University of Bath

MPHIL

Spatio-temporal modelling of forest monitoring data

Griffiths, Alexandra

Award date:
2015

Awarding institution:
University of Bath

Link to publication

Spatio-temporal modelling of forest monitoring data

submitted by

Alex Griffiths

for the degree of Master of Philosophy

of the

University of Bath

Department of Mathematical Sciences

January 2015

COPYRIGHT

Signature of Author . . . .

Alex Griffiths

SUMMARY

This thesis is about developing models to investigate spatio-temporal trends in defoliation levels in European forests. The dataset used is provided by the International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests (ICP), which has been measuring crown condition on several thousand survey plots across Europe since the 1980s. Initially, a pre-existing generalised additive mixed model (GAMM) for defoliation aggregated at the survey plot level is adapted for use with the ICP data, and an improved method of displaying spatial patterns of change over time is developed. We then consider modelling spatio-temporal trends at the level of individual trees within the GAMM framework with suitable response distributions, before moving on to set up a cumulative logistic regression model (incorporating temporal autocorrelation) for defoliation as an ordinal response. Finally, we begin to develop an MCMC algorithm to estimate this model, and offer suggestions for future improvements.

ACKNOWLEDGEMENTS

Contents

List of Figures

List of Tables

Chapter 1

ICP forest health monitoring data: background and descriptive statistics

Forests are valuable to people in many ways, quantifiable or otherwise – as well as their economic importance as a source of timber and other products, they provide habitats for wildlife, prevent soil erosion, mitigate air pollution, and play a role in controlling the climate. The condition of European forests has been monitored since the 1980s, initially in response to the high levels of damage observed with increasing air pollution. There is an interest in investigating the direction and magnitude of temporal and spatial trends in damage to forest health, which may be the result of pollution, climate change or other factors.

1.1 Forest health monitoring in Baden-Württemberg, Germany

The Terrestrial Crown Condition Inventory (TCCI) in the Baden-Württemberg region of Germany is an ongoing annual survey collecting data on crown condition (defoliation levels) from forest plots on an irregular grid, often with different subsets of plots sampled in different years.

In analysing these data, it is necessary to consider the possibility of spatial correlations between nearby survey plots, and of temporal autocorrelation between measurements made on the same plot in consecutive years, as well as nonlinear effects of predictors such as tree age. Augustin et al. (2009) addressed these issues by using a generalized additive mixed model (GAMM) (Lin and Zhang, 1999) to model spatio-temporal trends in crown condition (defoliation) in Baden-Württemberg. Their model incorporated a multidimensional tensor product smooth (Wood, 2006a) of space and time, allowing temporal trends to vary across locations, and took the form given in Equation 1.1.

$$\text{logit}(\mathbb{E}(y_{it})) = f_1(\texttt{tree age}_{it}) + f_2(\texttt{no}_i, \texttt{e}_i, \texttt{year}_t)$$
$$y_{it} = \mathbb{E}(y_{it}) + \epsilon_{it}$$

(1.1)

Here, y_{it} is the mean proportion defoliation for trees on plot i in year t; f_1 is a smooth

function of tree age; f_2 is a multidimensional smooth function of the spatial coordinates, northing (no) and easting (e), together with the year. The errors ϵ_{it} are normally distributed with a first-order autoregressive moving average (ARMA) structure to allow for temporal correlation.

Using this model, Augustin et al. demonstrated a marked increase in defoliation beginning in the late 1990s, from a mean of about 18% in 1998 to about 30% in 2005-2007 (for trees with a median age of 75 years). There was also a strong effect of tree age, with defoliation increasing up to around 75 years and then levelling off.

1.2 Forest health monitoring across Europe

In 1986, the International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests (ICP) was set up to co-ordinate monitoring of the condition of forests across Europe. As of 2014, 42 countries are involved in this programme (Michel and Seidling, 2014). An extensive network of around 7500 Level I survey plots has been established, mainly on a 16×16 km grid. From these plots, data have been collected about many aspects of forest health, including defoliation, nutrient levels in foliage, deposition of atmospheric pollutants, soil condition and ground vegetation. There is also a smaller Level II network comprising several hundred plots, which are used to carry out more intensive ecosystem monitoring.

Forest monitoring in Baden-Württemberg generally makes use of a denser grid of survey plots than the ICP (up to 4×4 km, depending on available funding); however the issues involved in modelling spatio-temporal trends in crown condition are broadly similar. Additional problems may arise when attempting to combine data from different countries.

1.3 ICP crown condition data collection

Crown condition data are collected annually from survey plots covering 35 European countries (plus the Azores and the Canary Islands), beginning in 1987. Currently data covering the years to 2010 are available to us. Data collection is carried out according to the established ICP protocols (Eichhorn et al., 2010). Plots are not necessarily on a regular 16×16km grid, but it is intended that numbers of plots for each country should be chosen to give approximately the same density of coverage across the total forested area. At each grid location, a number of trees will have been selected to form part of the sample; selected trees are permanently marked and re-assessed during subsequent surveys, and trees removed are replaced by newly selected trees.

The main variable of interest to us is percentage defoliation in the crown, estimated by eye (using binoculars) for each individual tree, and recorded in 5% classes, with additional categories for 0% defoliation and 100% defoliation (=dead). The sampling design excludes all parts of crowns which are under direct influence of shadow from surrounding trees, and all trees with biotic damage (e.g. from insects). Only the upper crown is assessed. The intention is "to

quantify the reduction in foliage as an effect of stressors including air pollutants and not as an effect of long lasting site conditions" (Michel and Seidling, 2014) – to achieve this, defoliation may be assessed relative to a healthy and vigorous "local reference tree", representing the best possible tree which could grow at a particular site.

Other information recorded includes the mean age of the dominant trees on the plot. Some topographic characteristics of each plot are also recorded (orientation, altitude, etc).

1.4 Survey plot numbers and locations
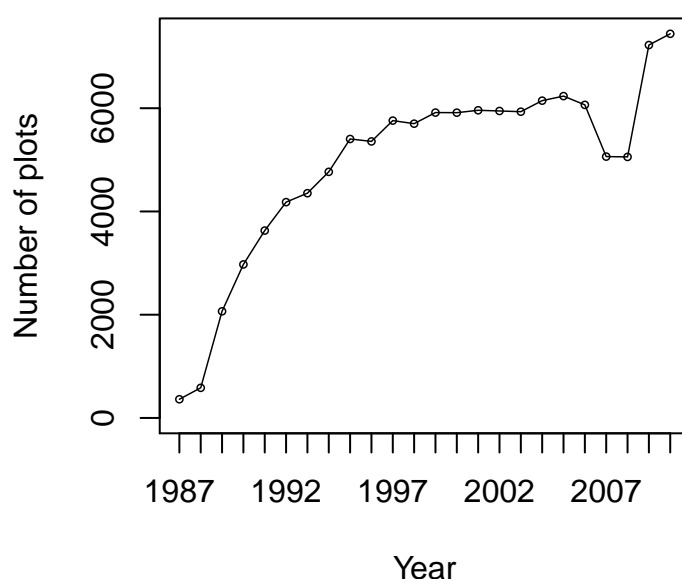
Figure 1-1 Total number of plots assessed in all countries, 1987-2010

Figure 1-1 and Table 1.1 illustrate where and when data were collected. Only Spain and the Czech Republic provided data in 1987, with Slovakia beginning to provide data in 1988, and a group of 12 countries joining in 1989. Russia provided data only in four years, Serbia began to collect data in 2003 and Turkey in 2007. Funding shortfalls led to missing data in the late 2000s in several countries – including Sweden (the country with the largest number of plots) and Austria – otherwise the number of plots in each country remained fairly stable or increased over time. A total of 13151 different plots were measured on at least one occasion; the median number of plots measured in any single year was 5552, with interquartile range 4311 to 5960.

Country	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Andorra	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3	.	3	3	3	3	3
Austria	.	.	.	72	79	77	76	76	76	130	130	130	130	130	130	133	131	136	136	135	.	.	.	135
Belarus	.	.	.	.	.	.	.	.	.	.	416	416	408	408	408	407	406	406	403	398	400	400	409	410
Belgium	.	.	33	29	29	29	29	29	29	29	29	29	30	29	29	29	29	29	29	27	27	26	26	9
Bulgaria	.	.	.	.	.	.	.	108	119	119	119	134	114	108	108	98	105	103	102	97	104	98	159	159
Croatia	.	.	.	.	.	.	84	88	82	83	86	89	84	83	81	80	78	84	85	88	83	84	83	83
Cyprus	.	.	.	.	.	.	.	.	.	.	.	.	.	15	15	15	15	15	15	15	15	15	15	15
Czech Republic	40	85	.	93	362	156	178	205	199	196	196	116	139	139	140	140	140	140	138	136	132	136	133	132
Denmark	.	.	25	25	25	25	25	25	24	23	22	23	23	21	21	20	20	20	22	22	19	19	16	17
Estonia	.	.	.	.	.	.	86	90	90	91	91	91	91	89	92	93	93	92	92	92	93	92	92	97
Finland	.	.	.	.	359	413	405	382	455	455	460	459	457	453	454	457	453	594	605	606	593	475	886	931
France	.	.	509	514	513	505	506	534	543	540	540	537	544	516	519	518	515	511	509	498	504	508	500	532
Germany	.	.	297	408	411	414	412	417	417	420	421	421	433	444	446	447	447	451	451	423	420	423	412	411
Greece	.	.	104	101	101	98	96	96	95	95	94	93	93	92	91	.	.	.	87	.	.	.	97	98
Hungary	.	.	.	67	66	65	65	62	63	60	58	59	62	63	62	.	62	73	73	73	72	72	73	71
Ireland	.	.	22	22	22	22	22	21	21	21	21	21	20	20	20	19	19	19	18	21	30	31	32	29
Italy	.	.	204	204	206	202	212	209	207	207	181	177	239	255	258	258	247	255	238	251	238	236	252	253
Latvia	.	.	.	80	101	100	101	94	94	99	96	97	98	94	97	95	95	95	92	93	93	92	207	207
Lithuania	.	.	.	.	.	73	74	73	73	67	67	67	67	67	66	66	64	63	62	62	62	70	72	75
Luxembourg	.	.	4	4	4	4	4	4	4	4	4	4	4	.	4	4	4	4	4	4	4	4	.	.
Moldova	.	.	.	.	.	.	12	12	11	10	10	10	10	10	10	.	.	.	.	.	.	.	.	.
Montenegro	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	49
Netherlands	.	.	14	14	14	14	13	13	13	.	11	11	11	11	11	11	11	11	11	11	.	.	11	11
Norway	.	.	.	.	.	387	390	384	386	387	386	386	381	382	408	414	411	442	460	463	476	481	487	491
Poland	.	.	.	474	476	476	476	441	432	431	431	431	431	431	431	433	433	428	432	376	458	453	376	374
Portugal	.	.	152	152	151	149	143	147	141	142	144	143	143	144	145	136	136	133	119	118	.	.	.	.
Azores (Portugal)	.	.	.	.	.	.	.	.	8	7	8	6	6	6	6	6	6	6	6	6	.	.	.	.
Romania	.	.	.	.	.	215	167	199	241	224	237	235	238	235	231	231	231	226	229	228	218	.	227	239
Russia	.	.	.	.	.	.	.	7	126	.	.	.	.	.	.	.	.	.	.	.	.	.	365	288
Serbia	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	103	103	130	129	127	125	123	122	121
Spain	322	388	457	447	436	462	460	444	454	447	449	452	598	607	607	607	607	607	607	607	607	607	620	620
Canaries (Spain)	.	.	.	.	.	.	.	12	12	12	12	13	13	13	13	13	13	13	13	13	13	13	.	.
Slovakia	.	111	111	111	111	111	111	111	111	110	110	109	110	111	110	110	108	108	108	107	107	108	108	108
Slovenia	.	.	.	38	45	67	59	34	42	42	42	41	41	41	39	41	41	42	44	45	44	44	44	44
Sweden	.	.	60	.	.	.	.	340	726	766	758	764	764	769	770	769	776	775	784	790	.	.	789	752
Switzerland	.	.	.	45	45	45	45	45	47	49	49	49	49	49	49	48	48	48	48	48	48	48	48	48
Turkey	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	43	396	560	554
UK	.	.	74	74	74	72	69	66	63	79	82	88	85	89	86	86	86	85	84	82	32	.	.	76
Total	362	584	2066	2974	3630	4181	4354	4768	5404	5357	5760	5701	5916	5914	5960	5947	5933	6147	6235	6065	5063	5057	7224	7442

Table 1.1 Numbers of plots with defoliation measured, by country and year
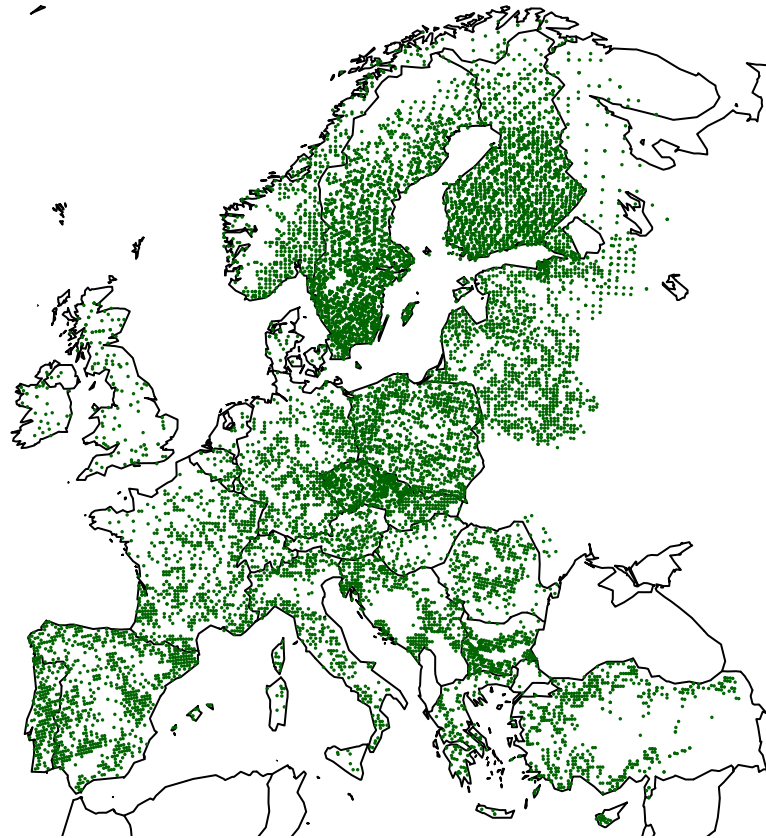
Figure 1-2 Survey plot locations: all species

Figure 1-2 shows the locations of all plots included at least once in the survey (excluding the Canary Islands and the Azores). Plot density varies between countries, due partly to differing levels of forest cover and partly to levels of funding available to carry out the survey.

1.5   Survey plot characteristics

In addition to crown condition, a number of other variables were measured, mainly relating to the landscape characteristics of each plot (altitude, orientation, water budget, humus type). The mean age of trees on the plot was also recorded (the ages of individual trees are not available). This information is summarised in Table 1.2 (for variables which may change over time, such as mean age of dominant trees, the values given correspond to the first occasion this plot was

included in the survey). It should be noted that water budget was also calculated on only one occasion for each plot and is assumed not to change over the period of monitoring; this may be misleading.

Table 1.2 Summary of plot characteristics

Latitude	range $27°44'$ to $70°28'$		Orientation	north	1171
Longitude	range $-28°24'$ to $42°45'$			northeast	985
Altitude	<50m	1243		east	962
	51-100m	1660		southeast	736
	101-150m	2128		south	1065
	151-200m	1493		southwest	876
	201-300m	1635		west	992
	301-400m	527		northwest	831
	401-500m	990		flat ground	5456
	501-1000m	2057		missing	77
	1001-1500m	825	Mean age of	0-20 years	818
	1501-2000m	219	dominant trees	21-40 years	2306
	>2000m	18		41-60 years	3059
	missing	53		61-80 years	2294
Water budget	insufficient	1593		81-100 years	1880
	sufficient	9601		101-120 years	1050
	excessive	680		>120 years	1290
	missing	1277		irregular stands	430
Humus type	Moder	3084		missing	24
	Mor	4898			
	Mull	2315			
	Peat/Anmoor	394			
	Histomull	218			
	Amphi	97			
	Roh/Histomoder	84			
	Histomor	629			
	missing	1432			

Although the ICP protocol was used to assess crown condition, the sample plots were not selected in the same way in all countries, and the number of trees monitored on each plot varied between one and over twenty. Over the whole survey period to 2010, 290,239 trees were measured at least once, of which 178,263 were conifers and 112,098 broadleaved. The most common genera were *Pinus* (pine, 35.8%), *Picea* (spruce, 21.7%), *Quercus* (oak, 13.5%), *Fagus* (beech, 8.0%) and *Betula* (birch, 6.7%), – Table 1.3 gives the full list.

