Analysis of spatial structure of epidermal nerve entry point patterns based on replicated data

Mari Myllymäki, Ioanna Panoutsopoulou and Aila Särkkä

May 9, 2012

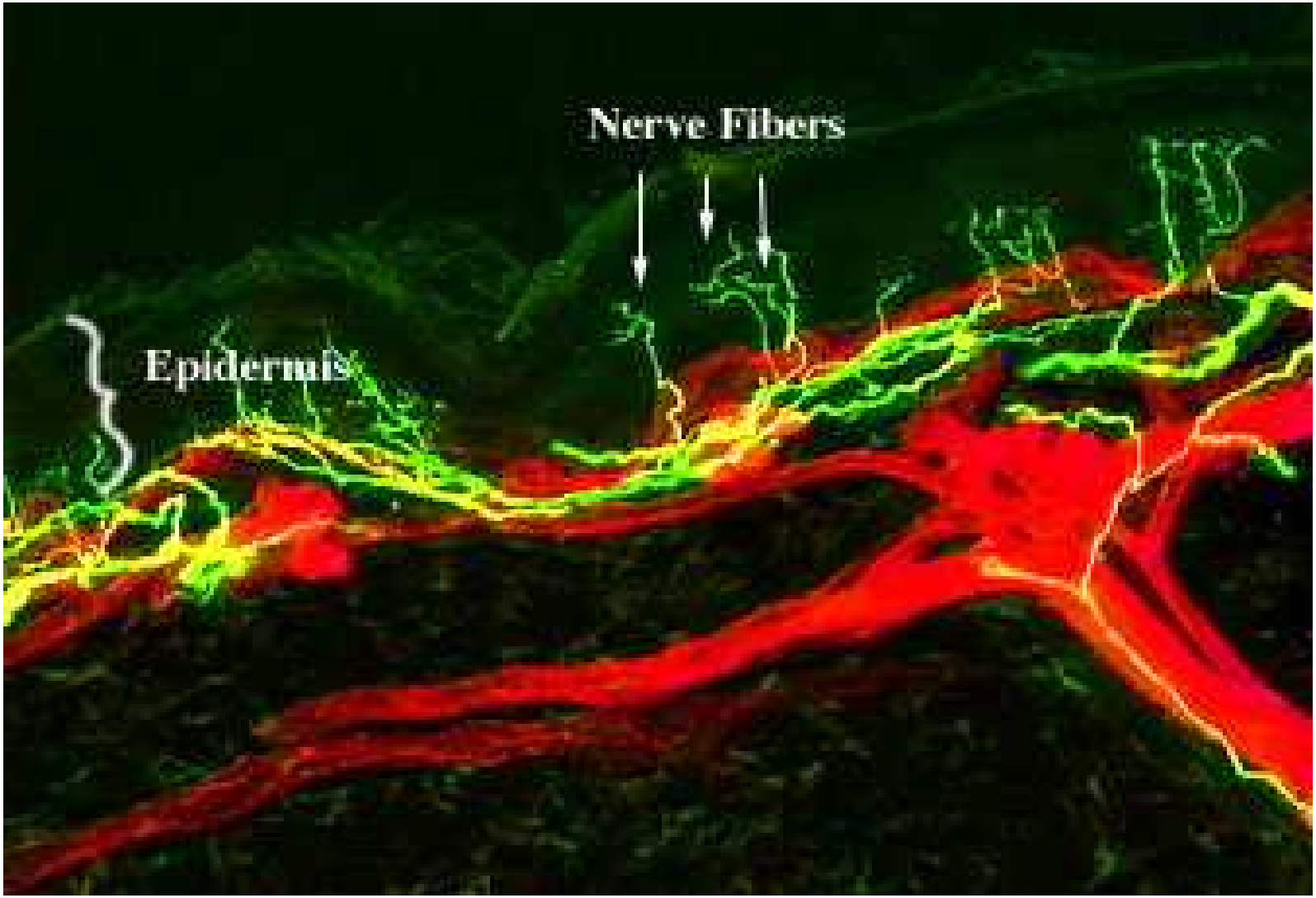
Outline

- What are epidermal nerve fibers?
- Spatial second-order analysis based on replicated data
- Linear mixed models
- Results/recommendations
- Future plans

Epidermal nerve fibers (ENFs)

