Environ Ecol Stat (2007) 14:149–159
DOI 10.1007/s10651-007-0009-0

ORIGINAL ARTICLE

Accounting for large-scale factors in the study of understory vegetation using a conditional logistic model

Sharon Kühlmann-Berenzon · Urban Hjorth

Received: September 2003 / Revised: October 2005 / Published online: 3 March 2007

Abstract Local-scale and large-scale factors can affect the presence of a species of understory vegetation in the forest. Local-scale factors may be the influence of surrounding trees, while climate and latitude are typically considered large-scale factors. A model for the presence of a species needs to take into account both scales. A conditional logistic model is proposed for those studies where only the local-scale factors are of interest and that avoids estimating the large-scale parameters. Conditioning is carried out by the number of quadrats in the plot where the vegetation is found. As the latter is a sufficient statistic for the large-scale factors, a model free from these parameters is obtained. Data gathered in the permanent sample plots of the 1985–1986 National Forest Inventory of Finland is used for illustration, where the local-scale factor of interest is the influence of the trees, quantified by an index based on the size and location of the trees. The model fitted to *Vaccinium vitis-idaea* showed a significant and positive influence of Scots pine on the presence of this species, while for *Calamagrostis arundinacea*, a decrease in the odds ratio was observed due to the influence of Norway spruce.

Keywords Ecology · Forestry · Influence potential · Local scale · Logistic regression · Nuisance parameters · Odds ratio · Sufficient statistics

S. Kühlmann-Berenzon (✉)· U. Hjorth
Department of Mathematical Statistics, Chalmers University of Technology and Göteborg University, S-412 96 Gothenburg, Sweden
e-mail: Sharon.Kuhlmann@smi.ki.se

Present address:
S. Kühlmann-Berenzon
Department of Epidemiology, Swedish Institute for Infectious Disease Control, 171 82 Solna, Sweden

U. Hjorth
Department of Business Administration, Computer Science, Economics and Statistics, Örebro University, S-701 82 Örebro, Sweden

1 Introduction

The layer of vegetation that grows underneath the canopy of the trees, or understory, plays an important role in the forest by giving shelter to animals, and protecting and enriching the soils through the nutrient cycle. Reinikainen et al. (2000) compared the abundance of understory vegetation species in Finland between the 1950s and 1990s and reported a significant decrease of many species. Some of the obvious causes were the changes in forest management and agricultural practices that have modified the site conditions, age class distributions, and tree species composition, but which have also led to an increase in the timber production (see also Mäkipää and Heikkinen 2003).

How the understory responds to the effect of the overstory, large-scale environmental factors, and competition is expected to be complex, but limited efforts have been done to study it (McKenzie and Halpern 1999). In particular, models that quantify the effect of the trees on the undestory at a local scale could help to further understand the ecological dynamics that occur in the forest between trees and vegetation, which is important for the biodiversity of the forests as well as for forest management practices. Such models may also help to improve the classification of sites which is currently carried out according to understory species; to comprehend the regeneration of stands where understory species such as grasses affect the survival of saplings; and to study how local changes in the tree stand affect the understory.

Kühlmann et al. (2001) analyzed the correlation between single-tree influences and the abundance of plant species by considering only the variation around local averages, in this way circumventing the large scale factors. Their analysis was purely explorative and the approach is not directly applicable to statistical modeling, which would provide a better insight into the relationship between tree influence and understory vegetation.

The aim of this paper is to develop an approach for modeling the local variation of factors on the presence of a specific understory species in such way that the large-scale variation due to other environmental factors is taken into account. One way would be to fit a logistic regression that requires estimating the parameters from both local and large scales. Instead we propose a conditional logistic model that avoids computing the large-scale parameters but still provides estimates for the local-scale factors of interest. This is achieved by including in the model nuisance parameters that are statistically sufficient for the large-scale factors. Our examples in particular illustrate how the model can be applied for quantifying the influence of species of trees on a single species of vegetation.

