# Invited Session: Intensive Longitudinal Data

The future in many scientific domains lies in devices recording in real-time and at high temporal frequency environmental, biological, physical and behavioral information. Such intensive longitudinal data require the development of new statistical methods that can for instance handle their functional nature. They also offer the possibility of exploring the dynamic of temporal changes by investigating for instance the heterogeneity both at the mean and variability levels. The speakers will illustrate such issues and developments in environmental studies with climate change effects on forests and in biomedical studies with the study of heart rate variability.

The invited speakers will be

- Donald Hedeker, Department of Public Health Sciences, University of Chicago, USA
- Nicole Augustin, Department of Mathematical Sciences, University of Bath, UK
- Paul Eilers, Department of Biostatistics, Erasmus University Medical Center, NL

## Shared parameter mixed-effects location scale models for intensive longitudinal data

Donald Hedeker, University of Chicago, hedeker@uchicago.edu

Abstract: Intensive longitudinal data are increasingly encountered in many research areas. For example, ecological momentary assessment (EMA) and/or mobile health (mHealth) methods are often used to study subjective experiences within changing environmental contexts. In these studies, up to 30 or 40 observations are usually obtained for each subject over a period of a week or so, allowing one to characterize a subject's mean and variance and specify models for both. In this presentation, we focus on an adolescent smoking study using EMA where interest is on characterizing changes in mood variation. We describe how covariates can influence the mood variances and also extend the statistical model by adding a subject-level random effect to the within-subject variance specification. This permits subjects to have influence on the mean, or location, and variability, or (square of the) scale, of their mood responses. The random effects are then shared in a modeling of future smoking levels. These mixed-effects location scale models have useful applications in many research areas where interest centers on the joint modeling of the mean and variance structure.

## Modelling tree health and mortality for mitigation of climate change effects on forests

Nicole Augustin, University of Bath

Abstract: Forest health is monitored in Europe by The International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects (ICP Forests) in cooperation with the European Union. More recently climate change has contributed to the decline in forest health and monitoring data are

increasingly being used to investigate the effects of climate change on forests in order to decide on forest management strategies for mitigation.

We present two applications on modelling extensive yearly monitoring forest health survey data from Germany.

The first application is for official reporting of forest health. We use a Generalised additive mixed model to model defoliation, an indicator for tree health, of 5 different tree species in Germany. The temporal trend of defoliation differs between areas because of site characteristics, climate and pollution levels, making it necessary to allow for space-time interaction in the model.

In the second application we combine forest health data on mortality and crown defoliation. On a changing grid, defoliation, mortality and other tree and site specific variables are recorded. We are interested in the process leading to tree mortality and this requires the inclusion of all potential drivers of tree mortality in the model. We use a smooth additive Cox model which allows for random effects taking care of dependence between neighbouring trees and non-linear smooth functions of time varying predictors.

This is joint work with Axel Albrecht, Stefan Meining, Heike Puhlmann (Forest Research Institute, Freiburg, Germany),
Karim Anaya-Azquierdo, Alice Davis (University of Bath),
Nadine Eickenscheidt, Nicole Wellbrock (Thuenen Institute, Germany)
and Simon Wood (University of Bristol).

## From the bottom of my heart. Models for pseudo-cardiograms obtained with mobile phones.

Paul Eilers, Erasmus University Medical Center, Rotterdam, The Netherlands, p.eilers@erasmusmc.nl

Abstract: The blood flow in our skin fluctuates with the beating of our heart, and light reflection of the skin fluctuates with it. This makes it possible to measure heart rate with the camera of a modern mobile phone. This is called photoplethysmography (PPG). On Google Play or in Apple’s App Store many applications can be found that can do this.

A step further is to record the (spatially averaged) camera signal over a longer time, say 60 seconds, roughly equivalent to 60 heart beats. It is a proxy for a true electrocardiogram (ECG) and can, in principle, provide important information on the condition and the functioning of the heart. A typical goal is the detection of arrhythmia, such as arterial fibrillation.

The PPG signal does not show sharp peaks like a true electrocardiogram (ECG). It looks more like a sine wave, with a small amplitude, on a strongly drifting baseline. Almost exclusively, researchers try to locate local peak maxima and use the distances between peaks to determine (aberrations in) the heart rate. Because of the low quality of the signal, the precision of this procedure is poor. I will present two statistical models as an improvement.

After trend removal, the signal looks like a sine wave with varying frequency and amplitude. Its logarithm can be modeled as a smooth series of complex numbers. The real part represents the amplitude and the imaginary part the phase. Both series are modeled with P-splines. The derivative of the phase gives the momentary frequency. This model is highly nonlinear, but it is possible to find good starting values for a fast, linearized, fitting procedure.

The second model assumes that sharp spikes are distorted by an impulse response function and that we measure a superposition of those strongly smeared out spikes. The goal is then to estimate the input spikes and the impulse response from a data series. This can be done, using regression with a so-called L0 penalty.

I will present theory and implementation of both models and apply them to real data. This is joint work with Hae-Won Uh (UMC Utrecht).