0.09
5	Disturbed Sleep	0.82	0.18
6	Fatigue	0.81	0.19
7	Headache	0.87	0.13
8	Heavy legs	0.91	0.09
9	High temperature	0.92	0.08
10	Illness	0.96	0.04
11	Increased breathing	0.96	0.04
12	Injury	0.98	0.02
13	Irritability	0.86	0.14
14	Muscle ache	0.89	0.11
15	Nausea	0.94	0.06
16	Poor concentration	0.92	0.08
17	Stomach cramps	0.78	0.22
18	Stressed	0.89	0.11
19	Tender breasts	0.88	0.12
20	Weak	0.91	0.09
	Levels of flow amount:		
21	Heavy	0.78	0.22
22	Medium	0.17	0.83
23	light	0.70	0.30
24	Spotting	0.81	0.19
25	None	0.89	0.11
	Levels of training Intensity:		
26	Low	0.93	0.07
27	Medium	0.92	0.08
28	High	0.93	0.07

Methods

Statistical Model

$$y_{ij} = m_{ij} + \gamma_{ij} + r_{ij} + c_{ij}$$

$$m_{ij} = m_{i,j-1} + \eta_{ij}, \text{ with } \eta_{ij} \sim N(0, \sigma^2_\eta), \text{ with } m_{i1} \sim N(\beta_0, \sigma^2_\eta) \text{ and } \eta_{ij} \sim N(0, \sigma^2_\eta),$$

$$\gamma_{ij} = \phi\gamma_{i,j-1} + \theta\epsilon_{i,j-1} + \epsilon_{ij}, \text{ with } \epsilon_{ij} \sim N(0, \sigma^2_\epsilon)$$

$$r_{ij} = \lambda_{ij} w_{ij}, \; w_{ij} \sim N\left(0, \sigma^2_w\right), \; \lambda_{ij} \sim \text{Bernoulli}\left(\pi\right), \; \pi \sim \text{Uniform}\left(0, 1\right)$$

$$c_{ij} = \sum_{k=1}^{K} \alpha_k C_{ijk}$$

- c_{ij} represents additional symptoms predictors (C_{ij}) that may be useful menstrual cycle length prediction

Methods

Statistical Model - State-Space Formulation

$$
\begin{aligned}
y_{ij} &= m_{ij} + \gamma_{ij} + \theta x_{ij} + r_{ij} + c_{ij}, \\
m_{ij} &= m_{i,j-1} + \eta_{ij}, \text{ with } \eta_{ij} \sim N(0, \sigma^2_\eta), \\
\gamma_{ij} &= \phi \gamma_{i,j-1} + \epsilon_{ij}, \text{ with } \epsilon_{ij} \sim N(0, \sigma^2_\epsilon), \\
x_{ij} &= \epsilon_{i,j-1}, \\
r_{ij} &= \lambda_{ij} w_{ij},\ w_{ij} \sim N\left(0, \sigma^2_w\right),\ \lambda_{ij} \sim \text{Bernoulli}\left(\pi\right),\ \pi \sim \text{Uniform}\left(0, 1\right) \\
c_{ij} &= \textstyle\sum_{k=1}^{K} \alpha_k C_{ijk}
\end{aligned} \tag{4}
$$

- Initial value $m_{i1} \sim N\left(\beta_0, \sigma^2_\eta\right)$ for the local level model and $\gamma_{ij} = \sum_{t=0}^{j-1} \phi^t \epsilon_{i,j-t}$, with $j \geq 1$

Linear mixed effect state-space model

1. Assuming a specific linear regression by woman can be an important issue:
   1. Random intercept account for a natural variability relative to a uncertainty about how the non-observed variables affected the first observed menstrual cycle;
   2. Random slope assumes that each woman may have different menstrual cycle length trends relative to her age
2. We can use the LMM approach to verify if the random walk model have the necessary flexibility to capture different possible trends
3. $m_{ij} = \beta_0 + b_{0i} + (\beta_1 + b_{1i})\, Age_{ij}$

$$\boldsymbol{b}_i = \begin{bmatrix} b_{0i} \\ b_{1i} \end{bmatrix} \sim N_2\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \boldsymbol{G} = \begin{bmatrix} \sigma^2_{b_0} & \sigma_{b_{01}} \\ \sigma_{b_{01}} & \sigma^2_{b_1} \end{bmatrix} \right).$$

Model selection

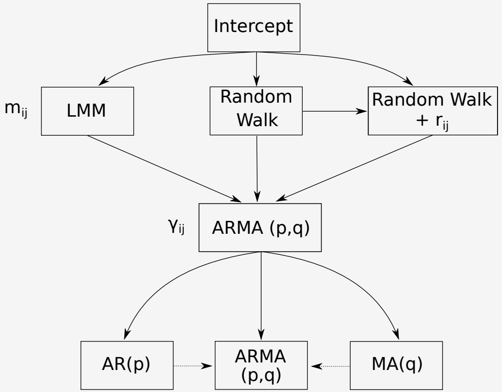


Figure: Stages 1 and 2 of the model selection procedure

- Selection of predictors based on 95% Bayesian credible interval.