Table 1.3 Genera of trees assessed

Genus	Common name(s)	n	Genus	Common name(s)	n
Acer	maple	1944	*Prunus*	plum, cherry, peach, almond, apricot	517
Alnus	alder	3756			
Arbutus	strawberry tree	129	*Pyrus*	pear	42
Betula	birch	19343	*Quercus*	oak	39120
Buxus	box	27	*Rhamnus*	buckthorn	14
Carpinus	hornbeam	3775	*Robinia*	locust	1974
Castanea	chestnut	2731	*Salix*	willow	340
Ceratonia	carob	10	*Sorbus*	rowan, whitebeam	337
Cercis	redbud	11	*Tilia*	lime	1049
Corylus	hazel	51	*Ulmus*	elm	233
Crataegus	hawthorn	14	Other broadleaves		1362
Erica	heath	122	*Abies*	fir	5447
Eucalyptus	eucalyptus	3378	*Cedrus*	cedar	292
Fagus	beech	23351	*Cupressus*	cypress	174
Fraxinus	ash	2412	*Juniperus*	juniper	1550
Ilex	holly	23	*Larix*	larch	2443
Juglans	walnut	33	*Myrica*	bayberry	36
Laurus	bay, laurel	34	*Myrtus*	myrtle	3
Malus	apple	7	*Picea*	spruce	62870
Olea	olive	267	*Pinus*	pine	104132
Ostrya	ironwood	774	*Pseudotsuga*	Douglas fir	1045
Phillyrea	phillyrea	123	*Taxus*	yew	1
Pistacia	pistachio	72	*Thuya*	red cedar	5
Platanus	sycamore, plane	172	*Tsuga*	hemlock	18
Populus	poplar	4429	Other conifers		247

1.6 Norway spruce (*Picea abies*)

Any spatio-temporal trends in crown condition are likely to vary between different species, and certainly between deciduous and evergreen trees. For some species, such as common beech (*Fagus sylvatica*), there is the additional complication that fructification (heavy production of beechmast) in some years may cause resources to be diverted away from leaf growth (Mund et al., 2010), and this must be taken into account in modelling crown condition. It seems sensible to focus initially on a single species which is well-represented in the dataset; Norway spruce (*Picea abies*) is the obvious candidate, as previous work on the Baden-Württemberg data has focused on this species (together with beech).

In the ICP dataset, Norway spruce is the second most common species overall, accounting for around 20% of all trees measured, and is frequently found at the same locations as the most common species, Scots pine (*Pinus sylvestris*). Figure 1-3 shows the locations of plots with at least one Norway spruce monitored in one or more years; these plots are concentrated mainly in Scandinavia and central Europe.

The following chapter adapts the model used for Baden-Württemberg to investigate spatio-temporal trends in mean percentage defoliation across a wider geographical area. We restrict ourselves to data from the central contiguous area covered fairly densely by survey plots in Germany, Austria and Switzerland.
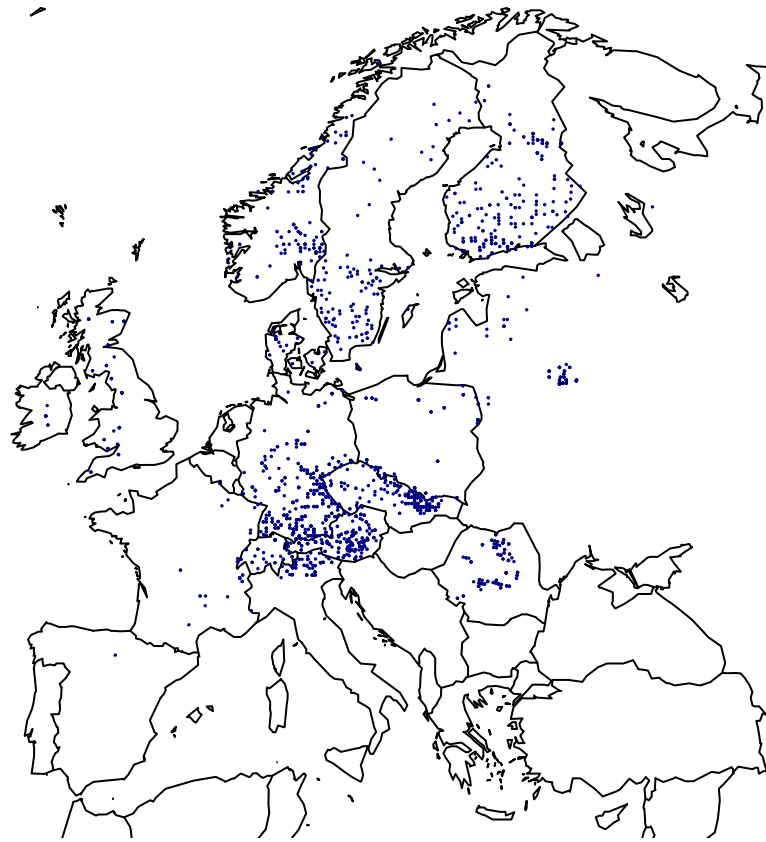
Figure 1-3 Survey plot locations: Norway spruce (*Picea abies*) only

Chapter 2

Modelling aggregated crown condition data

2.1 Descriptive statistics

Locations of German, Austrian and Swiss survey plots containing at least one Norway spruce tree and measured at least once are shown in Figure 1-3. Data are available for Germany from 1989 to 2010 and for Austria and Switzerland from 1990 to 2010; Austria did not provide data in the period 2007-2009. Numbers of plots in each country remained fairly steady over time, with the median number of plots used (in years when any monitoring was carried out) being 215 for Germany, 111 for Austria and 31 for Switzerland. The total number of plots measured in at least one year was: Germany 309, Austria 118, Switzerland 34.

Table 2.1 breaks down plot characteristics by country. The main differences appear to be that Switzerland tends to have plots dominated by much older trees, with a greater prevalence of Mull type humus rather than Moder, and a much greater proportion of plots with an excessive water budget. Plots in Austria and Switzerland are also mainly at a higher altitude than those in Germany.

Table 2.1 Characteristics of survey plots containing Norway spruce in Germany, Austria and Switzerland

	Germany	Austria	Switzerland
Mean age of dominant trees			
0-20 years	28	9	1
21-40 years	40	16	0
41-60 years	60	23	0
61-80 years	49	18	0
81-100 years	64	18	7
101-120 years	34	17	5
> 120 years	30	17	15
irregular stands	4	0	6
Humus type			
Moder	135	94	10
Histomull	74	2	0
Mor	39	11	0
Mull/Amphi/Peat/Anmoor	59	11	18
Roh/Histomoder	2	0	6
Water budget			
sufficient	268	108	20
insufficient	33	5	0
excessive	8	5	14
Orientation			
flat	97	8	2
N	39	23	9
NE	23	14	6
E	19	6	3
SE	22	8	4
S	25	14	2
SW	22	17	0
W	5	17	5
NW	37	11	3
Altitude (metres)			
median (IQR)	475 (375-625)	1000 (725-1375)	1150 (650-1450)

2.1.1 The "country effect"

Although protocols have been developed to determine how crown condition should be assessed, there is considerable scope for subjectivity leading to variation between different observers. An ICP exercise was carried out in 2004, which requested 79 experts in 31 countries to assess the apparent defoliation levels of trees in each of a set of 144 photographs (Lorenz et al., 2005). For Norway spruce (one of six species used in the exercise), the mean absolute deviation from the median defoliation assessment for each tree lay mostly in the range from seven to nine percentage points. This is a fairly substantial difference, indicating discrepancies between individual observers in crown condition assessment.

It is also possible that differences between countries in training, or in the reference photographs used, may lead to defoliation being recorded as systematically higher or lower in one country compared to another. Figure 2-1 shows much lower defoliation being recorded in every year in Austria compared to Germany and Switzerland. This could be explained by different site conditions; however, if only the subset of survey plots close to the Austrian-German border is considered (where topographical characteristics ought to be similar), the gap still exists. The presence of a "country effect" can be investigated by simply including country in the proposed model as a covariate, on the assumption that the magnitude of such an effect, if it exists, is constant over time.

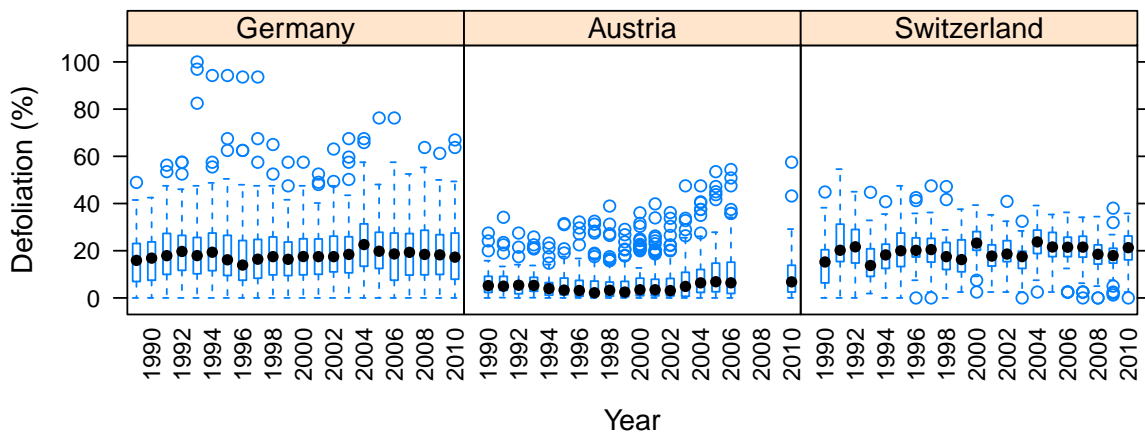**Figure 2-1** Boxplots of mean % defoliation in each year for Germany, Austria and Switzerland

2.2 Structure and fitting of the model

Following Augustin et al. (2009), the response "mean proportion defoliation for trees on plot i in year t" (y_{it}) is modelled using a generalised additive mixed model (GAMM) (Lin and Zhang, 1999; Fahrmeir and Lang, 2001; Wood, 2006a, 2011) of the following form:

$$\text{logit}(\mathbb{E}(y_{it})) = f_1(\texttt{latitude}_i, \texttt{longitude}_i, \texttt{year}_t) + f_2(\texttt{altitude}_i) + \mathbf{X}_{it}\boldsymbol{\beta}$$
$$y_{it} = \mathbb{E}(y_{it}) + \epsilon_{it}$$

(2.1)

Here, f_1 is again a tensor product smooth of the two spatial dimensions together with time; f_2 is a smooth function of plot altitude; and $\mathbf{X}_{it}\boldsymbol{\beta}$ represents the contribution made to the linear predictor by the effects of the other (categorical) covariates – tree age, country, humus type, water budget and plot orientation.

Parameters are estimated by penalized quasi-likelihood (Breslow and Clayton, 1993), with approximate restricted maximum likelihood (REML) estimation of the smoothness parameters as described in Wood (2011). This scheme is implemented via the `gamm` function of the `mgcv` package in R (R Core Team, 2014).

2.2.1 Choice of response distribution

The errors ϵ_{it} are assumed to follow a Gaussian distribution. A quasi-likelihood approach was also considered, in order to allow the error variance to vary with the mean rather than being fixed (pp 74–76 in Wood, 2006b). This did not improve the fit of the model, as assessed by residual plots, so the Gaussian errors are retained.

2.2.2 Choice of k for spatial smooth

The value of `k` specified in the call to the `gamm` function places a constraint on the maximum degrees of freedom associated with the tensor product smooth, but larger values of `k` require more computing time to fit the model. To check that this constraint was not too severe, the value of `k` for the spatial component of the smooth was increased from 49 to 80. As a result, the effective degrees of freedom in the fitted model increased from 163.3 to 202; however, none of the covariate effects were noticeably changed and the effective degrees of freedom were still far below the upper limit imposed by the choice of `k` (which would be the product of the `k` values for each component of the smooth, minus 1, or $(49-1) \times (49-1) \times (10-1) \simeq 20000$). Hence, the value of `k` was kept at 49, which allows models to be fitted in a reasonable amount of time (around 15 minutes) while allowing enough flexibility for the spatial smooth.

2.2.3 Temporal autocorrelation

The model allows for temporal autocorrelation between successive crown condition measurements on each plot. A simple AR(1) process was found to be insufficient, as displayed by the correlogram in Figure 2-2, which shows significant correlations at lags of up to six years. An ARMA(2,1) process appears to model the autocorrelation adequately (Figure 2-3).
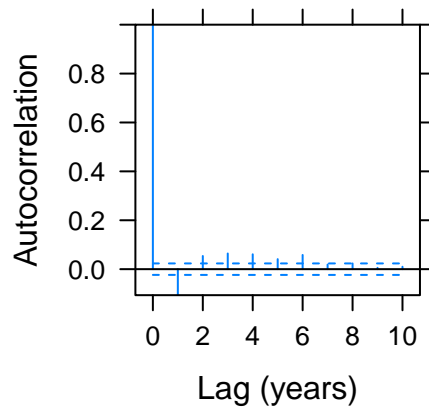
Figure 2-2 Autocorrelation plot for model with AR(1) residuals
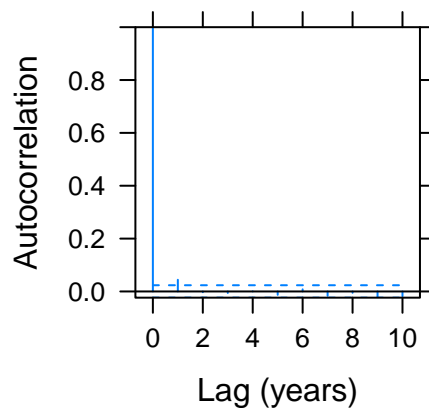
Figure 2-3 Autocorrelation plot for model with ARMA(2,1) residuals

2.3 Covariate effects

Some covariates were highly correlated with the country where the trees were located (specifically, excessive water budgets and Mull-type humus were fairly common on Swiss plots but very rare in Austria and Germany). To avoid over-estimating the size of the country effect, which is hypothesised to be the result of different levels of adherence to the ICP measurement protocols for crown condition, these covariates are not removed from the model. The coding of orientation is simplified to compare plots facing south-west (into prevailing winds) with all other plots, and altitude is included as a linear effect rather than a smooth term (since the estimated smooth function in the full model appears very close to linear).

Table 2.2 shows the parameter estimates from the full model which was the starting point for covariate selection ($R^2 = 58.5\%$). Table 2.3 shows the parameter estimates from the final

model ($R^2 = 58.4\%$). These coefficients are on the logit scale as a result of the link function used in the model, so are not directly interpretable in terms of percentage defoliation; however, positive values correspond to worsening crown condition. The spatio-temporal tensor product smooth was found to be significant in both models, with effective degrees of freedom equal to 163.3 in the full model and 160.0 in the final model ($p < 0.001$ in both cases).

The mean age of dominant trees on the plot is strongly associated with crown condition, with defoliation worsening with age sharply at first and then levelling out (see Figure 2-6 for further illustration of this trend). Improved crown condition was also associated, though less strongly, with Mull type humus, excessive water budget, south-west plot orientation, and lower altitude. Although most of these survey plot characteristics are fixed, it should be noted that water budget (which may vary from year to year) was assessed at only one point in time; thus the observed association with defoliation does not reflect the possible effect of annual variations in rainfall, though it may be due to the susceptibility of some plots to waterlogging.

The country effect is strongly present, even after adjusting for available covariates known to be confounded with the country of measurement (water budget, humus type). For Austria, the strength of this effect is on a similar scale to the effect of age, while for Switzerland the impact is substantially smaller; hence there is evidence that in both countries crown condition has been systematically recorded more optimistically than in Germany. Based on the available data, we cannot completely rule out the possibility that unobserved differences in survey plot characteristics (climate, topography, soil) may explain this, but it seems unlikely given the size of the discrepancy (particularly for Austria).

Table 2.2 Parameter estimates (on the logit scale) from the full model.

