Interacting cluster point process model for epidermal nerve fibers

Nancy L. Garcia[1]

nancyg@unicamp.br

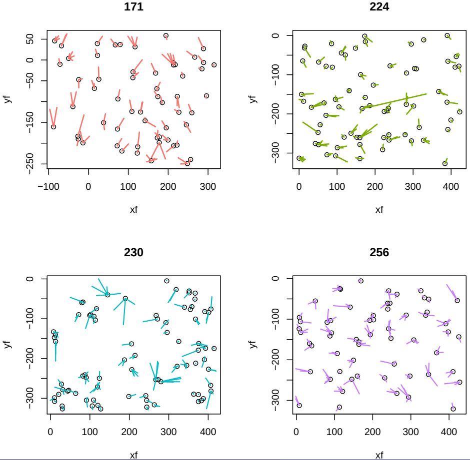
University of Campinas, Brazil

UFRJ – December 6, 2021

[1]Joint work with Peter Guttorp and Guilherme Ludwig.
Supported by FAPESP grants 2016/09390-4 and 2018/09877-6, Brazil.

Epidermal nerve fibers

Thin bundles of fibers branching from root ganglion cells and terminating at all levels of the epidermis, the outmost part of the skin. (Waller et al., 2011; Myllymäki et al., 2012; Olsbo et al., 2013; Andersson et al., 2016).



Thighs of four healthy patients, data from Wendelschafer-Crabb et al. (2005).

Interacting Cluster point processes

- The central idea of this work is to obtain attractive/repulsive cluster point process as invariant measures of birth and death cluster processes.
- Cluster process = marked point process

$$\mathcal{N} = \{(x_1, n_{x_1}), (x_2, n_{x_2}), \ldots, (x_L, n_{x_L})\}$$

$\{x_1, \ldots, x_L\}$ represents the germ process
n_x the cluster process (fingers) associated with germ x.

Inference

Likelihood for point processes

Definition: The likelihood of a realization $n = \{x_1, ..., x_L\}$ of a regular point process on a bounded Borel set $A \subset R^d$ is the local Janossy density.

$$p_L \, j_L(x_1, ..., x_L)$$

$j_L(x_1, ..., x_L) dx_1 ... dx_L \approx$ Given that there are L points in the process, the probability that they are located at $(x_i, x_i + dx_i)$.

Usually too hard.

Likelihood ratio

- Compare our process with a standard process.
- **Example:** Non-homogeneous Poisson process with intensity λ has likelihood ratio wrt the unit Poisson process:

$$\frac{L_{NH}(n|\lambda(\cdot))}{L_H(n|1)} = \frac{\frac{e^{-\int_A \lambda(u)du}}{L!}\prod_{j=1}^{L}\lambda(x_j)}{\frac{e^{-|A|}}{L!}\times\prod_{j=1}^{L}\mathbf{1}_A(x_j).} = e^{-\int_A(\lambda(u)-1)du}\prod_{j=1}^{L}\lambda(x_j)\mathbf{1}_A(x_j).$$

- The independence property characterizes the Poisson process.
- Most of the applications: point processes having interaction between points.

Process characterization

- Point processes with law (restricted to a finite box W) which are absolutely continuous with respect to the law of a homogeneous Poisson point process.
- The set of possible configurations are the same for the interacting process and the Poisson process (locally finite configurations).
- Radon-Nikodym derivative

$$\mu_W(\mathrm{d}n) = \frac{1}{Z_W} e^{-H(N,W)} \mu_W^0(\mathrm{d}n)$$

where $H(n, W)$ is the *energy function* and Z_W is a normalizing constant.

- The Radon-Nikodym derivative: how much more likely is the configuration n in this process than in the Poisson process.