Sections 2 and 3 of the paper describe the data and the influence potential measure used to summarize the effect of the trees and that later serve as examples. The statistical models in terms of unconditional and conditional logistic regression are provided in Sects. 4 and 5. An application to the presence of *Vaccinium vitis-idaea* and *Calamagrostis arundinacea* using the influence of Scots pine (*Pinus sylvestris*), Norway spruce (*Picea abies*), and birch (*Betula pendula* and *B. pubescens*) serves as illustrations in Sect. 6.

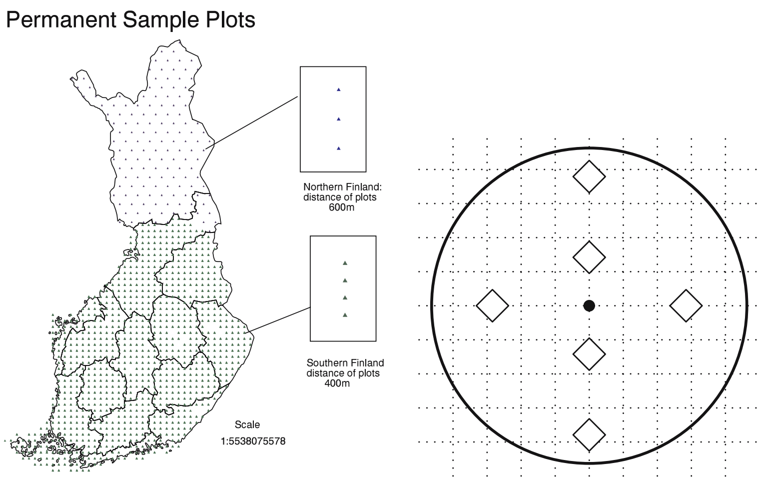
2 Data

The Finnish Forest Research Institute (METLA) gathered data on trees and understory vegetation on the permanent sample plots (PSP) during the 1985–1986 National Forest Inventory. The PSP were established for monitoring purposes and consisted of 2,905 circular plots located on forestry land. The extensive study area and the systematic sampling ensured that data was gathered from different tree stand compositions, ecological conditions, and management practices. This data has been analyzed for monitoring the health of the forest and for studying particular understory species (Tonteri et al. 1990; Korpela and Reinikainen 1996; Mäkipää and Heikkinen 2003).

The sample plots had a radius of 9.77 m (area = 300 m^2) and were distributed in clusters on a grid over Finland (area = 337,000 km^2) as shown in Fig. 1a. The clusters in Southern Finland consisted of four plots on a north–south transect, with 400 m between plots and 16 km in every direction between the clusters; in Northern Finland, the clusters were formed by three plots each, with 600 m between plots, and a distance of 24 km in north–south direction and 32 km in east–west direction separated the clusters.

Six quadrats of 2 m^2 were systematically assigned in each plot. The quadrats were located at 3 and 8 m north and south of the plot center, and at 6 m east and west; see Fig. 1b. Not all six quadrats, however, were consistently measured during the field work, and therefore the number of observed quadrats in a plot varied between one and six.

For each species of understory vegetation, the percentage of the area of the quadrat covered by it was determined visually by the field workers. This provided the data on the presence and absence of the vegetation. Information on trees with diameter at breast height (DBH) larger than 10.5 cm was also recorded; of particular interest to this study were species, relative location in the plot, and DBH. Furthermore, the



(a) Sampling grid over Finland **(b)** Location of quadrats in the plot

Fig. 1 Permanent sample plots of 1985–1986 National Forest Inventory of Finland. *Right*: Every *dot* represents a cluster of plots; greatest distance north–south is 1,160 km, and west–east is 540 km west–east. *Left*: *Grid* is 2 m^2 and included here for scale purposes type of soil of the quadrat was also ascertained, i.e., whether it was mineral soil or any of eight subtypes of peatlands. For this paper, only quadrats in mineral soil were selected, since the conditions in peatlands were expected to be more heterogeneous; 68% of the 10,929 quadrats were on mineral soil. A set of conditions on the number of quadrats where the understory species was found in the plot were also required; the reasons for this will become clear when the model is presented. Nevertheless, even after these selection criteria, the set of plots was still distributed over a large area of the country, so that the large-scale factors were still a concern.

