Disposition Models for the Analysis of Dynamic Changes in Forest-Ecosystems

Joachim Kötting, George E. Bonney and Wolfgang Urfer

Abstract. An important aim in forest-ecosystem investigation is to analyse the development of forest damages and to quantify changes in the damage-states over time of individual trees by influential factors. In addition to the ordinal measurement in such longitudinal studies one has to consider spatial correlations of the trees within an ecosystem. We present a practical method to include such dependency structures using logistic regression models. The strategy is to adopt the disposition model for correlated binary data (Bonney (1998)) and extend it to an ordinal-disposition-transitional model (ODT-model). This includes proportional-odds-transitional models (POT-model) as a special case, assuming independence over time and space given a markov model of first order. The ODT-model is used to analyse dynamic changes of damage in forest-ecosystems. The analysed data was sampled with infrared aerial photos by the Swiss Federal Institute for Forest, Snow and Landscape Research (Forschungsanstalt Wald, Schnee und Landschaft (WSL), Switzerland). A comparison of the independent case (POT-model) with the dependent case (ODT-model) shows that spatial correlations in forest-ecosystem should not be neglected.

Key words: Disposition model; Correlated categorical observation; Forest-ecosystem; Infrared aerial photos; Longitudinal data; Ordinal outcome; Proportional-odds model; Regression model; Spatial correlation; Transitional model.

1. Introduction

Many ecological projects are concerned with the historical development of organism and ecosystems in their environment over a prescribed period of time. So, an important aspect in the research of forest-ecosystems consists of watching for the development of forest-damages and to trace back temporal changes in the damage-states of individual trees to several covariates. The response-variable damaged-state is usually measured in ordinal categories. To account for time-dependencies of the individual trees such investigations are often designed as longitudinal studies, sampling repeated observations over time. This opens the possibility to model these changes with markov models, investigating the present state of the trees given their individual history.

Moreover, in the analysis of data measured on groups (clusters) of individuals, e.g. families, ecosystems, longitudinal measures, the possibility of dependence of the outcome under study can be understood as the probability that a group member has the outcome is not necessarily the same as that of an individual randomly selected from the population. We shall therefore speak of disposition to denote the tendency to manifest the outcome, and distinguish between the group disposition which is determined by group-specific covariates, and individual disposition which is determined by both the group- and individual covariates. Then, the disposition captures dependence, if any.

To analyse the effect of explanatory variables in markov models Bonney (1987) introduced the regressive logistic model for binary dependent outcomes. There, previous states were included in the set of regressors, obtaining conditional independent outcomes over time. An application of the regressive logistic model to forest-ecosystem data is described in Urfer (1993). An extension to ordinal regression analysis for modeling changes can be achieved by using the cumulative logit regression analysis. Ware et al. (1988) introduced such a model as a transitional model, refering to modeling transitions of individual states over time. Furthermore, they proposed to model first-order markov-chains even if the underlying markov-process is of higher order, using the methods of Stram et al. (1988). The first-order transitions are there modeled within the regression-matrix, and excess individual tree correlations over time are accounted by the empirical covariance matrix proposed by Stram et al. Reviews for the analysis of longitudinal analysis with ordinal outcomes are given in Diggle et al. (1994) and Fahrmeir and Tutz (1994).

Another problem occuring with ecosystem-data is the phenomenon of individual trees within an ecosystem forming a community. One has to expect inter-individual tree correlations within the ecosystem, apart from the dependencies of individuals over time. Spatial correlations in forest-ecosystems have been investigated by Quednau (1989) or

Fahrmeir and Pritscher (1996). Fahrmeir and Pritscher apply the GEE-method of Liang and Zeger (1986) and included global cross ratios.

In theory, the GEE-methods can be used to combine such spatial correlations with time dependencies. However, the high dimension of the working covariance matrix in most applications makes estimation often inpracticable. An interesting alternative is the full likelihood procedure for the regression analysis of correlated binary data proposed by Bonney (1995, 1998). By a decomposition of the logistic regression model one obtains there the ratio of dispositions of the independent to the dependent case. Different correlation structures, such as occuring in time- and spatial dependent data, can be included by nesting.

In this paper we develop an extension of the nested version of Bonney's disposition model to ordinal outcomes and apply it to our data. The project that motivated the study is introduced in section 2. This is followed by a description of the statistical issues raised by the data. In the fourth section the statistical models are described. Non-stationary changes over time of the ordinal response-variable damaged-state are explained by a modified proportional-odds-transitional model in section 4.1. The nested disposition model is introduced in section 4.2. and extended to an ordinal-disposition-transitional model in section 4.3. In chapter 5 applications of the model to the project are described for the cases of independence given a markov-model and dependence over time and space.

2. The project "Monitoring of forest damage with infrared aerial photos 1:3000"

Preceding researches of the Swiss Federal Institute for Forest, Snow and Landscape Research ("Forschungsanstalt Wald, Schnee und Landschaft" (WSL)), Birmensdorf, Switzerland, for the extent and the development of forest damage in Switzerland have shown, that the extent and intensity of forest damage varies between regions, locations and species of the trees. Although, there are areas with different degrees of damage, these areas often show a same behaviour with regard to improvements and deteriorations over time.