- Ripley (1977) showed that μ_W is the invariant measure of a spatial birth and death process.
- In fact, there is more than one process that has the same invariant measure,

$$\lambda(x, n)e^{-H(n,W)} = \delta(x, n)e^{-H(n\cup\{x\},W)}, \text{ if } n \cup x \in \mathcal{S}$$

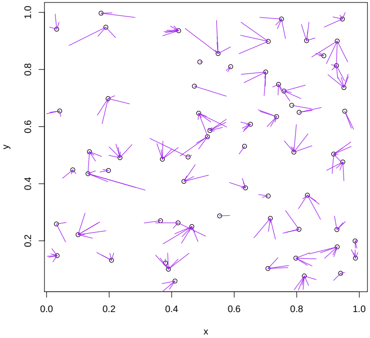
- We can always take $\delta(x, n) = 1$,

  whenever a point is added to the configuration it lives an exponential amount of time independently of the configuration of the process.

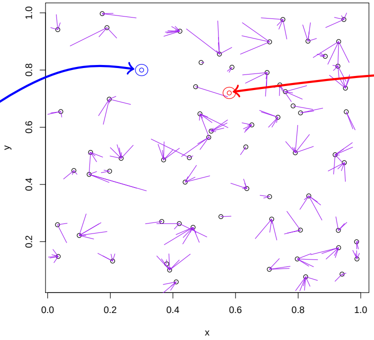
Dynamics of birth

- Germ: Given that at time t the configuration of the process is $\mathcal{N} = \{(x_1, n_{x_1}), (x_2, n_{x_2}), \ldots, (x_L, n_{x_L})\}$, the birth rate for a germ at x is
$$\lambda(x, \mathcal{N}) = \beta \phi^{-\sum_{i=1}^{L} \sum_{z \in n_{x_i}} \mathbf{1}\{d(x,z) < R_c\}}.$$
- We will treat the cluster n_x as marks.

Configuration at time t



$\beta = 70$, $\phi = 2$, $\gamma = 5$, $\sigma = 0.05$,
and $\kappa = 3$ (simulation # 1 setup).

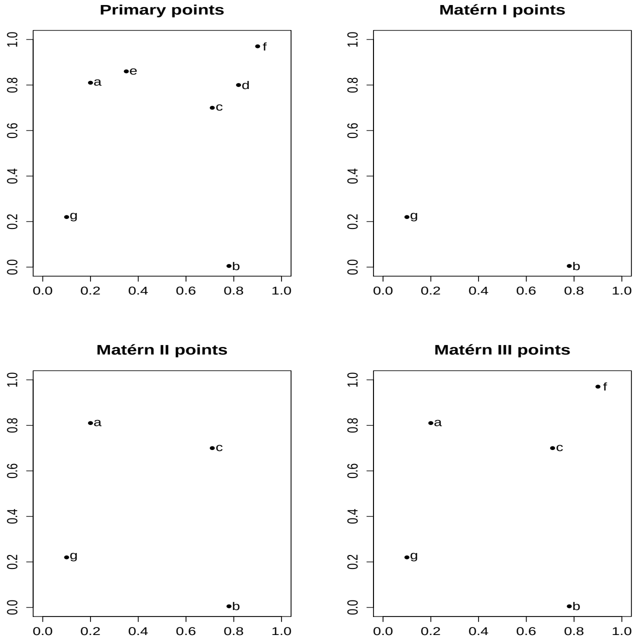


Accepted with $p = 1$.

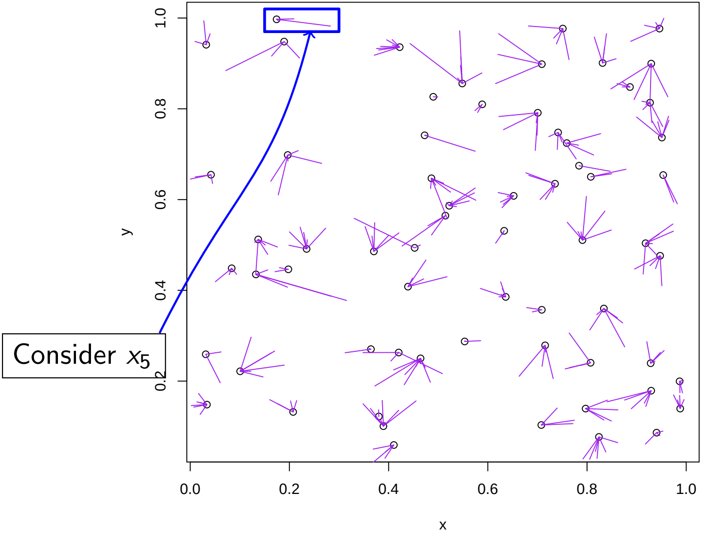
Accepted with $p = 1/\phi$.