- ENFs are thin nerve fibers in the epidermis (outmost part of the skin)
- Existence of ENFs has been theorized for over 130 years but still in the late 1980's some doubted their existence
- Kennedy and Wendelschafer-Crabb (1993) first conclusively established the existence of ENFs by confocal microscope studies

Epidermal nerve fibers



Diagnostic value of ENFs

- Kennedy *et al.* (1996)
  1) diminished number of ENFs per surface area
  2) reduced summed length of ENFs per volume

  in subjects with diabetic neuropathy
- Kennedy *et al.* (1999): Nerve fiber loss due to neuropathy does not seem to result in random removal of nerve trunks, rather the remaining nerves seem arranged in clusters

Original question and data

Original question: Is the spatial pattern of (the entry points of) ENFs from subjects with diabetic neuropathy more clustered than the pattern from healthy subjects?

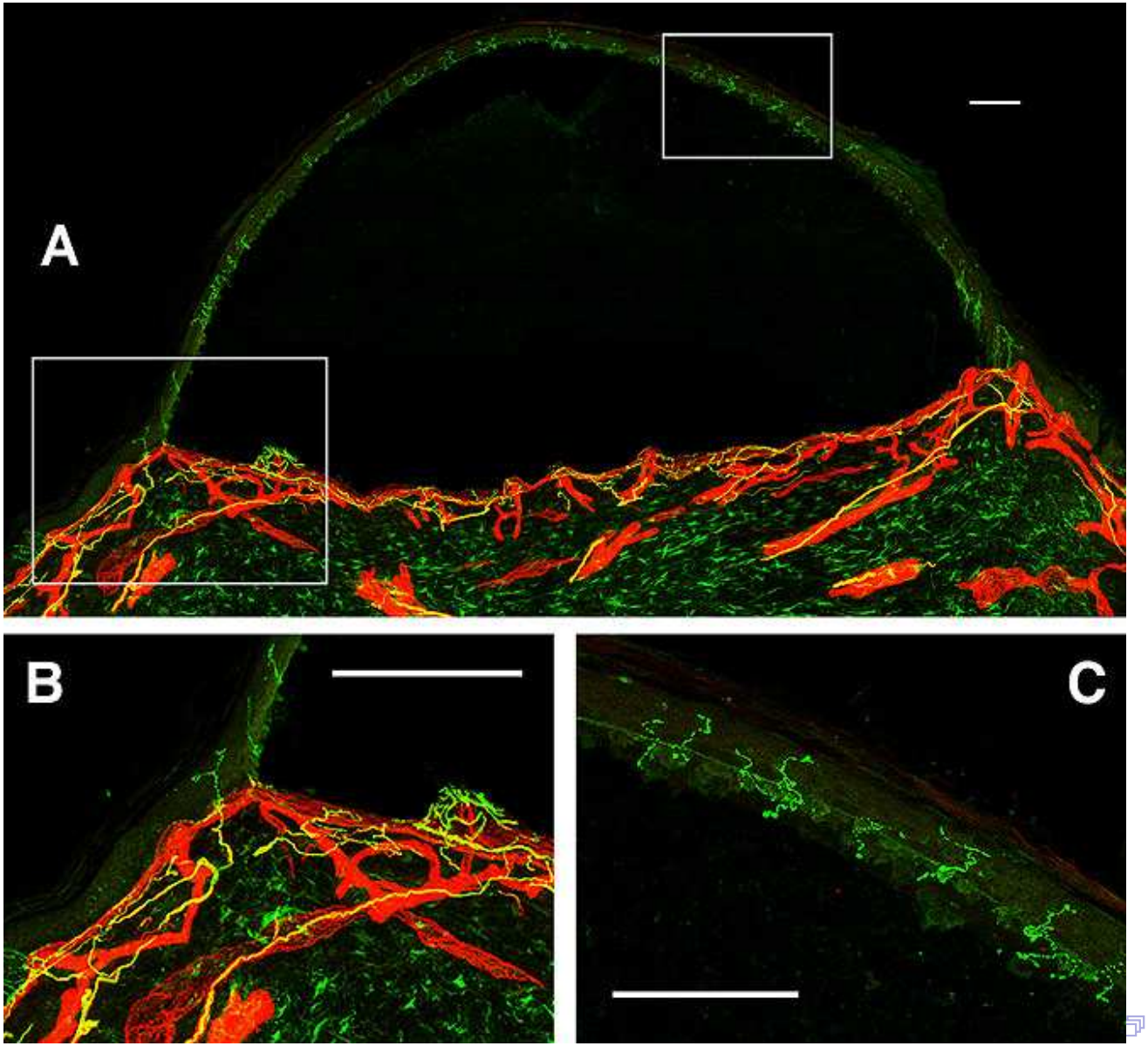
Data: Seven images taken from thighs: one normal (healthy), two with mild, two with moderate and two with severe diabetic neuropathy