Besides, the changes of damage state can vary strongly within a stand of trees. So, it is thoroughly possible to observe deteriorations and improvements of neighbouring trees within a forest-ecosystem in the same time period. Therefore, one must keep in mind, that the development of forest damage over time depends essentially on changes of the individual trees.

Such reflections are the starting-points for the project "Monitoring of forest damage with infrared aerial photos 1:3000" directed by Dr. F. H. Schwarzenbach and Dr. B. Oester of the WSL, Switzerland. To observe the dynamic changes of the damage state over time, infrared aerial photos on a scale 1:3000 of the forests were taken once a year in July/August from 1984 to 1991 in the following three locations in Switzerland: Altdorf (canton Uri, 536 - 870 m above sea-level), Flims (canton Graubünden, 920 - 1010 m above sea-level) and Zofingen (canton Aargau, 540 - 615 m above sea-level).

The aim of the project was to analyse the development of forest damage through repeatedly observed individual trees and tracing back changes, of different proportions, to several covariates. The judgement of the damaged-states of the individual trees results from interpretation keys, developed separately for each species by Dr. Oester, with five ordinal categories: 0 - healthy; 1 - slightly damaged; 2* - moderately damaged; 3 - strongly damaged; 4 - dead. In doing so, the appearance of the whole tree on the infrared aerial picture is judged, with the most important attributes concerned with the mass of the conifer-needles or leaves (by tone of the tree on the aerial infrared picture, marbling colouration of the tree, crown transparency). The interpretation keys are given in Oester (1991). The classification of the damages is somewhat subjective, since present biological knowledge cannot exactly distinguish healthy and damaged trees, including their damaged-states. So, one does not know the real boundary between healthy and damaged trees (if they lost 5%, 10% or perhaps 15% of its leaves or needles). The categories "moderately damaged", "strongly damaged" and "dead" are summarized to a class ‘2’ - "severely damaged", because for the present statistical analysis, in these groups are not sufficiently enough observations.

Apart from the damaged-state, further variables have been measured on different scales, which should describe the living conditions of the individual trees in the ecosystem. These are investigated as influential factors. The inclusion of these covariates is restricted by the fact that only the aerial pictures were to be used. By close interdisciplinary cooperation of biologists and statisticians, the following are finally chosen for the statistical analysis:

Variable	Code	Meaning
Altitude above sea-level (grouped)	0	< 601 m
	1	601-700 m
	2	701-800 m
	3	801-900 m
	4	901-1000m
	5	> 1000m
Height of the tree (grouped)	1	< 15 m
	2	15-30 m
	3	> 30 m
Social status of the tree within a stand	1	uppermost stratum of the canopy, dominant position
	2	uppermost stratum, rather dominant position
	3	uppermost stratum, codominant position
	4	uppermost stratum, subdominant position
	5	middle stratum of the canopy
	6	lowest stratum of the canopy
Degree of development of the stand	2	very thin stem
	3	small diameter of stem
	4	medium diameter of stem
	5	large diameter of stem
	6	mixed 'degree of development' within a stand

The location was transformed into two dummy-variables "location1" and "location2":

Location	**location1**	**location2**
Altdorf	0	0
Flims	0	1
Zofingen	1	0

Further covariates were removed after careful consideration and discussion among biologists and statisticians before the statistical analysis, for example such as exposition, forest type or mean height of the canopy. A detailed description of the variables and of the species specific interpretation keys is given in Oester (1991).

3. Some statistical issues raised by the given project

The aim of a statistical analysis is to quantify time-dependend individual changes in the "damaged-state" of individual trees in connection with several covariates. The response-variable damaged-state is ordinal, so that classical linear regression analysis seems to be unsuitable. The classification of the states is somewhat subjective and it would be advantageous, to have a statistical model that is invariant against shifting of the damaged-state-boundaries. In the data to be analysed, the last three categories were combined, so an invariant model is definitely of interest. Such considerations lead to proportional-odds models (McCullagh (1980), McCullagh and Nelder (1989)). Assuming the proportional-odds-assumption is fulfilled, there are additional difficulties, arising from spatial and temporary dependencies of the response-variables.

First, one has to take into consideration the longitudinal nature of the observations on individual trees and therefore the need to account for intra-individual tree correlations. To model the changes of the state, transitional models are a possible solution. For our data these can be assumed to be time-discrete and equidistant. However, the possibilities to model transitions are restricted by the fact that forests form ecosystems. There can be non-stationary behaviour from sources such as a unique historical development of an ecosystem or long-term adaptations to changing conditions of the environment. Furthermore, a high order of the underlying markov-chain is possible, because of time-lags between a stimulus and a visible reaction of the ecosystem or by long-term adaptations.

Second, the trees within an ecosystem form a community, having inter-individual tree correlations of the trees within this ecosystem. Trees from different ecosystems can be assumed to be independent. These spatial correlations should be included in the statistical modeling. The relations of dependencies between locations, trees and time is shown in figure 1.

As for the attribute of forests being ecosystems, compare for example Schwarzenbach (1987), Urfer (1993) or Urfer et al. (1994). The following statistical analysis is restricted to spruces, because only they are sufficiently represented in all three regions. The sample size amounts to 321.