Covariate	Estimate	Std. error	p-value
Mean age of dominant trees (ref. cat. 0-20 years):			
21-40 years	0.225	0.076	0.003
41-60 years	1.085	0.071	<0.001
61-80 years	1.426	0.074	<0.001
81-100 years	1.504	0.070	<0.001
101-120 years	1.622	0.073	<0.001
> 120 years	1.798	0.076	<0.001
irregular stands	1.290	0.122	<0.001
Country (ref. cat. Germany):			
Austria	-1.092	0.083	<0.001
Switzerland	-0.231	0.106	0.029
Humus type (ref. cat. Moder):			
Histomull	-0.050	0.055	0.362
Mor	-0.022	0.053	0.674
Mull/Amphi/Peat/Anmoor	-0.130	0.042	0.002
Roh/Histomoder	-0.211	0.157	0.180
Water budget (ref. cat. sufficient):			
insufficient	-0.050	0.058	0.384
excessive	-0.290	0.078	<0.001
Orientation (ref. cat. flat):			
N	-0.052	0.053	0.329
NE	0.054	0.062	0.377
E	-0.151	0.079	0.057
SE	0.028	0.069	0.684
S	-0.025	0.060	0.675
SW	0.166	0.062	0.007
W	0.009	0.059	0.872
NW	-0.042	0.060	0.483

Table 2.3 Parameter estimates (on the logit scale) from the final model after backwards selection.

Covariate	Estimate	Std. error	p-value
Mean age of dominant trees (ref. cat. 0-20 years):			
21-40 years	0.238	0.074	0.001
41-60 years	1.092	0.070	<0.001
61-80 years	1.430	0.073	<0.001
81-100 years	1.514	0.070	<0.001
101-120 years	1.638	0.073	<0.001
> 120 years	1.813	0.075	<0.001
irregular stands	1.349	0.122	<0.001
Country (ref. cat. Germany):			
Austria	-1.084	0.082	<0.001
Switzerland	-0.296	0.104	0.005
Humus type (ref. cat. Moder):			
Histomull	-0.046	0.054	0.403
Mor	-0.039	0.052	0.460
Mull/Amphi/Peat/Anmoor	-0.137	0.042	0.001
Roh/Histomoder	-0.143	0.155	0.356
Water budget (ref. cat. sufficient):			
insufficient	-0.040	0.057	0.485
excessive	-0.283	0.077	<0.001
SW orientation	0.184	0.053	<0.001
altitude (per 100 metres)	0.013	0.006	0.028

2.4   Overall temporal trend

To obtain confidence intervals for non-linear functions of the model parameters, such as the temporal trend, we use the approach suggested by Wood (2006a), generalising from Silverman (1985). We take the final model and obtain a random draw from the (multivariate normal) joint distribution of the parameter vector. We use this set of parameters to calculate predicted defoliation values at every survey plot in every year, and then – for the temporal trend – take the mean over all spatial locations to give a single predicted value for each year. This process is then repeated with a new random draw from the parameter distribution, for a total of 1000 draws, until we have a predictive distribution for defoliation in each year. The medians of these distributions, together with the 2.5% and 97.5% percentiles, provide estimates of the temporal trend and a 95% credible interval.

All predictions for locations in Austria and Switzerland are adjusted to remove the country effect, so that the defoliation values are (hypothetically) those which would have been obtained if all measurements had been carried out according to the protocols used in Germany. Other covariates are kept at their observed values for each survey plot.

Figure 2-4 shows the temporal trend averaged across all locations – median defoliation levels have risen from around 16% in 1989 to 20% in 2010, with a peak at 22% in 2005-6 and

a smaller peak in the early 1990s. The width of the 95% credible interval is in the region of 1–1.5 percentage points either side of the median, indicating that the upward trend in defoliation is beyond what could be explained by chance.

Figure 2-5 shows the temporal trends separately for each country in the model. In Germany (where the majority of the survey plots are located), the pattern is very similar to Figure 2-4. In Austria, the peak in the early 1990s is less noticeable, while the second spike in defoliation in around 2005 is much more pronounced than in Germany; there is also some suggestion that defoliation is beginning to rise again after 2009. Switzerland has the smallest number of survey plots and so uncertainty around the estimates is greater (roughly 4 percentage points either side of the median) – there is little evidence for a clear trend in either direction.

Figure 2-6 illustrates the impact of tree age on defoliation. Predictions are made for each location in each year using a fixed age for the trees on every plot (instead of the observed age), and then averaged over all locations as before. Three of the possible age groups are shown – defoliation is predicted to be much lower than average if the mean age of dominant trees on all plots is 21-40 years (around 5-8%), rising to around 20-25% if predictions are based on trees being aged 61-80 years, and then levelling off.
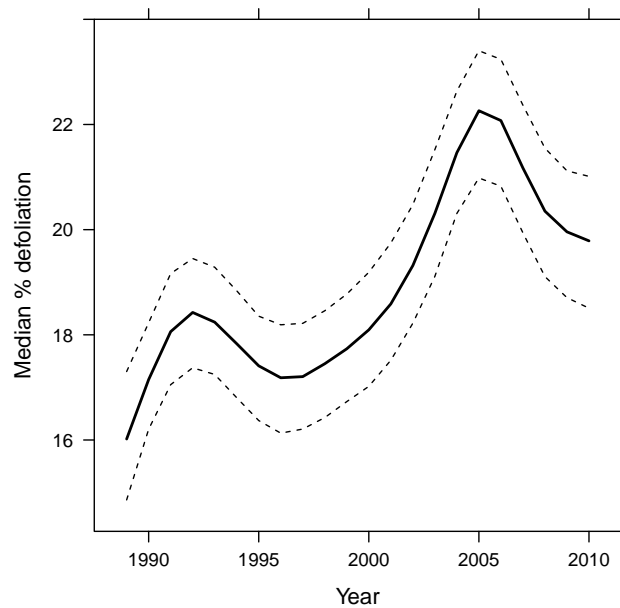

Figure 2-4 Overall temporal trend in defoliation, adjusted to remove country effect

18
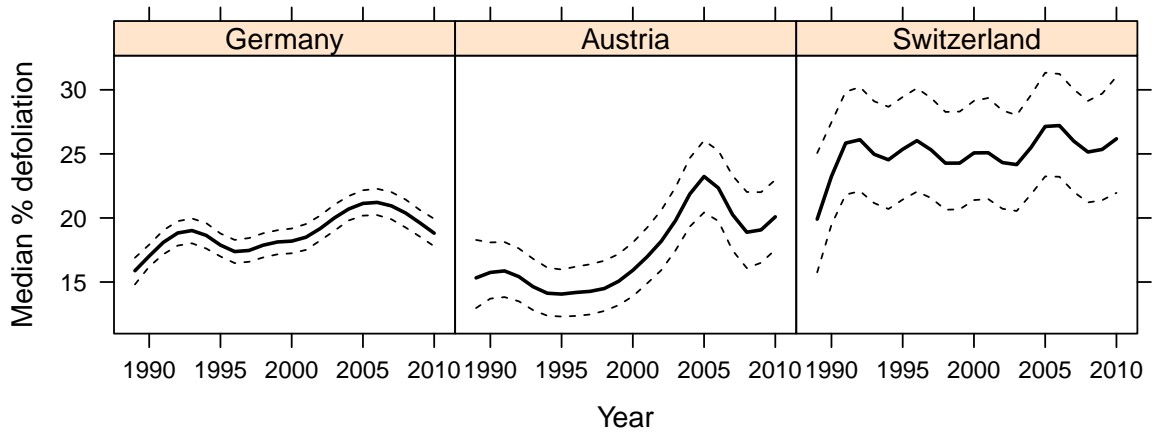
Figure 2-5 Temporal trends in defoliation by country, adjusted to remove country effect
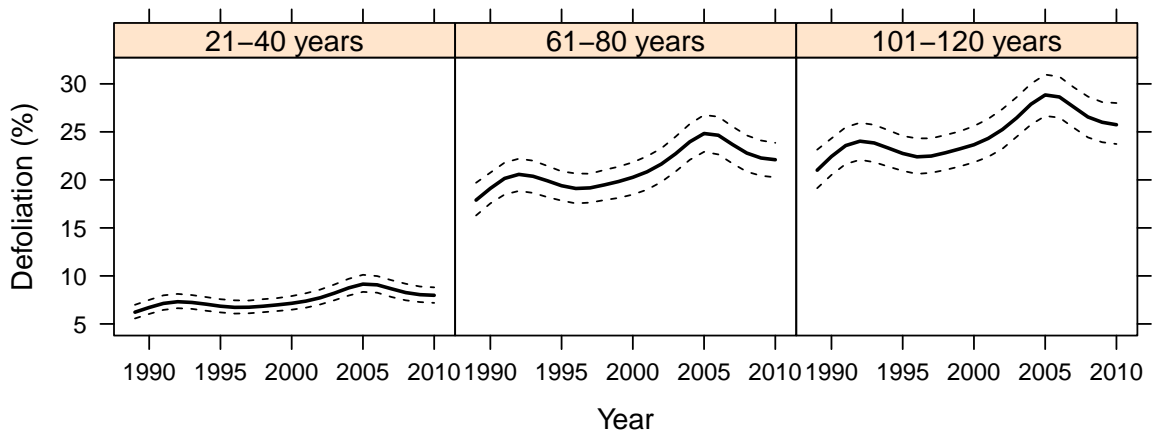
Figure 2-6 Temporal trends in defoliation for plots with differently-aged dominant trees, adjusted to remove country effect

2.5 Spatial patterns in crown condition

Figure 2-7 displays spatial patterns in defoliation in individual years (every five years from 1990 to 2010), with yellow corresponding to areas of poor crown condition and green/blue to less severely affected areas. Black dots show the locations of survey plots measured in that year, and red lines show contours of equal predicted defoliation levels (on a scale from 0 to 1 rather than as percentages).

These plots are created using the `vis.gam` function from the `mgcv` package – predictions are made for points on a regular grid covering Germany, Austria and Switzerland, with the restriction that these grid points must not be too far from an actual survey plot location. Because plot-level covariate values are not known for these grid points, a reduced version of the final model is used to make the predictions, with humus type, water budget, orientation and altitude removed, so that their effects are absorbed into the spatial smooth. (The alternative would be to use the final model as it stands and fix each covariate at a single value across the whole area of the map, which is unrealistic.) All predictions are made assuming a mean dominant tree age of 61-80 years, and adjusted for the country effect as with the temporal trend.

As expected, these maps agree with the temporal trend plot in Figure 2-4 in showing an general increase in defoliation between 1990 and 2010, with the highest levels being in 2005. There is some suggestion that areas in the western half of the map are more severely affected by this increase.

1990 **1995**

2000 **2005**

2010

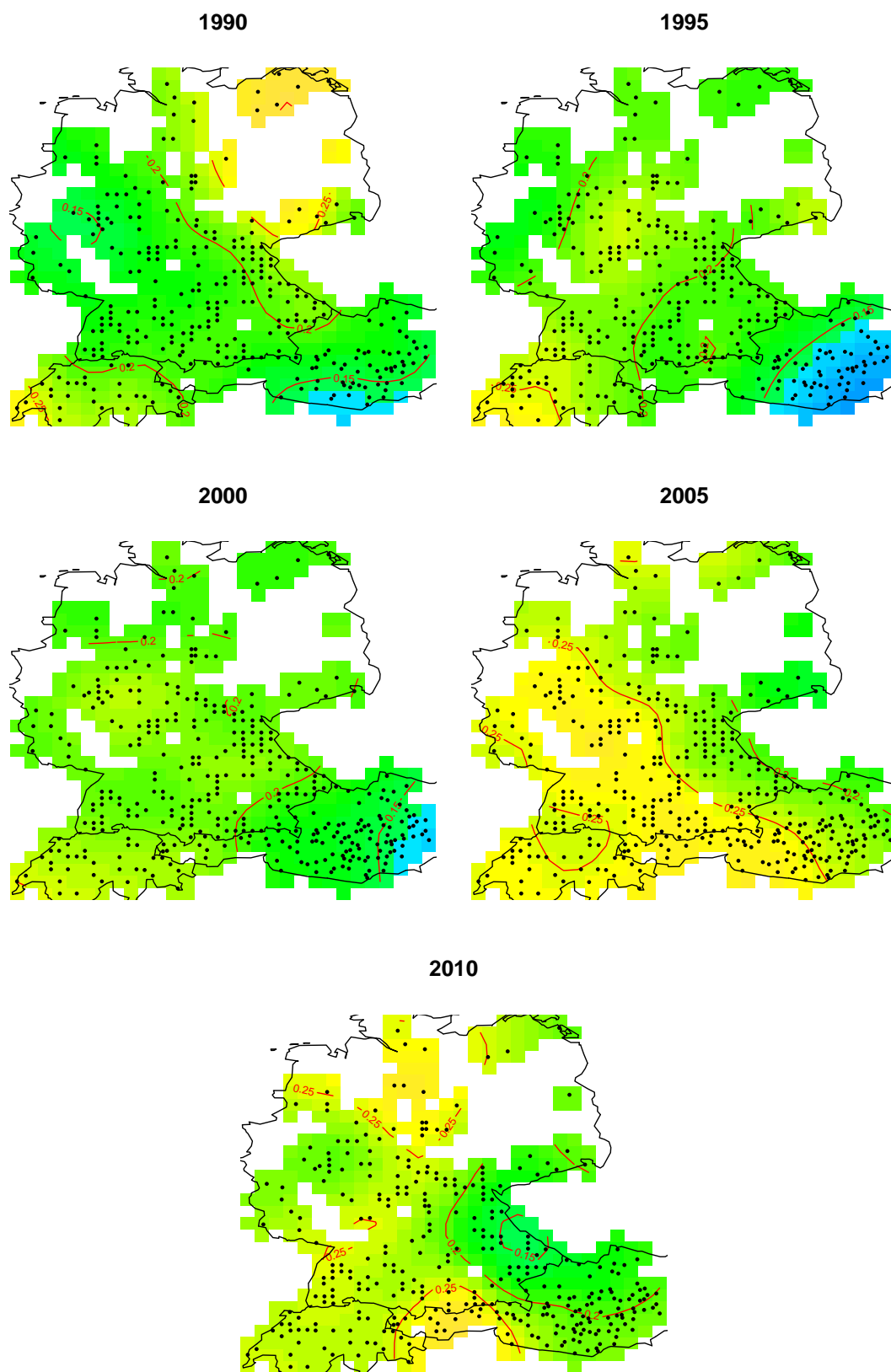**Figure 2-7** Spatial maps of defoliation levels at five-year intervals, adjusted to remove country effect

2.6 Spatial patterns of change over time

Spatial patterns of change in crown condition over time are displayed as "fogplots", which show on a single map both the estimated change and the degree of uncertainty about that change. The `fogplot` function (see appendix) was written for this purpose, using the capabilities of the `rgl` package (Adler et al., 2014) in R.

As previously described for the overall time trend, sampling from the parameter vector of the final model is used to obtain a Bayesian posterior predictive distribution for the change in defoliation between the two years of interest at each point on a regular grid. The median of this distribution is used as a summary measure of the change at that grid point, and these changes are then plotted on a map. Black dots show the locations of survey plots. Colour, on a scale from blue (improvement in crown condition) to red (deterioration) is used to denote the magnitude and direction of change, while greater transparency corresponds to greater uncertainty about the direction. Transparency in OpenGL, as used by the `rgl` package, is determined by the alpha value, which runs from 0 (complete transparency, so that the background colour shows through) to 1 (solid colour).

2.6.1 Choosing an appropriate alpha value

The fogplot function was originally intended to be used to display spatial patterns in crown condition in each individual year of measurement, as well as changes from year to year. The alpha values are then used to indicate the level of uncertainty about the predicted defoliation at each grid point. One way to achieve this is to set the alpha value equal to the probability (under the posterior) that the true crown condition measurement is within some specified distance δ of the predicted value, i.e.:

$$
\begin{aligned}
\texttt{alpha} &= \mathbb{P}(|X - \mathrm{median}(X)| > \delta) \\
&= 1 - F_X(\delta) + F_X(-\delta)
\end{aligned}
\tag{2.2}
$$

Here the probability is taken under the posterior predictive distribution of X (the defoliation level at a particular grid point), and F_X is the corresponding cumulative distribution function. In practice, the precision of the estimates does not vary greatly across the map and there is little advantage in using fogplots for this purpose.

For plots showing change between two years, it is more useful to display the uncertainty about the sign of the change, rather than its magnitude. This is achieved by calculating the p-value from a test of the null hypothesis that the true change between years is equal to zero; alpha is then set as equal to one minus the p-value, so that higher alpha (more solid colour) corresponds to stronger evidence against the null hypothesis. In practice this p-value is estimated from the posterior predictive distribution of change at each grid point, as given below.

$$\texttt{alpha} = 1 - 2 \times \min(p_1, 1 - p_1)$$
$$\text{where } p_1 = \mathbb{P}(\text{change} < 0) \tag{2.3}$$
$$= F_{change}(0)$$

Here F_{change} is the cumulative distribution function corresponding to the posterior predictive distribution of change in crown condition between the years of interest, at a particular grid point. This choice of alpha has the additional advantage that there is no need to specify a particular distance δ which is considered to be "close" to zero.

