Spatial modeling of epidermal nerve fiber patterns

Received: 4 March 2021 | Revised: 17 August 2021 | Accepted: 28 August 2021

DOI: 10.1002/sim.9194

RESEARCH ARTICLE

Statistics in Medicine WILEY

Spatial modeling of epidermal nerve fiber patterns

Konstantinos Konstantinou[1,2] | **Aila Särkkä**[1,2]

[1]Department of Mathematical Sciences, Chalmers University of Technology, Gothenburg, Sweden

[2]Department of Mathematical Sciences, University of Gothenburg, Gothenburg, Sweden

Correspondence
Konstantinos Konstantinou, Department of Mathematical Sciences, Chalmers University of Technology, Gothenburg, Sweden.
Email: konkons@chalmers.se

Peripheral neuropathy is a condition associated with poor nerve functionality. Epidermal nerve fiber (ENF) counts per epidermal surface are dramatically reduced and the two-dimensional (2D) spatial structure of ENFs tends to become more clustered as neuropathy progresses. Therefore, studying the spatial structure of ENFs is essential to fully understand the mechanisms that guide those morphological changes. In this article, we compare ENF patterns of healthy controls and subjects suffering from mild diabetic neuropathy by using suction skin blister specimens obtained from the right foot. Previous analysis of these data has focused on the analysis and modeling of the spatial ENF patterns consisting of the points where the nerves enter the epidermis, base points, and the points where the nerve fibers terminate, end points, projected on a 2D plane, regarding the patterns as realizations of spatial point processes. Here, we include the first branching points, the points where the nerve trees branch for the first time, and model the three-dimensional (3D) patterns consisting of these three types of points. To analyze the patterns, spatial summary statistics are used and a new epidermal active territory that measures the volume in the epidermis that is covered by the individual nerve fibers is constructed. We developed a model for both the 2D and the 3D patterns including the branching points. Also, possible competitive behavior between individual nerves is examined. Our results indicate that changes in the ENFs spatial structure can more easily be detected in the later parts of the ENFs.

KEYWORDS

branching point, competition, end point, epidermal active territory, point process, two-step model

1 | INTRODUCTION

Epidermal nerve fibers (ENFs) are thin sensory fibers in the epidermis, the outermost living layer of the skin, that are responsible for transferring signals of, for example, pain and heat to the brain. Their existence was theorized for many years before it was conclusively established by Kennedy and Wendelschafer-Crabb[1] via confocal microscopy studies. The diagnostic capabilities of the ENFs are discussed in many studies. In particular, Kennedy et al,[2] observed reduced fiber density in subjects suffering from diabetic neuropathy compared to healthy controls, as well as reduced "coverage" of the epidermis by ENFs. It is also well established, that the 2D ENF patterns of subjects with diabetic neuropathy tend to be more clustered than healthy patterns.[3-6]

The spatial pattern of ENF patterns on different body locations has been investigated in several studies by representing the ENF patterns consisting of base points, that is, points where the nerves enter the epidermis, and end points where the nerve fibers terminate, as realizations of spatial point processes. Data from thighs of the patients were considered in a small study.[3] Even though significant differences in the spatial structure of the ENFs between healthy and diabetic patients were found, healthy subjects and patients with mild diabetic neuropathy were indistinguishable. In a larger study,[5] data from feet of healthy subjects and subjects with mild diabetic neuropathy were compared. The results indicated that end points of the ENFs in the mild patients are more clustered than the end points of healthy patients.

In earlier research on the spatial structure of ENFs, base points and end points, are projected on a two-dimensional (2D) plane and the interest has been on the distribution of the ENFs across the skin and the total area of the skin covered by the end point clusters. This is especially interesting since it is the end points that sense, for example, heat and pain and send the signal to the brain. In this article, the main objective is to study the three-dimensional (3D) distribution of ENFs and to construct 3D spatial models for the complete ENF structure represented by the base points, first branching points, and end points which are regarded as realizations of 3D spatial point processes.

The spatial structure of individual nerve trees is investigated by using the lengths and angles of the segments of nerve fibers. The spatial structure of end points and the competitive behavior of nerve trees are studied by using some second-order point process summary statistics. As a part of our analysis, a new tool, called epidermal active territory (EAT), which measures the volume of the epidermis covered by the nerve trees is developed. This generalizes the earlier proposed 2D reactive territory[5] to 3D.

Some 2D point process models for the spatial structure of ENFs can be found in the literature. Using models describing the spatial structure of the base points,[7,8] base points were found to be clustered which suggests that the nerves begin to branch before entering the epidermis. Models for the 2D locations of end points conditioned on the base points have also been developed.[5,6] The main purpose was to investigate how the end points cover the skin and therefore, only the 2D projections of the locations of the base and end points were used. In this article, we study the whole 3D structure of the nerve trees and construct a 3D two-step model. By conditioning on the observed base points, we propose a model for the first branching points and then, a cluster model for the end points having the simulated first branching points as parent points. For this purpose, appropriate 3D angular distributions as well as distributions for the segment lengths and the number of end points per nerve tree were chosen. In addition, we fitted a 2D version of the model to the projected data and concluded that it provides a better fit to the 2D projections compared to the existing models.

The article is organized as follows. In Section 2, we give a description of the epidermal nerve data. In Section 3, we briefly present the theory of point processes and define the EAT. In Section 4, we analyze and compare the ENF data from the healthy subjects and subjects with mild diabetic neuropathy. In Section 5, the two-step model is introduced. Our results are presented in Section 6 and further discussed in Section 7.

2 | DATA

The ENF data set consists of skin samples taken from 32 healthy volunteers and 20 diabetic subjects. The diabetic subjects are divided into three groups based on the severity of the neuropathy. Those groups consist of patients with mild, moderate, and severe diabetic neuropathy. From three to six samples from each subject are obtained using suction induced skin biopsies, where a portion of the epidermis is removed, mounted on a slide and stained for imaging. The nerves are then manually traced using confocal microscopy and the locations of the points, where the points enter the epidermis, branch, and terminate are recorded.[9,10]

The data are in 3D and the samples are in boxes of size 320 μm × 432 μm × z where z varies from 70 to 200 μm depending on the local thickness of the epidermis. Samples are taken from six body parts. The three diabetic groups have substantial differences and cannot be treated as a single group. In particular, it is straightforward to detect the neuropathy in the moderate and severe groups based on the ENF counts alone. However, the differences are not as obvious for the mild diabetic group. Therefore, we here compare ENF samples from healthy subjects and subjects with mild neuropathy. We concentrate on data taken from feet of the subjects as the nerve loss in this body part starts in an early stage.[5] In total, 28 skin samples from 8 subjects with mild diabetic neuropathy and 112 samples from 32 healthy controls are considered in our analysis. Also, non-spatial covariates such as age, gender, and BMI for every subject are available. However, according to Myllymaki et al,[11] it is not clear if these covariates influence the spatial structure of the ENFs and hence, they are not considered in our analysis.

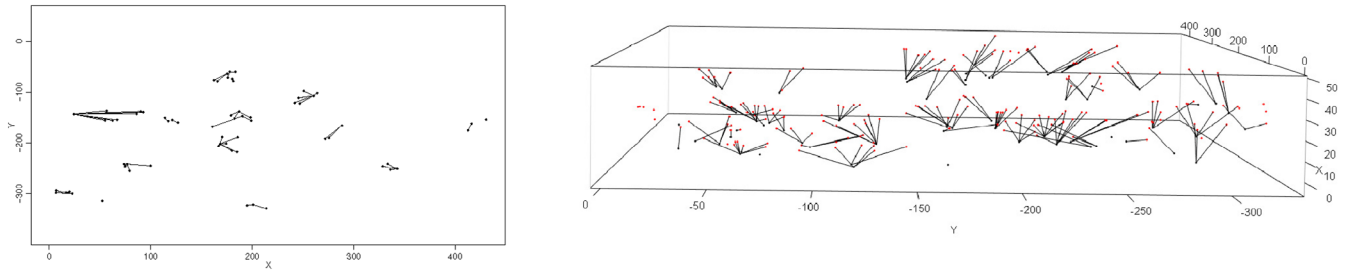

FIGURE 1 Structure of the nerve trees in 2D (left) and in 3D (right). Red points correspond to the locations of the end points and black points to their corresponding base points [Colour figure can be viewed at wileyonlinelibrary.com]

Our analysis below is based on the locations of base points, the locations where the nerve tree enters the epidermis, end points, the locations where the nerve fibers terminate, and first branching points, the locations where the nerve tree branches for the first time. From now on, we call the first branching points as branching points. An example of an ENF sample is illustrated in Figure 1. Only the locations of base points and end points are displayed.

3 | METHODS FOR SPATIAL POINT PROCESSES

We regard the ENF samples consisting of base points, branching points, and end points as realizations of spatial point processes. In this section, we define spatial point processes and introduce some summary statistics for spatial point processes. For more rigorous treatment of the subject, the reader is referred to the literature.[12-15] The definitions and notations given here mainly follow the book by Illian et al.[12] Throughout this work, $\mathbb{R}^3$ and $\mathcal{B}(\mathbb{R}^3)$ denote the 3D Euclidean space and its corresponding Borel sets, respectively. The indicator function is denoted by $\mathbb{I}\{\cdot\}$.

3.1 | Spatial point processes

Spatial point processes describe the spatial arrangement of objects under study. A spatial point process X is defined as the set of locations in a spatial domain D where an event of interest occurred. Even though the point process exists in the whole spatial domain, it is usually observed in an observation window $W \subset D$. We refer to this as a realization of the point process in the observation window W. In this work, the locations, where the nerves enter the epidermis, branch, and terminate, are treated as realizations of spatial point processes in a 3D box $W \subset \mathbb{R}^3$. We assume that the point processes are *locally finite*, that is for every bounded subset $B \in \mathcal{B}(\mathbb{R}^3)$, the number of points of the process that lie in B, $N_X(B)$, is finite. We further assume that the point processes are *simple*, that is at any location at most one point of the process is observed. We also assume that the point processes are stationary (translation invariant). In many applications, specific features of the points of the process might be of interest. For instance, in our study the type of the point: base, branching, or end, or the number of branching layers might be of interest. In order to analyze such information, those features are attached to the points as marks. Hence, the process $Y = \{(x, m_x) : x \in X\}$ is called a *marked point process* with a mark $m_x \in M$, where M is an appropriate mark space, attached to the spatial location $x \in X$.

3.2 | Summary statistics

Below, we recall two summary statistics to describe a point process, Ripley's K function to describe the second-order property and mark correlation to describe the correlation between the marks.