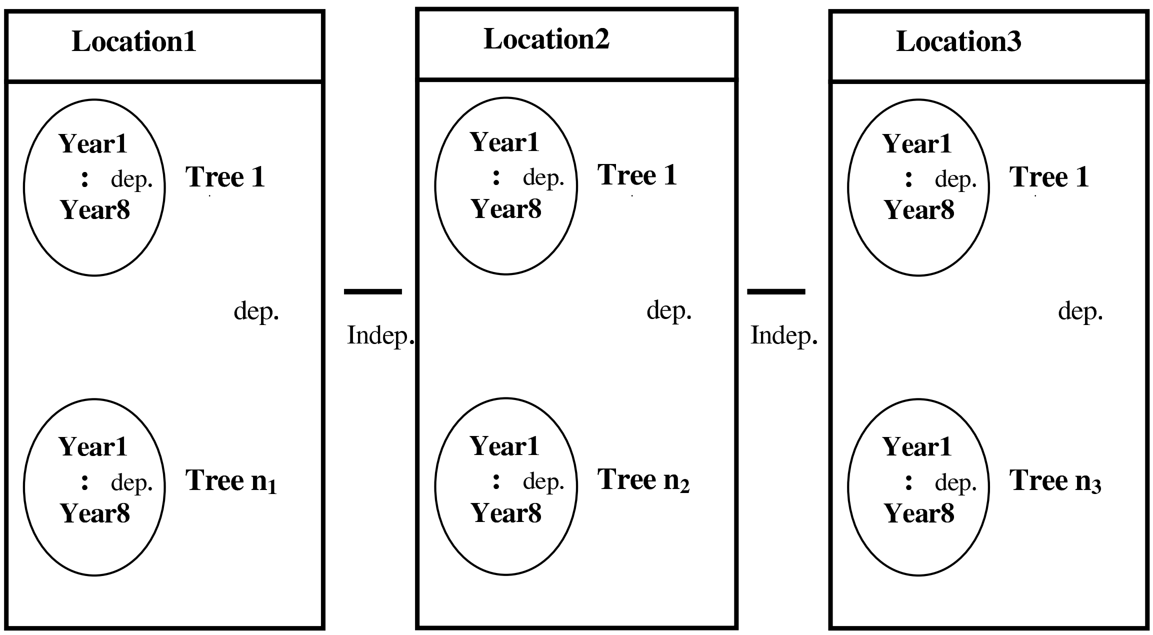

Figure 1. Dependence structures between locations, trees and time

4. Statistical models

4.1 Proportional-odds-transitional-model

To take account of ordinal outcomes and to model changes we will use a cumulative logit regression model and define it as a transitional model. The previous states of the response-variable are then included into the set of regressors, as proposed by Bonney (1987) or Ware et al. (1988). For the time-points i=1,..,D let $y \in \{0,1,..q\}$ represent the ordinal response-variable having q+1 ordinal categories. The vector of covariates is given by $\mathbf{x} \in \mathbb{R}^p$. Because of the possibility of non-stationarity over time we define a covariate vector $\mathbf{T} \in \mathbb{R}^{D-2}$, belonging to the time-points i=3,..,D as indicator variables. Because we are dealing with transitions of first order we have a look to D-1 time-points. The transition between the first and the second time-point is used as a reference category to avoid identyfiability problems. Therefore we need only to model D-2 transitions. Thus we define

$$\mathbf{T} := \begin{cases} (0,..,0,1,0,..,0) \in \{0,1\}^{D-2}, & \text{'1' at position '(i-2)'} \\ (0,..,0) \in \{0\}^{D-2}, & \text{if } \; i = 2 \end{cases}.$$

The additional regression-vector describing the previous damaged-states at time i-1 is represented by indicator-variables, too. At time i = 3,..,D , we have

$$\mathbf{b} := \begin{cases} (0,..,0,1,0,..,0) \in \{0,1\}^{q}, & \text{'1' at position 'm', if } \; y_{i-1} = m, \; m = 0,..,q-1 \\ (0,..,0) \in \{0,1\}^{q}, & \text{if } y_{i-1} = q \end{cases}$$

Then, a first order proportional-odds-transitional model at times i = 3,..,D is defined by

$$\ln \frac{P(y \le u \mid \mathbf{x}, \mathbf{T}, \mathbf{b})}{1 - P(y \le u \mid \mathbf{x}, \mathbf{T}, \mathbf{b})} = \lambda_u + \mathbf{x}^T \beta + \mathbf{T}^T \tau + \mathbf{b}^T \kappa, \quad u = 0,1,..,q\text{-}1$$

with thresholds $\lambda_0 < \lambda_1 < ... < \lambda_{q-1}$ and $\beta \in \mathbb{R}^p$, $\tau \in \mathbb{R}^{D-2}$, $\kappa \in \mathbb{R}^q$ as unknown parameters.