Bayesian parameter estimation

1. Here the inverse-gamma(κ, κ) prior distributions are used for:
    1. the random walk model
    2. variance components of mixed effect model
2. An attempt at non-informativeness within the conditionally conjugate family, with κ set to a low value such as 0.1^3:

$$\sigma_{\epsilon}^{-2}, \sigma_{\eta}^{-2}, \sigma_{b_0}^{-2}, \sigma_{b_1}^{-2}, \sigma_{\beta}^{-2}, \sigma_{ar}^{-2} \sim \text{Gamma}\left(0.1^3, 0.1^3\right).$$

Likelihood ratio test under $H_0 : \sigma_{b01} = 0$ vs $H_a : \sigma_{b01} \neq 0$

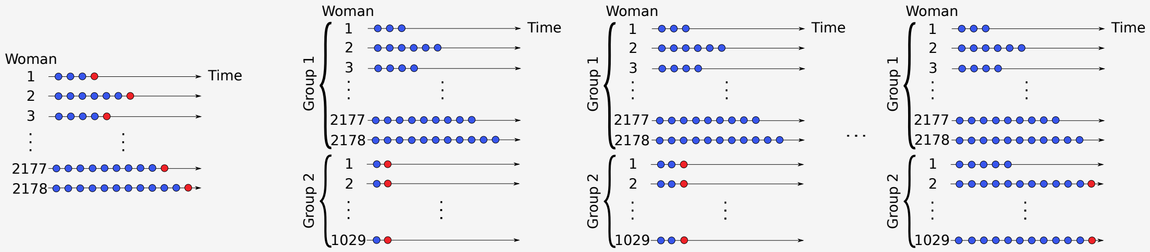
- σ_{b01} did not plays a crucial role and then can be removed from the model (p-value $= 0.33$)

Bayesian parameter estimation

1. Prior distribution for fixed effect parameters is given by
   - $\beta_0 \sim N\left(\mu_{\beta_0}, \sigma^2_\beta\right)$, with $\mu_{\beta_0} \sim$ Uniform $(24, 32)$
   - $\beta_1 \sim N\left(\mu_{\beta_1}, \sigma^2_\beta\right)$, with $\mu_{\beta_1} \sim$ Uniform $(-2, 2)$
   - $\phi_0, \theta_0 \sim \mathcal{N}\left(\mu_{ar}, \sigma^2_{ar}\right)$, with $\mu_{ar} \sim \mathcal{N}\left(0, 100\right)$
   - $\alpha_k \sim \mathcal{N}\left(0, 100\right)$ (vague normal density prior)
2. All assumptions were checked to make sure that results were not sensitive to specific choices of prior parameters.

Evaluating the forecasting prediction

1. We are interested in the one-step ahead predictive value for each woman
2. Split our data set in train and test data
3. 561 cycles in the test data
4. Compute the RMSE, CCC and ρ between the fitted and true values (test data)



(a) (b)