3 Influence potential of trees

In this study, we used the definition of influence potential by Kühlmann et al. (2001) to measure the effect of surrounding trees, and that adapts to the structure of the PSP data and the information available there. Here we opted to call the index "influence potential on a quadrat" or IPQ. Its formulation is

$$\mathrm{IPQ}(q_{ij}; T) = \sum_{t=1}^{N_{iT}} d_t \exp\left(-\frac{r^2_{t(ij)}}{c_T}\right), \tag{1}$$

where q is the quadrat indexed by the plot i and quadrat number within the plot $j = 1, \ldots, n_i$; d_t the DBH of tree $t = 1, \ldots, N_{iT}$; N_{iT} the total number of trees of species T in the plot; r the Euclidean distance between the tree t and the quadrat ij; and c_T reflects the range of influence of the tree species. The range of influence is defined as $\sqrt{\log(100)\ c_T}$, i.e., the distance at which the effect of a tree, independent of its DBH, reaches 0.01. Thus, any tree beyond that distance is not expected to have a significant influence on the vegetation on the quadrat. One generalization is to let c_T depend on DBH. Although this is biologically reasonable, it also introduces additional complexity in the estimates, and we have therefore opted for the current simpler form. The absence of a tree species in the plot results in IPQ equal to zero.

Similar tree influence indices as the one we applied have previously been used by Kuuluvainen and Pukkala (1989), Kuuluvainen et al. (1993), Økland et al. (1999), and Saetre (1999) to relate it also to understory vegetation; however, in those cases the understory species were analyzed in groups, and the data was collected in small and relatively homogeneous boreal stands where large-scale factors were not an issue.

4 Logistic regression model

In developing the model, we assume that measurements are collected in quadrats located in plots, and that plots are distributed over a large study area, which is a typical setup in forestry and ecology. We are interested in modeling the probability of the presence of an understory species in a quadrat. This probability may be affected by two processes at different scales. The first is at large-scale and affects the vegetation throughout the study area. In the data from the PSP, for example, latitude would be obviously such a factor since certain species prefer warmer conditions and are present more often in plots situated in the south of Finland (see, e.g., Reinikainen et al. 2000). When the plots are relatively small compared to the study area, the large-scale factors affect all the quadrats in the plot in similar fashion and thus can be considered as plot-level covariates. Of these factors, those that have been measured can in principle be specified explicitly in the model; other factors, unknown or unmeasurable, can be included through a plot-level intercept.

Furthermore, we assume that there is a second process that works at a local scale. Covariates related to that process differ for each quadrat, as is the case of IPQ that depends on the distance between the trees and the particular quadrat. They can therefore be also called quadrat-level covariates.

To study the presence of a species of understory vegetation, one possible model is the logistic model (McCullagh and Nelder, 1989, Chap. 4). When it includes both the large- and local-scale factors, it can be written as

$$P(W_{ij} = 1) = \frac{\exp(\alpha_i + \mathbf{y}'_i\boldsymbol{\gamma} + \mathbf{x}'_{ij}\boldsymbol{\beta})}{1 + \exp(\alpha_i + \mathbf{y}'_i\boldsymbol{\gamma} + \mathbf{x}'_{ij}\boldsymbol{\beta})}, \tag{2}$$

where W_{ij} is an indicator variable coded 1 when the understory species is present in quadrat j in plot i, and 0 otherwise; α_i the plot-level intercept; $\mathbf{y}'_i$ the row vector of explicit large-scale factors for plot i; $\mathbf{x}'_{ij}$ the row vector of quadrat-level measurements for the quadrat j in plot i; and $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ are vectors of the corresponding coefficients. To simplify the discussion, the terms $\mathbf{y}'_i\gamma$ can be absorbed into α_i, i.e.,

$$P(W_{ij} = 1) = \frac{\exp(\alpha_i + \mathbf{x}'_{ij}\boldsymbol{\beta})}{1 + \exp(\alpha_i + \mathbf{x}'_{ij}\boldsymbol{\beta})}.$$