It should be emphasized that by including **b**, we are not comparing the probability distributions of the response-variable for different time-points and different values in the covariates **x**, but we are modeling a comparison between the rows of transition-matrices (of first order). The vector of parameters κ describes the effects of the previous states on the present state; τ can be interpretated as the year effect in comparison with the baseline year category; β is the vector of regression-coefficient. By using proportional-odds, these relative effects correspond to the logarithm of cumulative odds ratios and they are independent of the response-category. McCullagh (1980), McCullagh and Nelder (1989) and Agresti (1990) provide more detailed accounts of proportional-odds models.

This model can be used to cover time-dependent changes in the individual damaged-states of trees. The proposed first-order model can be generalized to higher order models by extensions of **b** to the farther states. However, problems often arise from linear dependencies and the increasing number of parameters.

4.2 Disposition model for correlated response-variables

The model proposed in the preceding section account for the longitudinal structure. However, figure 1 shows that we also need to consider dependence among trees. To do this, we now introduce the nested version of the disposition model for correlated binary outcomes developed by Bonney (1998). The n_k trees within the ecosystem in location k, k = 1,..,K, are assumed to be spatially dependent, because they share common attributes observed as in location-specific covariates $\mathbf{Z}=(Z_1,..,Z_q)^T$ or unobserved as in latent or missing variables. Simultaneously, we consider measures on a given tree as a subgroup nested within location, where i=1,.., t_{n_k} over the time repeated observations of the single trees are nested within. In this section we are concerned with the outcome of y dichotomized as binary response-variable $y_j^b=1$, if tree j is damaged ($y\in\{1,2\}$) and $y_j^b=0$ if the j-th tree is observed healthy. Now, we distinguish three types of covariates: location-specific covariates $\mathbf{Z}=(Z_1,.., Z_q)^T$, tree-specific covariates $\mathbf{x}=(x_1,.., x_p)^T$ and time-specific covariates $\mathbf{T_{kij}}$. In analogous, we differentiate location-disposition δ_{k0}, tree-disposition δ_{kj} and individual tree disposition over time δ_{kij}. As for disposition see also the introduction of this paper. δ_{kij} is considered to be the probability observing the outcome $y_j^b=1$ for the j-th tree in location k at time i. δ_{k0} resp. δ_{kj} are the location- resp. the tree-specific tendency to observe a damaged tree. The dispositions also account for the strength of the correlation over time not accounted for by the observed covariates. Now, as a measure of dependence Bonney (1998) defines α's as ratios of disposition of the independent to the dependent cases:

$$\alpha_{k0} = \frac{\text{disposition to tree damage in location k assuming independence}}{\text{disposition to tree damage in location k assuming dependence}} =: \frac{\mu_{k0}}{\delta_{k0}}$$

$$\Leftrightarrow \delta_{k0} = \frac{\mu_{k0}}{\alpha_{k0}} \quad ,$$

with α_{k0} as measure of dependence within an ecosystem and analogous $\alpha_{kj} = \mu_{kj}/\delta_{kj}$ as measure of dependence within a tree over time. Then $\alpha_{k0} = 1$ means independence of the trees within location k and $\alpha_{kj} = 1$ means there is no excess dependence across time over that between trees at the same location. There can not be within independence when there is between dependence. For regression analysis some function of δ_{k0} can be related to **Z** and some function of δ_{kj} can be related to $\mathbf{x_{kj}}$. Modeling the effects of the covariates results from the following decomposition for the individual disposition over time:

$$\ln\frac{\delta_{kij}}{1-\delta_{kij}} = \ln\frac{\delta_{kj}}{1-\delta_{kj}} + \tau^T\mathbf{T_{kij}}$$

$$=\ln\frac{\mu_{k0}}{1-\mu_{k0}} +(\ln\frac{\delta_{k0}}{1-\delta_{k0}} -\ln\frac{\mu_{k0}}{1-\mu_{k0}})+(\ln\frac{\mu_{kj}}{1-\mu_{kj}} -\ln\frac{\delta_{k0}}{1-\delta_{k0}})+(\ln\frac{\delta_{kj}}{1-\delta_{kj}} -\ln\frac{\mu_{kj}}{1-\mu_{kj}})+W_{kij}(\mathbf{T_{kij}})$$

$$= M_{k0}(\mathbf{Z_{k0}}) + D_{k0}(\mathbf{Z_{k0}}) + M_{kj}(\mathbf{x_{kj}}) + D_{kj}(\mathbf{x_{kj}}) + W_{kij}(\mathbf{T_{kij}})$$

$$= \gamma^T\mathbf{Z_{k0}} + \zeta^T\mathbf{Z_{k0}} + \beta^T\mathbf{x_{kj}} + \rho^T\mathbf{x_{kj}} + \tau^T\mathbf{T_{kij}} \; ,$$

where

$M_{k0}(\mathbf{Z_{k0}}) = \ln\frac{\mu_{k0}}{1-\mu_{k0}}$ -is the location logit mean risk

$D_{k0}(\mathbf{Z_{k0}}) = (\ln\frac{\delta_{k0}}{1-\delta_{k0}} - \ln\frac{\mu_{k0}}{1-\mu_{k0}})$ -is the excess location logit disposition because of dependencies between the trees

$M_{kj}(\mathbf{x_{kj}}) = (\ln\frac{\mu_{kj}}{1-\mu_{kj}} - \ln\frac{\delta_{k0}}{1-\delta_{k0}})$ -is the excess on the logit scale of mean risk for the j-th tree of the location disposition