2.6.2 Example fogplots
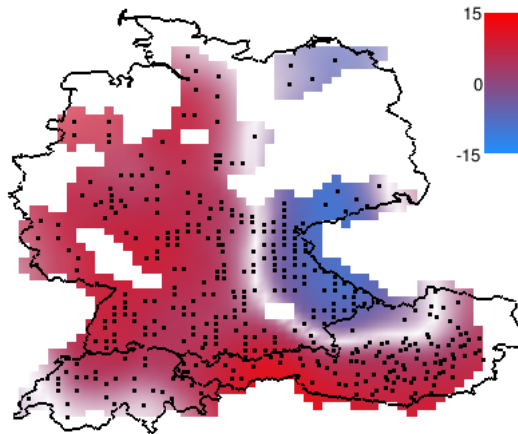
Figure 2-8 Spatial pattern of changes in crown condition between 1990 and 2010

Figure 2-8 shows the estimated change in defoliation between 1990 and 2010 (the whole period for which data were available, excluding 1989 when fewer survey plots were measured). There is a clear division of the map into an area of worsening defoliation in the west and south, and a smaller area of modest improvement in the southeast of Germany.

Figures 2-9, 2-10 and 2-11 show spatial patterns of change in defoliation levels for overlapping ten-year periods. The trend in each case is towards poorer crown condition (red); however, the greater transparency across large areas of Figure 2-9 illustrates that the change in this period was not significantly different from zero. As expected from earlier results, the ten years leading up to 2005 show the most severe change for the worse (Figure 2-10, with areas in the south being most strongly affected and almost no areas of improvement. Figure 2-11 is, again, mainly negative, with two small areas of (non-significant) improvement.
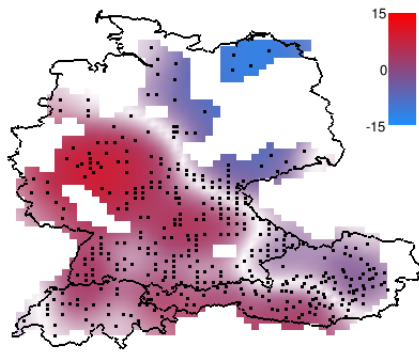
Figure 2-9 Spatial pattern of changes in crown condition between 1990 and 2000
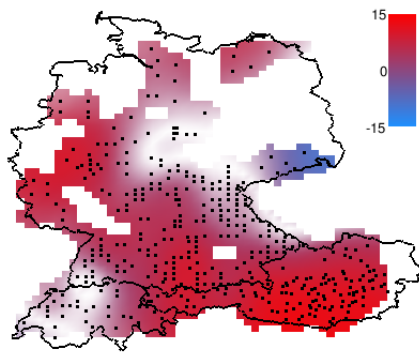
Figure 2-10 Spatial pattern of changes in crown condition between 1995 and 2005
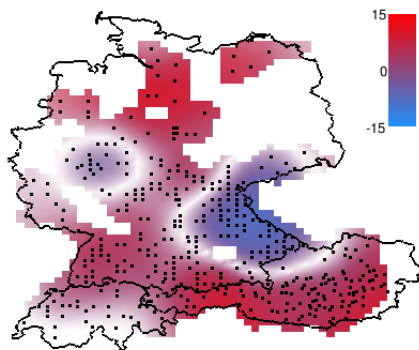
Figure 2-11 Spatial pattern of changes in crown condition between 2000 and 2010

Chapter 3

Modelling defoliation in individual trees

This chapter considers some possible approaches to modelling defoliation in individual trees, as opposed to the mean percentage defoliation aggregated within each plot, which has been considered so far. Since trees included in the ICP monitoring programme are generally retained from year to year and hence have repeated measurements associated with them, it may be informative to model their individual trends in defoliation. (It should be noted that, owing to the availability of data at the time this work was carried out, all models for tree-level data only use data up to 2008.)

3.1 Features required in a model of individual tree defoliation

Although for the plot-aggregated data, a Gaussian distribution fitted reasonably well, this is not the case for the tree-level data. As illustrated in Figure 3-1, we require a response distribution that is skewed, and which allows for a possible second (smaller) peak representing extremely damaged/dead trees.

Ideally, the final model would also include the following features:

- smooth/non-linear temporal and spatial trends in defoliation;

- temporal autocorrelation between repeated measurements on the same tree in different years;

- spatial correlation between trees on the same plot, and possibly between neighbouring plots; and

- adjustment for the country effect, as discussed in the previous chapter.

In this chapter, two continuous distributions (negative binomial and inflated beta) are briefly investigated but ultimately rejected, the first because the model fit is still poor, and the second
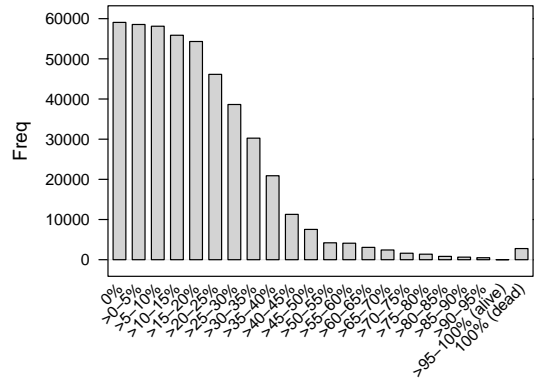
Figure 3-1 Distribution of defoliation measured in individual spruce trees (all countries, all years to 2008)

due to practical limitations of the software available at the time to fit such a model. We then move on to considering defoliation as an ordinal categorical variable, exploring ways of fitting ordinal logistic models to the data using pre-existing R packages (`polr` and `R2BayesX`).

3.2 Negative binomial response

A negative binomial response ought to allow the skewness in the response to be modelled, though not the bimodality. This distribution is already supported by the `mgcv` package in R, so smooth terms can be included in the model without difficulty.

We begin by restricting to a roughly rectangular area covering central Europe – France, Belgium, the Netherlands, Germany, Austria, Switzerland, Poland, Slovakia and the Czech Republic – in order to avoid interpolating over large areas of water, though this could be dealt with using soap-film smoothing (Wood et al., 2008) or similar. Again, for simplicity, we consider only Norway spruce (*Picea abies*), which is widespread in all the countries under consideration. All these countries have almost complete data from 1990 to 2008 (Austria in 2007 and 2008 is the only exception). At this stage, dead trees were not added back into the survey sample. This gives a dataset containing 228404 measurements of 26894 individual trees on 1099 plots over the 19 years from 1990 to 2008.

Using this restricted dataset, models of the form given by equation (3.1) are fitted, with the response y_{it} following a negative binomial distribution.

$$\log(\mathbb{E}(y_i t)) = f_1(\texttt{year}_t) + \mathbf{X}_{it}\boldsymbol{\beta} \tag{3.1}$$

Here, f_1 is a smooth function of `year`, with smoothing parameters estimated by REML. Observation y_{it} corresponds to the crown condition of tree i in year t. The negative binomial distribution requires a second ("size") parameter, θ, which is estimated in `mgcv` by finding the

value minimizing the AIC of the model. In this case, θ was found to be very close to 1, which actually reduces the negative binomial distribution to a geometric distribution.

Attempts were also made to fit models which incorporated a spatial smooth (a two-dimensional smooth function of latitude and longitude) and/or a random effect term to take account of the correlation between the levels of defoliation of trees on the same plot. However, as expected, none of these models were able to fit well to the data, as a result of the cluster of values around 100% defoliation (Figure 3-1). This is clearly shown in Figure 3-2, which is the quantile-quantile plot for the residuals of the model given by equation (3.1) with no additional covariates beyond the mean age of trees on the plot.
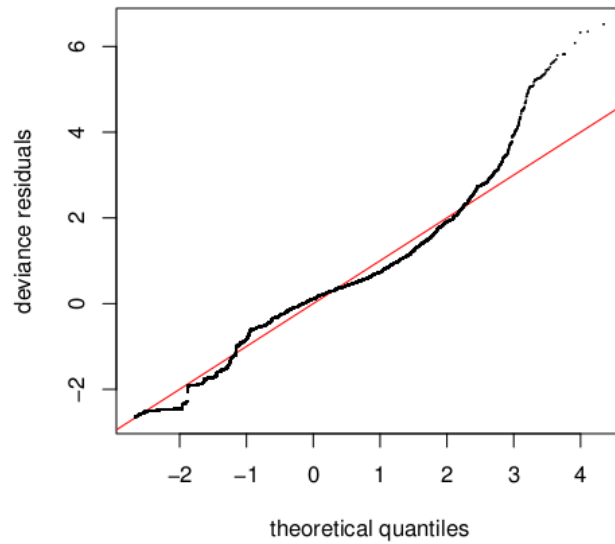
Figure 3-2 Quantile-quantile plot for model with negative binomial response

3.3 Inflated beta response

A more flexible distribution is required, which can adequately model the bimodality present in the data. In fact, if the data are broken down by country and year (Figures 3-3, 3-4), it appears that there may be three peaks present. These occur at 0% defoliation (completely healthy trees), 100% defoliation (dead trees), and at a point in between representing the modal level of defoliation among trees with some degree of damage (see for example the data for Germany 1989-1991 in Figure 3-3). It can also be seen that the general form of the distribution may be quite different depending on the country – in Germany, defoliation levels follow a smooth curve, while in Slovakia there is a much sharper peak and very few completely healthy trees.
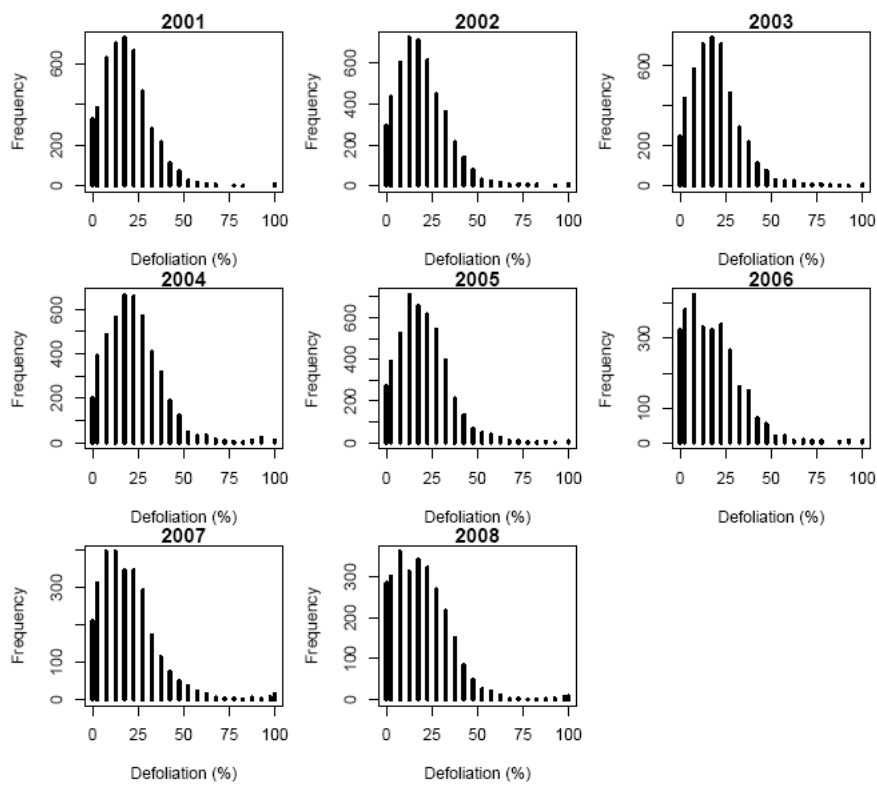
Figure 3-3 Observed distribution of defoliation levels by year: Germany, 2001-2008
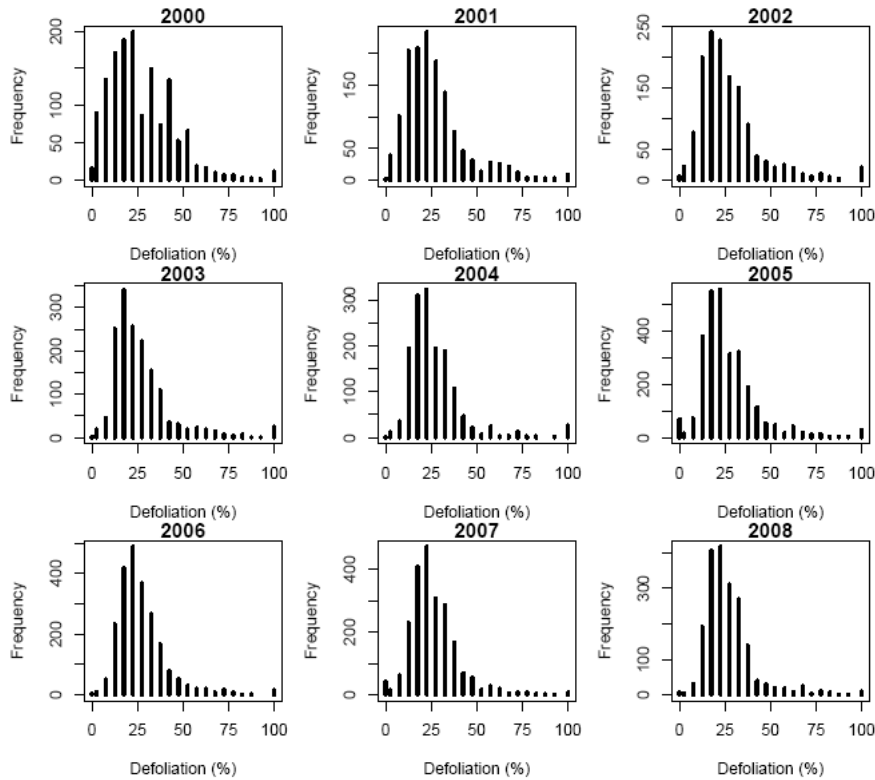
Figure 3-4 Observed distribution of defoliation levels by year: Slovakia, 2000-2008

If defoliation is to be modelled as a continuous variable, some form of mixture distribution is required in order to incorporate the point masses at 0% and 100%. The remaining part of the mixture (i.e. the part representing trees with some damage) also needs to be flexible enough to allow for different levels of skewness. The beta distribution would satisfy this requirement, and has the additional advantage of being defined only on the interval [0,1] (corresponding to the possible range of values for defoliation, [0%, 100%]).

For $x \in [0,1]$ and shape parameters $\alpha > 0$, $\beta > 0$, the beta distribution is defined by the probability density function:

$$f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \tag{3.2}$$

This can also be reparametrized in terms of the mean $\mu = \frac{\alpha}{\alpha+\beta}$ and the sum of the shape parameters $\nu = \alpha + \beta$.

Beta distributions are not currently supported by the `mgcv` package which we have been using so far to fit GAMs to the defoliation data; this is because the second (shape) parameter is not simply a function of the mean and must be estimated separately. However, the beta distribution (with the μ, ν parametrization) is included in the `gamlss` package (Rigby and Stasinopoulos, 2005). This fits "generalized additive models for location, scale and shape" (GAMLSS), which allow the location, scale, skewness and kurtosis parameters for the distribution of the response

29

to be modelled as smooth functions of one or more explanatory variables, using maximum (penalized) likelihood estimation with a backfitting algorithm for the smoothing parameters.

Ospina and Ferrari (2010) describe two possible parameterisations for a zero-and-one-inflated beta distribution, the second of which is implemented (with modifications) as a response distribution in the R package `gamlss`.

Both versions parameterise the beta distribution in terms of its mean, $0 < \mu < 1$, and precision, $\phi > 0$. The density function is:

$$f(y; \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} \; y^{\mu\phi-1}(1-y)^{(1-\mu)\phi-1}, \qquad y \in (0,1) \qquad (3.3)$$

3.3.1 Parameterisation 1: mixture of beta and Bernoulli distributions

In this parameterisation, $y \sim BEINF(\alpha, \gamma, \mu, \phi)$, where $0 < \alpha < 1$ is the mixture parameter, $0 < \gamma < 1$ is the success probability in the Bernoulli part of the mixture, and μ and ϕ are the mean and precision of the beta part of the mixture as in equation (3.3).