In similar way, the probability of a vegetation species being absent from the quadrat can be modeled as

$$P(W_{ij} = 0) = \frac{1}{1 + \exp(\alpha_i + \mathbf{x}'_{ij}\boldsymbol{\beta})}$$

and in general, the probability of W_{ij} can be written as

$$P(W_{ij} = w_{ij}) = \frac{w_{ij}\ \exp(\alpha_i + \mathbf{x}'_{ij}\boldsymbol{\beta})}{1 + \exp(\alpha_i + \mathbf{x}'_{ij}\boldsymbol{\beta})}, \tag{3}$$

where $w_{ij} = \{0, 1\}$. Models such as (3) are only estimable if we have restrictions on α_i or regard these parameters as random effects.

5 Conditional logistic regression

One way to allow for the large-scale factors in (3) but to avoid estimating them is to condition on another event that is subject to the same large-scale factors. In this way, if the model is appropriate, the large-scale factors will be canceled, leaving the local-scale characteristics for further study. For this solution, we need a sufficient statistic for the large scale parameters (Casella and Berger 1990, Chap. 6).

To find such a statistic, we turn to the vector of indicators for plot i, represented as

$$\mathbf{w}'_i = (w_{i1}, \ldots, w_{ij}, \ldots, w_{in_i}),$$

where $w_{ij} = 1$ if the vegetation is present in the quadrat, and $w_{ij} = 0$ if not, and $n_i = \{1, \ldots, N_i\}$ is the number of quadrats measured in the plot; $\mathbf{w}_i'$ will be referred to as a pattern.

By assuming that the quadrats are independent in terms of the local-scale effects, then the probability of observing the pattern $\mathbf{w}_i$ is obtained from the probabilities of the individual quadrats in (3) as

$$P(\mathbf{w}_i) = \prod_{j=1}^{n_i} P(w_{ij}) = \prod_{j=1}^{n_i} \frac{\exp\left[w_{ij}(\alpha_i + \mathbf{x}_{ij}'\boldsymbol{\beta})\right]}{1 + \exp(\alpha_i + \mathbf{x}_{ij}'\boldsymbol{\beta})}. \tag{4}$$

Furthermore, let $z_i = \sum_j w_{ij}$ be the number of quadrats where the understory species is present in plot i. There are $(M_i + 1) = \binom{n_i}{z_i}$ ways of arranging a vector of length n_i with z_i elements equal to 1. In this context it means that in a particular plot i, there are $(M_i + 1)$ different patterns where the vegetation is present in z_i out of the n_i quadrats. The probability of z_i can be calculated from the probability of each of these patterns. Using again (3), then

$$P(Z_i = z_i) = \sum_{s=0}^{M_i} P\left(\mathbf{w}_i^{(s)}\right) = \sum_{s=0}^{M_i} \prod_{j=1}^{n_i} \frac{\exp\left[w_{ij}^{(s)}(\alpha_i + \mathbf{x}_{ij}'\boldsymbol{\beta})\right]}{1 + \exp(\alpha_i + \mathbf{x}_{ij}'\boldsymbol{\beta})}, \tag{5}$$

where s designates the pattern. In order for $(M_i + 1) > 1$, however, it is necessary that $1 < z_i < n_i$. The observed pattern is identified here as $\mathbf{w}_i^{(0)}$, while the other possible patterns use superscripts $s = 1, 2, \ldots, M_i$.