$D_{kj}(\mathbf{x_{kj}}) = (\ln\frac{\delta_{kj}}{1-\delta_{kj}} - \ln\frac{\mu_{kj}}{1-\mu_{kj}})$ -is the excess on the logit scale of the dispositions over time that cannot be explained by location logit mean risk and the differences in μ_{kj}

$W_{kij}(\mathbf{T_{kij}})$ -is a function of the individual specific covariates

Then the values of the dispositions and correlations can be calculated from the parameters as:

$$\delta_{k0}=\frac{1}{1+e^{-\{M_{k0}(\mathbf{Z_{k0}})+D_{k0}(\mathbf{Z_{k0}})\}}}, \qquad \alpha_{k_0}=\frac{1+e^{-\{M_{k0}(\mathbf{Z_{k0}})+D_{k0}(\mathbf{Z_{k0}})\}}}{1+e^{-\{M_{k0}(\mathbf{Z_{k0}})\}}},$$

$$\delta_{kj}=\frac{1}{1+e^{-\{M_{k0}(\mathbf{Z_{k0}})+D_{k0}(\mathbf{Z_{k0}})+M_{kj}(\mathbf{x_{kj}})+D_{kj}(\mathbf{x_{kj}})\}}},$$

$$\alpha_{kj}=\frac{1+e^{-\{M_{k0}(\mathbf{Z_{k0}})+D_{k0}(\mathbf{Z_{k0}})+M_{kj}(\mathbf{x_{kj}})+D_{kj}(\mathbf{x_{kj}})\}}}{1+e^{-\{M_{k0}(\mathbf{Z_{k0}})+D_{k0}(\mathbf{Z_{k0}})+M_{kj}(\mathbf{x_{kj}})\}}},$$

$$\delta_{kij}=\frac{1}{1+e^{-\{M_{k0}(\mathbf{Z_{k0}})+D_{k0}(\mathbf{Z_{k0}})+M_{kj}(\mathbf{x_{kj}})+D_{kj}(\mathbf{x_{kj}})+W_{kij}(\mathbf{T_{kij}})\}}}$$

Using a moment series representation of the shared risk within the locations and within the trees (compare Bonney (1998)) the joint likelihood function for all three locations of the proposed model is given by:

$$P(Y_1^b = y_1^b, \ldots, Y_n^b = y_n^b \mid \mathbf{Z_{k0}}, \mathbf{x_{kj}}, \mathbf{T_{kij}}) = \prod_{k=1}^{K} \Bigg[(1-\alpha_{k_0}) \prod_{j=1}^{n_k} \prod_{i=1}^{t_{n_k}} (1 - y_{kij}^b) + \alpha_{k0} \prod_{j=1}^{n_k} \Bigg\{ (1-\alpha_{kj}) \prod_{i=1}^{t_{n_k}} (1 - y_{kij}^b) + \alpha_{kj} \prod_{i=1}^{t_{n_k}} \delta_{kij}{}^{y_{kij}^b} (1-\delta_{kij})^{1-y_{kij}^b} \Bigg\} \Bigg]$$

4.3 Disposition model for the analysis of changes in ordinal outcomes

For ordered outcome with three or more categories we propose a combination of the models of section 4.1 and 4.2. As in transitional models, the vector **b** is included in the set of the time-specific covariates to model the changes of damage states. Furthermore, the ordinal nature of the response-variable can be modeled by replacing δ_{kij} by δ^*_{kiju}, where $\delta^*_{kiju} = P(y_{kij} \le u \mid \mathbf{Z_{k0}}, \mathbf{x_{kj}}, \mathbf{T_{kij}}, \mathbf{b_{kij}})$, u = 0,1,..q. Thus, one obtains an ordinal-disposition-transitional model for spatial- and time dependent data. The case of the cumulative logit will be considered here:

$$\ln \frac{\delta^*_{kiju}}{1-\delta^*_{kiju}} = \lambda_u + \gamma^T \mathbf{Z_{k0}} + \zeta^T \mathbf{Z_{k0}} + \beta^T \mathbf{x_{kj}} + \rho^T \mathbf{x_{kj}} + \tau^T \mathbf{T_{kij}} + \kappa^T \mathbf{b_{kij}} =: \theta_{\mathbf{kij}}(\mathbf{u})$$

Defining $s_{kiju} := \begin{cases} 1, & y_{kij} = u \\ 0, & \text{otherwise} \end{cases}$ the likelihood function of the used model is given by

$$P(Y_1=y_1,.., Y_n=y_n \mid \mathbf{Z_{k0}}, \mathbf{x_{kj}}, \mathbf{T_{kij}}, \mathbf{b_{kij}}) = \prod_{k=1}^{K}\Bigg[(1-\alpha_{k_0}) \prod_{j=1}^{n_k} \prod_{i=1}^{t_{n_k}} (1-s_{kij0}) +$$

$$\alpha_{k0} \prod_{j=1}^{n_k} \left\{ (1-\alpha_{kj}) \prod_{i=1}^{t_{n_k}} (1-s_{kij0}) + \alpha_{kj} \prod_{i=1}^{t_{n_k}} \prod_{u=0}^{q} \left(\frac{e^{\theta_{kij}(u)}}{1+e^{\theta_{kij}(u)}} - \frac{e^{\theta_{kij}(u-1)}}{1+e^{\theta_{kij}(u-1)}} \right)^{s_{kiju}} \right\} \Bigg]$$