3.2.1 | Ripley's K function

Ripley's K function[16] is defined as follows. For a stationary and isotropic point process in $\mathbb{R}^3$, $\lambda K(r)$ gives the expected number of further points of the process in the sphere with radius r centered at an arbitrary point x of the process. For example,

in the case of the Poisson point process in $\mathbb{R}^3$,

$$K(r) = \frac{4}{3}\pi r^3, \quad r \geq 0. \tag{1}$$

Since Poisson point process corresponds to the complete spatial randomness, it is used as a reference model. In particular, the structure of a point process in $\mathbb{R}^3$ over a distance r is regular if $K(r) < \frac{4}{3}\pi r^3$ and clustered if $K(r) > \frac{4}{3}\pi r^3$. As the points of the process might have r-close neighbors outside the observation window, a naive estimate for the K function is biased. Hence, an edge correction term is introduced to make the estimate of the K function unbiased. An unbiased estimate of the K function can be computed as

$$\hat{K}(r) = \frac{1}{\hat{\lambda} n} \sum_{i=1}^{n} \sum_{j \neq i} w(x_i, x_j) \mathbb{I}\{||x_i - x_j|| \leq r\}, \quad r \geq 0, \tag{2}$$

where n is the number of points observed in W, $\hat{\lambda}$ is an estimate of the process intensity, that is, the mean number of points per unit volume, $||x_i - x_j||$ denotes the Euclidean distance between the points x_i and x_j, and $w(x_i, w_j)$ is an edge correction term. In this work, we use the variance stabilized and centered variant of the K function in $\mathbb{R}^3$, defined by

$$L(r) - r = \sqrt[3]{\frac{K(r)}{\frac{4\pi}{3}}} - r, \quad r \geq 0, \tag{3}$$

which for a Poisson process equals zero.[12] We use the translation correction having $w(x_i, x_j) = \frac{1}{|W_{x_i} \cap W_{x_j}|}$, where W_{x_i} is the translated window $W_{x_i} = \{z + x_i : z \in W\}$ and $|\cdot|$ denotes the 3D Lebesgue measure. The K function can be extended for multi-type point processes. If X_a and X_b are realizations of two point processes in W, an estimate for the bivariate $K_{a,b}(r)$ function can be computed as

$$\hat{K_{a,b}}(r) = \frac{1}{\hat{\lambda}_b \hat{\lambda}_a |W|} \sum_{i=1}^{n_a} \sum_{j=1}^{n_b} w(x_{a_i}, x_{b_j}) \mathbb{I}\{0 < ||x_{a_i} - x_{b_j}|| \leq r\}, \quad r \geq 0, \tag{4}$$

where λ_a, λ_b are estimates for the intensities of X_a and X_b respectively. For a stationary process, $\lambda_b K_{a,b}(r)$ gives the expected number of further points of the process X_b in the sphere with radius r centered at an arbitrary point x of the process X_a. When $X_a = X_b$, the bivariate $K_{a,b}(r)$ function is equivalent to Ripley's K function.

Ripley's K and L functions are suitable for characterizing the spatial structure of one point pattern. However, in many applications, including ours, the average spatial structure of replicated point patterns might be of interest. Our data are hierarchically structured into groups (normal and mild), subjects within the groups, and samples from the subjects, and we are especially interested in comparing the spatial structure of the ENFs between the two groups. Using the aforementioned methodology, samplewise summary functions $K_{ij}(r)$ for sample $j \in \{1, \ldots, m_i\}$ of subject i can be estimated. Then, subject specific $\overline{K}_i(r)$ functions can be obtained as a weighted mean of the $K_{ij}(r)$ functions for all subjects $i \in \{1, \ldots, N\}$ as

$$\overline{K}_i(r) = \sum_{j=1}^{m_i} w_{ij} \hat{K}_{ij}(r). \tag{5}$$

Finally, the subjectwise $\overline{K}_i(r)$ functions are weighted to obtain the groupwise $\overline{K}_g(r)$ function for the group g.

$$\overline{K}_g(r) = \sum_{i=1}^{N} w_i \overline{K}_i(r). \tag{6}$$

To compute the subjectwise and groupwise K functions, square point number weights are chosen as point patterns from different samples and subjects cannot be assumed to have the same intensity.[17] Let n_{ij} denote the number of points

in sample j of subject i, and let $n_i = \sum_{j=1}^{m_i} n_{ij}$ be the total number of points in the samples from subject $i \in \{1, \dots, N\}$. Then the square point number weights for the groupwise $\overline{K}_g(r)$ and subjectwise $\overline{K}_i(r)$ are given by

$$w_i = \frac{n_i^2}{\sum_{k=1}^{N} n_k^2}, \quad w_{ij} = \frac{n_{ij}^2}{\sum_{k=1}^{m_i} n_{ik}^2}. \tag{7}$$

An explicit and detailed description of this procedure can be found in the literature.[4,5,18]

3.2.2 | Mark correlation function

Mark correlation functions can be used to capture possible spatial dependencies between the marks of a marked point process. A vital part of the mark correlation function is the so called test function $f(m_1, m_2)$. Here, we use the commonly used test function $f(m_1, m_2) = m_1 m_2$ and hence the mark correlation function is given by

$$K_{mm}(r) = \frac{E_{o,r}(m_o \cdot m_{\mathbf{r}})}{\mu^2}, \quad r \geq 0, \tag{8}$$

where μ is the mean mark. The expectation $E_{o,r}$ is a conditional expectation given that there is a point of the process in the origin and a point distance r away. If there is no correlation between the marks, the mark correlation function equals one. Values less than one indicate negative correlation and values larger than one positive correlation. Permutation tests can be used to study whether the marks are randomly labeled. They are performed in an iterative fashion, where in each iteration the marks of the process are randomly permuted and the mark correlation function for the permuted point pattern is computed. By definition, this procedure creates marked point processes with randomly labeled marks, which are used to construct envelopes[19] for the mark correlation function of the point process under random labeling.

3.3 | Epidermal active territory

Two-dimensional reactive territories[5] approximate the area of the skin covered by the nerve trees. The reactive territory of a nerve tree is defined as the convex hull determined by the locations of the projected end points and base points belonging to the same nerve tree. In this study, a 3D counterpart of the reactive territory, called EAT, is introduced.

The EAT is designed to capture the 3D proportion of the epidermis covered by nerve fibers, as the volume of the 3D shape created by the Delaunay triangulation[12,13,20] of the points in the set B. In our case, B is the set containing the 3D locations of end and base points of a nerve tree. Note that the EATs do not define a partition of the 3D space as they can overlap and do not necessarily cover the whole space. It is also important to note that as it is defined the EAT function produces positive values only when the set of distinct points has cardinality $\text{card}(B) \geq 4$, given that those points do not lie in a plane. Since we want all the trees to have a positive EAT, we slightly modified the definition when $\text{card}(B) \leq 3$. In the case $\text{card}(B) = 3$, that is, when there are only two end points and one base point in the nerve tree, we compute the volume of the shape created by the Delaunay triangulation of the set $B \bigcup \{z_1, z_2\}$, where

$$z_i = \frac{x_1 + x_2}{2} + (-1)^i \begin{bmatrix} 0 \\ 0 \\ \frac{||x_1 - x_2||}{4} \end{bmatrix}, \quad i \in \{1, 2\}, \tag{9}$$

where x_1 and x_2 are the end point locations. The motivation for this is to create 3D shapes that approximate the epidermal volume covered by the nerve tree. In the case when $\text{card}(B) = 2$, that is, when there is only one end point x_1 and one base point x_2 in a nerve tree we set $E = ||x_1 - x_2||$. Therefore, the EATs of the two and three point nerve trees are negligible compared to the EATs of larger trees. Examples of the shapes created in the cases when $\text{card}(B) = 3$ and $\text{card}(B) = 5$ are displayed in Figure 2.

(i) card(B) = 3 (ii) card(B) = 5

FIGURE 2 Examples of the shapes of EATs [Colour figure can be viewed at wileyonlinelibrary.com]

3.4 | Shift plot

The shift function[21] is a statistical tool for the graphical comparison of two random variables $X \sim F$ and $Y \sim G$. The shift function is defined as the function $\Delta(x)$ such as $F(x) = G(x + \Delta(x))$ which can be rewritten as $\Delta(x) = G^{-1}(F(x)) - x$. Therefore, $\Delta(\cdot)$ expresses the amount of "shift" required so that X and Y coincide. If X and Y have the same distribution, then $\Delta(x) \equiv 0$. The shift function is closely related to the quantile quantile (qq) plot. When comparing two random samples with a qq-plot, the quantiles of one sample are plotted against the quantiles of the other. If they have the same distribution then the points are expected to lie on the line $y = x$. The shift function $\Delta(x)$ is the smallest distance from the points in the qq-plot to the line $y = x$.

Simultaneous 95% confidence bands for $\Delta(x)$ based on the Kolmogorov-Smirnov statistic can be constructed. Hence, if the line $y = 0$ lies within the confidence bands of estimated shift function $\hat{\Delta}(x)$ then F and G are statistically indistinguishable. An advantage of this tool is that if F and G differ, information on how the distributions differ can be obtained from visual inspection of the shift plot.

4 | PRELIMINARY ANALYSIS OF THE DATA

In this section, we perform some preliminary analysis of the ENF data. First, we investigate the individual nerve trees and then, the spatial arrangement of the branching and end point locations.

4.1 | Characteristics of trees

Here, the length and direction of the first segments and the segments connecting the branching point with its end points are studied more carefully both within and between the two groups.

4.1.1 | Tree segment lengths

The lengths of the first segments and the segments connecting the first branching point with its end points are compared within the groups. The first segments tend to be on average longer than the later segments, see Table A1 in Appendix, but the difference is not significant. However, when comparing the distributions of the lengths using shift plots as shown for the healthy group in Figure 3 (left), significant differences can be found. In particular, the first branching points tend to be further away from their base points than from their end points. The respective results for the mild group are similar and are shown in the middle part of Figure 3.

Furthermore, we compared the tree segment lengths between the groups. Based on the mean values and standard deviations shown in Table A1, the groups do not differ significantly from each other. A shift plot comparing the whole distribution of this distance shown in Figure 3 (right) reveals that mild subjects seem to have end points closer to the branching point than healthy subjects and the difference is significant.

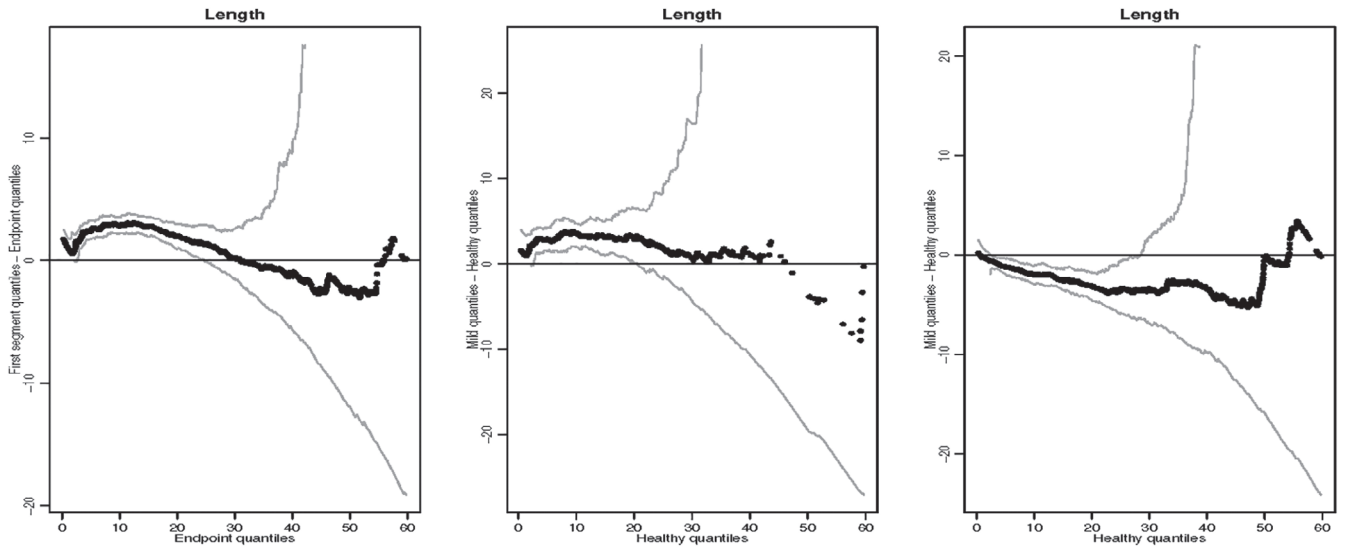

FIGURE 3 Shift plot of the empirical tree segment lengths. Quantities for the first segment (y axis) are compared to the corresponding quantities for the segments connecting the first branching points and their end points (x axis) using data from the healthy (left) and mild (middle) group. Shift plot of the segment lengths (right) of the segments connecting the first branching point with end points comparing the two groups

