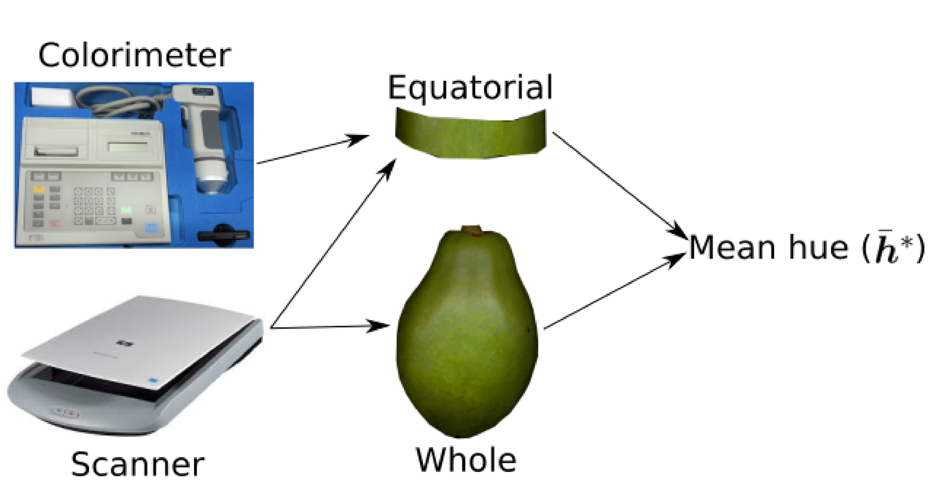
The longitudinal concordance correlation: lcc package

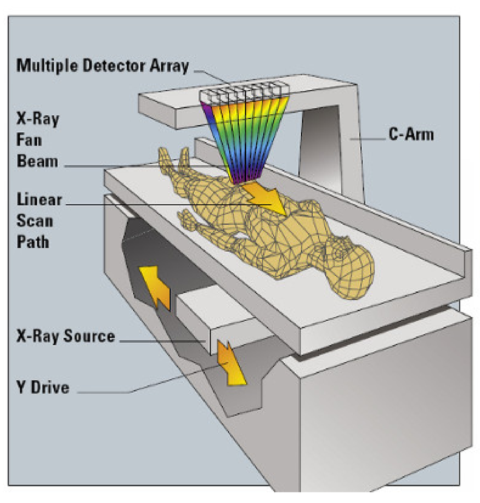
Thiago P. Oliveira, Rafael A. Moral; John Hinde, Silvio S. Zocchi Clarice G. B. Demetrio
NUIGalway; ESALQ/USP; Insigth Centre for Data Analytics; and ORRECO
Email: thiago.paulaoliveira@insigth-centre.org

Introduction

- Package implements estimation and inference procedures for the LCC
- LCC is a quantity that measures the extent of agreement between two (or more) methods
- Flexibility to accommodate unbalanced experimental designs
- Allows for different within-group error and random effects structures

Motivation





82 Females adolescent
Initial visit occured at 12
Response: mean percentage of body fat
Data: cccrm package

- Assess the agreement of the results from blood sampling performed hourly or every two hours
- The study conducted by the Asthma Clinical Research Network (ACRN) in 2001
- 144 subjects under study
- Blood draw data is available in the cccrm package

The longitudinal concordance correlation (LCC)

- Let Y_{ijlk} denote the measurement on the ith subject ($i = 1, 2, \ldots, N$) by the jth method ($j = 1, 2, \ldots, m$) at time t_{ik} ($k = 1, 2, \ldots, n_i$), where n_i is the total number of observations taken on the ith subject over time.
- Multiple mixed-effects regression model for longitudinal data

$$Y_{ijk} = \sum_{h=0}^{p} \beta_{hj} t_{ik}^{h} + \sum_{h=0}^{q} b_{hi} t_{ik}^{h} + \epsilon_{ijk}, \qquad (1)$$

with $\boldsymbol{b}_i \sim \text{MVN}\left(\boldsymbol{0}, \boldsymbol{G}\right)$ and $\boldsymbol{\epsilon}_i \sim \text{MVN}\left(\boldsymbol{0}, \boldsymbol{R}_i\right)$,

- According to [3], under the model (1), we can define the LCC based on variance components for observations measured from different unique combinations of two factors at time t_{ik} as

$$\rho_{jj'}\left(t_k\right) = \frac{\boldsymbol{t}_k \boldsymbol{G} \boldsymbol{t}_k^T}{\boldsymbol{t}_k \boldsymbol{G} \boldsymbol{t}_k^T + \frac{1}{2}\left\{\sigma_\epsilon^2\left[g\left(t_k, \boldsymbol{\delta}_j\right) + g\left(t_k, \boldsymbol{\delta}_{j'}\right)\right] + S_{jj'}^2\left(t_k\right)\right\}} = \rho_{jj'}^{(p)}\left(t_k\right) C_{jj'}\left(t_k\right) \qquad (2)$$