The density of y is given by:

$$beinf(y; \alpha, \gamma, \mu, \phi) = \begin{cases} \alpha(1-\gamma), & \text{if } y = 0, \\ \alpha\gamma, & \text{if } y = 1, \\ (1-\alpha)f(y; \mu, \phi), & \text{if } y \in (0,1). \end{cases} \qquad (3.4)$$

3.3.2 Parameterisation 2: point masses at zero and one

If $P(y = 1) = \delta_1 = \alpha\gamma$ and $P(y = 0) = \delta_0 = \alpha - \delta_1$, then the density of y is instead given by:

$$beinf(y; \delta_0, \delta_1, \mu, \phi) = \begin{cases} \delta_0, & \text{if } y = 0, \\ \delta_1, & \text{if } y = 1, \\ (1-\delta_0-\delta_1)f(y; \mu, \phi) & \text{if } y \in (0,1). \end{cases} \qquad (3.5)$$

This requires that $0 < \delta_0 + \delta_1 < 1$ (this is unlikely to be a problem when modelling crown condition data, as the proportion of trees which are either completely healthy or completely dead is never close to 0% or 100%).

3.3.3 Implementation in `gamlss`

Distribution families in `gamlss` may have up to four parameters: `mu` (location), `sigma` (scale), `nu` and `tau` (shape). The relationships between these parameters and the parameters in equation (3.5) are given below.

$$\mu = \texttt{mu} \qquad\qquad \phi = \frac{1 - \texttt{sigma}}{\texttt{sigma}}$$
$$\delta_0 = \frac{\texttt{nu}}{1 + \texttt{nu} + \texttt{tau}} \quad \delta_1 = \frac{\texttt{tau}}{1 + \texttt{nu} + \texttt{tau}} \tag{3.6}$$

So the requirement that $0 < \delta_0 + \delta_1 < 1$ is equivalent to requiring that $\texttt{nu} + \texttt{tau} > 0$.

3.3.4 Very simple model for crown condition

We apply this zero-and-one-inflated beta distribution to the measurements of crown condition in individual spruce trees in Germany between 1989 and 2008, beginning with a very simple model containing no covariates. Table 3.1 summarises the number and percentage of trees falling into each category of crown condition (measured in 5% steps with additional categories for 0% defoliation and "100% defoliation (dead)".

Table 3.1 Defoliation levels in spruce trees in Germany

Needle loss	Frequency	%	Needle loss	Frequency	%
0%	6977	8.9%	>55-60%	380	0.5%
>0-5%	7885	10.1%	>60-65%	285	0.4%
>5-10%	10347	13.2%	>65-70%	154	0.2%
>10-15%	11286	14.4%	>70-75%	87	0.1%
>15-20%	10687	13.7%	>75-80%	81	0.1%
>20-25%	9874	12.6%	>80-85%	40	0.1%
>25-30%	7728	9.9%	>85-90%	37	0.0%
>30-35%	5303	6.8%	>90-95%	45	0.1%
>35-40%	3412	4.4%	>95-100% (alive)	8	0.0%
>40-45%	1878	2.4%	100% (dead)	190	0.2%
>45-50%	1085	1.4%	Total	78270	100%
>50-55%	501	0.6%			

The simplest zero-and-one-inflated beta model is given by:

$$\text{logit}(\texttt{mu}) = \eta_1$$
$$\text{logit}(\texttt{sigma}) = \eta_2$$
$$\log(\texttt{nu}) = \eta_3 \tag{3.7}$$
$$\log(\texttt{tau}) = \eta_4$$

where the linear predictors η_k, $k = 1, 2, 3, 4$, are constants. This model was fitted using the `BEINF` family in `gamlss`, and the resulting parameter estimates are given in Table 3.2.

This is not very useful unless covariates can be added to the model; unfortunately, attempts to do so, even in a very simple way, resulted in the model failing to converge. The `BEINF` family of distributions in the `gamlss` package was a fairly recent development at the time this work was carried out, and the authors do not, in their original paper (Ospina and Ferrari,

Table 3.2 Parameter estimates for the zero-and-one-inflated beta model for crown condition

Parameter	Estimate	Standard error
η_1	-1.407	0.003
η_2	-0.716	0.003
η_3	-2.312	0.013
η_4	-5.925	0.073
mu, μ	0.197	
sigma	0.329	
nu	0.098	
tau	0.003	
ϕ	2.047	
δ_0	0.089	
δ_1	0.003	

2010), demonstrate applications using real data which involve covariate effects. It was felt that attempting a detailed investigation of the possible causes of non-convergence in a model implementation that was still under development might not be the best use of time, and this approach was not pursued further.

(A further point to consider is that any well-fitting model for defoliation would have to account for temporal autocorrelation. This is theoretically possible within the `gamlss` package, using random effects terms, although no attempt has so far been made to combine this with a zero-and-one-inflated beta model.)

Rather than investigating increasingly complicated continuous distributions for defoliation, a more practical option might be to consider it as a categorical variable and fit an ordinal response model instead.

3.4 Defoliation as an ordinal response

For the remainder of this chapter, defoliation is considered as an ordinal categorical variable with four levels: "none", $\leq 10\%$ defoliation; "slight", $>10\text{-}25\%$; "moderate, $>25\text{-}60\%$; and "severe/dead", $>60\text{-}100\%$. This corresponds to the UNECE-EU standards used in the ICP technical report, with the "severe" and "dead" categories combined – numbers of dead trees are very low, and in some countries (including Germany), severely damaged trees may be preemptively felled before they actually die.

For simplicity, analysis is initially restricted to spruce trees in Germany only.

3.4.1 Relationships between defoliation and possible predictors

Looking at Tables 3.3 to 3.7, it appears that higher levels of defoliation may be associated with:

- greater mean stand age

- being on a site facing north or east

- higher altitudes

- "Moder" or "Mor" humus type.

Table 3.3 Defoliation and mean age of dominant trees: frequency table

Defoliation category	Mean age of dominant trees on plot (years)							
	0-20	21-40	41-60	61-80	81-100	101-120	over 120	irregular stands
none	3567	7764	7382	2942	2407	654	289	204
slight	895	2118	6057	7253	8779	3926	2679	140
moderate	288	388	1544	3741	5894	3689	4686	57
severe/dead	38	42	98	111	182	165	280	11

Table 3.4 Defoliation and water budget: frequency table

Defoliation category	Water budget on plot		
	insufficient	sufficient	excessive
none	2524	21377	1308
slight	2688	27996	1163
moderate	1987	17686	614
severe/dead	100	794	33

Table 3.5 Defoliation and plot orientation: frequency table

Defoliation category	Orientation of plot								
	N	NE	E	SE	S	SW	W	NW	flat
none	2548	2343	1763	2413	2369	1216	1574	4022	6961
slight	4299	3862	1429	2082	2182	1850	2493	4190	9460
moderate	2097	2342	619	1162	940	1400	2704	2888	6135
severe/dead	76	119	37	57	96	49	127	91	275

Table 3.6 Defoliation and altitude

Defoliation category	Mean plot altitude (m)
none	499.2
slight	554.1
moderate	575.7
severe/dead	609.9

Table 3.7 Defoliation and humus type: frequency table

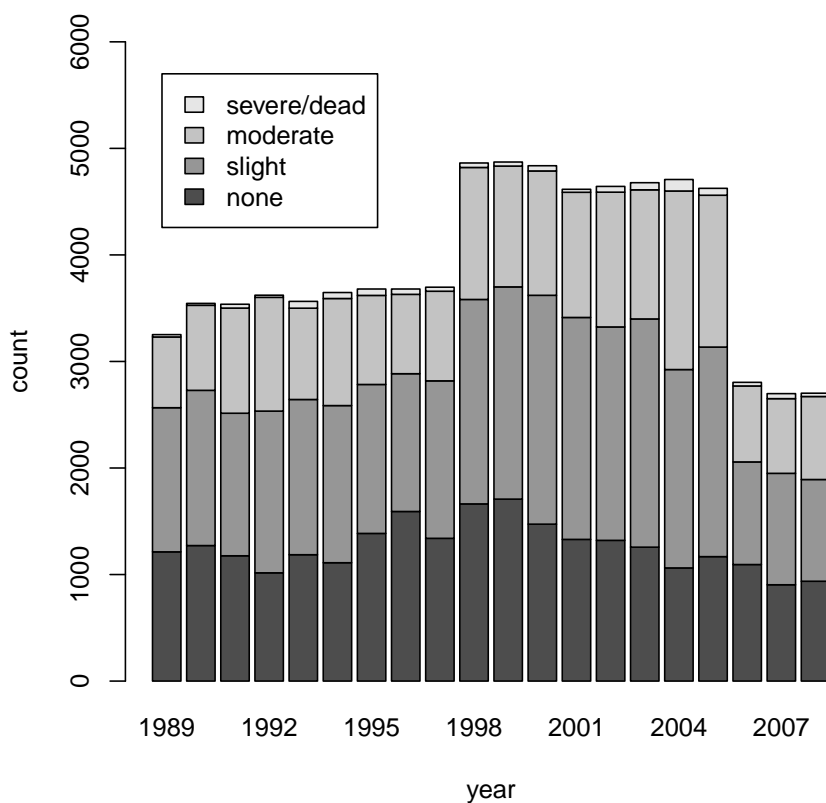| Defoliation category	Orientation of plot						
	Mull	Moder	Mor	Amphi	Anmoor	Histomull	Histomoder
none	4731	10759	1832	311	102	7025	339
slight	4906	13678	2295	178	114	10189	338
moderate	2075	8407	1608	255	76	7489	182
severe/dead	70	510	51	26	1	258	7

Figure 3-5 Numbers of spruce trees in each defoliation category, by year

34

3.5 Structure of models for ordinal responses

3.5.1 Cumulative threshold models

Let U be a continuous latent variable (unobservable), and let $Y \in \{1, ..., k\}$ be an observable categorical variable, with U and Y connected by:

$$Y = r \iff \theta_{r-1} < U < \theta_r, \quad r = 1, ..., k \tag{3.8}$$

with thresholds θ_r satisfying:

$$-\infty = \theta_0 < \theta_1 < ... < \theta_k = \infty \tag{3.9}$$

U is related to the explanatory variables $\mathbf{X}$ by:

$$U = \mathbf{X}^T \boldsymbol{\beta} + \epsilon \tag{3.10}$$

where $\boldsymbol{\beta} = (\beta_1, ..., \beta_p)^T$ is a vector of coefficients and ϵ is a random variable with distribution function F.

So Y is related to the explanatory variables by:

$$\mathbb{P}(Y \leq r \mid \mathbf{X}) = \mathbb{P}(U \leq \theta_r \mid \mathbf{X}) = F(\theta_r - \mathbf{X}^T \boldsymbol{\beta}) \tag{3.11}$$

If F is the logistic function $F(x) = \frac{1}{1+e^{-x}}$, equation (3.11) describes a cumulative logistic model. This is also known as a proportional odds model, since the ratio of the odds of observing $Y \leq r$ rather than $Y > r$ for two different combinations of covariates ($\mathbf{X_1}$ and $\mathbf{X_2}$) does not depend on the category r, i.e.:

$$\frac{\mathbb{P}(Y \leq r \mid \mathbf{X_1})/\mathbb{P}(Y > r \mid \mathbf{X_1})}{\mathbb{P}(Y \leq r \mid \mathbf{X_2})/\mathbb{P}(Y > r \mid \mathbf{X_2})} = \exp((\mathbf{X_2} - \mathbf{X_1})^T \boldsymbol{\beta}) \tag{3.12}$$

Cumulative logistic models can be fitted in R using either the `polr` function from the `MASS` package (Venables and Ripley, 2002) (no smooth terms) or `R2BayesX` (Umlauf et al., 2013), which provides an R interface to the standalone package BayesX (Belitz et al., 2012). BayesX also allows the proportional odds assumption to be relaxed so that separate coefficients are estimated for each category boundary of the ordinal response.

3.5.2 Sequential models

Sequential models may be more appropriate than cumulative models where, in order to move from a lower to a higher category, subjects must pass through all intervening categories (as is the case with defoliation), though without necessarily being observed in each one. Let $U_r, r = 1, ..., k-1$, be latent variables, one for each threshold between categories of the

response Y, with the U_r and Y now connected by:

$$Y = r \mid Y \geq r \iff U_r \leq \theta_r, \quad r = 1, ..., k-1 \tag{3.13}$$

The thresholds θ_r no longer need to satisfy equation (3.9), since each one now relates to a different latent variable U_r. U_r and Y are now related to the explanatory variables by:

$$U_r = \mathbf{X}^T \boldsymbol{\beta} + \epsilon_r \tag{3.14}$$

$$\mathbb{P}(Y = r \mid Y \geq r, \mathbf{X}) = F(\theta_r - \mathbf{X}^T \boldsymbol{\beta}) \tag{3.15}$$

for $r = 1, ..., k-1$, where each ϵ_r is a random variable with distribution function F.

If F is the logistic function, equation (3.15) is referred to as a sequential logistic model. This has a similar property to the cumulative logistic model, in that:

$$\frac{\mathbb{P}(Y = r \mid Y \geq r, \mathbf{X_1})/\mathbb{P}(Y > r \mid Y \geq r, \mathbf{X_1})}{\mathbb{P}(Y = r \mid Y \geq r, \mathbf{X_2})/\mathbb{P}(Y > r \mid Y \geq r, \mathbf{X_2})} = \exp((\mathbf{X_2} - \mathbf{X_1})^T \boldsymbol{\beta}) \tag{3.16}$$

Sequential logistic models can be fitted using `R2BayesX`; category-specific coefficients can also be estimated.

3.6   Exploring cumulative logistic models for defoliation

As before, analysis at this stage is restricted to Norway spruce trees (*Picea abies*) in Germany. First, a model of the form described by equation 3.11 is fitted using the `polr` function in R, including all available covariates except spatial location (i.e. mean age of dominant trees on the plot, water budget, plot orientation, plot altitude, humus type and year). Since `polr` does not allow for smooth functions of covariates, the year of measurement is entered into the model as a categorical variable with twenty levels corresponding to the years from 1989 to 2008. The main aim of this simple model was to assess whether the proportional odds assumption appeared to be consistent with the data.

Table 3.8 displays the results of this model in the form of odds ratios, with ratios greater than one indicating increased odds of defoliation. Tree age is still a major factor, even though the measure available in this dataset is only indirectly related to the age of the individual tree. A water budget classified as "insufficient" is associated with increased odds of higher defoliation compared to a "sufficient" waterbudget, while an "excessive" water budget seems to have a protective effect. Contrary to the impression given by Table 3.5, it appears (after adjusting for other covariates in the model) that only trees on west-facing slopes are more vulnerable to defoliation, while all other orientations are at lower risk compared to trees on level ground; this effect is probably confounded with altitude, however, and will require further investigation. The humus types "Moder" and "Mor" are also associated with increased defoliation compared to the "Mull" type.