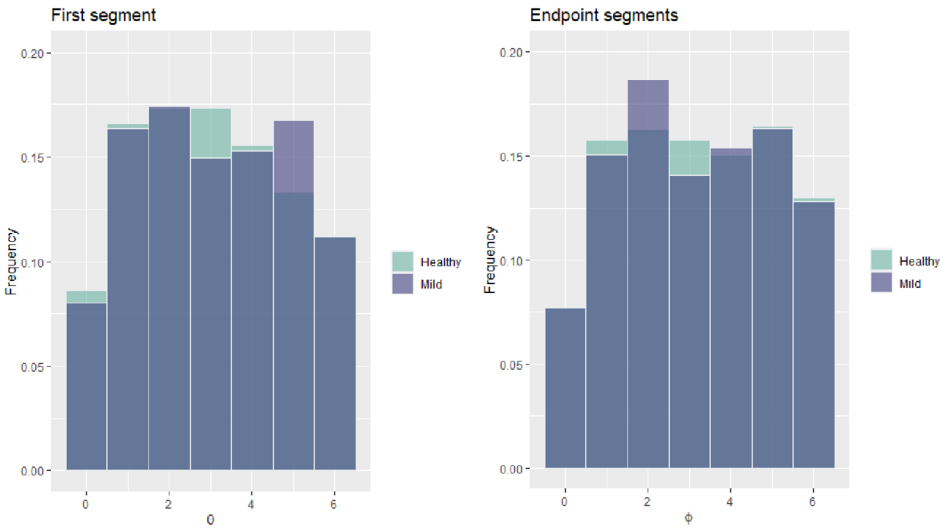

FIGURE 4 Distribution of the planar angles of the first (left) and later (right) segments in the healthy and mild groups [Colour figure can be viewed at wileyonlinelibrary.com]

4.1.2 | Tree segment angles

First, the angular distribution of the first segment and the later segments is compared within the groups. The planar angle for both the first segment and the later segments is approximately uniform on the interval $[0, 2\pi]$ for both groups as displayed in Figure 4. However, significant differences are found between the colatitude of the first and later segments of the healthy group as can be seen in the shift plot in Figure 5 (left). More specifically, the first segment grows more vertically than the further segments. Results for the mild group are similar and shown in the middle part of Figure 5.

Further, the angle in the z direction of the segments connecting the branching point and its end points is compared between the two groups. Our findings suggest that the nerve trees tend to grow from the first branching point to the end points more vertically in the mild group than in the healthy group. This result is displayed in Figure 5

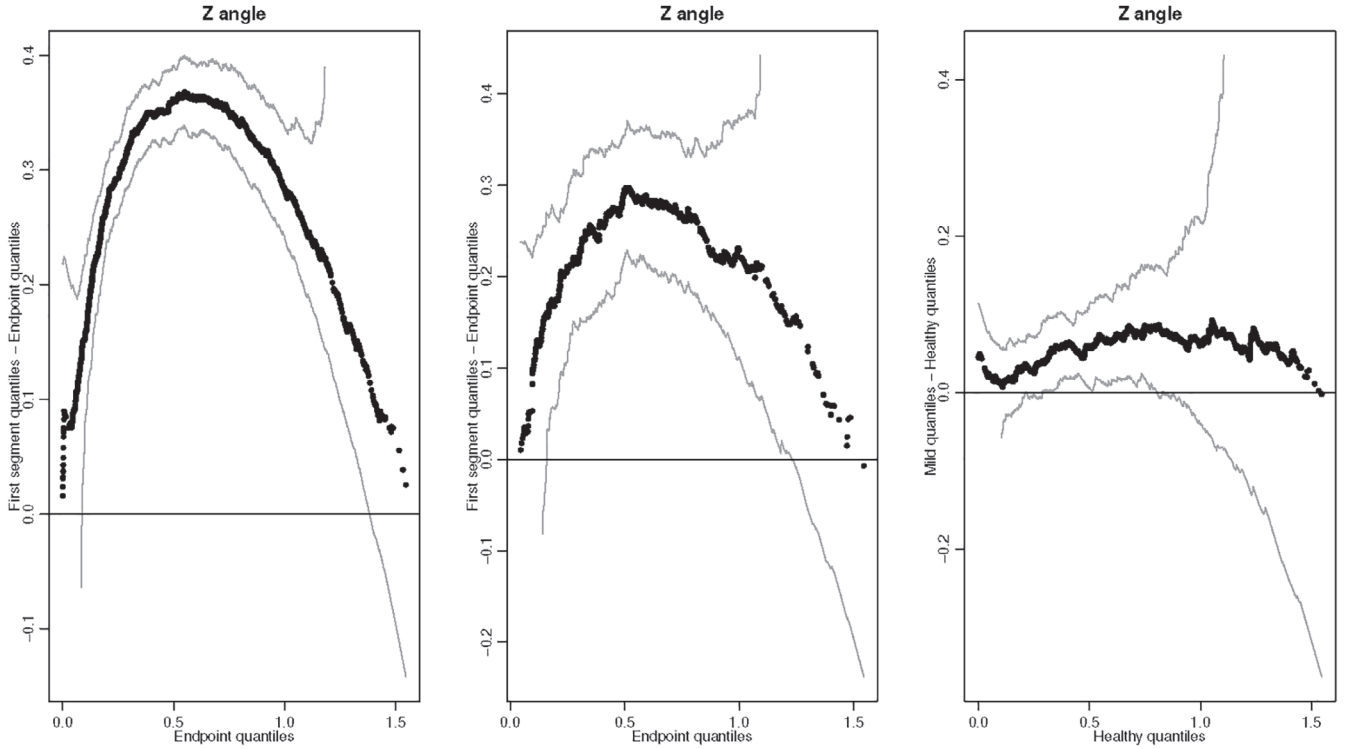

FIGURE 5 Shift plot of the empirical angles between the z axis and the xy plane where the quantities for the first segment (y axis) are compared to the corresponding quantities for the segments connecting the branching points and their end points (x axis) using data from the healthy (left) and mild (middle) group. Shift plot of the z angles (right) of the segments connecting the first branching points with end points comparing the two groups

(right). Therefore, even if diminishing ENF counts cannot be observed at this stage of the neuropathy (see Table A1 in Appendix) the individual nerve tree skin coverage is decreased. No differences between the groups concerning the z angles of the first segment were found (results not shown here). This indicates that, in the early stage of peripheral neuropathy, changes in the spatial structure of an individual nerve tree can more easily be identified in later segments.

4.2 | Epidermal active territories

In this section, we compare the total volume of the EATs in the two groups. As displayed in the boxplot and shiftplot in Figure 6, the volume of the epidermis covered by the nerve trees is significantly larger in samples from the healthy group than samples from the mild group. This result was expected, as it is well established that peripheral neuropathy reduces the nerve tree density as well as the total coverage of the epidermis by ENFs,[2] see also Table A1 in Appendix. This result, however, enhances the validity of the EAT as a tool for approximating the volume in the epidermis covered by nerve trees.

A well-established phenomenon in forestry applications is that large trees discourage competition in their neighborhood by consuming the majority of available resources. As a result, small trees are more likely than larger trees to appear close to large trees. As our application consists of nerve trees living in the epidermis, we examine if such competitive behavior can be justified. To capture possible dependencies between the marks, that is, the size of EATs, the mark correlation function is utilized. Here, the EAT volumes are attached to the base points as marks and the correlation between the marks is investigated. The resulting mark correlation functions given the test function $f(m_1, m_2) = m_1 m_2$ and the corresponding envelopes from the random labeling test are presented in Figure 7. We can conclude that no spatial correlation is detected between the marks. Here we note that other characteristics of the nerve trees, such as the maximum tree order, that is, the number of branches that needs to be crossed from the base point to reach the farthest end point, can also be attached to the base point process as marks to investigate possible correlations.

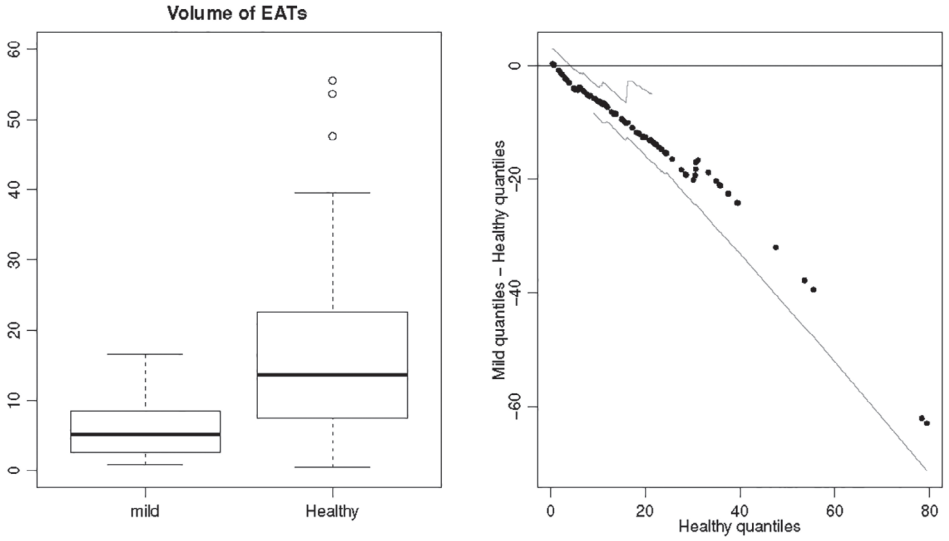

FIGURE 6 Boxplot (left) and shift plot (right) comparing the volumes of the epidermal active territories, in the scale of thousand μm^3, of the normal and mild groups

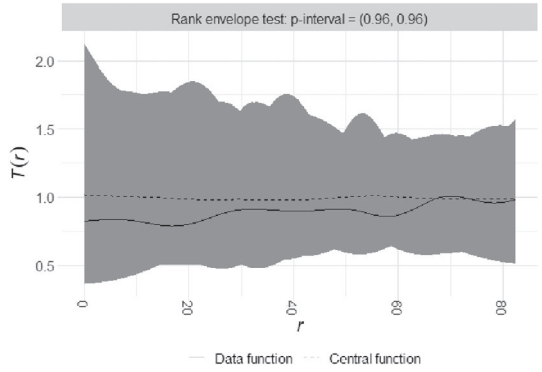

FIGURE 7 Mark correlation function with global envelopes based on random labeling for the EATs of the nerve trees in the normal group

4.3 | Spatial pattern of branching points and end points

It has earlier been reported that the 2D spatial patterns of the base points and the end points of the same data are clustered.[5-8] The clustering of end points seems quite natural and the clustering of base point can be explained by the nerve fibers branching prior to entering the epidermis. To fully understand the mechanism guiding the nerve fibers within the epidermis, we analyze the 3D spatial structure of the first branching points and the end points using Ripley's K function. In all cases below, the K function was first estimated for each sample, then the estimates from a subject were pooled to obtain an estimate for the subject, and finally, the subjectwise estimates were pooled to get the groupwise envelopes (as explained in Section 3.2.1). The transformed and centered versions of the estimates (Formula 3) with pointwise 95% bootstrap envelopes[4,17] are then presented in the figures below. We initially point out that the first branching points are a better choice as the end point cluster centers than the base points, and then describe their spatial structure. As the branching points have not been studied earlier, we study both their 2D and 3D structure. Finally, we compare the spatial arrangement of the 3D end point clusters between the two groups.