The conditional probability of $\mathbf{w}_i$ given z_i and n_i is obtained by dividing (4) by (5), giving

$$P(\mathbf{w}_i^{(0)} \mid z_i, n_i) = \frac{\prod_{j=1}^{n_i} \exp\left[w_{ij}^{(0)}\ (\alpha_i + \mathbf{x}_{ij}'\boldsymbol{\beta})\right]}{\sum_{s=0}^{M_i} \prod_{j=1}^{n_i} \exp\left[w_{ij}^{(s)}\ (\alpha_i + \mathbf{x}_{ij}'\boldsymbol{\beta})\right]} = \frac{\exp\left[\left(\mathbf{X}_i'\mathbf{w}_i^{(0)}\right)'\boldsymbol{\beta}\right]}{\sum_{s=0}^{M_i} \exp\left[\left(\mathbf{X}_i'\mathbf{w}_i^{(s)}\right)'\boldsymbol{\beta}\right]}, \tag{6}$$

where $\mathbf{X}_i$ is the matrix of covariates for the quadrats in the plot i and has size $(n_i \times l)$, and l is the number of covariates. Moreover $\mathbf{x}_{ki}'\mathbf{w}_i^{(0)}$ represents the sum of the kth local-scale covariate (e.g., IPQ) in the quadrats where the vegetation is present in plot i. We have used that the denominators in (4) and (5) are equal, and that $\sum_j w_{ij}^{(s)} = z_i$ for every s; the explicit intermediate steps to achieve (6) are included in the Appendix.

Since this conditional distribution (6) is free from the α_i parameters, it follows that Z_i is sufficient for the large scale parameters (Casella and Berger 1990, p. 247). Estimates for $\boldsymbol{\beta}$ may be obtained using the maximum likelihood method. Inference on a particular continuous covariate $\mathbf{x}_k$ is done with the odds ratio calculated as $\exp(\beta_k)$ (Collett 1991, p. 263). Since the coefficients $\boldsymbol{\beta}$ in the conditional model in (6) are the same as those in the unconditional model in (3), the estimated odds ratio also applies to the unconditional probabilities, and thus also the interpretation of the effect of a covariate is valid for both models.

6 Application

The conditional logistic regression model was applied to the study of the presence of *V. vitis-idaea L.*, and of *C. arundinacea L.* using the data from the PSP. For the analysis, only quadrats situated on mineral soils were considered. As local-scale covariates, IPQ of Scots pine, Norway spruce, and birch (hairy and silver birch) were analyzed; these tree species represented the dominating species in the study area. The IPQ measurements were corrected for edge-effects in order to account for the ignored trees outside the plot. The correction consisted of an estimate of the expected IPQ outside the plot; see Kühlmann-Berenzon et al. (2005) for details on the method.

The analysis was carried out with the statistical package R v. 1.7.0 (Ihaka and Gentelman 1996), using the package survival v. 2.8-2, which contains a function for fitting conditional logistic models. As no previous information on possible values of c_T were available, the likelihood for the model including IPQ of spruce, pine, and birch was numerically maximized with the constrain that c_T should be positive. The obtained values of c_T for each species were used in the model during the fitting.

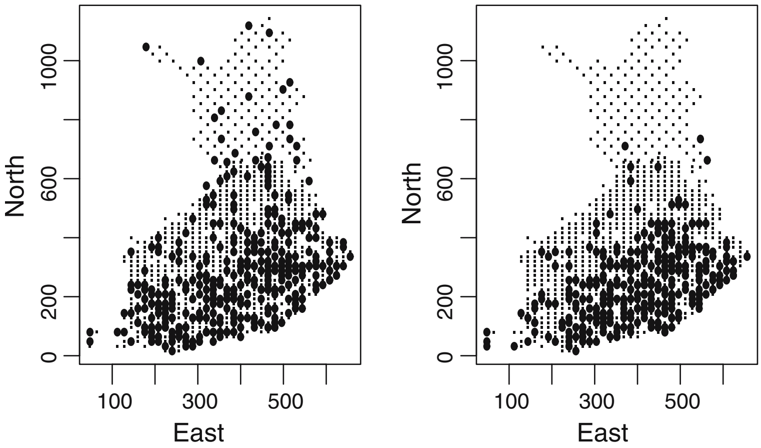
The goodness of fit of the models was tested with the likelihood ratio test at the 5% level of significance. Main effects were first included and subsequently interactions. Residual plots were also used to check for outliers, in particular Pearson residuals (Collett 1991, Chap. 5) and delta-beta graphs of Pregibon (1984); the latter show the change in the estimate of a coefficient when a plot is ignored. No particular outliers were observed in any of the cases.