Results & Discussion

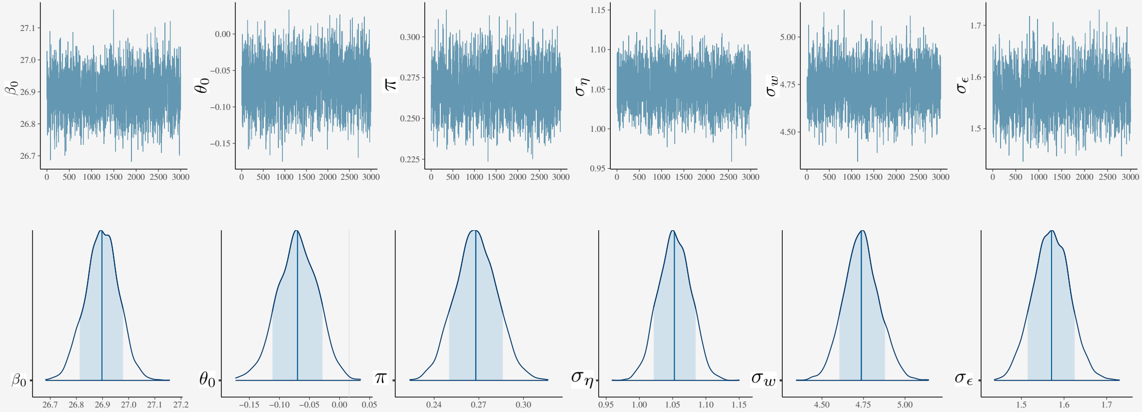
Model Selection

Model	N. Par.	Forecasting RMSE	Forecasting CCC	Forecasting r	BIC
$y_{ij} = m_{ij} + \epsilon_{ij}$	3	1.6104	0.7433	0.7606	16,961.19
$y_{ij} = m_{ij} + \text{AR}(1)$	4	1.6740	0.7546	0.7618	9,733.27
$y_{ij} = m_{ij} + \text{MA}(1)$	4	1.6640	0.7564	0.7637	9,726.88
$y_{ij} = m_{ij} + \text{ARMA}(1,1)$	5	1.6782	0.7560	0.7630	9,314.73
$y_{ij} = m_{ij} + r_{ij} + \epsilon_{ij}$	5	1.6153	0.7362	0.7479	8259.67
$y_{ij} = m_{ij} + r_{ij} + \text{AR}(1)$	6	1.6098	0.7410	0.7510	7797.06
$y_{ij} = m_{ij} + r_{ij} + \text{MA}(1)$	6	1.6126	0.7401	0.7501	**7739.40**
$y_{ij} = m_{ij} + r_{ij} + \text{ARMA}(1,1)$	7	1.6031	0.7405	0.7521	7995.92
$y_{ij} = \beta_0 + b_{0i} + (\beta_1 + b_{1i})\, Age_{ij} + \epsilon_{ij}$	5	1.6481	0.7265	0.7457	18,099.64
$y_{ij} = \beta_0 + b_{0i} + (\beta_1 + b_{1i})\, Age_{ij} + \text{AR}(1)$	6	1.7796	0.6964	0.7087	13,314.69
$y_{ij} = \beta_0 + b_{0i} + (\beta_1 + b_{1i})\, Age_{ij} + \text{MA}(1)$	6	1.7608	0.7005	0.7137	13,381.90
$y_{ij} = \beta_0 + b_{0i} + (\beta_1 + b_{1i})\, Age_{ij} + \text{ARMA}(1,1)$	8	1.7769	0.6998	0.7113	13,289.62