4.3.1 | Second order analysis of the first branching point patterns

Some nerve trees in terms of their base points (dots), branching points (crosses), and end points (triangles) are shown in Figure 8. In addition, the geometrical cluster centers (+ signs) of each end point cluster are shown. It can be seen that

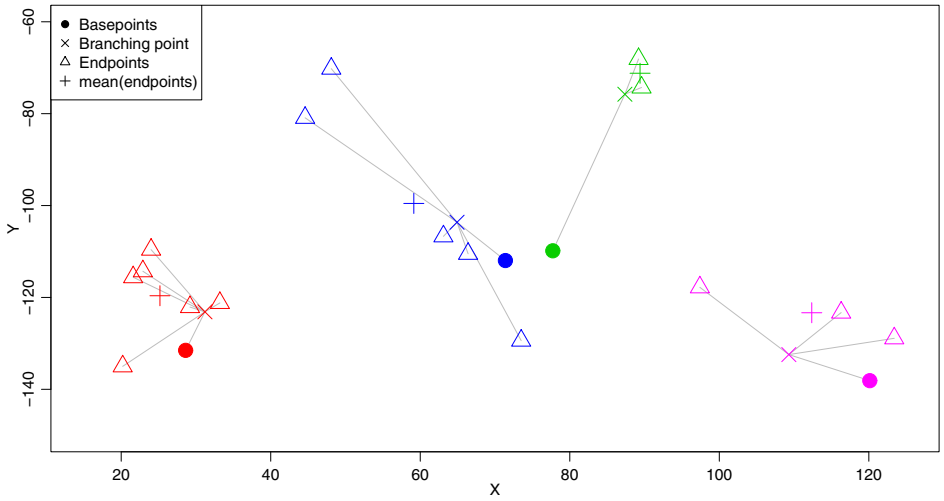

FIGURE 8 Example of base points (dots), branching points (x), end points (triangles), and the mean of the end point clusters (+). Different colors represent different nerve trees [Colour figure can be viewed at wileyonlinelibrary.com]

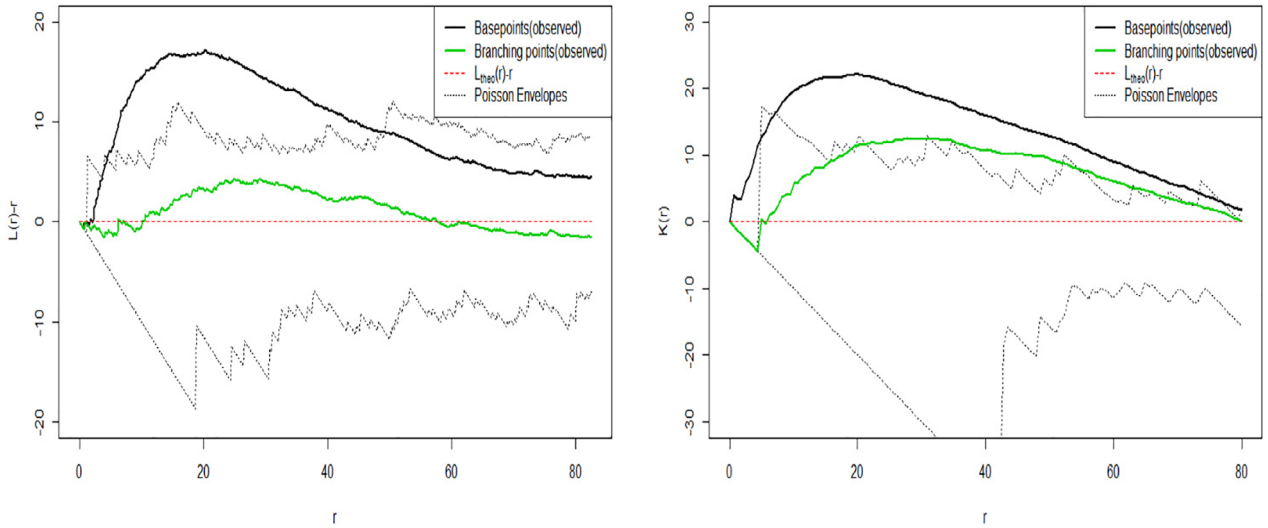

FIGURE 9 Ripley's overall $L(r) - r$ function for the base points (black) and the branching point patterns (green) of the healthy group compared with the theoretical value of a Poisson point process (red) with 95% pointwise envelopes (dashed lines). The results are shown for the planar point patterns in the left and for the 3D point patterns in the right [Colour figure can be viewed at wileyonlinelibrary.com]

the branching points are closer to the cluster centers than the base points are and therefore, they would be a better choice as the parent points for the end point clusters. Therefore, we include the branching points in our model introduced in Section 5. Here, we analyze them in terms of their second order properties in 2D and 3D (see Figure 9). We can see that the branching point patterns are less clustered than the base point patterns. In fact, our results suggest that the branching point patterns are completely spatially random in $\mathbb{R}^2$, even though their parent process (base points), are clustered. This finding is in agreement with neurologists' belief that the nerve fibers are guided toward open space in order to cover as much area as possible. This trend is observed in both the healthy and the mild neuropathy groups (results shown only for the healthy group).

4.3.2 | Second order analysis of the end point patterns

The estimated 3D centered L functions for the end point patterns can be seen in Figure 10. In both the healthy and mild groups, the end points are clearly clustered. At small distances, the L functions are almost identical in the two

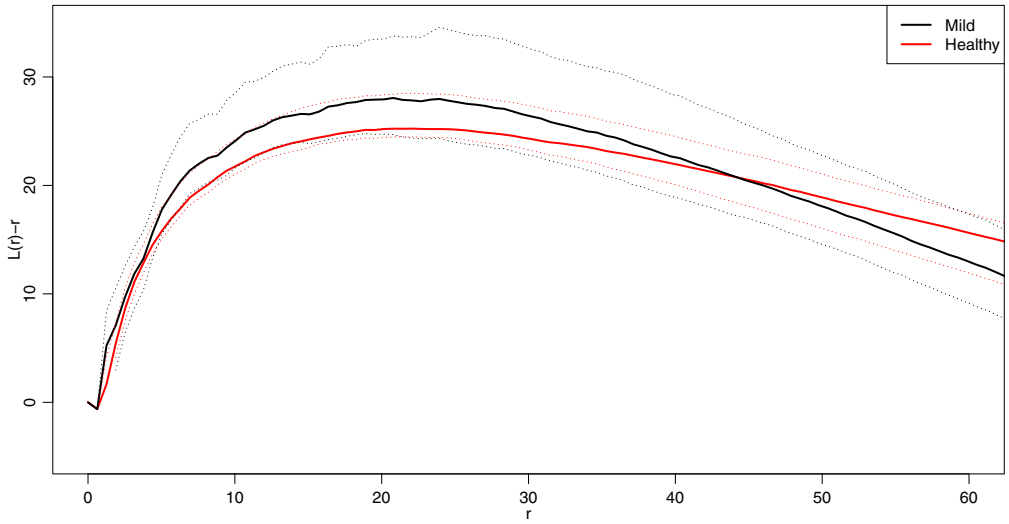

FIGURE 10 Overall mean $L(r) - r$ function of the point patterns of the two groups (thick solid lines) with 95% pointwise bootstrap envelopes (dashed lines) [Colour figure can be viewed at wileyonlinelibrary.com]

groups, while at distances $r \geq 40$, the healthy patterns appear to be more clustered. However, the bootstrap envelopes for the mean L functions did not show significant differences between the groups. We even performed a nonparametric statistical test for distances $r \leq 100$ as described in Diggle et al,[17] where the null hypothesis is that the mean K functions are the same in the two groups. To create a realization of the approximate sampling distribution of the test statistic we repeated the procedure $N = 9999$ times and obtained an approximate p-value of 0.63 which indicates that there is not enough evidence to reject the null hypothesis, that the mean K functions are the same in both groups.

5 | MODELING THE ENF SPATIAL STRUCTURE

Previous research has focused on modeling the spatial arrangement of the termination points of the nerve trees conditioned on their corresponding base points using 2D projections of the points on the skin. Examples of such models are the non-orphan cluster (NOC) model[6] and the uniform cluster center (UCC) model,[5] where the end points are modeled as clusters around the base points. The models consist of three independent components: number of termination points per cluster or nerve tree, branch length, that is, distance between a termination point and its base point, and angle of the termination point. In the NOC model, a Jonquiére distribution was suggested for the number of end points per base, a gamma distribution for the branch length, and a von Mises distribution with mean direction defined as the direction away from the closest other base point. In the UCC model, a negative binomial distribution was suggested for the number of end points per base, a gamma distribution for the branch length, and a von Mises distribution with uniform mean direction for the angles between base and end points. Even though the UCC model describes the end point patterns quite well at short distances, neither of the models is able to give a good description of the data at longer distances.

In this study, we construct a NOC-like model in three dimensions, that is, we model the 3D locations of the end points instead of their 2D projections where the nerve trees are guided toward open space. Furthermore, we include the first branching points in the model. In addition, the length distribution of the initial tree segments is different from the length distribution of the later segments[22] and according to our initial investigation even the angular distributions of the initial and later segments are significantly different, see the shift plots presented in Figure 3. Therefore, we propose a two-step model, where, as the first step, the locations of the first branching points are modeled conditioned on their corresponding base points, the locations of which are known and fixed. As the second step, the locations of the termination points of the nerve trees are modeled conditioned on the locations of the simulated branching points. We also give a 2D version of the model.

5.1 | Locations of the first branching points

First, the points where the nerve trees begin to branch are modeled conditioned on their base points. The model consists of two independent components, a positive random variable L that measures the distance between the base point and the branching point and a random direction A that measures the direction of the first segment relative to the base point. For the lengths of the first segments, we used a Gamma distribution which has the density

$$f_L(x;\alpha,b) = \frac{b^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-bx}, \quad x > 0, \tag{10}$$

where $\alpha > 0$, $b > 0$ are the shape and rate parameters, respectively. The distribution of the direction $A = (\Theta, \Phi)$ of the first tree segment is constructed in two steps. First, the planar angle Φ is assumed to follow a von Mises distribution with mean direction that favors directions toward open space as in the NOC model,[6] that is, the mean direction is the direction opposite to the direction to the nearest other base point. Here, we note that both UCC-like and NOC-like models were tested, however as presented in Figure A1, the UCC-like model cannot capture the spatial randomness of the first branching point patterns. The von Mises distribution has the density

$$f_\Phi(\phi;\kappa,\mu) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa\cos(\phi-\mu)}, \quad \phi \in [-\pi,\pi), \tag{11}$$

where $\kappa \geq 0$, with $\kappa = 0$ corresponding to the uniform distribution, and $\mu \in \mathbb{R}$ are the concentration parameter and mean direction, respectively, and $I_0(\kappa)$ denotes the modified Bessel function of order zero.