- **Longitudinal Pearson Correlation** (LPC): $\rho_{j,j'}^{(p)}\left(t_k\right)$ measures how far each observation deviated from the best-fit line at a fixed time $t_k = t$ (precision measure).
- **Longitudinal accuracy** (LA): $C_{j,j'}\left(t_k\right)$ measures how far the best-fit line deviates from the 45° line at a fixed time $t_k = t$ (accuracy measure).
- $S_{j,j'}\left(t_{ik}\right) = E\left(Y_{ijk}\right) - E\left(Y_{ij'k}\right) = \boldsymbol{t}_{ik}\left(\boldsymbol{\beta}_j - \boldsymbol{\beta}_{j'}\right)$, with $h = 1, 2, ..., p$ and $j \neq j'$.
- $Var\left(\epsilon_{ijk}\right) = \sigma_\epsilon^2 g\left(t_{ik}, \boldsymbol{\delta}_j\right)$, where $g(.)$ is a general variance function.
- Non-parametric bootstrap confidence interval — using simple case-resampling.

Specifying models in the lcc function

Model:

$$Y_{ijk} = \beta_{0j} + b_{0i} + \beta_{1j} t_k + \epsilon_{ijk}, \text{ with}$$
$$b_{0i} \sim N(0, \sigma_{b_0}^2) \quad \text{and} \quad \epsilon_{ijk} \sim N\left(0, \sigma_\epsilon^2\right)$$

Data structure:

```
R> library(dplyr); glimpse(data)
Observations:  554
Variables:  4
$ y <dbl> 116.5380, 115.4055, 116.9801, ...
$ Method <fct> Col, Col, Col, , ...
$ Time <int> 0, 0, 0, ...
$ Subject <fct> 1, 2, 3, , ...
```

Code:

```
R> library(lcc)
R> m1 = lcc(dataset=data, subject="Subject",
+           resp="y", method="Method",
+           time="Time", qf=1, qr=0)
```

Summary:

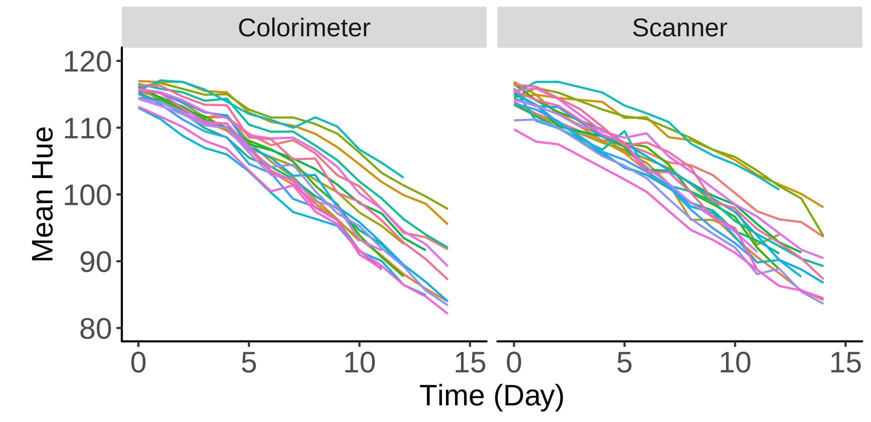
```
R> summary(m1, type = "model")
R> summary(m1, type = "lcc")
```

Plot:

```
R> lccPlot(m1,type = "lcc")
R> lccPlot(m1,type = "lpc")
R> lccPlot(m1,type = "la")
```