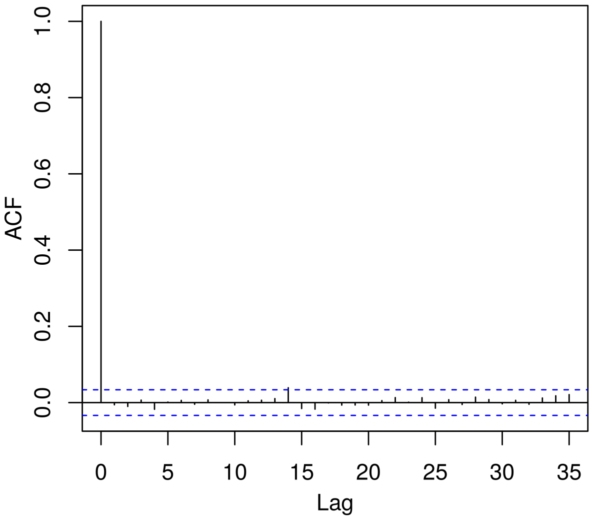
Estimates

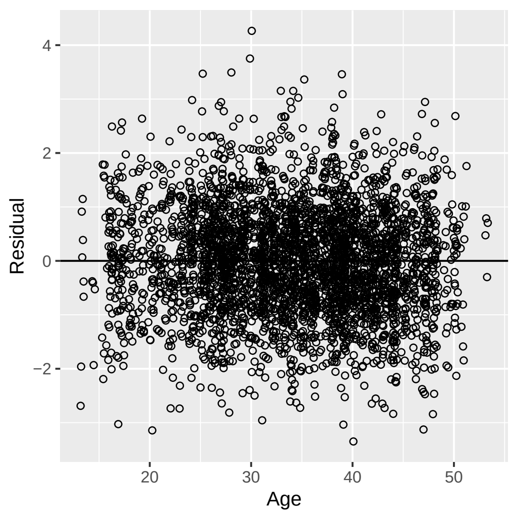
			95% Credible Interval	
Parameter	Estimate	SE	Lower	Upper
β_0	26.8974	0.0639	26.7704	27.0197
π	0.2681	0.0139	0.2414	0.2958
θ_0	-0.0669	0.0307	-0.1287	-0.0088
α_1 (Heavy Legs)	-0.0619	0.0363	-0.1320	0.0084
α_2 (Increased Breathing)	0.1240	0.0447	0.0353	0.2114
α_3 (Injury)	0.1438	0.0486	0.0498	0.2382
α_4 (Stomach Cramps)	0.1315	0.0303	0.0717	0.1911
α_5 (Tender Breasts)	-0.1255	0.0248	-0.1734	-0.0767
α_6 (Flow Amount: Heavy)	0.0881	0.0410	0.0061	01679
α_7 (Flow Amount: Medium)	0.1462	0.0140	0.1199	0.1742
α_8 (Flow Amount: Light)	0.0940	0.0302	0.0357	0.1544
α_9 (Flow Amount: Spotting)	0.0977	0.0330	0.0332	0.1611
α_{10} (Flow Amount: None)	0.0308	0.0131	0.0051	0.0565
σ_w	4.7386	0.1058	4.5379	4.9492
σ_η	1.0530	0.0241	1.0060	1.0526
σ_ϵ	1.5702	0.0428	1.4864	1.6545