Dynamics of birth

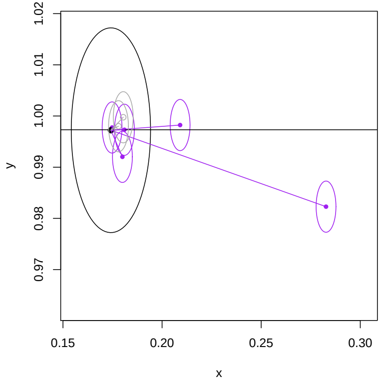
- Germ: Given that at time t the configuration of the process is $\mathcal{N} = \{(x_1, n_{x_1}), (x_2, n_{x_2}), \ldots, (x_L, n_{x_L})\}$, the birth rate for a germ at x is
$$\lambda(x, \mathcal{N}) = \beta\phi^{-\sum_{i=1}^{L}\sum_{z \in n_{x_i}} \mathbf{1}\{d(x,z)<R_c\}}.$$
- Fingers: Given a germ is born at site x, a cluster n_x appears according to a Matérn-III repulsive process with radius R from a inhomogeneous Poisson process with intensity γ and density given by
  1. Direction: von Mises distribution in the circle with center x and dispersion parameter κ,
  2. the distances between the point on the cluster and the center are iid half-normal with variance σ random variables restricted to the set A.



Primary points
Matérn I points
Matérn II points
Matérn III points
a
b
c
d
e
f
g
0.0
0.2
0.4
0.6
0.8
1.0



Consider x_5
y
x
0.0
0.2
0.4
0.6
0.8
1.0



Circles around extremites have radii R, greyed-out fingers were thinned (born within shadow of an older finger)

Bayesian estimation methods for Matérn type-III processes were introduced in Rao et al. (2017)

Dynamics of death

- Germ: Given that at time t the configuration of the process is given by $\mathcal{N} = \{(x_1, n_{x_1}), (x_2, n_{x_2}), \ldots, (x_L, n_{x_L})\}$ the probability that a germ at x_i dies in the interval $(t, t + \Delta_t)$ is approximately Δ_t. That is, all germs die at rate 1.
- Fingers: Once the germ at x dies, all the points in the cluster n_x die.

Likelihood

First: assume random directions $\mathbf{u}$ and the random birth times $\mathbf{t}_1, \ldots, \mathbf{t}_L$ from the Matérn III processes are known.

$$\mathcal{N}_{ut} = \{(x_1, u_1, n_{x_1}, \mathbf{t}_1), (x_2, u_2, n_{x_2}, \mathbf{t}_2), \ldots, (x_L, u_L, n_{x_L}, \mathbf{t}_L)\}.$$

In this case, in order to satisfy the detailed balance conditions

$$\lambda((x, u_x, n_x, \mathbf{t}_x), \mathcal{N}_{ut}) \; = \; \frac{\mu\left(\mathcal{N}_{ut} + \{(x, u_x, n_x, \mathbf{t}_x)\}\right)}{\mu(\mathcal{N}_{ut})},$$

The likelihood ratio with respect to the unit-homogeneous Poisson process is

$$\begin{aligned} L(\beta, \phi, \gamma, \zeta | \mathcal{N}_{ut}) \; &= \; \frac{1}{Z(\beta, \phi, R_c, R)} \beta^L \phi^{-\sum_i \sum_{j \neq i} \sum_{z \in n_{x_j}} \mathbf{1}(d(x_i, z) < R_c)} \\ &\times \prod_{i=1}^{L} \exp(-\gamma) \, \gamma^{\#(n_{x_i})} \prod_{z \in n_{x_i}} \bar{\phi}_{u_i, x_i, \zeta}(z) \exp(\gamma \underbrace{A(n_{x_i}, u_i, \mathbf{t})}_{\text{area of the shadow}}), \end{aligned}$$