The angle between the z axis and the xy plane, Θ, is assumed to be uniform, that is, $\Theta = \arccos(1-\rho)$, where $\rho \sim \mathrm{U}(0,1)$. This choice is justified by our exploratory analysis (see Figure 4). Here, it is important to note that the concentration parameter estimate κ for both groups is close to zero which implies large dispersion, that is, the distribution of the planar angle Φ is close to a uniform distribution.

5.2 | Locations of the end points

In the second part of the model, the end points are modeled conditionally on the locations of the simulated branching points. The model consists of three independent components, a positive integer N that describes the tree size, that is, the number of end points per nerve tree, and the distance L and direction A_2 between termination points and their mother point, in this case the location of the branching point. The Poisson distribution is not flexible enough to fit the offspring counts,[5,6] hence a negative binomial distribution with two parameters is used to model the tree size minus one since the tree size is at least one. The negative binomial distribution has the density given by

$$f_N(n;k,p) = \frac{\Gamma(n+k)}{n!\Gamma(k)} p^k (1-p)^k, \quad n = 0,1,2,\ \ldots\ , \tag{12}$$

where $p \in [0,1]$ and k is a non-negative integer. The distribution of the distance between the branching point and the end points connected to it is modeled by a Gamma distribution with density given by (10). As the 3D von Mises distribution was not able to capture the direction $A_2 = (\Theta_2, \Phi_2)$ of the segments connecting the termination points with the branching point, the Schladitz distribution[23] was used. The Schladitz distribution is a 3D angular distribution, a special case of the angular central Gaussian distribution, which has a density given by

$$f_{\Theta_2,\Phi_2}(\theta,\phi;\theta_0,\phi_0,\beta) = \frac{1}{4\pi} \frac{\beta\sin\theta}{(1+(\beta^2-1)(\sin\theta_0\sin\theta\cos(\phi_0-\phi) - cos\theta_0\cos\theta)^2)^{\frac{3}{2}}} \tag{13}$$

for $\theta \in [0,\pi)$ and $\phi \in [0,2\pi)$, where $\theta_0 \in [0,\pi)$ and $\phi_0 \in [0,2\pi)$ are the altitude and planar angle of the mean direction, respectively, and $\beta > 0$ is the anisotropy parameter, with $\beta = 1$ corresponding to the isotropic case. Bipolar angular distributions, that is, distributions with density concentrated in the poles appear when $\beta < 1$, and girdle distributions, that is, distributions with density concentrated in the equator when $\beta > 1$. In the case where we have rotational symmetry around the z-axis, that is, $(\theta_0,\phi_0) = (0,0)$, the density is simplified to

$$f(\theta, \phi; \beta) = \frac{1}{4\pi} \frac{\beta \sin\theta}{(1 + (\beta^2 - 1)\cos^2\theta)^{\frac{3}{2}}}. \tag{14}$$

As the distribution of the planar angle Φ_2 in our work is approximately uniform this simplification can be justified.

5.3 | Two-dimensional version of the model

The 3D model defined above is a natural extension of a simple 2D model inspired by both the NOC and UCC models. We suggest to model the 2D locations of the branching points (the distance from the base point L_1 and the planar angle Φ_1) by using the NOC model, that is, the first branching points are sent toward open space. The end points (having distance L_2 and direction Φ_2 from the branching point) are then uniformly clustered around the simulated branching points. The tree size (the number of end points per cluster) has the negative binomial distribution as in the UCC model. The 2D version of our two-step model is given by

$$\begin{aligned} L_1 &\sim \Gamma(\alpha_1, \beta_1), \\ L_2 &\sim \Gamma(\alpha_2, \beta_2), \\ \Phi_1 &\sim \text{VonMises}(\mu, \kappa), \\ \Phi_2 &\sim \text{Uniform}(0, 2\pi), \\ S &\sim NB(r, p), \end{aligned} \tag{15}$$

where μ is the NOC-like direction and S denotes the size of the nerve trees.

5.4 | Parameter estimation

This part briefly describes the procedure followed to estimate the parameters of the model. For simplification in the parameter estimation procedure, we make the same assumptions as in the earlier work,[5,6] that is all individual parts of the model are independent, hence parameter estimation can be performed independently for every component. For the estimation of the parameter of the Gamma and the negative binomial distributions, the method of maximum likelihood is used.[5,6] For the von Mises distribution, we followed the NOC model approach that assumes that the mean direction μ is known, hence we only have to estimate the concentration parameter κ. For this purpose, we used the simple approximation of the maximum likelihood estimate for κ.[24] Similarly, for the parameters estimation of the Schladitz distribution, we used the simplified version that assumes rotational symmetry around the z-axis. Therefore, only the anisotropy parameter β needs to be estimated. In this regard, the method, which uses the Newton's method to maximize the likelihood of the Schladitz distribution, was used.[25]

6 | ENF DATA ANALYSIS

Here, the 3D model introduced in Section 5 is fitted to the ENF data from healthy and neuropathy subjects. We give the parameter estimates for both groups and investigate the performance of the models. In addition, the goodness-of-fit of the 2D version of the model for the projected data is investigated.

6.1 | Parameter estimates and model fit

The parameters of the model are estimated as described in Section 5.4. The subjectwise estimates are obtained by using all the samples from the specific subject, and groupwise estimates by using all the subjects and samples in the healthy and neuropathy groups, respectively. The groupwise estimates are shown in Table A2 and the subjectwise estimates for the healthy group in Figure A4 and for the mild group in Figure A5 in Appendix. The subject specific estimates in the healthy group are distributed around the estimated groupwise parameters, except in the case of the concentration parameter κ

of the von Mises distribution. The mean of the subject-wise parameter estimates for κ is 0.31 which implies that the planar directions of the first segment of the nerve trees from the same subject are clustered around the mean direction, which in our model is assumed to be random. However, when we consider the whole healthy group the estimate for κ is 0.07 which indicates that the planar angle of the first segment is approximately uniformly distributed. This is expected since the parameter in this case is estimated from the superposition of von Mises distributions with similar concentration parameters and random mean directions. The same applies to the mild group. The estimates are very similar in the two groups, except the parameter β which is somewhat larger in the healthy group indicating that there are differences in the angular distribution of the later segments. In particular, this indicates that the direction of the later segments of the nerve trees is more vertical in the mild diabetic patients than in the healthy controls.

As reported earlier,[5] the negative binomial distribution describes well the number of end points in both groups. The empirical length distribution and the fitted gamma distribution are compared in Figure A6, where the length of the first segment is plotted on the left and the distance between the branching point and its end points on the right for the healthy group. A gamma distribution seems to describe the data very well. The goodness-of-fit of the proposed angular distributions for the healthy group are illustrated in Figure A2 by using qq-plots.[26] Both the planar von Mises distribution and the 3D Schladitz distribution for the angle of the end points fit well to the data. The length and angle distributions fit well also for the neuropathy group (results not shown here).

6.2 | Spatial goodness-of-fit of the model

In this section, we investigate whether the suggested model and its 2D counterpart are able to produce 3D and 2D end point patterns that have similar spatial structures as the observed patterns. We estimate the centered L function from the data and compare it to the same function estimated from simulations of the model. The groupwise L function for one simulation round is estimated as follows. By using the subjectwise parameter estimates, we simulate as many samples per subject as observed in the data. Each realization is simulated by conditioning on the observed base point pattern. The groupwise centered L function is then computed by first combining the samplewise and then, the subjectwise functions as described in Section 3.2.1. The procedure is repeated 2500 times and based on these simulations, global envelopes for the group-wise second order summary statistics are constructed.[19] The main advantage of this approach is that, not only the empirical summary functions can be graphically compared with their simulated counterparts from the null model but also a p-value in the form of a p-interval can be computed. A p-interval denotes the interval $(p-, p+)$ where $p-$ denotes the most liberal and $p+$ the most conservative p-value. The interpretation of a p-interval as a test at a level α is the following: If $p+ \leq \alpha$ the null hypothesis is rejected and if $p- \geq \alpha$ there is not enough evidence to reject the null hypothesis. If $p- \leq \alpha \leq p+$ it is not clear if we should reject the null hypothesis or not.

6.2.1 | 3D model

The estimated groupwise centered L functions together with global envelopes can be seen in Figure 11 for the healthy group (left) and for the mild group (right). The clustering behavior of the observed and simulated end point patterns can easily be verified for all subjects as the $L(r) - r$ functions are clearly above zero. The model succeeds in capturing the spatial structure of the end points at very small distances and at distances above 30 μm. However, it fails to capture the behavior at intermediate distances. The model tends to create less clustered patterns than observed in the data. This trend is common in both groups. Further, we estimated the subjectwise centered L functions for all the subjects within the two groups. The results for the healthy and the mild subjects together with global envelopes can be found in the supplementary material. In most cases the subjectwise fit of the model cannot capture the spatial structure for intermediate distances. This is an indication that there is some interaction between the end points that cannot be captured by the model.

Moreover, also related to the spatial structure, we studied how well the model generated EATs correspond to the observed ones. In Figure 12, the total volume of the EATs computed from the observed patterns is compared to the model based volume by using shift plots. The model seems to capture the total nerve tree coverage in the healthy group well. However, in the mild group, the model does not capture the total coverage as well but produces slightly larger nerve trees, with respect to the EAT volume, than the nerve trees observed in the data.

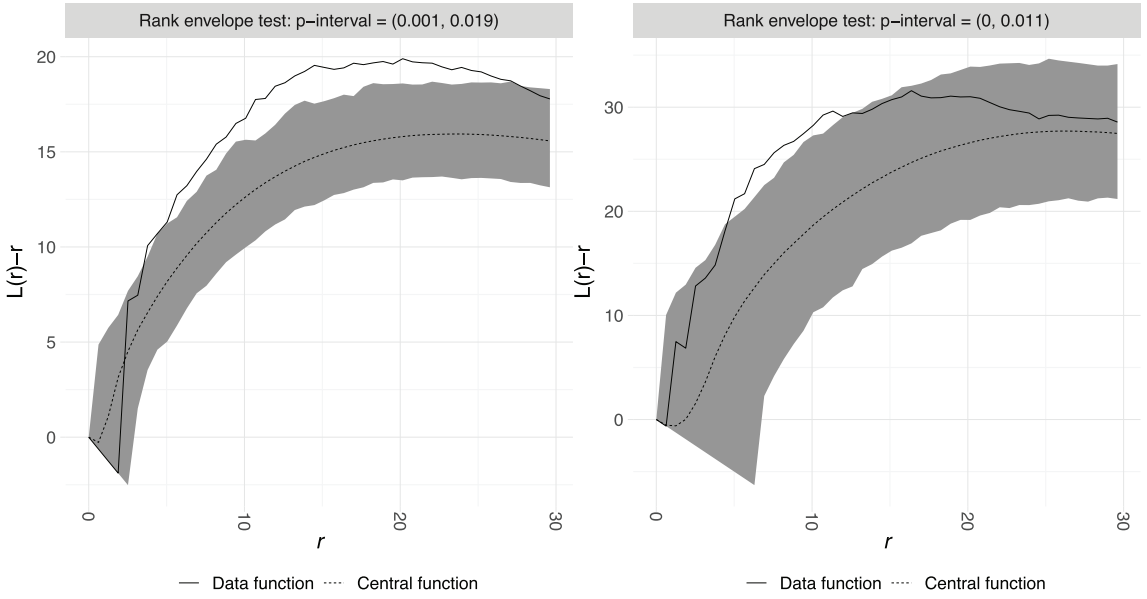

FIGURE 11 The groupwise pooled $L(r) - r$ functions with global envelopes for the end points for the healthy group (left) and the mild group (right)