6.1 Vaccinium vitis-idaea

According to Reinikainen et al. (2000), *V. vitis idaea*, known as cowberry, is one of the most frequent field layer species, and it is most abundant on relatively dry and poor sites with open canopy. Older Scots pine stands are known to be preferred habitat type for this berry. Cowberry was found in 410 plots in mineral soils where the plot also fulfilled the criteria that $1 < z_i < n_i$. The locations of the plots can be seen in Fig. 2a. In the analyzed plots, 89% of the trees were spruce, pine or birch. Some relevant additional statistics are given in Table 1.

Even though there was a larger number of spruce trees and their mean size was also larger than the other species, this species did not provide any influence, since its optimal c_T was 0.00, same as for birch. For pine, the parameter obtained was 5.80, corresponding to an influence range of 5.16 m. The average total IPQ of pine in a plot was 12.40, with a minimum of 0 and a maximum of 150.0.

The influence of pine was significant in the conditional logistic model, with a log likelihood of −599.53 (log likelihood of null model was −604.10, $P = 0.0025$). The estimated odds ratio indicated that with every increase of one unit in the sum of IPQ of Scots pine in the plot, the odds of finding cowberry in at last one quadrats increased



(a) *V. vitis-idaea* **(b)** *C. arundinacea*

Fig. 2 Location of plots analyzed: The *small dots* are the sampling grid shown in Fig. 1a. The *bigger dots* identify the clusters that included plots used in the conditional logistic analysis. The scale and origin of the maps are arbitrary

Table 1 *Vaccinium vitis-idaea*: Mean (SD) number of trees per plot (area = 300 m^2), DBH per tree, and influence range of optimal c_T

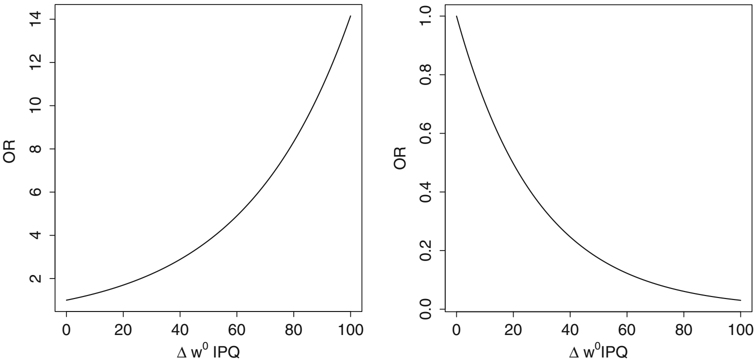
	Tree species					
	Pine		Spruce		Birch	
Number of trees per plot	6.88	(6.80)	10.92	(8.57)	4.25	(4.25)
DBH of tree (cm)	18.93	(7.39)	19.14	(6.76)	17.59	(6.30)
Influence range (m)	5.16		0.00		0.00	

3.0% (see Fig. 3a), with a 95% confidence interval of 1–5%; the estimated coefficient was 0.0265 with a standard error of 0.00901.

6.2 Calamagrostis arundinacea

Calamagrostis arundinacea is a type of small-reed grass most frequent in southern Finland and known to be abundant on relatively fertile site types (Reinikainen et al. 2000). For the analysis 350 plots were used that fulfilled the criteria of *C. arundinacea* not being present in all the quadrats. The spatial location of the analyzed plots can be seen in Fig. 2b. The optimal c_T for Scots pine, Norway spruce, and birch were 6.53, 3.05, and 0.00, and the corresponding influence ranges for the first two were 5.48 and 3.75 m. In average, the total IPQ from pine in the plot averaged 11.31 (range 0–130.6), and from spruce 8.97 (range 0–91.2); see Table 2.