Result: By using Ripley's K function, we were able to show that the nerve entry point pattern from subjects with moderate or severe diabetic neuropathy is significantly more clustered than the pattern from the healthy subject

New data from healthy subjects

- 25 healthy volunteers with information on gender, age, and body mass index (BMI)
- Two skin blister specimens were taken from the right calf and from the right foot of each subject
- ENFs were immunostained, imaged confocally, and traced to determine entry point coordinates for each image
- Three to six images (usually four) per each body location of each subject
- Blisters of approximately the size 330 microns by 432 microns (in fact 3D)

Skin blister method

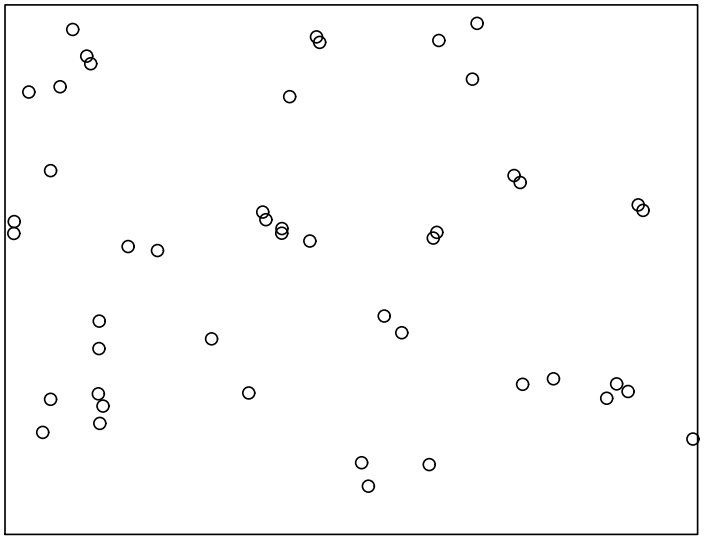


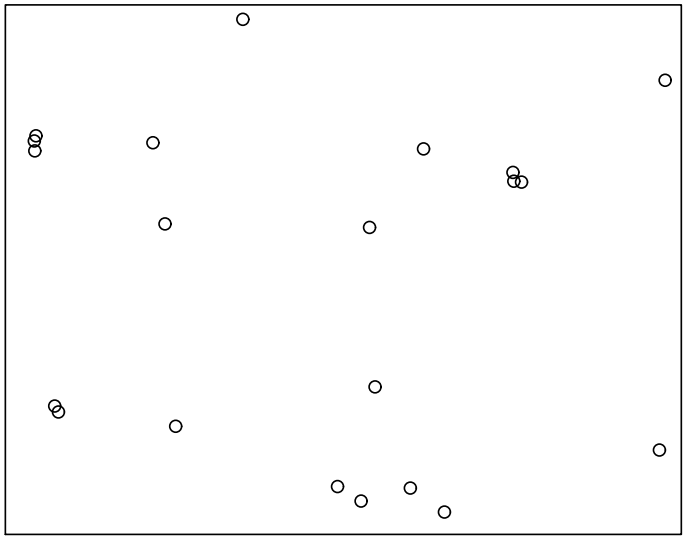
Main question

- How is the spatial pattern of ENFs affected by the body location (calf, foot) and the covariates (gender, age, BMI)?
- The point pattern of ENF entry points is regarded as a realization of a (stationary) spatial point process
- The spatial structure is investigated by using Ripley's K (or L) function

Spatial pattern of the ENF entry points

Foot blister images from two individuals: the pattern on the left has 41 entry points and the one on the right 21 entry points