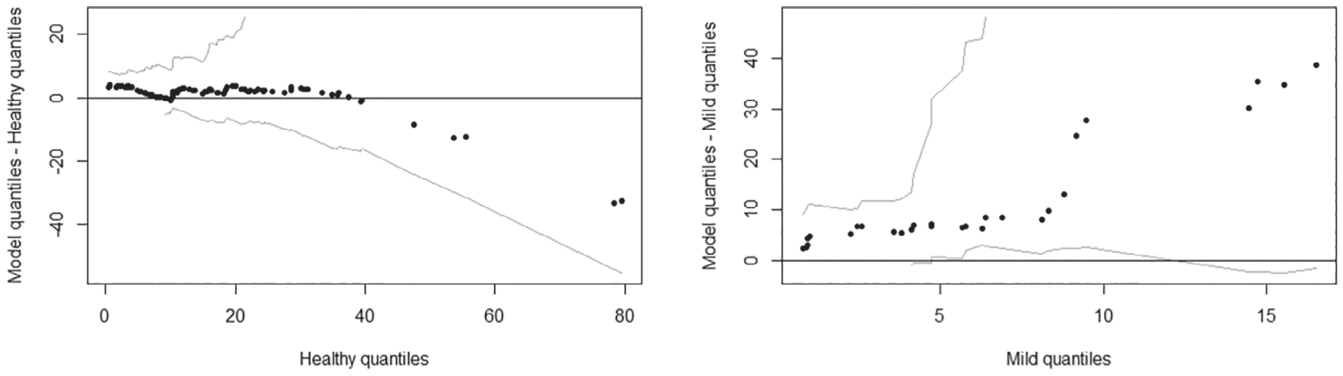

FIGURE 12 Shift plots of the epidermal active territory volumes produced by the model for the normal group (left) and for the mild group (right)

6.2.2 | 2D model

Finally, we studied whether the 2D structure of the projected data can be better captured by the model with branching points introduced in Section 5.3 than by the previously introduced models.[5,6] The centered L functions for the end points for the healthy and mild groups are plotted in Figure 13 (left). The proposed NOC-like model for the branching points given base points and UCC-like model for the end points given branching points is able to describe the spatial distribution of the planar end point patterns quite well. The wider envelopes in the mild case are due to the smaller number of samples in that group. We also plotted the bivariate $L_{bp,ep}(r) - r$ and $L_{br,ep}(r) - r$ functions describing the relationship between the base points and end points and branching points and end points, respectively, and they are shown in Figure 13 (middle and right). The model captures the clustering behavior of the end points with respect to the base points and branching points in the mild group but not as well in the healthy group. This is at least partly due to the wider envelopes in the mild case. The fit of the model with respect to the groupwise summary statistics reveals that there is large variance between the point patterns of different subjects within the groups.

We plotted even the subjectwise $L(r) - r$ summary functions for all the 32 healthy controls and the 8 subjects with mild neuropathy. Our findings suggest that our 2D model captures the end point structure of the individual patterns very well. Moreover, the model captures the clustering behavior of the end points with respect to both the base points and branching

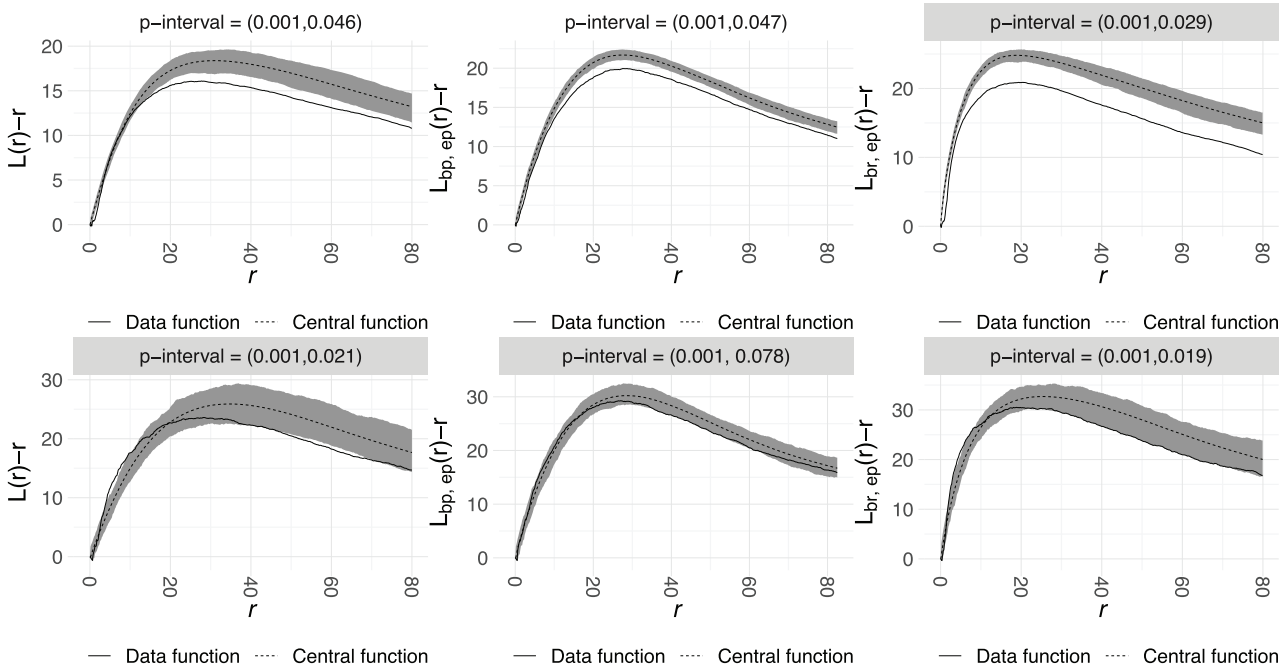

FIGURE 13 Groupwise pooled two-dimensional $L(r) - r$ (left) for the end points, $L_{bp,ep}(r) - r$ (middle) and $L_{br,ep}(r) - r$ (right) functions with global envelopes for the healthy (top) and mild (bottom) groups

points as the bivariate empirical functions $L_{bp,ep}(r) - r$ and $L_{br,ep}(r) - r$ are within the 95% global envelopes constructed by the simulated patterns in most of the cases. The respective plots can be found in the supplementary material. We can conclude that the 2D model that includes the branching points describes the spatial structure much better than the previously suggested models.

We would also like to point out that the end point patterns could also be modeled by using Neyman-Scott processes, such as Matérn process or Thomas process, but those models fail to capture the end point structure at small distances. The goodness-of-fit of the models for the healthy group is shown in Figure A3 in Appendix. Most importantly, our model takes the whole spatial structure of the nerve trees into account while the Neyman-Scott models use no information about the connections between the base, branching, and end points in the nerve trees. Therefore, we prefer the proposed NOC-like model over the Neyman-Scott models as it is a model for the complete ENFs structure.

7 | DISCUSSION

The spatial structure of ENFs of healthy subjects and patients with mild diabetic neuropathy was investigated. An earlier study of this nature can be found in Andersson et al,[5] where a comparison of the spatial structure in healthy and mild diabetic groups was performed considering the base point and the end point patterns of the same data in 2D. Their results indicated that the end point locations projected on 2D of healthy subjects are less clustered than the ones of mild subjects. Moreover, they found that the coverage of the skin by nerve fibers is higher in the healthy group than in the mild group. Furthermore, 2D point process models for the end points given the base point patterns have been proposed.[5,6]

Here, we study the whole 3D structure, not only the coverage on the skin, by representing each nerve tree by the base point, first branching point, and ENF end points. We studied the spatial structure of ENFs by using second order summary statistics and the structure of the individual nerve trees by means of the length and angle of the first segment and the segment connecting the first branching point and its end points. We extended the concept of reactive territory, here called EAT for 2D point patterns,[5] to 3D point patterns. In addition, we constructed a 3D model for the first branching points and end points given the base points. We also fitted a 2D counterpart of the model to the projected patterns. To evaluate the fit of the angular distributions, qq-plots were used and for the comparison of the properties of the data, shift plots were used. The main advantage of those methods is that they compare the whole distributions, not only the mean values, and hence we may be able to identify significant differences that we are not able to see by only comparing the means.

Direct comparison of the performance between our model and the 2D models in the literature[5,6] should be avoided since our model is more complex, with more parameters and aims to capture a more complete nerve tree structure rather than capturing the coverage of the skin by the end points.

Unlike the end points projected on 2D,[5] no evidence was found to suggest that the 3D end points from neuropathic subjects would be more clustered than the end points from healthy subjects. However, the segments connecting the end points with the first branching points are longer in healthy patients than in mild patients, and the direction of that segment in mild patients is more vertical compared to the corresponding direction in healthy patients. It was also found that the total volume of the epidermis covered, measured by the EAT, in healthy patients was larger than the volume in the mild group. These findings are in agreement with the results in 2D that indicate that the nerve tree density and the coverage of the epidermis by ENFs are reduced due to neuropathy. Finally, we studied the competitive behavior of the nerve trees by using the mark correlation function of base points having the EAT as the mark. The sizes of the territories seem not to be spatially correlated.

The 3D model that we constructed can be considered an extension of the UCC and NOC models.[5,6] The suggested Gamma distribution seems to fit well to the distance between the base point and its first branching point and for the distance between the first branching point and its end points. Also, the angle distributions that were chosen fitted well to the data. The empirical and model based distributions of the volume of the active territory match well for healthy subjects but not as well for subjects suffering from mild diabetic neuropathy, where the model tends to create larger volumes than those in the data. Despite that the different components of the model seem to fit reasonably well to the data, the model fails to capture the 3D spatial structure of the end points. However, the 2D version of the model fitted well to the projected end point patterns, much better than the previously suggested models, indicating that it is important to include the first branching points in the model. This result was expected since the end point clusters tend to be distributed around their respective branching points rather than around their base point.

The analysis is based on a rather small data set since we had only eight subjects in the mild diabetic group. Since we have replicates from each subject, we have in total 28 samples in the group, still a rather small number. However, we should point out that the parameters of the models are estimated based on the individual nerve tree characteristics such as the distance between an end point and its branching point or the direction of an end point, increasing the number of observations from each group. In addition, even though the number of points per sample may not always be so large, we can increase the power by having replicates. The main difficulty is that there is so much variation within each group that it is difficult to find statistically significant differences between them. It should also be mentioned that our study should be regarded as a pilot study since the control group consists of healthy volunteers and the neuropathy group may not be a random sample from the population of subjects having mild neuropathy. Therefore, our results could be used as guidelines for a larger study. We also believe that the differences between the groups with respect to the volume of the epidermal active territories and between the angular distributions of the later segments between the groups could contribute to the development of more accurate methods for early detection of the disease.