The best possible model only included the influence potential of Norway spruce as significant covariate ($\hat{\beta} = -0.035$, SE$(\hat{\beta}) = 0.01034$); the log likelihood of the model was -502.98, and of the null model, -509.10 ($P < 0.001$). As the estimated coefficient was negative, this meant that an increase of one unit in the total IPQ of Norway spruce in the plot led to a decrease by 0.965 in the odds ratio, with a 95% confidence interval



(a) *V. vitis-idaea* : effect of pine **(b)** *C. arundinacea* : effect of spruce

Fig. 3 Estimated odds ratio according to the change in $\mathbf{w}^{(0)}$ IPQ

of (0.946, 0.985); see Fig. 3b. The effect of pine nor the interaction between pine and spruce were significant.

7 Discussion

Our objective was to develop a model for the presence of an understory species that estimated the local-scale variation but that also took into account large-scale factors. In order to avoid estimating the latter, we conditioned on the number of quadrats where the vegetation was present, which was sufficient in the statistical sense. Consequently, the conditional model was free from the large-scale parameters. The relative change in the odds ratio due to the local-scale covariates was the same in terms of the unconditional and the conditional models since the coefficients for both logistic models were the same.

The conditional model was fitted to the data from the permanent sampling plots of the 1985–1986 National Forest Inventory of Finland where the interest was on the influence of species of trees on two particular species of understory vegetation. The results obtained from analyzing *V. vitis-idaea* showed that an increasing influence of pine increased the odds of finding cowberry in the forest. In other words, the larger the number and size of the pine trees in the plot, the more likely cowberry is to be found. In the case of the grass *C. arundinacea*, an increasing influence of spruce caused a decrease in the odds ratio. The results from the fit provided information not available earlier, i.e., quantifiable effects of the tree species on the vegetation.

Table 2 *Calamagrostis arundinacea*: Mean (SD) number of trees per plot (area = 300 m^2), DBH per tree, and influence range of optimal c_T

	Tree Species					
	Pine		Spruce		Birch	
Number of trees per plot	7.59	(7.84)	10.11	(7.18)	3.80	(3.30)
DBH (cm)	19.42	(7.47)	18.75	(6.98)	17.99	(6.83)
Influence range (m)	5.48	3.75	0.00			

In order to conduct a thorough study of local-scale factors that affect the presence of understory species, other covariates beside the influence of the trees should be included. Independently of what quadrat-level covariates are used, the model presented is applicable to ecological studies of vegetation, where the aim is to estimate local-scale factors while taking into account the variation from large-scale environmental factors.

Acknowledgements The authors wish to thank Juha Heikkinen, Aila Särkkä, and Raisa Mäkipää for many useful comments; and to two anonymous referees for helping to improve the manuscript.

Appendix

$$
\begin{aligned}
P(\mathbf{w}_i^{(0)} \mid z_i, n_i) &= \frac{\prod_{j=1}^{n_i} \exp\left[w_{ij}^{(0)}\ (\alpha_i + \mathbf{x}_{ij}'\boldsymbol{\beta})\right]}{\sum_{s=0}^{M_i} \prod_{j=1}^{n_i} \exp\left[w_{ij}^{(s)}\ (\alpha_i + \mathbf{x}_{ij}'\boldsymbol{\beta})\right]} \\
&= \frac{\exp\left(\alpha_i\ \sum_j w_{ij}^{(0)}\right) \prod_j \exp\left(w_{ij}^{(0)}\ \mathbf{x}_{ij}'\boldsymbol{\beta}\right)}{\sum_s\left[\exp\left(\alpha_i \sum_j w_{ij}^{(s)}\right) \prod_j \exp\left(w_{ij}^{(s)}\ \mathbf{x}_{ij}'\boldsymbol{\beta}\right)\right]} \\
&= \frac{\exp(\alpha_i\ z_i) \prod_j \exp\left(w_{ij}^{(0)}\ \mathbf{x}_{ij}'\boldsymbol{\beta}\right)}{\sum_s\left[\exp(\alpha_i\ z_i) \prod_j \exp\left(w_{ij}^{(s)}\ \mathbf{x}_{ij}'\boldsymbol{\beta}\right)\right]} \\
&= \frac{\prod_j \exp\left(w_{ij}^{(0)}\ \mathbf{x}_{ij}'\boldsymbol{\beta}\right)}{\sum_s \prod_j \exp\left(w_{ij}^{(s)}\ \mathbf{x}_{ij}'\boldsymbol{\beta}\right)} \\
&= \frac{\exp\left[\left(\mathbf{X}_i'\mathbf{w}_i^{(0)}\right)'\boldsymbol{\beta}\right]}{\sum_{s=0}^{M_i} \exp\left[\left(\mathbf{X}_i'\mathbf{w}_i^{(s)}\right)'\boldsymbol{\beta}\right]}.
\end{aligned}
$$