Log-likelihood

$$
\begin{aligned}
\ell(\beta,\phi,\gamma,\zeta|\mathcal{N}_{ut}) &= -\log(Z(\beta,\phi,R_c,R)) + L\log(\beta) \\
&\quad + \log(\phi)\left[-\sum_{i}\sum_{j\neq i}\sum_{z\in n_{x_j}} \mathbf{1}(d(x_i,z)<R_c)\right] - L\gamma \\
&\quad + \sum_{i=1}^{L}\prod_{z\in n_{x_i}}\left(\#(n_{x_i})\log(\gamma) + \sum_{z\in n_{x_i}}\log(\bar{\phi}_{x_i,u_i,\zeta}(z)) + \gamma A(n_{x_i},u_i,\mathbf{t})\right).
\end{aligned}
$$

Marginal likelihood

$$
\begin{aligned}
\ell(\beta, \phi, \gamma, \zeta | \mathcal{N}) \quad = \quad & -\log(Z(\beta, \phi, R_c, R)) + L \log(\beta) \\
& + \log(\phi) \left[-\sum_{i} \sum_{j \neq i} \sum_{z \in n_{x_j}} \mathbf{1}(d(x_i, z) < R_c) \right] - L\gamma + \sum_{i=1}^{L} \#(n_{x_i}) \log(\gamma) \\
& + \sum_{i=1}^{L} \log \left[\int_0^{2\pi} \frac{1}{2\pi} \prod_{z \in n_{x_i}} \bar{\phi}_{x_i, u_i, \zeta}(z) \int_{[0,1]^{\#(n_{x_i})}} \exp(\gamma A(n_{x_i}, u_i, \mathbf{t})) d\mathbf{t}_i du_i \right].
\end{aligned}
$$

Marginal likelihood

$$
\begin{aligned}
\ell(\beta,\phi,\gamma,\zeta|\mathcal{N}) &= -\log(Z(\beta,\phi,R_c,R)) + L\log(\beta) \\
&+ \log(\phi)\left[-\sum_i \sum_{j\neq i} \sum_{z\in n_{x_j}} \mathbf{1}(d(x_i,z) < R_c)\right] - L\gamma + \sum_{i=1}^{L} \#(n_{x_i})\log(\gamma) \\
&+ \sum_{i=1}^{L} \log\left[\int_0^{2\pi} \frac{1}{2\pi} \prod_{z\in n_{x_i}} \bar{\phi}_{x_i,u_i,\zeta}(z) \int_{[0,1]^{\#(n_{x_i})}} \exp(\gamma A(n_{x_i},u_i,\mathbf{t}))d\mathbf{t}_i du_i\right].
\end{aligned}
$$

Two hard steps in evaluating this expression: integral of A over $\mathbf{u},\mathbf{t}$ and evaluation of Z.

Some approximations: Z

Following Geyer and Thompson (1992), we consider ℓ as a function of β and ϕ only, it can be written as

$$\ell(\beta, \phi|\mathcal{N}) \; = \; -\log(Z(\beta, \phi)) + L\log(\beta) + \log(\phi)\left[-\sum_i \sum_{j \neq i} \sum_{z \in n_{x_j}} \mathbf{1}(d(x_i, z) < R_c)\right].$$

which belongs to the exponential family with natural parameter vector $\theta = (\log(\beta), \log(\phi))$ and sufficient statistics:

$$T_1(\mathcal{N}) = \#(\mathcal{N}) \quad \text{and} \quad T_2(\mathcal{N}) = \left[-\sum_i \sum_{j \neq i} \sum_{z \in n_{x_j}} \mathbf{1}(d(x_i, z) < R_c)\right].$$

We use Monte Carlo methods to approximate the constant $Z(\beta, \phi)$. In fact, denoting

$$Z(\boldsymbol{\theta}) = \int \exp(<\boldsymbol{\theta}, \mathbf{T}(\mathcal{N})>) d\mu_0(m)$$

where μ_0 is the reference measure. Then,

$$Z(\boldsymbol{\theta}) = Z(\boldsymbol{\psi}) \int \exp(<\boldsymbol{\theta} - \boldsymbol{\psi}, \mathbf{T}(\mathcal{N})>) d\mu_{\boldsymbol{\psi}}(m)$$

where $\mu_{\boldsymbol{\theta}}$ is the probability measure with parameter $\boldsymbol{\psi}$ with respect to μ_0.