Table 3.8 Results of cumulative logistic regression model with proportional odds assumption (Norway spruce, Germany)

Covariate	Odds ratio	95% CI
Mean age of dominant trees: reference category = 0-20 years		
21-40 years	1.08	1.01, 1.15
41-60 years	3.48	3.29, 3.68
61-80 years	12.18	11.49, 12.91
81-100 years	18.17	17.16, 19.25
101-120 years	27.41	25.68, 29.25
over 120 years	55.25	51.57, 59.19
irregular stands	4.10	3.71, 4.54
Water budget: reference category = sufficient		
insufficient	1.06	1.01, 1.12
excessive	0.77	0.70, 0.84
Plot orientation: reference category = flat ground		
N	0.77	0.73, 0.81
NE	0.77	0.73, 0.81
E	0.80	0.74, 0.86
SE	0.86	0.81, 0.92
S	0.89	0.84, 0.95
SW	1.00	0.94, 1.07
W	1.15	1.08, 1.21
NW	0.88	0.84, 0.92
Altitude (per 100m)	1.07	1.06, 1.08
Humus type: reference category = Mull		
Moder	1.53	1.47, 1.60
Mor	1.65	1.54, 1.77
Amphi	1.12	0.95, 1.33
Anmoor	1.10	1.07, 1.13
Histomull	1.27	1.22, 1.33
Histomoder	1.28	1.10, 1.48

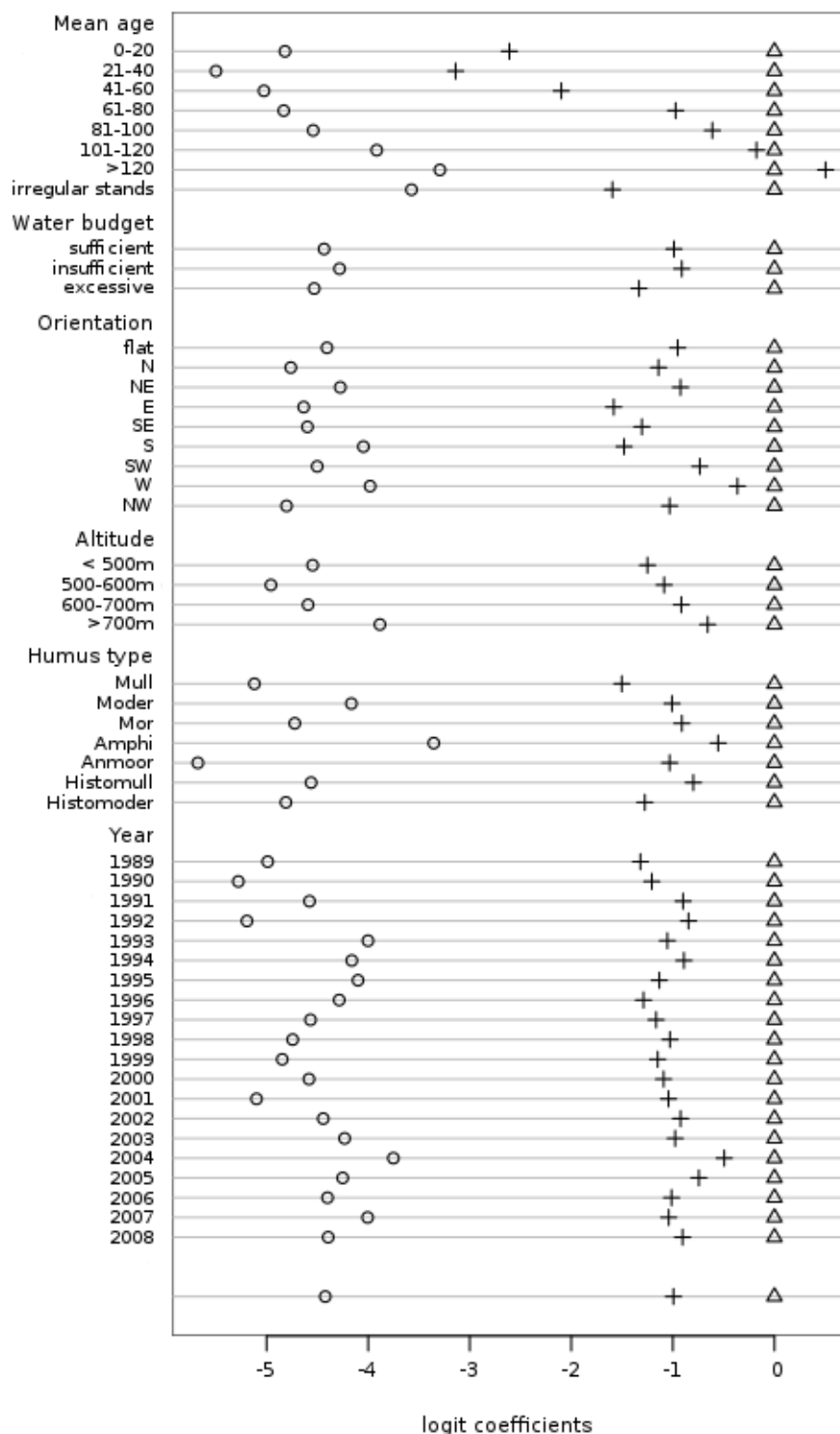| Covariate	Odds ratio	95% CI
Year: reference category = 1989		
1990	1.15	1.08, 1.23
1991	1.53	1.43, 1.63
1992	1.88	1.76, 2.00
1993	1.40	1.31, 1.49
1994	1.52	1.43, 1.62
1995	1.04	0.97, 1.10
1996	0.81	0.76, 0.87
1997	1.09	1.02, 1.16
1998	1.09	1.03, 1.15
1999	0.99	0.94, 1.05
2000	1.12	1.06, 1.19
2001	1.21	1.14, 1.28
2002	1.36	1.29, 1.44
2003	1.36	1.29, 1.44
2004	2.17	2.05, 2.29
2005	1.64	1.55, 1.73
2006	1.28	1.19, 1.38
2007	2.37	2.19, 2.57
2008	1.38	1.28, 1.49

Figure 3-6 Checking the proportional odds assumption: odds ratios from separate logistic regressions for each threshold

A relatively straightforward method for checking the proportional odds assumption graphically has been suggested by Harrell (2001). For a given covariate, each threshold of the ordinal response is considered separately and a binary logistic regression model is fitted where the outcome is defined according to whether the original response was above or below that threshold; the distances between the log odds ratios for the different thresholds should be equal across all categories of a variable if the proportional odds assumption is satisfied. Figure 3-6 is the result of applying this method to the defoliation data, and it suggests that the proportional odds assumption does not hold for mean stand age, orientation, altitude or humus type. This will need to be addressed in whichever model is finally chosen, perhaps by allowing threshold-specific covariate effects.

A second cumulative logistic regression model was subsequently fitted using the `R2BayesX` package so that smooth functions of time and altitude could be included; a third model was also fitted in which the proportional odds assumption was relaxed and threshold-specific coefficients were estimated for all covariates (results not shown). Associations with defoliation were broadly similar in the second model to those shown in Table 12, except that the effect of age was reduced substantially. As expected, the results of the third model confirmed that the proportional odds assumption does not hold for these data, with covariate effects varying substantially depending on which threshold of the ordinal response was being considered.

Clearly the final model will need to in some way deal with the fact that the proportional odds assumption is not satisfied. Another feature of the data which must be addressed by the model is temporal autocorrelation, which (for ordinal responses) cannot be modelled in BayesX.

Chapter 4

Modelling defoliation as an ordinal response with temporal autocorrelation

Defoliation displays fairly high correlation between measurements in successive years, even if observations on individual trees are aggregated at plot level before analysis (Augustin et al., 2009). It seems reasonable to assume that this temporal autocorrelation will be even stronger when considering tree-level data, so any sensible model for these data should be able to incorporate it. Models with an ordinal response and temporal autocorrelation are not currently implemented in BayesX.

Some work has been done (Wang and Kockelman, 2009; Higgs and Hoef, 2012; Stegmueller, 2013) taking a Bayesian approach to fit cumulative probit threshold models with temporal autocorrelation, with the authors writing their own R or Matlab code. So far, it does not appear that anyone has developed a cumulative logistic threshold model with temporal autocorrelation – one advantage of the logistic model is that the coefficients are (log) odds ratios and so more easily interpreted than probit coefficients – and nor has there been any work to date on a sequential logistic threshold model (which is probably more appropriate for the defoliation data).

Using Wang and Kockelman (2009) as a starting point, we now construct a cumulative logistic threshold model with temporal autocorrelation.

4.1 Cumulative logistic threshold model with temporal autocorrelation

4.1.1 Setting up the model

Defoliation (% leaf/needle loss) is considered to be an unobserved (latent) continuous variable U_{it}, where $i = 1, ..., N$ indexes trees and $t = 1, ...T$ indexes time in years. The observed defoliation, y_{it}, is an ordinal variable with k categories, which is determined from U_{it} by:

$$y_{it} = r \iff \theta_{r-1} < U_{it} \leq \theta_r, \quad r = 1, ..., k \tag{4.1}$$

with the thresholds θ_r satisfying:

$$-\infty = \theta_0 < \theta_1 < ... < \theta_k = \infty \tag{4.2}$$

Adding an AR(1) temporal autocorrelation term to the standard form of the cumulative threshold model gives the following:

$$U_{it} = \rho U_{i(t-1)} + \mathbf{X}_{it}^T \boldsymbol{\beta} + \epsilon_{it} \tag{4.3}$$

where ρ is the temporal autocorrelation coefficient, with $|\rho| < 1$; $\boldsymbol{\beta} = (\beta_1, ..., \beta_p)^T$ is a vector of coefficients; $\mathbf{X}_{it}$ is a vector of explanatory variables; and ϵ_{it} is a random variable with distribution function F (the ϵ_{it} are assumed to be i.i.d.). If F is the standard logistic function $F(x) = \frac{1}{1+e^{-x}}$, this is a cumulative logistic threshold model.

This method of including temporal autocorrelation in the model follows the approach of Wang and Kockelman (2009); other, more natural, parametrisations are possible (see Section 4.2).

It should be noted that $y_{i1} : i = 1, ..., N$, the defoliation values observed in the first year of measurement, depend (through U_{i1}) on the values of U_{i0}, which have no corresponding observed y_{i0}. This is also the result of following Wang and Kockelman (2009) at this stage, and improvements may be possible.

We have:

$$\begin{aligned}
\mathbb{P}(U_{it} \leq \theta_r) &= \mathbb{P}(\rho U_{i(t-1)} + \mathbf{X}_{it}^T \boldsymbol{\beta} + \epsilon_{it} \leq \theta_r) \\
&= \mathbb{P}(\epsilon_{it} \leq \theta_r - \rho U_{i(t-1)} - \mathbf{X}_{it}^T \boldsymbol{\beta}) \\
&= F(\theta_r - \rho U_{i(t-1)} - \mathbf{X}_{it}^T \boldsymbol{\beta})
\end{aligned} \tag{4.4}$$

and:

$$\begin{aligned}
\mathbb{P}(y_{it} = r) &= \mathbb{P}(U_{it} \leq \theta_r) - \mathbb{P}(U_{it} \leq \theta_{r-1}) \\
&= F(\theta_r - \rho U_{i(t-1)} - \mathbf{X}_{it}^T \boldsymbol{\beta}) - F(\theta_{r-1} - \rho U_{i(t-1)} - \mathbf{X}_{it}^T \boldsymbol{\beta})
\end{aligned} \tag{4.5}$$

So the distribution of $\mathbf{y} = (y_{11}, ..., y_{N1}, ..., y_{NT})$ depends on $\boldsymbol{\beta}, \boldsymbol{\theta}, \rho, \mathbf{U} = (U_{11}, ..., U_{N1}, ..., U_{NT})$,

41

and $\mathbf{U}_0 = (U_{10}, ..., U_{N0})$.

4.1.2 Posterior joint distribution of the parameters

According to Bayes' rule, posterior $\propto$ likelihood $\times$ prior, we have:

$$p(\boldsymbol{\beta}, \boldsymbol{\theta}, \rho, \mathbf{U}, \mathbf{U_0} \mid \mathbf{y}) \propto p(\mathbf{y} \mid \boldsymbol{\beta}, \boldsymbol{\theta}, \rho, \mathbf{U}, \mathbf{U_0})\pi(\boldsymbol{\beta}, \boldsymbol{\theta}, \rho, \mathbf{U}, \mathbf{U_0}) \quad (4.6)$$

$\mathbf{U}$ depends on $\boldsymbol{\beta}$, ρ and $\mathbf{U}_0$ (by equation (4.3)) and θ is independent of the other parameters, so the prior joint density can be broken down as follows:

$$\begin{aligned} \pi(\boldsymbol{\beta}, \boldsymbol{\theta}, \rho, \mathbf{U}, \mathbf{U_0}) &= \pi((\boldsymbol{\beta}, \rho, \mathbf{U}, \mathbf{U_0})\pi(\boldsymbol{\theta}) \\ &= \pi(\mathbf{U} \mid \boldsymbol{\beta}, \rho, \mathbf{U}_0)\pi(\boldsymbol{\beta})\pi(\rho)\pi(\mathbf{U}_0)\pi(\boldsymbol{\theta}) \end{aligned} \quad (4.7)$$

Also, $\mathbf{y}$ depends on $\boldsymbol{\beta}$, ρ and $\mathbf{U}_0$ only through $\mathbf{U}$, so the likelihood can be simplified:

$$p(\mathbf{y} \mid \boldsymbol{\beta}, \boldsymbol{\theta}, \rho, \mathbf{U}, \mathbf{U_0}) = p(\mathbf{y} \mid \mathbf{U}, \boldsymbol{\theta}) \quad (4.8)$$

Taking equations (4.6), (4.7) and (4.8) together, the posterior joint density for the parameters of interest can be expressed as:

$$p(\boldsymbol{\beta}, \boldsymbol{\theta}, \rho, \mathbf{U}, \mathbf{U_0} \mid \mathbf{y}) \propto p(\mathbf{y} \mid \mathbf{U}, \boldsymbol{\theta})\pi(\mathbf{U} \mid \boldsymbol{\beta}, \rho, \mathbf{U}_0)\pi(\boldsymbol{\beta})\pi(\rho)\pi(\mathbf{U}_0)\pi(\boldsymbol{\theta}) \quad (4.9)$$

4.1.3 Choosing priors

Non-informative prior distributions are needed for $\boldsymbol{\beta}$, $\boldsymbol{\theta}$, ρ, $\mathbf{U}_0$ and $\mathbf{U} \mid \boldsymbol{\beta}, \rho, \mathbf{U}_0$.

A suitable prior for $\boldsymbol{\beta}$ would be multivariate normal, $\boldsymbol{\beta} \sim N(\mathbf{c}, h\mathbf{I}_Q)$ (where Q is the number of parameters). This is non-informative for small $\mathbf{c}$ and large h.

Similarly, a multivariate normal prior could be used for $\boldsymbol{\theta}$, with the additional condition that each threshold must be strictly greater than the previous one (equation (4.2)); i.e. $\boldsymbol{\theta} \sim N(\mathbf{q}, g\mathbf{I}_{k-1}) \cdot \delta(\theta_1 < ... < \theta_{k-1})$ for small $\mathbf{q}$ and large g, where δ is an indicator function.

(Alternatively, Albert and Chib (2001) suggest using the transformation $\alpha_1 = log(\theta_1), \alpha_r = log(\theta_{r-1} - \theta_r), 2 \leq r \leq k - 1$ and assigning a multivariate normal prior to $\boldsymbol{\alpha}$.)

The uniform distribution on (-1,1) would be a suitable prior for ρ.

From equation (4.3), we have that $U_{it} \mid U_{i(t-1)}, \boldsymbol{\beta}, \rho \sim F(\rho U_{i(t-1)} + \mathbf{X}_{it}^T\boldsymbol{\beta})$, where F is the cdf of the standard logistic function. The corresponding pdf would be $F'(x) = \frac{e^{-x}}{(1+e^{-x})^2}$. Assuming the U_{it} are (conditionally) i.i.d., the conditional prior for $\mathbf{U}$ would be:

$$\pi(\mathbf{U} \mid \boldsymbol{\beta}, \rho, \mathbf{U}_0) = \prod_{t=1}^{T}\prod_{i=1}^{N} \frac{e^{-(U_{it} - \rho U_{i(t-1)} - \mathbf{X}_{it}^T\boldsymbol{\beta})}}{(1 + e^{-(U_{it} - \rho U_{i(t-1)} - \mathbf{X}_{it}^T\boldsymbol{\beta})})^2} \quad (4.10)$$

To match the distributions of $\mathbf{U}_1$, $\mathbf{U}_2$, and so on, the prior for $\mathbf{U}_0$ could also be a logistic distribution. In its general form, the logistic distribution has the cdf:

$$F(x) = \frac{1}{1 + e^{-\frac{x-\mu}{s}}} \tag{4.11}$$

and corresponding pdf:

$$f(x) = \frac{e^{-\frac{x-\mu}{s}}}{s(1 + e^{-\frac{x-\mu}{s}})^2} \tag{4.12}$$

where μ is the mean and s is proportional to the standard deviation (in fact $s = \frac{\sqrt{3}}{\pi}$ s.d.). For small μ and large s, this would represent a non-informative prior for $\mathbf{U}_0$, with each U_{i0} being drawn independently from the same distribution.