Examples

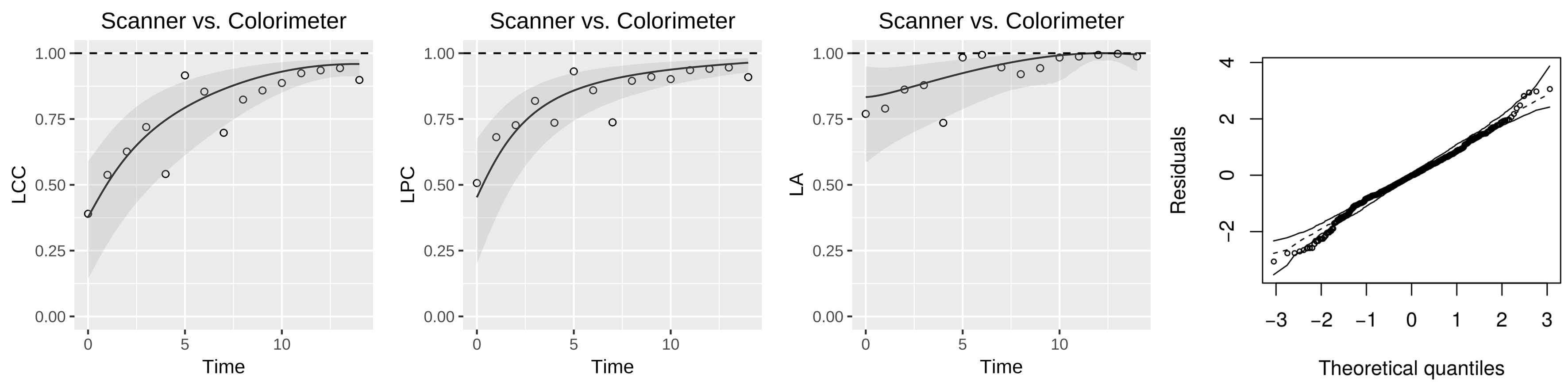
1 Hue color: equatorial region



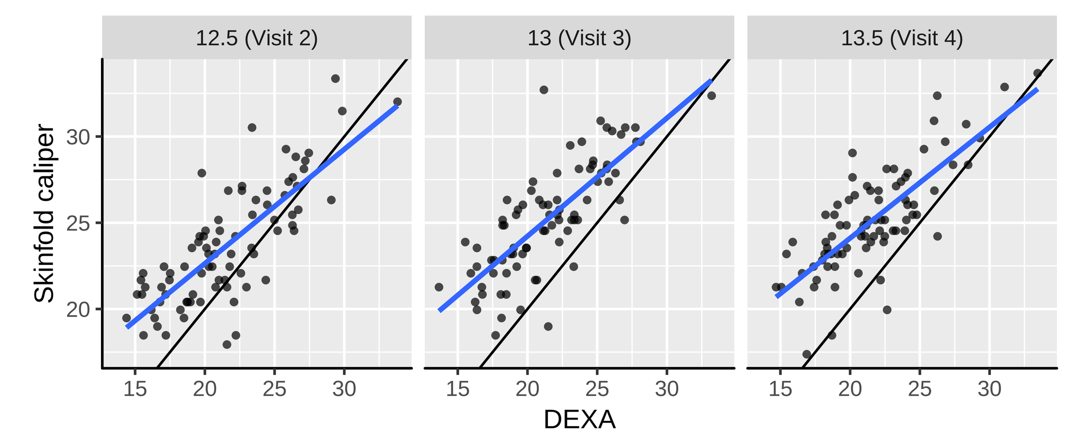
Model:

$$y_{ijk} = \beta_{0j} + b_{0i} + \left(\beta_{1j} + b_{1i}\right) t_k + \left(\beta_{2j} + b_{2i}\right) t_k^2 + \epsilon_{ijk},$$

$$\boldsymbol{b} = \left[b_{0i}, b_{1i}, b_{2i}\right]^T \sim N_3\left(\boldsymbol{0}, \boldsymbol{G}\right)$$
$$\text{and} \quad \epsilon_{ijk} \sim N\left(0, \sigma_\epsilon^2\right),$$