Some approximations: Z (cont.)

Therefore, having $M_1, M_2, \ldots, M_J$ an iid sample of the cluster processes generated by $\mu_{\boldsymbol{\psi}}$, it is possible to approximate

$$d(\boldsymbol{\theta}) = \frac{Z(\boldsymbol{\theta})}{Z(\boldsymbol{\psi})}$$

by

$$d_J(\boldsymbol{\theta}) = \frac{1}{J}\sum_{j=1}^{J} \exp(< \boldsymbol{\theta} - \boldsymbol{\psi}, \mathbf{T}(M_j) >).$$

Approximate likelihood

$$\ell(\beta, \phi | \mathcal{N}) \;=\; -\log d_J(\theta) + L\log(\beta) + \log(\phi)\left[-\sum_i \sum_{j \neq i} \sum_{z \in n_{x_j}} \mathbf{1}(d(x_i, z) < R_c)\right]$$

Some approximations: A

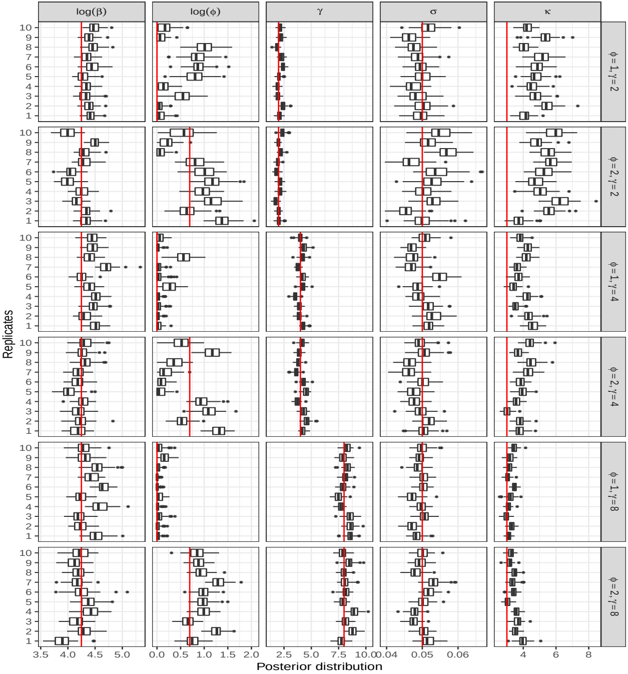
In this case, we follow Rao et al. (2017):

- The missing times are imputted by sampling from thinned events; given initial times t_0, we draw new fingers with rate γ within the region of the shadow of the finger data, with density $\bar{\phi}_{u_i,x_i,\zeta}$.
- Time imputation is made by sampling uniformly on $[0, t_{\min})$, where $t_{\min}$ is the first deletion of one thinned event exclusively within that shadow.
- These steps are iterated for some time.
- The area (volume) of the shadow is estimated by Monte Carlo integration.

Simulation

Simulation experiment

- Using window $W = [0,1]^2$, $R = 0.005$ and $R_c = 0.02$.
- $\beta = 70$, $\phi = 1$ or $\phi = 2$, $\gamma = 2$ or $\gamma = 4$, $\sigma = 0.05$, and $\kappa = 3$.
- Priors:
$$\begin{aligned} \beta &\sim \log N(\log \hat{\beta}, 1), \\ \phi &\sim \Gamma(8/3, 3/4), \\ \gamma &\sim \Gamma(\hat{\gamma}, 1), \\ \sigma^2 &\sim \Gamma^{-1}(\hat{\sigma}^2/10, 10/\hat{\sigma}^2), \\ \kappa &\sim \Gamma(\hat{\kappa}^2/10, 10/\hat{\kappa}^2). \end{aligned}$$
- Initial values for estimates are $\hat{\beta} = L/|W|$, $\hat{\phi} = 2$, $\hat{\gamma} = \max_i\{\#(n_{x_i})\}$, $\hat{\sigma}^2$ is the variance of length of fingers and $\hat{\kappa}$ is the maximum likelihood estimate of the von Mises parameter κ for pooled angle data.
- `mat3c` R package is available for simulation and MCMC fitting, with customizable priors. See `www.github.com/guiludwig/mat3c`.
- Settled on $b = 2000$ burn-in period.
- Replicated $B = 20$ independent experiments on each case.