4.1.4 Conditional posterior distributions for the parameters

From the joint posterior density in equation (4.9), the forms of the conditional posterior densities for the individual parameters can be derived as follows:

β (covariate effects)

$$
\begin{aligned}
p(\boldsymbol{\beta} \mid \boldsymbol{\theta}, \rho, \mathbf{U}, \mathbf{U}_0, \mathbf{y}) &\propto \pi(\mathbf{U} \mid \boldsymbol{\beta}, \rho, \mathbf{U}_0)\pi(\boldsymbol{\beta}) \\
&= \prod_{t=1}^{T}\prod_{i=1}^{N} \frac{exp(-(U_{it} - \rho U_{i(t-1)} - \mathbf{X}_{it}^T\boldsymbol{\beta}))}{(1 + exp(-(U_{it} - \rho U_{i(t-1)} - \mathbf{X}_{it}^T\boldsymbol{\beta})))^2} \\
&\quad \cdot \frac{1}{\sqrt{(2\pi)^Q h^Q}} exp(-\frac{1}{2}(\boldsymbol{\beta} - \mathbf{c})^T h^{-Q}(\boldsymbol{\beta} - \mathbf{c})) \\
&\propto exp(-\frac{1}{2}(\boldsymbol{\beta} - \mathbf{c})^T h^{-Q}(\boldsymbol{\beta} - \mathbf{c})) \\
&\quad \cdot \prod_{t=1}^{T}\prod_{i=1}^{N} \frac{exp(-(U_{it} - \rho U_{i(t-1)} - \mathbf{X}_{it}^T\boldsymbol{\beta}))}{(1 + exp(-(U_{it} - \rho U_{i(t-1)} - \mathbf{X}_{it}^T\boldsymbol{\beta})))^2}
\end{aligned} \tag{4.13}
$$

θ (threshold parameters)

$$
\begin{aligned}
p(\boldsymbol{\theta} \mid \boldsymbol{\beta}, \rho, \mathbf{U}, \mathbf{U}_0, \mathbf{y}) &\propto p(\mathbf{y} \mid \mathbf{U}, \boldsymbol{\theta})\pi(\boldsymbol{\theta}) \\
&= \prod_{t=1}^{T}\prod_{i=1}^{N}\sum_{r=1}^{k} \delta(y_{it} = r) \cdot \delta(\theta_{r-1} < U_{it} \le \theta_r) \\
&\quad \cdot (2\pi g)^{-\frac{k-1}{2}} exp(-\frac{1}{2}(\boldsymbol{\theta} - \mathbf{q})^T(g\mathbf{I}_{k-1})^{-1}(\boldsymbol{\theta} - \mathbf{q}))\delta(\theta_1 < ... < \theta_{k-1})
\end{aligned} \tag{4.14}
$$

This is equivalent to equation (A.23) in Wang and Kockelman (2009), and reduces similarly to a set of truncated normal distributions if each threshold parameter θ_r is considered separately.

$$p(\boldsymbol{\theta}_r \mid \boldsymbol{\beta}, \boldsymbol{\theta}_{\neq r}, \rho, \mathbf{U}, \mathbf{U}_0, \mathbf{y}) \propto \delta(\theta_r^{inf} < \theta_r < \theta_r^{sup}) \cdot exp(-\frac{1}{2g}(\theta_r - q_r)^2)$$

$$\text{where } \theta_r^{inf} = max\{max\{U_{it} : y_{it} = r\}, \theta_{r-1}\}$$

$$\text{and } \theta_r^{sup} = min\{min\{U_{it} : y_{it} = r + 1\}, \theta_{r+1}\} \quad (4.15)$$

ρ (temporal autocorrelation parameter)

$$p(\rho \mid \boldsymbol{\beta}, \boldsymbol{\theta}, \mathbf{U}, \mathbf{U}_0, \mathbf{y}) \propto \pi(\mathbf{U} \mid \boldsymbol{\beta}, \rho, \mathbf{U}_0)\pi(\rho)$$

$$= \prod_{t=1}^{T}\prod_{i=1}^{N} \frac{exp(-(U_{it} - \rho U_{i(t-1)} - \mathbf{X}_{it}^T\boldsymbol{\beta}))}{(1 + exp(-(U_{it} - \rho U_{i(t-1)} - \mathbf{X}_{it}^T\boldsymbol{\beta})))^2} \quad (4.16)$$

$$\cdot \frac{1}{2}\delta(\rho \in (-1, 1))$$

$\mathbf{U}_0$ (initial values of the latent variable)

$$p(\mathbf{U}_0 \mid \boldsymbol{\beta}, \boldsymbol{\theta}, \rho, \mathbf{U}, \mathbf{y}) \propto \pi(\mathbf{U} \mid \boldsymbol{\beta}, \rho, \mathbf{U}_0)\pi(\mathbf{U}_0)$$

$$= \prod_{t=1}^{T}\prod_{i=1}^{N} \frac{exp(-(U_{it} - \rho U_{i(t-1)} - \mathbf{X}_{it}^T\boldsymbol{\beta}))}{(1 + exp(-(U_{it} - \rho U_{i(t-1)} - \mathbf{X}_{it}^T\boldsymbol{\beta})))^2} \quad (4.17)$$

$$\cdot \prod_{i=1}^{N} \frac{exp(-(U_{i0} - a)d^{-1})}{d(1 + exp(-(U_{i0} - a)d^{-1}))^2}$$

Taking each U_{i0} separately and ignoring factors not involving U_{i0} (which can be considered as part of the normalising constant), we have:

$$p(U_{i0} \mid \boldsymbol{\beta}, \boldsymbol{\theta}, \rho, \mathbf{U}_{\neq i,0}, \mathbf{U}, \mathbf{y}) \propto \frac{exp(-(U_{i1} - \rho U_{i0} - \mathbf{X}_{i1}^T\boldsymbol{\beta}))}{(1 + exp(-(U_{i1} - \rho U_{i0} - \mathbf{X}_{i1}^T\boldsymbol{\beta})))^2}$$

$$\cdot \frac{exp(-(U_{i0} - a)d^{-1})}{d(1 + exp(-(U_{i0} - a)d^{-1}))^2} \quad (4.18)$$

U (values of the latent variable at all other time points)

$$p(\mathbf{U} \mid \boldsymbol{\beta}, \boldsymbol{\theta}, \rho, \mathbf{U}_0, \mathbf{y}) \propto p(\mathbf{y} \mid \mathbf{U}, \boldsymbol{\theta})\pi(\mathbf{U} \mid \boldsymbol{\beta}, \rho, \mathbf{U}_0)$$

$$= \prod_{t=1}^{T}\prod_{i=1}^{N}\sum_{r=1}^{k} \delta(y_{it} = r) \cdot \delta(\theta_{r-1} < U_{it} \leq \theta_r) \quad (4.19)$$

$$\cdot \prod_{t=1}^{T}\prod_{i=1}^{N} \frac{exp(-(U_{it} - \rho U_{i(t-1)} - \mathbf{X}_{it}^T\boldsymbol{\beta}))}{(1 + exp(-(U_{it} - \rho U_{i(t-1)} - \mathbf{X}_{it}^T\boldsymbol{\beta})))^2}$$

Again considering each U_{it} separately and ignoring factors which could be absorbed into the normalising constant, we have:

$$p(U_{it} \mid \boldsymbol{\beta}, \boldsymbol{\theta}, \rho, \mathbf{U}_0, \mathbf{U}_{\neq i, \neq t}, \mathbf{y}) \propto \sum_{r=1}^{k} \delta(y_{it} = r) \cdot \delta(\theta_{r-1} < U_{it} \leq \theta_r)$$

$$\cdot \frac{exp(-(U_{it} - \rho U_{i(t-1)} - \mathbf{X}_{it}^T \boldsymbol{\beta}))}{(1 + exp(-(U_{it} - \rho U_{i(t-1)} - \mathbf{X}_{it}^T \boldsymbol{\beta})))^2} \quad (4.20)$$

$$\cdot \frac{exp(-(U_{i(t+1)} - \rho U_{it} - \mathbf{X}_{i(t+1)}^T \boldsymbol{\beta}))}{(1 + exp(-(U_{i(t+1)} - \rho U_{it} - \mathbf{X}_{i(t+1)}^T \boldsymbol{\beta})))^2}$$

The conditional posterior density for $\boldsymbol{\beta}$ is the product of i.i.d. logistic densities and a multivariate normal density, which does not simplify to any standard distribution. Posterior densities for the other parameters behave similarly.

As a first step towards fitting this model to the (very large) European defoliation dataset, we simulated a small dataset ($n = 100$ trees, $T = 20$ years) with a single covariate (time) having only a simple linear effect on the progress of defoliation:

$$U_{it} = \rho U_{i(t-1)} + \beta t + \epsilon_{it} \quad (4.21)$$

with y_{it} (the observed defoliation in ordinal categories) determined by U_{it} as given in equation 4.1. The parameter values used when simulating the data were $\beta = 0.3$, $\rho = 0.2$ and $\boldsymbol{\theta} = (-1, 5, 10)$; values of U_{i0} are randomly generated from a Logistic(2,3) distribution. The model given by equation 4.21is then fitted to the simulated data using Markov Chain Monte Carlo (MCMC) methods, with the Metropolis-Hastings algorithm employed to construct the chain in order to (temporarily) avoid the issues raised by the complex posterior distributions described above. The R code for simulating the data and fitting this model is reproduced in the Appendix.

Figures 4-1 to 4-3 are trace plots for β, ρ and $\boldsymbol{\theta}$ after running the algorithm for 1,000,000 iterations. After a fairly lengthy burn-in period, the distributions of $\theta_1, \theta_2, \theta_3$ appeared stable and were reasonably close to the original values used in the simulation; however, β was slightly too high (around 0.4 rather than 0.3) while the temporal autocorrelation coefficient ρ had collapsed to zero. Figures 4-4 and 4-5 are trace plots for the values of the latent variables $U_{9,0}$ and $U_{9,4}$ respectively, i.e. for tree 9 in year 0 (the year before data collection started) and for the same tree in year 4. Other simulated trees in other years showed very similar traces. Variability is much greater than for the other parameters, and for $U_{9,0}$ it is not clear whether or not the chain is mixing well.
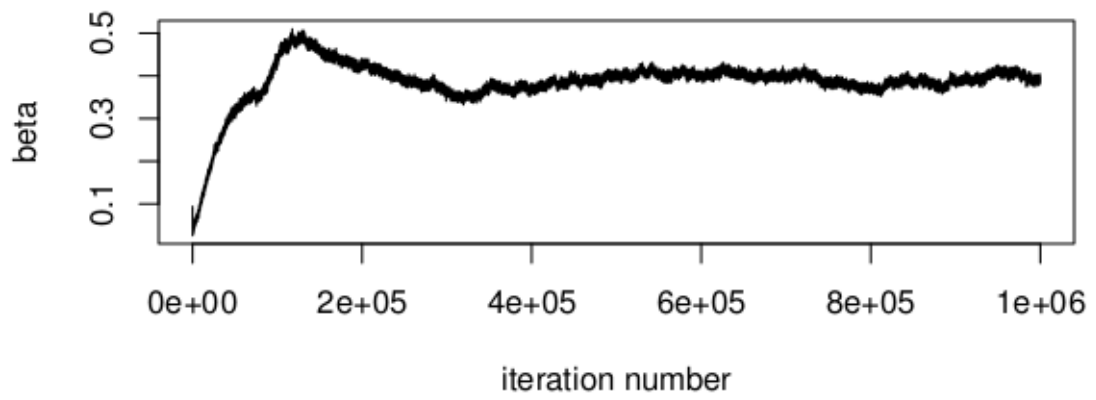
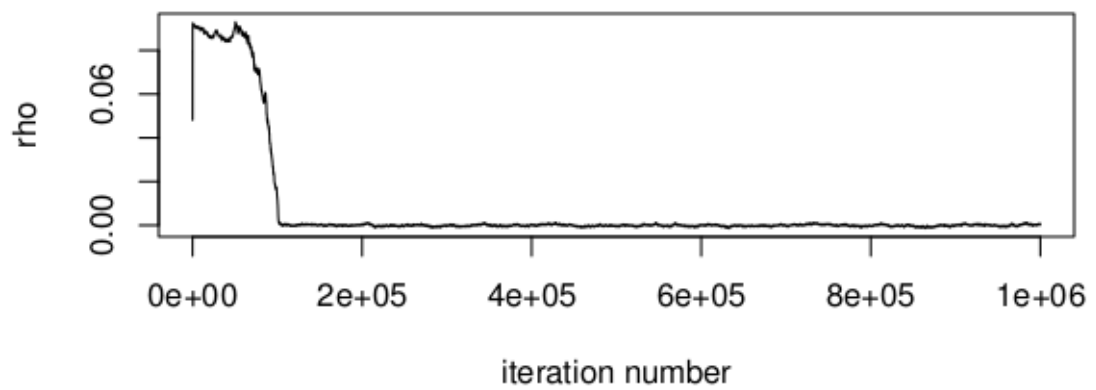**Figure 4-1** MCMC trace plot for β

Figure 4-2 MCMC trace plot for ρ
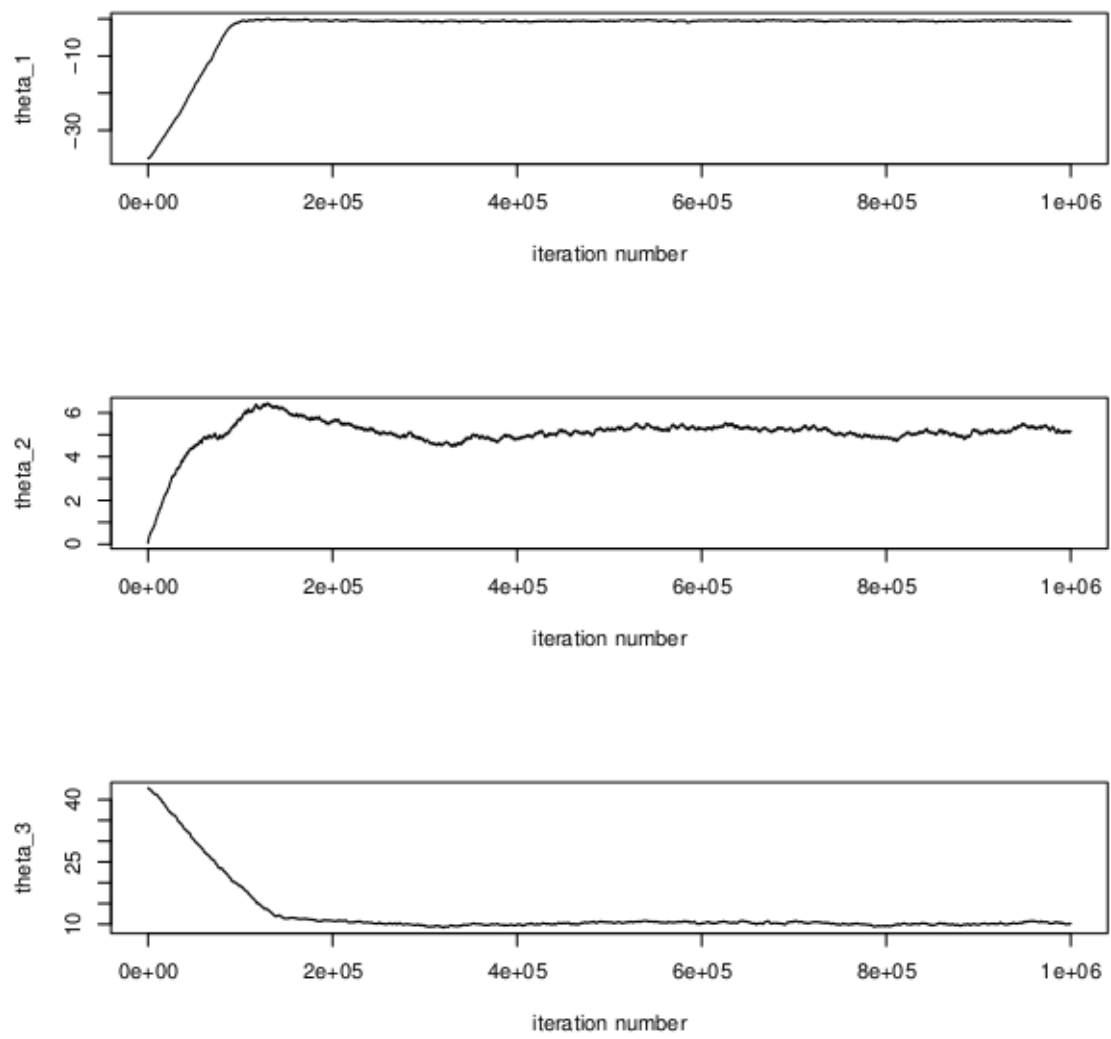
Figure 4-3 MCMC trace plots for $\theta_1, \theta_2, \theta_3$
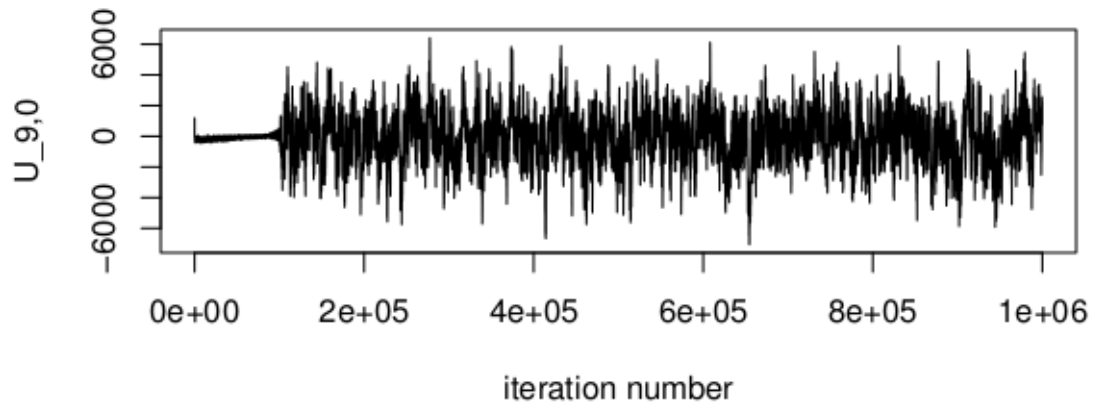
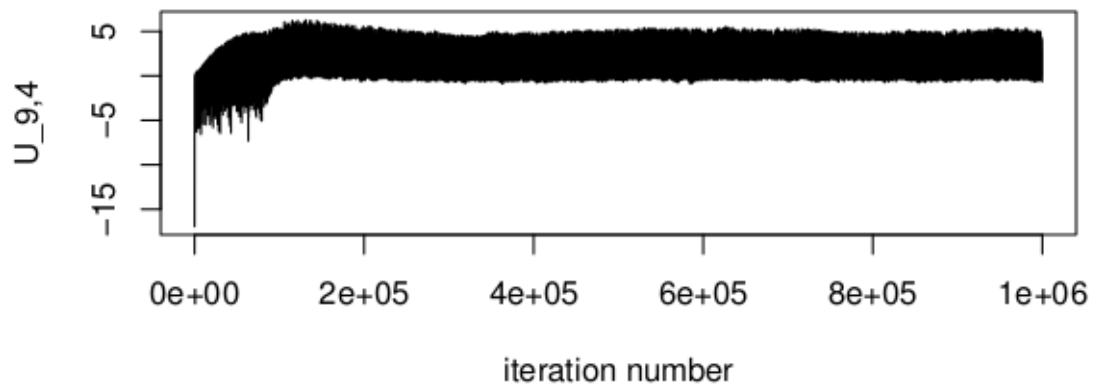**Figure 4-4** Example MCMC trace plot for U_{i0} (a single tree)