Diagnostic plots

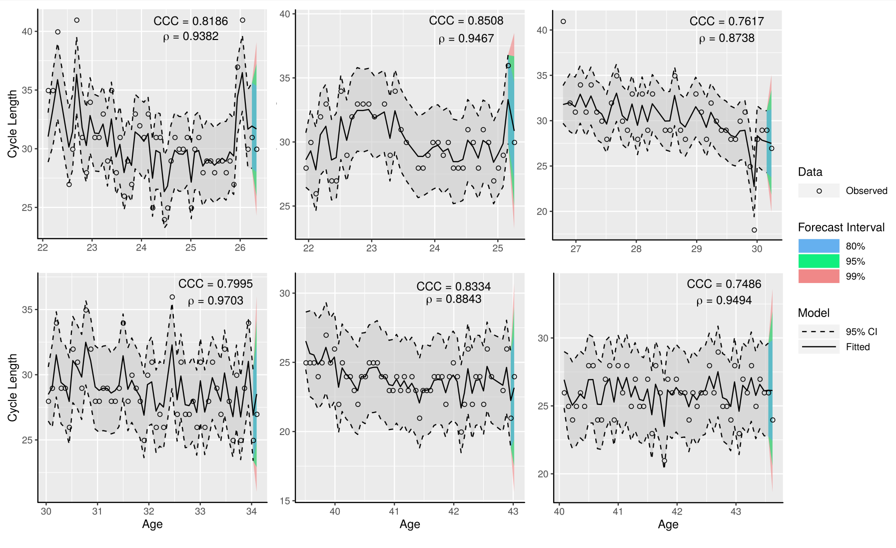


Diagnostic plots

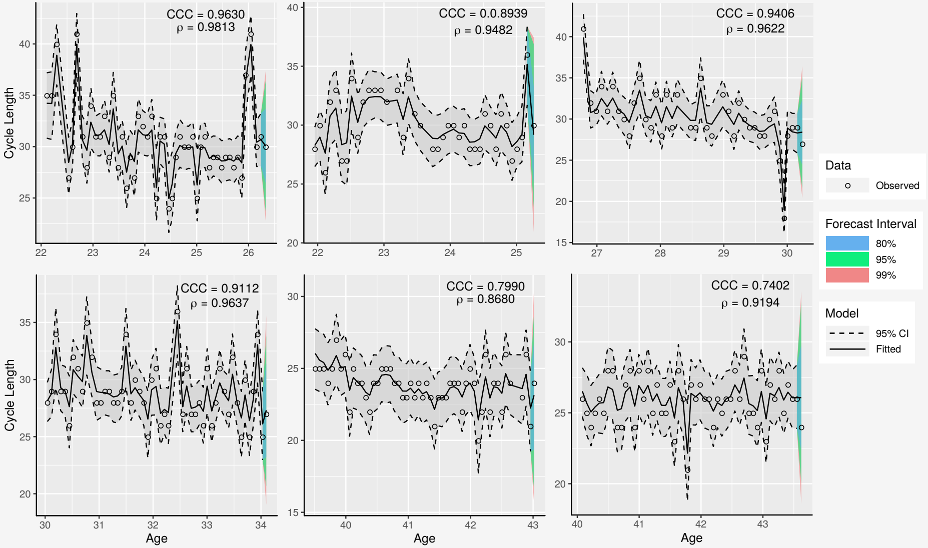




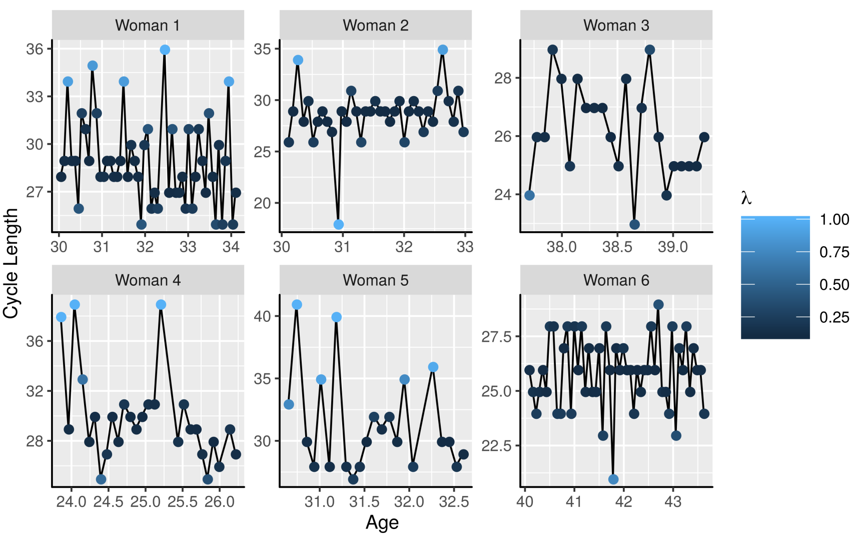
State-Space Model without consider the extra-variability



State-Space Model considering the extra-variability (r_{ij})



$$r_{ij} = \lambda_{ij} w_{ij}, \ \ w_{ij} \sim N\left(0, \sigma_w^2\right)$$



One-step-ahead point forecast prediction for a new group

Table: Evaluation of one-step-ahead forecast prediction based on root mean square error (RMSE), concordance correlation (CCC), Pearson correlation (r), and accuracy (C_b) coefficients between the predict and observed values of a new group with N women whom have n_i cycles

			CCC				
N	n_i	RMSE	Est	Lower	Upper	r	C_b
1029	3	5.2654	0.2544	0.2072	0.3005	0.3066	0.8297
760	4	5.0889	0.3113	0.2567	0.3640	0.3657	0.8514
603	5	5.1767	0.2670	0.1989	0.3326	0.2940	0.9079
434	6	5.0879	0.2688	0.1855	0.3482	0.2869	0.9367
324	7	5.2182	0.2236	0.1231	0.3195	0.2348	0.9521
248	8	5.5807	0.1783	0.0620	0.2898	0.1882	0.9472

Conclusions

- A random walk with overdispersion parameter and MA(1) model was useful to describe the complex dynamics of menstrual cycle length over time
- Results shows high values of CCC and Pearson correlation between fitted and observed values by woman
  - Prediction for a new group of female athletes is still problematic
- We demonstrated the importance in consider a overdispersed parameter to capture the variability of non-standard cycle by woman
  - 26.81% $[24.14\%, 29.58\%]$ of cycles in our database are overdispersed
- Random walk standard deviation:
  - under a non-overdispersed cycle is $\sigma_\eta = 1.0530$ $[1.0060, 1.0526]$ days
  - under a overdispersed cycle it increases to $\sigma_w = 4.7386$ $[4.5379, 4.9492]$ days

References

- Bull, J. R.et al. Real-world menstrual cycle characteristics of more than 600,000 menstrual cycles. **Digit.Medicine**, 2, 2019. URL: http://dx.doi.org/10.1038/s41746-019-0152-7.
- Scherwitzl, E. B., Hirschberg, A. L. & Scherwitzl, R. Identification and prediction of the fertile window using Natural Cycles. **Eur. J. Contracept. Reproductive Heal Care**, 20, 403-408, 2015.
- Li, K.et al.Characterizing physiological and symptomatic variation in menstrual cycles using self-tracked mobile-health data. **Digit. Medicine**, 3, 1-13, 2020.
- Fukaya, K., Kawamori, A., Osada, Y., Kitazawa, M. & Ishiguro, M. The forecasting of menstruation basedon a state-space modeling of basal body temperature time series. **Stat. Medicine**, 36, 3361-3379, 2017
- Gelman, A. & Rubin, D. B. Inference from iterative simulation using multiple sequences. **Stat. Sci.**, 7, 457–472, 1992
- Plummer, M., Best, N., Cowles, K. & Vines, K. CODA: convergence diagnosis and output analysis for MCMC. **R News**, 6, 7-11, 2006