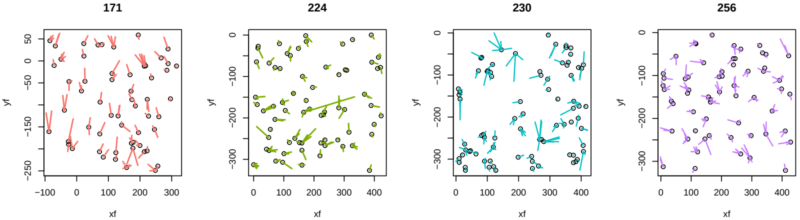

log(β)
log(ϕ)
γ
σ
κ
ϕ = 1, γ = 2
ϕ = 2, γ = 2
ϕ = 1, γ = 4
ϕ = 2, γ = 4
ϕ = 1, γ = 8
ϕ = 2, γ = 8
Replicates
Posterior distribution

Posterior covarage of true parameters

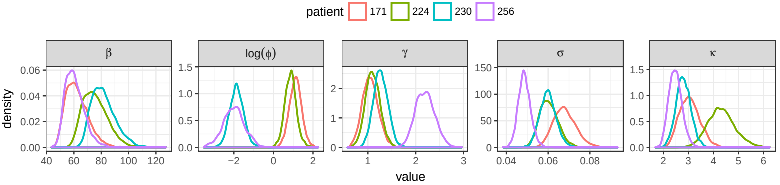
Simulated Scenario	β	$\log(\phi)$	γ	σ	κ
$\phi = 1$, $\gamma = 2$	0.80	0.25	0.90	0.95	0.00
$\phi = 2$, $\gamma = 2$	0.70	0.50	1.00	0.90	0.00
$\phi = 1$, $\gamma = 4$	0.70	0.30	1.00	0.95	0.25
$\phi = 2$, $\gamma = 4$	0.70	0.30	0.95	0.95	0.25

Data Analysis

Data analysis, using priors from simulation



Parameter Estimates



Predictive posteriors - statistics

For example,

- $T_1(\mathcal{N}) = \mathcal{N}(\times) = L$, the total number of base points;
- $T_2(\mathcal{N}) = \sum_i \sum_{\substack{z \in \eta_{x_j} \\ j \neq i}} \{d(x_i, z) < R_c\}$, the number of end points close to a base point by a distance less than R_c;
- $T_3(\mathcal{N}) = L^{-1} \sum_{i=1}^{L} \eta_{x_i}()$, the average size of the cluster;
- $T_6(\mathcal{N}) = L^{-1} \sum_{i=1}^{L} \mathrm{MDE}(\eta_{x_i})$, where MDE is the minimum distance between end points;

1. Sample $(\beta_1, \phi_1, \gamma_1, \sigma_1, \kappa_1), \ldots, (\beta_M, \phi_M, \gamma_M, \sigma_M, \kappa_M)$, from the posterior
2. Generate $n_1, \ldots, n_M$ from the likelihood $L(n|\theta_M)$
3. Compute $\{T_{1,\ell}, \ldots, T_{7,\ell}\}_{\ell=1}^{M}$.
4. Define the Posterior predictive assesment (Bayarri and Berger, 2000) as $(1/M)\sum_{\ell=1}^{M} \mathbf{1}(|T_{k,\ell}| > |T_k|)$.

We discarded burn-in period of $b = 2000$, and sampled $(\beta, \phi, \gamma, \sigma, \kappa)$ at every 100 steps.
Note patients 171 and 224 have $\hat{\phi} \approx 2$ while patients 230 and 256 have $\hat{\phi} \approx 1$.