To the best of our knowledge, this is the first 3D (point process) model suggested for the ENFs. The 2D and 3D models introduced here are also the first models that in addition to the base and end points, include the first branching points. In addition, the competitive behavior of the epidermal nerve trees in terms of their sizes in three dimensions was studied here, as far as we are aware, for the first time. Therefore, this article contributes to better understanding of the biological mechanisms guiding how diabetic neuropathy affects the structure of epidermal nerve trees. The individual distributions in the 3D model we constructed seem to fit well to the data but the model was not able to capture the spatial structure so well at group level. Therefore, a possible path for future research would be to develop a 3D model that takes into account possible interactions between the different types of points in the nerve trees. Also, it would be interesting to investigate whether a more complex model taking into account multiple neighboring base points to approximate the directions toward open space would compare to our proposed model, where the simple NOC-directions depending only on the nearest base point are used. Another possible topic for future research would be to develop methods to discriminate the patterns into healthy and mild groups with help of the epidermal active territories and other knowledge obtained in this study. Such a study can extend the earlier proposed discrimination methods,[22] where no spatial information was used.

ACKNOWLEDGEMENTS

The authors thank William R. Kennedy and Gwen Wendelschafer-Crabb (University of Minnesota) for blister immunostaining, quantification, and morphometry of the ENF data. The authors also thank the Swedish Research Council (VR 2018-03986) for financially supporting the project. We are also very grateful to the two anonymous referees who carefully read the article and suggested comments that helped us to improve the article.

DATA AVAILABILITY STATEMENT

Unfortunately, we are not able to share the data publicly.

ORCID

Konstantinos Konstantinou https://orcid.org/0000-0003-3549-5851

REFERENCES

1. Kennedy WR, Wendelschafer-Crabb G. The innervation of human epidermis. *J Neurol Sci*. 1993;115(2):184-190.
2. Kennedy WR, Wendelschafer-Crabb G, Johnson T. Quantitation of epidermal nerves in diabetic neuropathy. *Neurology*. 1996;47(4):1042-1048.
3. Waller LA, Särkkä A, Olsbo V, et al. Second-order spatial analysis of epidermal nerve fibers. *Stat Med*. 2011;30(23):2827-2841.
4. Myllymäki M, Panoutsopoulou IG, Särkkä A. Analysis of spatial structure of epidermal nerve entry point patterns based on replicated data. *J Microsc*. 2012;247(3):228-239.
5. Andersson C, Guttorp P, Särkkä A. Discovering early diabetic neuropathy from epidermal nerve fiber patterns. *Stat Med*. 2016;35(24):4427-4442.
6. Olsbo V, Myllymäki M, Waller LA, Särkkä A. Development and evaluation of spatial point process models for epidermal nerve fibers. *Math Biosci*. 2013;243(2):178-189.
7. Andersson C, Rajala T, Särkkä A. A Bayesian hierarchical point process model for epidermal nerve fiber patterns. *Math Biosci*. 2019;313:48-60.
8. Andersson C, Mrkvička T. Inference for cluster point processes with over-or under-dispersed cluster sizes. *Stat Comput*. 2020;30(6):1573-1590.
9. Wendelschafer-Crabb G. Epidermal nerve fiber densities in six body locations of normal and diabetic subjects. *J Periph Nerv Sys*. 2005;10:104.
10. Panoutsopoulou IG, Wendelschafer-Crabb G, Hodges JS, Kennedy WR. Skin blister and skin biopsy to quantify epidermal nerves: a comparative study. *Neurology*. 2009;72(14):1205-1210.
11. Myllymäki M, Särkkä A, Vehtari A. Hierarchical second-order analysis of replicated spatial point patterns with non-spatial covariates. *Spat Stat*. 2014;8:104-121.
12. Illian J, Penttinen A, Stoyan H, Stoyan D. *Statistical Analysis and Modelling of Spatial Point Patterns*. Hoboken, NJ: John Wiley & Sons; 2008.
13. Diggle PJ. *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*. Boca Raton, FL: CRC Press, Taylor Francis Group; 2014.
14. Møller J, Waagepetersen RP. *Statistical Inference and Simulation for Spatial Point Processes*. Boca Raton, FL: CRC Press; 2004.
15. Chiu SN, Stoyan D, Kendall WS, Mecke J. *Stochastic Geometry and Its Applications*. Hoboken, NJ: John Wiley & Sons; 2013.
16. Ripley BD. The second-order analysis of stationary point processes. *J Appl Probab*. 1976;13:255-266.
17. Diggle PJ, Mateu J, Clough HE. A comparison between parametric and non-parametric approaches to the analysis of replicated spatial point patterns. *Adv Appl Probab*. 2000;32(2):331-343.
18. Schladitz K, Särkkä A, Pavenstädt I, Haferkamp O, Mattfeldt T. Statistical analysis of intramembranous particles using freeze fracture specimens. *J Microsc*. 2003;211(2):137-153.
19. Myllymäki M, Mrkvička T, Grabarnik P, Seijo H, Hahn U. Global envelope tests for spatial processes. *J R Stat Soc Ser B Stat Methodol*. 2017;79(2):381-404.
20. Sibson R. Locally equiangular triangulations. *Comput J*. 1978;21(3):243-245.
21. Doksum KA, Sievers GL. Plotting with confidence: graphical comparisons of two populations. *Biometrika*. 1976;63(3):421-434.
22. Andersson C, Rajala T, Särkkä A. Hierarchical models for epidermal nerve fiber data. *Stat Med*. 2018;37(3):357-374.
23. Schladitz K, Peters S, Reinel-Bitzer D, Wiegmann A, Ohser J. Design of acoustic trim based on geometric modeling and flow simulation for non-woven. *Comput Mater Sci*. 2006;38(1):56-66.
24. Banerjee A, Dhillon IS, Ghosh J, Sra S. Clustering on the unit hypersphere using von Mises-Fisher distributions. *J Mach Learn Res*. 2005;6:1345-1382.
25. Franke J, Redenbach C, Zhang N. On a mixture model for directional data on the sphere. *Scand J Stat*. 2016;43(1):139-155.
26. Fisher NI, Lewis T, Embleton BJJ. *Statistical Analysis of Spherical Data*. Cambridge, MA: Cambridge University Press; 1993.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Konstantinou K, Särkkä A. Spatial modeling of epidermal nerve fiber patterns. *Statistics in Medicine*. 2021;1–22. https://doi.org/10.1002/sim.9194

APPENDIX . TABLES AND FIGURES

TABLE A1 Mean and standard deviation (from left to right) of the length of the first and later segments, number of nerve trees and end points per sample, and number of end points per nerve tree in the normal (top) and mild (bottom) groups

	Branch length		* per sample		
Group	**First segment**	**Later segments**	**Trees**	**End points**	**End points/cluster**
Normal	17.48 (9.15)	15.83 (9.64)	26.39 (13.02)	67.93 (36.12)	2.60 (1.69)
Mild	15.98 (8.86)	13.63 (9.19)	21.42 (10.72)	51.4 (30.28)	2.36 (1.41)

TABLE A2 Mean and standard deviations of the parameter estimates: α_1 and b_1 control the distance between the base points and the branching points, α_2 and b_2 the distance between the branching points and end points, and κ and β are the angular parameters

Group	$\boldsymbol{\alpha_1}$	$\boldsymbol{b_1}$	$\boldsymbol{\alpha_2}$	$\boldsymbol{b_2}$	$\boldsymbol{\kappa}$	$\boldsymbol{\beta}$
Normal	2.62 (0.15)	0.15 (0.01)	2.46 (0.22)	0.13 (0.01)	0.07 (0.05)	1.25 (0.07)
Mild	2.49 (0.18)	0.15 (0.01)	2.30 (0.18)	0.14 (0.01)	0.07 (0.04)	1.07 (0.06)

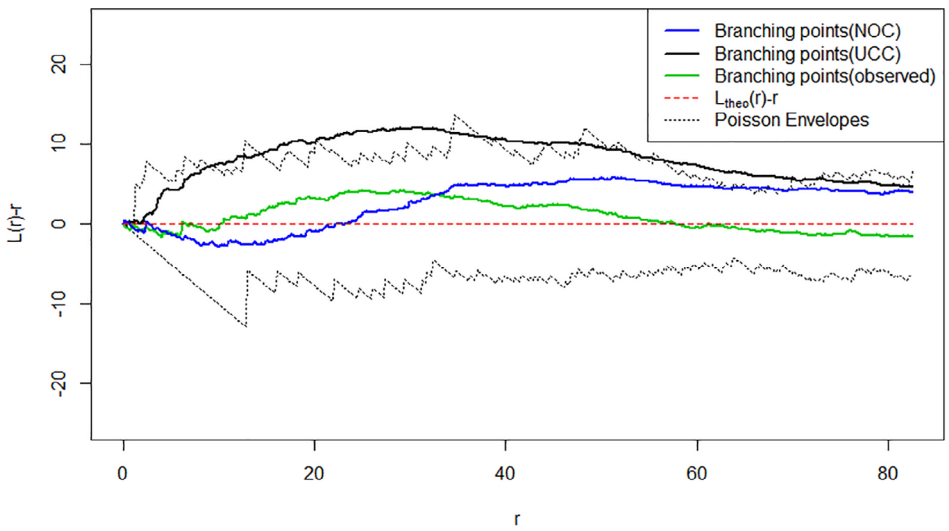

FIGURE A1 Ripley's overall $L(r) - r$ function for the locations of observed branching points (green) and the simulated branching points from the NOC model (blue) and UCC model (black) compared with the theoretical value of a Poisson point process (red) with simulation envelopes (dashed lines) [Colour figure can be viewed at wileyonlinelibrary.com]