5. Analysis of data on dynamic changes of damage in forest-ecosystems

To analyse the data of the project "Monitoring of forest damage with infrared aerial photos 1:3000" introduced in section 2, we use the following types of covariates

- ecosystem-specific covariates: $\mathbf{Z_k} := (Z_{loc1}, Z_{loc2})^T :=$ (location1, location2)^T;
- tree-specific covariates: $\mathbf{x} = (x_{Hgt}, x_{Soc})$, describing the height of the trees and the social status as shown in section 2;
- time-specific covariates: $\mathbf{T_{kij}} = (T_{k,1986,j},...,T_{k,1991,j})^T$ with $T_{kij}=1$ if year i is the actual year of the observation, i=1986,..,1991, and $T_{kij}=0$ otherwise; and the previous state $\mathbf{b_{kij}}$ with

$$\mathbf{b_{kij}} = \begin{cases} (1,0)^T & \text{healthy state} \\ (0,1)^T \text{ for a} & \text{slightly damaged state} \quad \text{of the } j-\text{th tree in the previous year } i-1 \\ (0,0)^T & \text{severely damaged state} \end{cases}$$

For a comparison we fit a first-order proportional-odds-transitional-model (POT-model) as in section 4.1 and an ordinal-disposition-transitional-model (ODT-model) as in section 4.3. In both cases the covariates ‘altitude above sea-level’ and ‘degree of development’ are removed during a variable selection using Akaike’s Information Criterion. The estimates are obtained by a likelihood optimization software MULTIMAX developed at the Fox Chase Cancer Center, Philadelphia, USA.

5.1 Application of the proportional-odds-transitional-model (POT-model)

First, we fit a POT-model as the special case of the ordinal-disposition-transitional-model, assuming independence, given a first-order markov model for the data. This is equivalent to the usual cumulative logit regressive or transitional models, and can be fit with standard-software (for example SAS PROC-Logistic). One obtains this POT-model in the proposed ODT-model, by setting $D_{k0}(\mathbf{Z_{k0}})=0$ and $D_{kj}(\mathbf{x_{kj}})=0$. The other functions are given as

$$M_{k0}(\mathbf{Z_{k0}}) = \gamma^T \mathbf{Z_{k0}} = \gamma_{00} + (\gamma_{loc1}, \gamma_{loc2})\, \mathbf{Z_k}$$

$$M_{kj}(\mathbf{x_{kj}}) = \beta^T \mathbf{x_{kj}} = (\beta_{Hgt}, \beta_{Soc})\, \mathbf{x_{kj}}$$

$$W_{kij}(\mathbf{T_{kij}}, \mathbf{b_{kij}}) = \lambda_1 + (\tau_{1986},\ldots, \tau_{1991})\, \mathbf{T_{kij}} + (\kappa_0,\kappa_1)\, \mathbf{b_{kij}}$$

In this model γ_{00} can be regarded as the usual threshold λ_0 separating healthy and damaged trees, whereas λ_1 represents the threshold separating severely damaged trees from the other ones. The result of maximum-likelihood estimation is shown in table 2. Level-α-tests are obtained by using Wald χ^2-tests. Here, significant estimates at the level α=5% can be observed for the parameters κ_0, κ_1, τ_{1987}, τ_{1988}, τ_{1989}, β_{Soc} and γ_{loc2}. The AIC value is calculated to 2121.0.

The high positive value of $\hat{\kappa}_0$ means that the odds of staying in a healthy state is very much higher ($e^{7.21}$-times) than the odds of a transition from damaged-state '2' into a damaged-state '0'. The positive value of $\hat{\kappa}_1$ indicates a higher probability for a severely damaged tree to stay in damage-class '2' relative to an one-year transition from state '1' into state '2'.

The negative values for $\hat{\tau}_{1987}$ and $\hat{\tau}_{1988}$ indicates more transitions into the damaged states in the one-year periods 1986/1987 and 1987/1988 than in 1984/1985, whereas the positive $\hat{\tau}_{1989}$ shows less deteriorations or more improvements for the period 1988/1989

Table 2: POT-model (independence given a markov chain of order one)**:** Estimates of the parameters (standard error)

$\hat{\lambda}_0$	$\hat{\lambda}_1$	$\hat{\gamma}_{loc1}$	$\hat{\gamma}_{loc2}$	$\hat{\beta}_{Hgt}$	$\hat{\beta}_{Soc}$	$\hat{\tau}_{1986}$	$\hat{\tau}_{1987}$	$\hat{\tau}_{1988}$	$\hat{\tau}_{1989}$	$\hat{\tau}_{1990}$	$\hat{\tau}_{1991}$	$\hat{\kappa}_0$	$\hat{\kappa}_1$
-6.13	-1.80	-0.03	-0.53	-0.14	0.40	0.09	-1.36	-0.58	0.48	-0.06	-0.06	7.21	3.76

(0.46)	(1.01)	(0.23)	(0.19)	(0.13)	(0.08)	(0.21)	(0.22)	(0.22)	(0.23)	(0.21)	(0.21)	(0.24)	(0.16)

in comparison with the period 1984/1985. The other time-periods are not significantly different from the reference-period 1984/1985.