References

Casella G, Berger RL (1990) Statistical inference. Duxbury Press, Belmont

Collett D (1991) Modelling binary data. Chapman and Hall/CRC, Florida

Ihaka R, Gentleman R (1996) R: a language for data Analysis and graphics. J Comput Grap Stat 5:299–314

Korpela L, Reinikainen A (1996) A numerical analysis of mire margin forest vegetation in South and Central Finland. Ann Bot Fenn 3:193–213

Kühlmann-Berenzon S, Heikkinen J, Särkkä A (2005) An additive edge correction for the influence potential of trees. Biom J 47:517–526

Kühlmann S, Heikkinen J, Särkkä A, Hjorth U (2001) Relating abundance of ground vegetation species and tree patterns using ecological field theory. In Rennolls K (ed) Proceedings of IUFRO 4.11 Conference: Forest Biometry, Modelling and Information science. University of Greenwich, Greenwich, UK. http:/cms1.gre.ac.uk/conferences/iufro/proceedings/

Kuuluvainen T, Pukkala T (1989) Effect of Scots pine seed trees on the density of ground vegetation and tree seedlings. Silva Fennica 23:159–167

Kuuluvainen T, Hokkanen TJ, Järvinen E, Pukkala T (1993) Factors related to seedling growth in a boreal Scots pine stand: a spatial analysis of a vegetation-soil system. Can J For Res 23:2101–2109

Mäkipää R, Heikkinen J (2003) Lanscape-scale changes in abundance of terricolous bryophyte and macrolichens in Finland. J Veg Sci, 16:497–508
McCullagh P, Nelder JA (1989) Generalized Linear Models, 2nd edn. Chapman & Hall, London
McKenzie D, Halpern CB (1999) Modeling the distributions of shrub species in Pacific northwest forests. For Ecol Manage, 114:293–307
Økland R, Rydgren K, Økland T (1999) Single-tree influence of understorey vegetation in a Norwegian boreal spruce forest. Oikos, 87:488–498
Pregibon D (1984) Data analytic methods for match case-control studies. Biometrics, 40:639–651
Reinikainen A, Mäkipää R, Vanha-Majamaa I, Hotanen JP (eds) (2000) Kasvit Muuttuvassa Metsäluonnossa, Kustannusosakeyhtiö Tammi, Helsinki
Saetre P (1999) Spatial patterns of ground vegetation, soil microbial biomass and activity in a mixed spruce-birch stand. Ecography, 22:183–192
Tonteri T, Mikkola K, Lahti T (1990) Compositional gradients in the forest vegetation. J Veg Sci, 1:691–698

Biographical sketches

Sharon Kühlmann-Berenzon received a M.Sc.Eng in Applied Environmental Measurement Techniques and a Ph.D. in Mathematical Statistics from Chalmers University of Technology, Sweden. Currently she works as a biostatistician at the Swedish Institute for Infectious Disease Control. Research interests have turned from spatial models for trees to spatial-temporal models for humans in connection with infectious diseases.

Urban Hjorth is Professor in Mathematical Statistics at Chalmers Univerity of Technology, Sweden. Research and teaching interests include prediction, model uncertainty, computer intensive methods, environmental, traffic science, meteorological and other applications. He is also author of *Computer intensive statistical methods*, Chapman and Hall/CRC, 1994 and some Swedish texts.