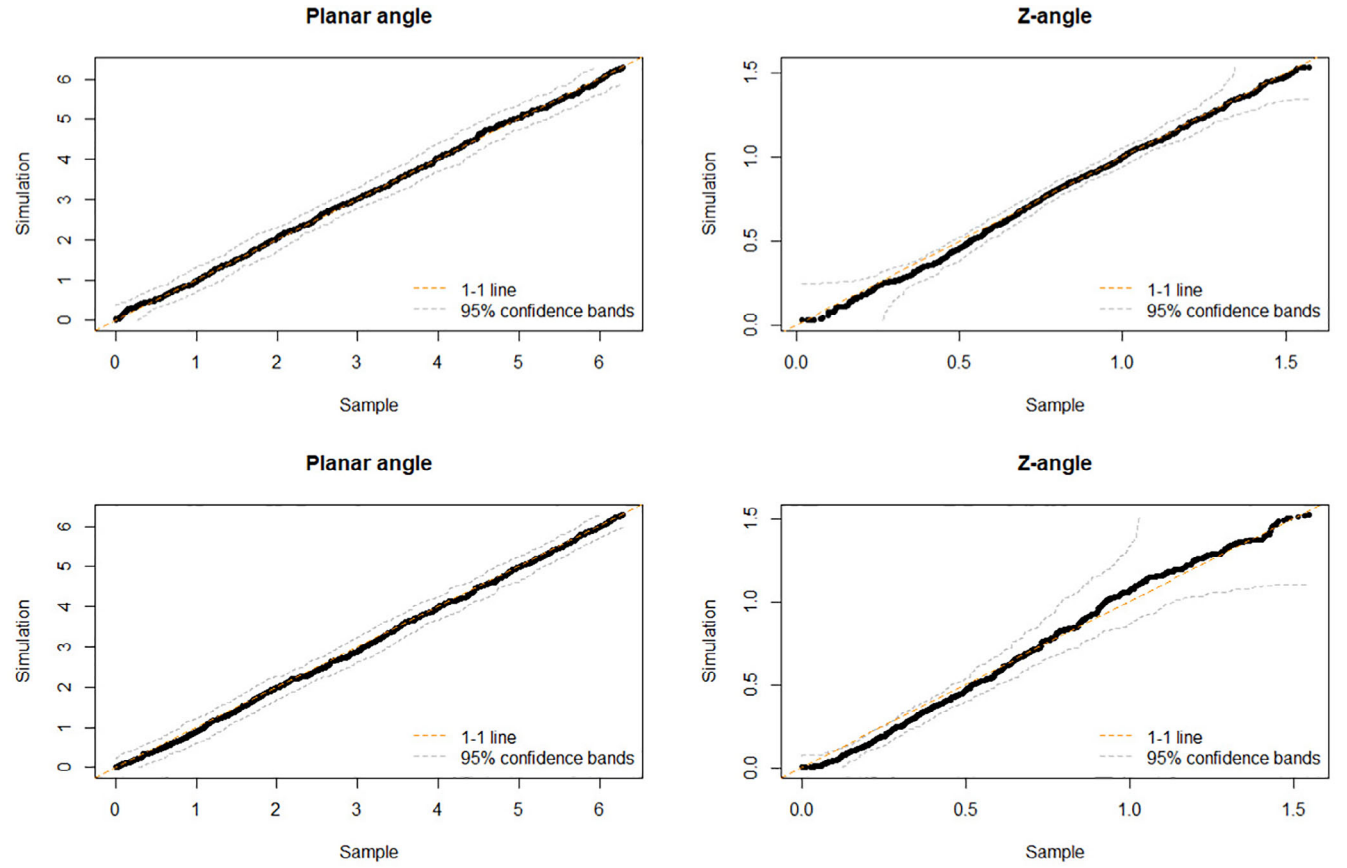

FIGURE A2 Goodness-of fit of the angular distributions estimated from the healthy group. qq-plots for the fit for the distribution for the branching point (top) and for the fit of the Schladitz distribution for the end point (bottom) [Colour figure can be viewed at wileyonlinelibrary.com]

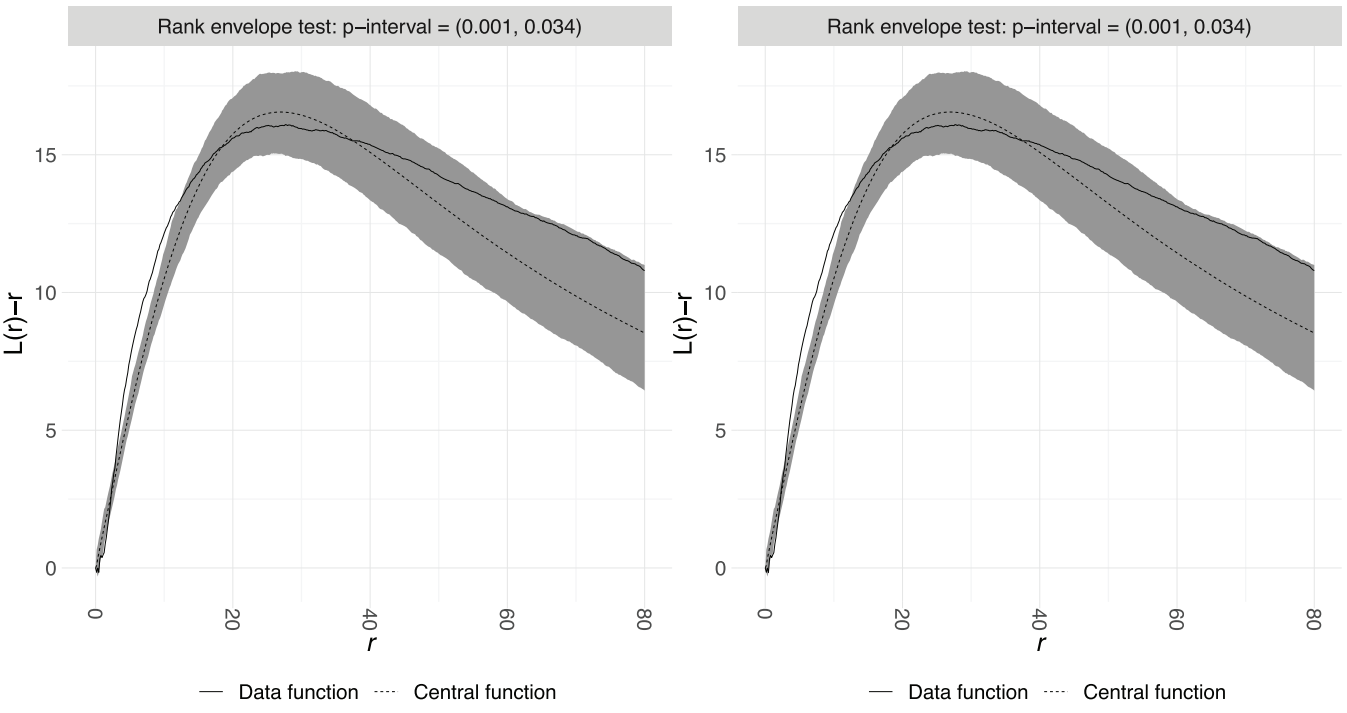

FIGURE A3 Groupwise pooled two-dimensional $L(r) - r$ functions with global envelopes for the end points for healthy subjects of the two Neyman-Scott models, Matérn (left) and Thomas (right)

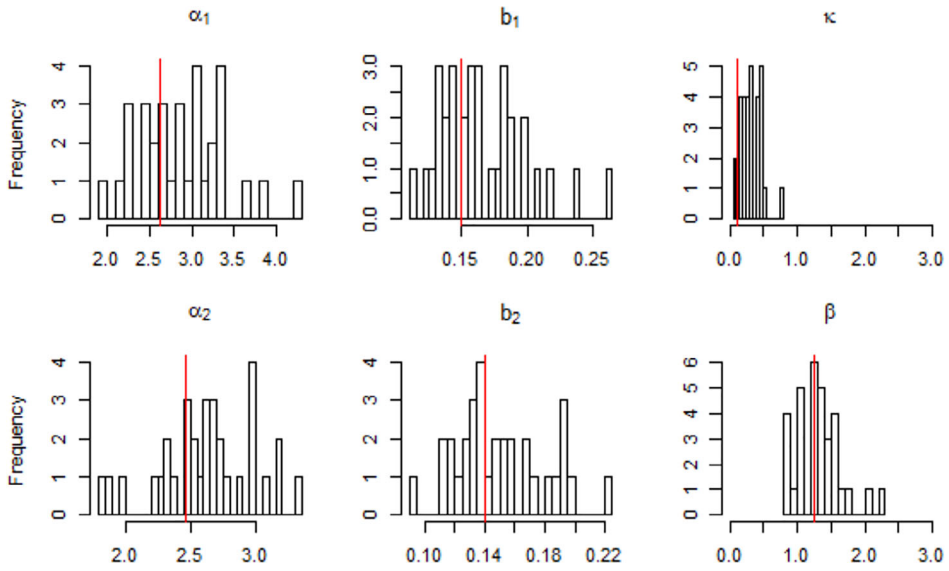

FIGURE A4 Histogram of the estimated parameters for the healthy subjects. The red line denotes the parameter estimate for the whole group. The parameter estimates of the first and second part of the model are presented on the first and second row, respectively. In each row, the first two plots correspond to the parameters of the Gamma distribution, and the right most plot to the parameter of the angular distribution [Colour figure can be viewed at wileyonlinelibrary.com]

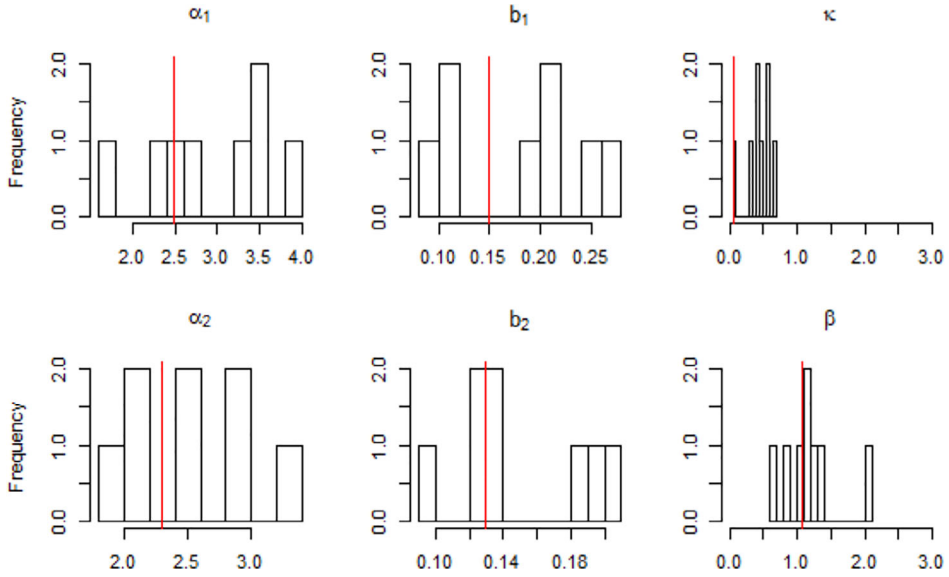

FIGURE A5 Histogram of the estimated parameters for the subjects in the mild group. The red line denotes the parameter estimate for the whole group. The parameter estimates of the first and second part of the model are presented on the first and second row, respectively. In each row, the first two plots correspond to the parameters of the Gamma distribution, and the right most plot to the parameter of the angular distribution [Colour figure can be viewed at wileyonlinelibrary.com]

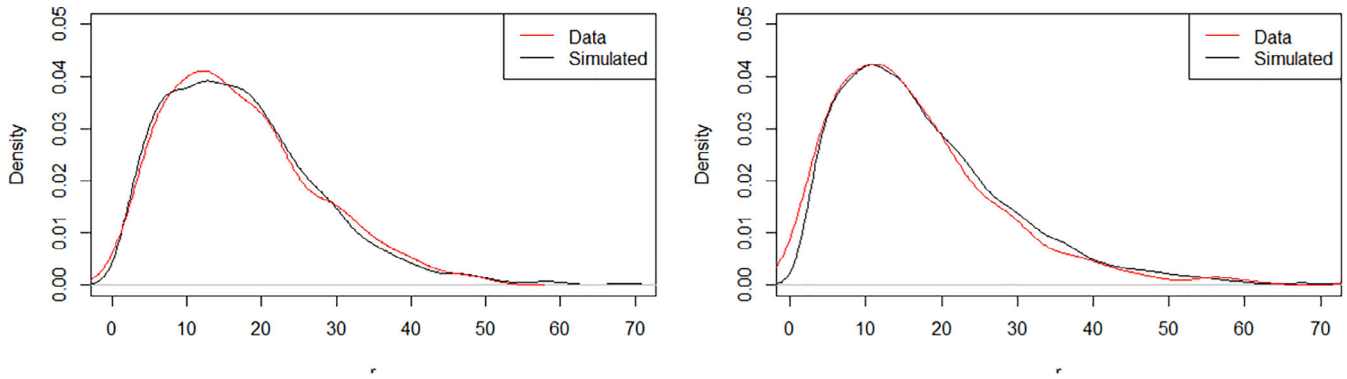

FIGURE A6 The empirical (red) and fitted (black) distributions for the first segment length (left) and for the branch length in the second part of the model (right) for the healthy group [Colour figure can be viewed at wileyonlinelibrary.com]