Posterior predictive assessment

Table: $\mathrm{mat3c}$ = proposed interacting cluster model, NI-$\mathrm{mat3c}$ = interacting cluster model with non-informative priors, $\mathrm{NOC\text{-}like}$ model similar to the NOC model of Olsbo et al. (2013).

Model	Subject	T_1	T_2	T_3	T_4	T_5	T_6	T_7
4*$\mathrm{mat3c}$	171	0.44	0.64	0.60	0.84	0.12	0.50	0.64
	224	0.43	0.38	0.50	0.80	0.19	0.68	0.73
	230	0.64	0.15	0.58	0.04	0.47	0.93	0.88
	256	0.55	0.05	0.55	0.33	0.02	0.29	0.31
4*NI-$\mathrm{mat3c}$	171	0.57	0.52	0.57	0.81	0.05	0.52	0.62
	224	0.38	0.24	0.43	0.81	0.14	0.62	0.81
	230	0.71	0.29	0.67	0.05	0.33	0.95	0.81
	256	0.62	0.10	0.62	0.29	0.00	0.24	0.33
4*$\mathrm{NOC\text{-}like}$	171	0.50	0.77	0.54	0.81	0.07	0.39	0.49
	224	0.54	0.79	0.70	0.78	0.21	0.55	0.73
	230	0.45	0.26	0.55	0.00	0.45	0.96	0.82
	256	0.49	0.11	0.57	0.28	0.02	0.19	0.21

Summary

- We propose a model that is rich in parameters but motivated by biological descriptors of the data.
- Estimation is done via Bayesian methods, which provides credibility intervals and other desirable tools. However, MCMC is somewhat slow (roughly 4 hours to run an experiment with 10000 samples and $J = 10000$ Geyer-Thompson steps).
- Choice of ψ in Geyer-Thompson approximation is quite tricky.
- Irregular parameters R and R_c affect estimates substantially (we examined simulated cases to choose case studies' parameters).
- Future work: examine patients with mild and severe neuropathy, develop model-based diagnostic descriptors.

IMECC - UNICAMP

Thank you!

References I

C. Andersson, P. Guttorp, and A. Särkkä. Discovering early diabetic neuropathy from epidermal nerve fiber patterns. *Statistics in Medicine*, 35(24):4427–4442, 2016.

M. Bayarri and J. O. Berger. P-values for composite null models. *Journal of the American Statistical Association*, 95(452):1127–1142, 2000.

C. J. Geyer and E. A. Thompson. Constrained Monte Carlo maximum likelihood for dependent data. *Journal of the Royal Statistical Society, Series B*, 54: 657–683, 1992.

M. Myllymäki, I. G. Panoutsopoulou, and A. Särkkä. Analysis of spatial structure of epidermal nerve entry point patterns based on replicated data. *Journal of Microscopy*, 247(3):228–239, 2012.

V. Olsbo, M. Myllymäki, L. A. Waller, and A. Särkkä. Development and evaluation of spatial point process models for epidermal nerve fibers. *Mathematical Biosciences*, 243(2):178–189, 2013.

References II

V. Rao, R. P. Adams, and D. D. Dunson. Bayesian inference for Matérn repulsive processes. *Journal of the Royal Statistical Society, Series B*, 79(3):877–897, 2017.

B. D. Ripley. Modelling spatial patterns. *Journal of the Royal Statistical Society, Series B*, pages 172–212, 1977.

L. A. Waller, A. Särkkä, V. Olsbo, M. Myllymäki, I. G. Panoutsopoulou, W. R. Kennedy, and G. Wendelschafer-Crabb. Second-order spatial analysis of epidermal nerve fibers. *Statistics in Medicine*, 30(23):2827–2841, 2011.

G. Wendelschafer-Crabb, D. Walk, S. Foster, and W. R. Kennedy. Epidermal nerve fiber densities in six body locations of normal and diabetic subjects. *Journal of the Peripheral Nervous System*, 10:104, 2005.