The significant positive value entirely for $\hat{\beta}_{Soc}$ means that the more dominant a spruce is within its social status the higher is the odds to change into the damaged states '1 or 2' respectively '2'. $\hat{\gamma}_{loc2} = -0.53$ indicates more changes into the damaged classes in Flims than in Altdorf, whereas for the comparison between Altdorf and Zofingen and for the height of the trees no significant differences can be shown.

5.2 Application of the ordinal-disposition-transitional-model (ODT-model)

Now, we are assuming dependencies over trees within an ecosystem and that a Markov-chain of order one will not fully describe time-dependencies. To consider these additional intra-individual tree correlations over time and the inter-individual tree correlations within locations we are using the nested model. The fitted model has the form

$$\ln \frac{\delta^*_{kiju}}{1-\delta^*_{kiju}} = M_{k0}(\mathbf{Z_{k0}}) + D_{k0}(\mathbf{Z_{k0}}) + M_{kj}(\mathbf{x_{kj}}) + D_{kj}(\mathbf{x_{kj}}) + W_{kij}(\mathbf{T_{kij}}, \mathbf{b_{kij}})$$

with

$M_{k0}(\mathbf{Z_{k0}}) = \gamma_{00}$

$D_{k0}(\mathbf{Z_{k0}}) = \zeta^T \mathbf{Z_{k0}} = (\gamma_{loc1}, \gamma_{loc2})\, \mathbf{Z_k}$

$M_{kj}(\mathbf{x_{kj}}) = \beta^T \mathbf{x_{kj}} = (\beta_{Hgt}, \beta_{Soc})\, \mathbf{x_{kj}}$

$D_{kj}(\mathbf{x_{kj}}) = \rho_0$

$W_{kij}(\mathbf{T_{kij}}, \mathbf{b_{kij}}) = \lambda_1 + (\tau_{1986},\ldots, \tau_{1991})\, \mathbf{T_{kij}} + (\kappa_0,\kappa_1)\, \mathbf{b_{kij}}$

The result of maximum likelihood estimation is given in table 3. An AIC value of 2095.7

indicates a better fit of the ODT-model than for the POT-model, which assumes independence (AIC=2121.0). The correlations of the trees within an ecosystem are contained in $D_{k0}(\mathbf{Z_{k0}})$. However, the strength of dependencies within location can not be determined, because $M_{k0}(\mathbf{Z_{k0}})$ and $D_{k0}(\mathbf{Z_{k0}})$ can not be separated for the same covariates of $\mathbf{Z_{k0}}$, due to the small numbers of location (K=3). Nevertheless, these dependencies are included because of modeling $\delta_{k0} = \mu_{k0} / \alpha_{k0}$, and in doing so, they influence the estimates of the parameters and variances. In contrast to the POT-model, using Wald χ^2-tests at the level α=5%, here is no longer a significant influence in the changes of damaged-states between Flims and Altdorf. Furthermore, we cannot show significant different transitions between Zofingen and Altdorf.

$\hat{\rho}_0$ shows no significant value for the excess of the mean risk of the trees explained by time dependencies not modeled by the first-order markov chain at the α-level of 5%. Therefore, it can be assumed, that for the analysed data-set a markov chain of order one describes the intra-individual correlation over time entirely.

The tree-specific covariate having significant influence with regard to the changes in damaged-states is the social-status variable. The positive value indicates higher odds for transitions into the damaged states '1 or 2' or '2' the higher the stratum of the canopy and the more dominant the tree within a stand is.

The positive values for $\hat{\kappa}_0$ resp. $\hat{\kappa}_1$ describe, that there are higher probabilities to observe healthy states of a tree, having healthy resp. slightly damaged states for the same tree one year before than for trees which are oberved as severely damaged trees in the previous year. The negative $\hat{\tau}_{1987}$-parameter describes more changes into the damaged states in the period 1986/1987 than in the period 1984/1985.

Table 3: ODT-model (nested correlations): Estimates of the parameters (standard error)

$\hat{\gamma}_{00}$	$\hat{\lambda}_1$	$\hat{\zeta}_{loc1}$	$\hat{\zeta}_{loc2}$	$\hat{\beta}_{Hgt}$	$\hat{\beta}_{Soc}$	$\hat{\rho}_0$	$\hat{\tau}_{1986}$	$\hat{\tau}_{1987}$	$\hat{\tau}_{1988}$	$\hat{\tau}_{1989}$	$\hat{\tau}_{1990}$	$\hat{\tau}_{1991}$	$\hat{\kappa}_0$	$\hat{\kappa}_1$
-5.71	-1.82	-0.22	-0.30	-0.27	0.58	-0.28	-0.11	-1.42	-0.62	0.38	0.40	-0.14	7.83	3.92

(1.36)	(1.01)	(0.25)	(0.32)	(0.37)	(0.21)	(0.49)	(0.49)	(0.31)	(0.32)	(0.40)	(0.39)	(0.38)	(0.41)	(0.19)

A comparison with the POT-model shows, that not considering the correlation of trees within an ecosystem leads to significant values for the parameters of the covariates Z_{loc1}, T_{1988} and T_{1989} in contrast to modeling the ODT-model. A markov chain of order one seems to be sufficient for time-dependencies over the individual trees.