Pooled K functions

Subject specific mean functions can be estimated by

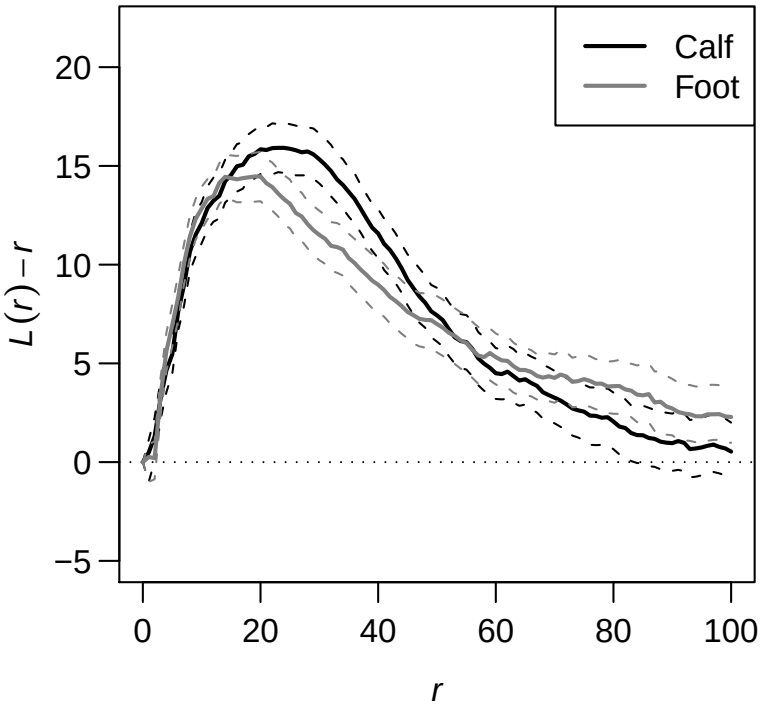
$$\bar{K}_i(r) = \sum_{j=1}^{m_i} w_{ij} \widehat{K}_{ij}(r)$$

where the replicate specific $\widehat{K}_{ij}$ functions are weighted by the squared number of points n_{ij}^2 in the point pattern in question (and $w_{ij} = n_{ij}^2 / \sum_{j=1}^{m_i} n_{ij}^2$)
Group ($\mathcal{G}$) specific (for example older women) mean functions can be estimated by

$$\bar{K}_{\mathcal{G},2}(r) = \frac{1}{n_{\mathcal{G},2}} \sum_{i=1}^{N} \mathbf{1}(i \in \mathcal{G}) n_i^2 \bar{K}_i(r),$$

where $n_{\mathcal{G},2} = \sum_{i=1}^{N} \mathbf{1}(i \in \mathcal{G}) n_i^2$ and $n_i = \sum_{j=1}^{m_i} n_{ij}$

Overall mean L functions for calf and foot



r-wise 95% envelopes constructed by using bootstrap (dashed lines)

Linear mixed models

How to include the covariates?

L function modeled by using linear mixed models usually used to model growth curves. Distance r is the "time variable".

Model for the centered L function

The model for the L function can be written as

$$L_{ijk} - r_k = \mathbf{x}_{ik}\beta + \mathbf{z}\mathbf{u}_j + \epsilon_{ijk},$$

for subjects $i = 1, ..., N$, repetitions $j = 1, ..., m_i$ within subject i and r_k values, $k = 1, ..., 26$.

Here, fixed effects are in β, $\mathbf{X}$ is a known matrix, $\mathbf{u}$ is a vector of random effects and $\mathbf{Z}$ is a known model matrix.

Assumptions

- Errors ϵ_{ijk} independent and $N(0, \sigma^2_{ijk})$, where $\sigma^2_{ijk} = \sigma^2 / n^2_{ij}$
- Random effects normally distributed with mean zero and some covariance structure
- Random effects independent between the experimental units