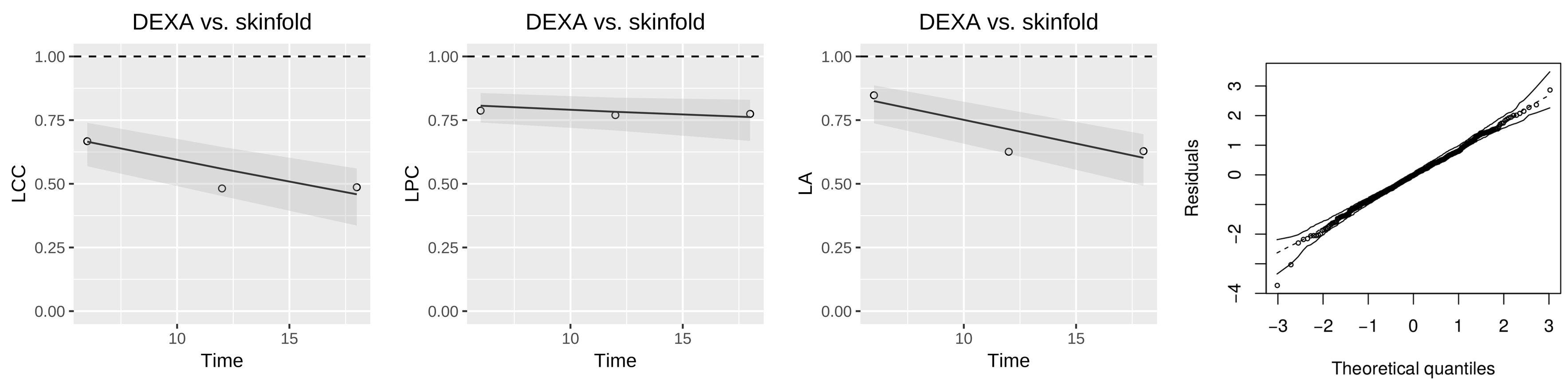
2 Percentage of body fat



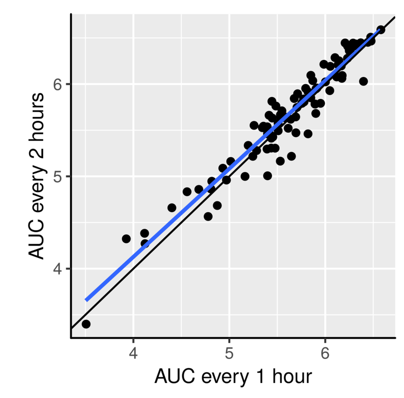
Model:

$$y_{ijk} = \beta_{0j} + b_{0i} + \left(\beta_{1j} + b_{1i}\right) t_k + \epsilon_{ijk}$$

$$\boldsymbol{b} = \left[b_{0i}, b_{1i}\right]^T \sim N_2\left(\boldsymbol{0}, \boldsymbol{G}\right)$$
$$\text{and} \quad \epsilon_{ijk} \sim N\left(0, \sigma_\epsilon^2\right),$$

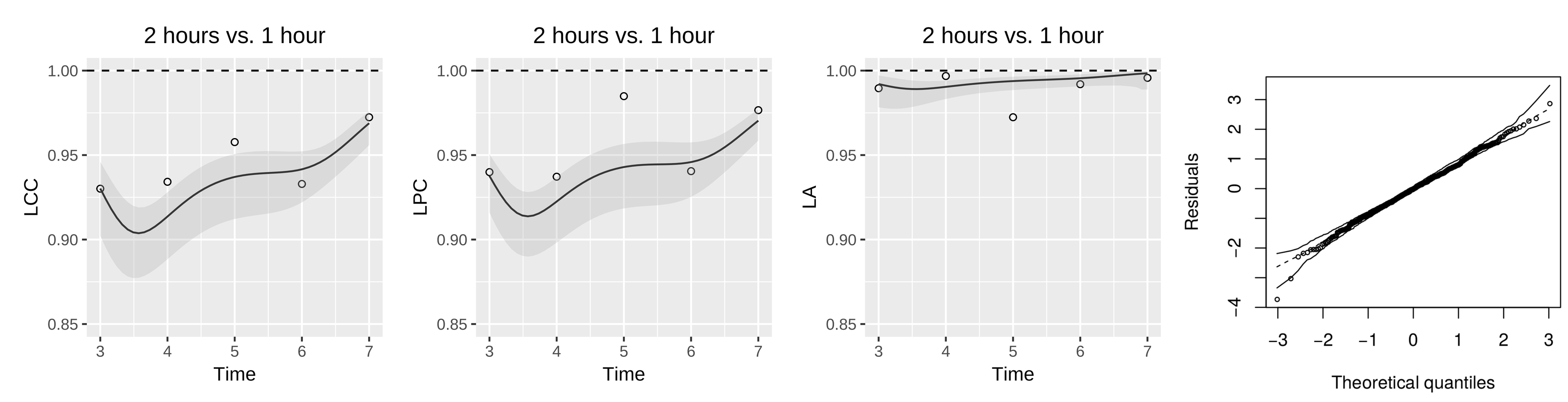


3 Blood draw sample



Model:

```
R> fit1 <- lcc(dataset = bdaw_subset, subject = "SUBJ",
+               resp = "AUC", method = "MET", time = "VNUM",
+               qf = 2, qr = 2, components = TRUE,
+               time_lcc = list(from = 3, to = 7, n = 50),
+               ci = TRUE, nboot = 10000)
```



References

[1] Lin, L. I. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, **45**, 255–268.

[2] Carrasco, J. L., King, T. S., and Chinchilli, V. M. (2009). The concordance correlation coefficient for repeated measures estimated by variance components. *Journal of Biopharmaceutical Statistics*, **19(1)**, 90–105.

[3] Oliveira, T.P., Hinde, J., Zocchi, S.S. (2018). Longitudinal Concordance Correlation Function Based on Variance Components: An Application in Fruit Color Analysis. *Journal of Agricultural, Biological, and Environmental Statistics*, **23(2)**, 233–254.

Acknowledgement

The authors are grateful to FAPESP (grant 2010/16955-1); CNPq; CAPES; and SFI under grant number SFI/12/RC/2289, co-funded by the European Regional Development Fund.