6. Discussion

For the analysis of dynamic changes of the damaged-states in forest ecosystems with regression models one has to take into consideration the ordinal measurement of the response-variable, intra-individual tree correlations over time, presumably with an underlying non-stationary markov-chain of high order, and inter-individual tree correlations of the trees forming an ecosystem. In this paper we developed an ordinal disposition model for changes, which take account of such feature simultaneously, by extending the nested disposition model of Bonney (1998) to ordinal-disposition-transitional models. A great advantage of this model is its applicability to real data with such complex dependency-structures.

In comparison with the POT-model of first order, assuming no further correlations over time and space, the ODT-model shows a better fit of the data, by Akaike's Information Criterion. The spatial correlation within the ecosystem seems not to be negligible. The small number of locations in the given project (K=3) causes numerical problems. So, it is questionable that the three locations are sufficient to assume an approximated normal distribution for the Wald χ^2-tests. In such cases maybe it could be interesting to develop methods for calculating the true underlying probability distribution for testing hypothesis. Usually, the number of locations in forest-ecosystem researches are much more higher. Besides, an increasing number of locations open the possibility to estimate M(**Z**) and D(**Z**) at the same time and to calculate the values of μ and α. Therefore, further researches of the model will be worthwhile.

Furthermore, in the development of the disposition model for correlated binary outcomes genetic applications were discussed by Bonney (1995, 1998), in which a large number of families were used. An application of the developed models to the analysis of human genetics are in the development.

Acknowledgements. This work was motivated by the research project 'Complexity Reduction in Multivariate Data Structures', supported by the Deutsche Forschungsgemeinschaft (DFG), Sonderforschungsbereich 475.

7. References

Agresti, A. (1990), *Categorical Data Analysis*, John Wiley & Sons, New York.

Bonney, G. E. (1987), "Logistic Regression for Dependent Binary Observations", *Biometrics*, 43, 951-973.

Bonney, G.E. (1995), "Some New Results on Regressive Models in Family Studies", in: *1995 Proceedings of the Biometric Section*, *American Statistical Association*, 177-182.

Bonney, G. E. (1998), *Regression Analysis of Disposition to Correlated Binary Outcomes*, Department of Biostatistics Technical Report No. 1998-001, Fox Chase Cancer Center, Philadelphia, PA, USA.

Diggle, P. J., Liang, K. L., and Zeger, S. L. (1994), *Analysis of Longitudinal Data*, New York, Oxford Press.

Fahrmeir, L., and Tutz, G. (1994), *Multivariate Statistical Modelling Based on Generalized Linear Models*, New York, Springer.

Fahrmeir, L., and Pritscher, E. (1996), "Regression analysis of forest damage by marginal models for correlated ordinal responses", *Environmental and Ecological Statistics*, 3, 257-268.

Liang, K.-Y., and Zeger, S. (1986), "Longitudinal data analysis using generalized linear models", *Biometrika*, 73, 13-22.

McCullagh, P. (1980), "Regression models for ordinal data (with discussion)", *Journal of the Royal Statistical Association, Series B*, 42, 109-142.

McCullagh, P., and Nelder, J. A. (1989), *Generalized Linear Models, 2nd edition*, Chapman and Hall, London.

Oester, B. (1991), *Erfassen der Waldschaden-Entwicklung anhand von großmassstäblichen Infrarot-Farbluftbildern*, Dissertation, Eidgenössische Technische Hochschule, Zürich.

Quednau, H. D. (1989), "Statistische Analyse von Waldschadensdaten aus Luftbildern mit Berücksichtigung von Nachbarschaftseffekten.", *Forstwissenschaftliches Centralblatt*, 108, 96-102.

Schwarzenbach, F. H. (1987), *Grundlagen für die Entwicklung einer allgemein anwendbaren Strategie zur Lösung ökologischer Probleme*, Eidgenössische Anstalt für das forstliche Versuchswesen, Birmensdorf (CH), Berichte Nr. 293.

Stram, D. O., Wei, L. J., and Ware, J. H. (1988), "Analysis of repeated ordered categorical outcomes with possibly missing observations and time-dependent covariates", *Journal of the American Statistical Association*, 83, 631-637.

Urfer, W. (1993), "Statistical Analysis of Climatological and Ecological Factors in Forestry; Spatial and Temporal Aspects", in Barnett, V., Turkman, F. K. (eds.): *Statistics for the Environment* , Wiley, New York, 335-345.

Urfer, W., Schwarzenbach, F. H., Kötting, J., and Müller, P. (1994), "Multistate models for monitoring individual trees in permanent observation plots", *Environmental and Ecological Statistics*, 1, 171-199.

Ware, J. H., Lipsitz, S., and Speizer, F. E. (1988), "Issues in the analysis of repeated categorical outcomes", *Statistics in Medicine*, 7, 95-107.