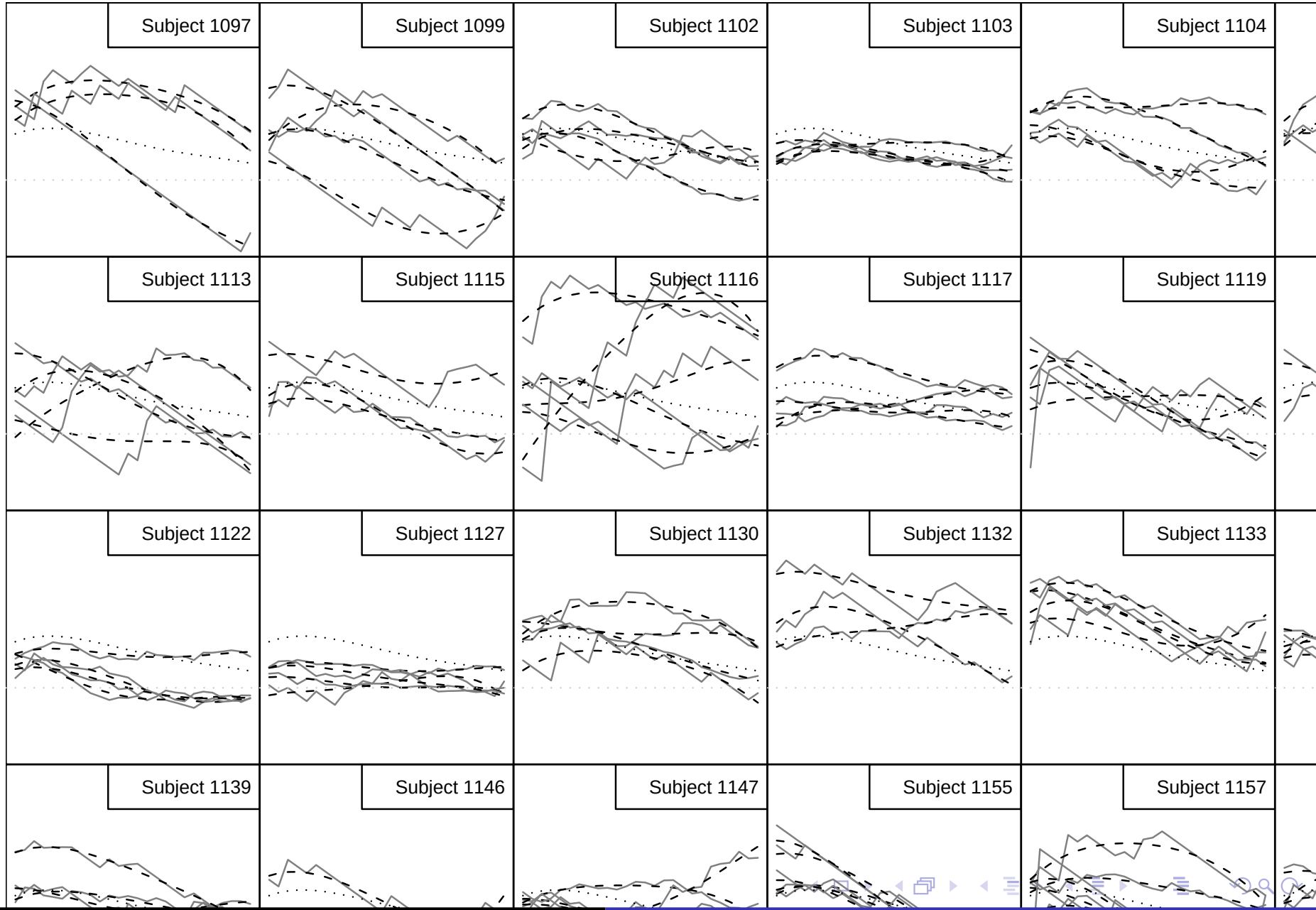
ENF data

- Modelling done separately for foot and calf
- Distance r (varies between 10 and 60 microns) included as a fourth order polynomial
- Fixed effects: age, gender, BMI, r, interactions between the covariates, interaction between the covariates and distance r (all powers)
- Sample specific random effects: intercept and r (all powers)
- Subject specific random effects: intercept and r (all powers)

Results

- The shape of the $L_{ijk} - r_k$ function can be modeled as a forth order polynomial, $10 \leq r_k \leq 60$
- Within subject (sample specific) random effect included (the level and scale of clustering vary within a subject)
- Foot: None of the covariates have significant effect on the curve
- Calf: Covariates have effect
  - Among men clustering more pronounced with low BMI than high BMI (two outliers which most likely affect the results)
  - Older people tend to have more pronounced clustering than younger

Observed and predicted centered L functions



Recommendations

May be preferable to take samples from foot since

- the spatial structure of ENFs not affected by covariates (all covariates easy to measure)
- number of entry points is larger in samples taken from foot than from calf
- in early stages of small fiber neuropathy the ENF density and distribution may be normal on the calf but abnormal on the foot

Future plans: Gaussian process approach

- Gaussian process models are flexible non-parametric models for making inferences about the relationship between covariates and our characteristics (centered L function)
- We do not need to assume linear or any other particular form of dependence between the characteristics and covariates, a priori
- Bayesian approach
- Both base points and end points considered
- New data: subjects with diabetic neuropathy included
  $\longrightarrow$ disease status can be added as a covariate into the model