Figure 4-5 Example MCMC trace plot for U_{it} (a single tree in a single year)

48

4.2 Improving the algorithm

As it stands, this algorithm is neither reliable nor computationally efficient enough to be applied to the real defoliation data; however, it provides a basis from which to work towards this goal, through a combination of improvements to the form of the current model. Some possibilities are listed below.

1. The model in its current form conflates the temporal trend (β) with the temporal autocorrelation (ρ), so that it is difficult for the MCMC algorithm to estimate both accurately. This was confirmed by simulating data from the model given by 4.21 with various different values of ρ and fixed β; increasing ρ induces an apparently greater temporal drift in the latent variable $\mathbf{U}$. This issue could be addressed by reformulating the model as given by Equation 4.22. The interpretation of $\mathbf{X}_{i(t-1)}^T \boldsymbol{\beta}$ would then be straightforward, as it would simply represent the expected value at each point (which is not the case currently).

$$U_{it} - \mathbf{X}_{it}^T \boldsymbol{\beta} = \rho(U_{i(t-1)} - \mathbf{X}_{i(t-1)}^T \boldsymbol{\beta}) + \epsilon_{it} \tag{4.22}$$

2. Rather than starting at $t = 0$ and making assumptions about the distribution of $\mathbf{U}_0$ with no information (since $\mathbf{y}_0$ is not observed), we could start instead at $t = 1$. This would speed up the algorithm, as there would be fewer parameters to estimate.

3. The intractable forms of some of the posterior distributions slow down the MCMC estimation considerably; finding workable approximations to some or all of these would allow the Metropolis-Hastings steps to be dispensed with in favour of faster methods.

4. At the moment, $\mathbf{U}$ and the threshold parameters $\theta_1, \theta_2, \theta_3$ (which together determine the observed values $\mathbf{y}$) are all free to vary. Fixing one of the threshold parameters (θ_1, say) would speed up the burn-in period without restricting the model in any meaningful way, since the scale of $\mathbf{U}$ and $\boldsymbol{\theta}$ doesn't matter.

Chapter 5

Conclusions and future work

5.1 Plot-level models of defoliation

This thesis began by applying previously-developed methods for modelling crown condition to part of the Europe-wide ICP Forests data set. Estimates of smooth, non-linear temporal and spatial trends in defoliation for Norway spruce in Germany, Austria and Switzerland were produced, together with measures of the uncertainty surrounding these trends.

As discussed in chapter 2, the overall temporal trend has been towards poorer crown condition, with median defoliation rising from approximately 16% in 1989 to a peak of 22% in 2005-6 and falling back slightly to 20% in 2010. The uncertainty around each of these estimates is less than 1.5 percentage points in either direction. Spatial patterns are less definite, although the fogplot method for displaying spatial trends in change over time will hopefully prove to be a useful tool. Tree age, as expected, is a major factor, with plots dominated by the oldest trees exhibiting defoliation levels up to 15 percentage points higher than those containing mainly young trees.

Models were adjusted for the "country effect" – sharp, and ecologically implausible, differences in average defoliation between countries – which is believed to be due to differences in interpretation of the protocols for assessing crown condition. It is clear that any future attempts to model ICP Forests data from more than one country must take this effect into account. (It may also be necessary to allow for the size of the effect to vary over time as adherence to protocols improves, although this should be done cautiously to avoid confusion with the real temporal trend in defoliation.)

Future work on plot-level models might include:

- extensions to a wider geographical area, possibly using soap-film smoothers or similar to avoid smoothing over the sea, and taking care to adjust appropriately for the various country effects;

- modelling defoliation in other common tree species, such as beech (*Fagus sylvatica*) or sub-temperate oak (*Quercus* species);

- investigating covariate effects, using, for example, data on soil characteristics which has also been collected from Level I plots as part of the ICP Forests programme.

5.2 Tree-level models of defoliation

Although plot-level models are adequate for monitoring large-scale trends in crown condition, it may also be informative to look at trends in individual trees. Some progress has been made towards this goal in the second half of this thesis, with the development of an ordinal logistic model with temporal autocorrelation, estimated using MCMC methods.

Specific suggestions for immediate improvements to the efficiency and interpretability of this current model are listed at the end of the previous chapter. If these prove successful, the model could then be extended to incorporate other essential features, such as:

- a more flexible, non-linear temporal trend;

- extra random effects terms to model spatial correlations between trees on the same plot;

- inclusion of other covariates besides tree age;

- allowance for missing data (not all trees are measured in all years).

Such a model would allow the crown condition trajectories of individual trees to be mapped and the factors influencing them – which might include tree-specific variables such as insect damage – to be studied.

Appendix A

R code for drawing fogplots

See Section 2.6.

```
### Function to draw maps displaying, on a regular grid, predicted change in
### defoliation between two time points, with transparency (alpha values)
### indicating the probability of a non-zero change.

### Values must be supplied for all covariates used in the model, even though these
### will cancel out in change plots.

### If label=TRUE (default), plots will be labelled with the appropriate year(s).

### If show.plots=TRUE, locations of survey plots will be marked on the map.

### If legend=TRUE, a legend will be created in the top right of the map.

### The plot will be saved in the directory given by the path argument.

fogplot <- function(model,covs=list(),ngrid=50,too.far=0.05,yearsdiff=NULL,
    colours=c("dodgerblue","red"),path,borders=c("Germany","Austria","Switzerland"),
    tag="GAS",maxchange=15,label=TRUE,show.plots=FALSE,legend=FALSE) {
    libraries <- c("rgl","maps","mapdata","MASS","mgcv","INLA")
    for (l in libraries) require(l,character.only=TRUE)

    if (is.null(yearsdiff)) stop("Must specify yearsdiff")
    if (!is.numeric(yearsdiff) || length(yearsdiff)!=2) {
        stop("Must specify yearsdiff as a vector of two years")
        }
    if (substring(path,nchar(path))!="/") path <- paste(path,"/",sep="")

    dat <- model$lme$data
    yearsdiff <- sort(yearsdiff)

    # Predict on a regular grid covering all observed locations
    ux <- unique(dat$x); uy <- unique(dat$y)
    gx <- seq(min(ux),max(ux),length=ngrid)
    gy <- seq(min(uy),max(uy),length=ngrid)

    ndat <- expand.grid(x=gx,y=gy)

    predterms <- attr(model$gam$terms,"term.labels")
```

52

```
predterms <- predterms[!predterms %in% c("x","y","year")]
for (term in predterms) {
    ndat[,term] <- covs[term]
}

M1 <- predict(model$gam,newdata=cbind(ndat,year=yearsdiff[1]),type="lpmatrix")
M2 <- predict(model$gam,newdata=cbind(ndat,year=yearsdiff[2]),type="lpmatrix")

# Simulate from posterior distribution of parameters to get posterior distributions
# for defoliation at each grid point for the year/s required
simcoef <- mvrnorm(n=1000,mu=coef(model$gam),Sigma=model$gam$Vp)
simfit1 <- M1%*%t(simcoef)
simfit2 <- M2%*%t(simcoef)
simfit <- 100*(plogis(simfit2)-plogis(simfit1))

# Mark grid points too far from data
has.data <- unique(dat[,c("x","y")])
ndat$exclude <- exclude.too.far(ndat$x,ndat$y,has.data$x,has.data$y,too.far)

# Function to extract medians and alpha values from simulated distributions
alpha.sim <- function(x) {
    m <- median(x)
    Fn <- ecdf(x) # empirical CDF
    p1 <- Fn(0)
    alpha <- 1 - 2 * min(p1,1-p1) # 1 - P(|change|>observed | H_0:true change = 0)
    list(median=m,alpha=alpha)
}

# Calculate alpha values for each grid point
ma <- unlist(apply(simfit,1,alpha.sim))
ma <- data.frame(matrix(ma,length(ma)/2,2,byrow=TRUE))
names(ma) <- c("median","alpha")
ndat$simmedians <- ma$median
ndat$alpha <- ma$alpha
ndat[ndat$exclude,c("simmedians","alpha")] <- NA

# Reshape medians and alphas into matrix forms needed for plots
xx <- reshape(ndat,direction="wide",v.names="simmedians",drop=c("alpha",
    "exclude",predterms),idvar="x",timevar="y")
xx <- as.matrix(xx[,-1])
aa <- reshape(ndat,direction="wide",v.names="alpha",drop=c("simmedians",
    "exclude",predterms),idvar="x",timevar="y")
aa <- as.matrix(aa[,-1])
aa[is.na(aa)] <- 0

# Misc. plot parameters
windowRect <- 50+c(0,0,400,400)
zoom <- 0.8
line.width <- 2

# Set up colour palette
cp <- colorRampPalette(colours)
cc <- as.array(inla.generate.colors(as.vector(xx),color.axis=c(-maxchange,maxchange),
    color.palette=cp)$colors,dim=dim(xx),dimnames=dimnames(xx))

# Warn if maximum absolute value of change in defoliation is greater than the
```

```
# maxchange argument; this will lead to all grid points with change >= maxchange
# being plotted as the same color (at the end of the colour ramp).
if (max(abs(xx),na.rm=TRUE) > maxchange) warning("Predicted changes outside range of
    colour ramp; increase 'maxchange' for all plots")


# Borders and labelling
maplines <- map("worldHires",regions=borders,exact=TRUE,plot=FALSE)
if (label) {
    tx <- (max(maplines$x,na.rm=TRUE)+min(maplines$x,na.rm=TRUE))/2
    ty <- min(maplines$y,na.rm=TRUE)-1
}


# Legend positioning
if (legend) {
    xrange <- range(maplines$x,na.rm=TRUE); yrange <- range(maplines$y,na.rm=TRUE)
    topright <- c(floor(xrange[2]),floor(yrange[2]))
    legendwidth <- (xrange[2]-xrange[1])/15
    legendheight <- (yrange[2]-yrange[1])/3
    ny <- 30; nx <- 3
    legendx <- seq(topright[1]-legendwidth,topright[1],length.out=nx)
    legendy <- seq(topright[2]-legendheight,topright[2],length.out=ny)
    legendcc <- inla.generate.colors(seq(-maxchange,maxchange,length.out=ny),
        color.axis=c(-maxchange,maxchange),color.palette=cp)$colors
    legendcc <- matrix(rep(legendcc,nx),nrow=nx,ncol=ny,byrow=TRUE)
}


# Draw fogplot
open3d(windowRect=windowRect)
view3d(0,0,fov=0,zoom=zoom)
surface3d(gx,gy,z=xx*0,color=cc,alpha=aa,specular="black")
lines3d(maplines$x,maplines$y,maplines$x*0,lwd=line.width)
if (show.plots) {
    plotlocs <- unique(dat[,c("x","y")])
    plot3d(plotlocs$x,plotlocs$y,z=0,add=TRUE)
}
labtext <- paste("Change ",as.character(yearsdiff[1]),"-",as.character(yearsdiff[2]),sep="")
if (label) {
    text3d(tx,ty,z=0,text=labtext,adj=c(0.5,0),cex=2)
}
if (legend) {
    surface3d(legendx,legendy,z=matrix(0,nx,ny),color=legendcc,specular="black")
    text3d(legendx[1],legendy[1],z=0,text=paste("-",as.character(maxchange),sep=""),adj=1.1,
        cex=0.8)
    text3d(legendx[1],mean(legendy[c(1,ny)]),z=0,text="0",adj=1.1,cex=0.8)
    text3d(legendx[1],legendy[ny],z=0,text=as.character(maxchange),adj=1.1,cex=0.8)
}
snapshot3d(paste(path,tag,gsub(" ","",tolower(labtext)),".png",sep=""))
}
```

Appendix B

R code for MCMC algorithm

See Chapter 4.

```
library(mvtnorm)

## Simulate data (linear time trend, no other covariates)
rho.sim <- 0.2          # temporal autocorrelation coefficient
beta.sim <- 0.3         # linear time trend
theta.sim <- c(-1,5,10) # thresholds
mm <- 2                 # U_i0 ~ Logistic(mm,ss) iid
ss <- 3
n <- 100                # sample size in each year
T <- 20                 # number of years
set.seed(13)

# simulate values of U_0 (latent variable at time 0)
U.0.sim <- rlogis(n,mm,ss)
# simulate residuals for t>0
eps <- matrix(rlogis(n*T,0,1),nrow=n,ncol=T)
# calculate latent variables U_t
U.t.sim <- matrix(0,nrow=n,ncol=T)
U.t.sim[,1] <- rho.sim*U.0.sim + beta.sim*1 + eps[,1]
for (t in 2:T) U.t.sim[,t] <- rho.sim*U.t.sim[,t-1] + beta.sim*t + eps[,t]
# calculate observed variables y_t
y <- matrix(findInterval(U.t.sim,vec=theta.sim),n,T)
# stack y into one column (obs in each year grouped together)
y <- c(y)
time <- rep(1:T,each=n)
simdat <- data.frame(cbind(y,time))

######################################################################
## Fit cumulative logistic threshold model with temporal autocorrelation

# Parameters for priors
# beta ~ N(c,h)  (MVN if more than one parameter)
c <- 0.0001; h <- 1000
# theta ~ MVN(q,gI)
q <- c(0.0001,0.0001,0.0001); g <- 1000
# rho ~ Uniform(-1,1)
# U.0 ~ Logistic(m,s)
m <- 0.0001; s <- 1000
```

```
# Initial parameter values (these will be overwritten by the MH loop)
beta0 <- 0
rho0 <- 0.1
theta0 <- sort(rmvnorm(1,mean=q,sigma=diag(rep(g,3))))
U.00 <- rlogis(n,m,s)
# Set U.t0 to match up with values of y, given theta0
U.t0 <- rep(0,length(y))
U.t0[y==0] <- theta0[1] - 0.1*(theta0[1]-min(U.00))
U.t0[y==1] <- mean(theta0[1:2])
U.t0[y==2] <- mean(theta0[2:3])
U.t0[y==3] <- theta0[3] + 0.1*(max(U.00)-theta0[3])
U.t0 <- matrix(U.t0,n,T)


# Storage
n.rep <- 1000000
thin <- 10
# If n.rep is large, store only 1/thin of the iterations (set to 1
# if n.rep is small enough for everything to be stored)
n.store <- n.rep/thin
n.accept <- list(beta=0,theta=0,rho=0,U.0=rep(0,n),U.t=matrix(0,n,T))
alpha <- list(beta=rep(0,n.store),theta=rep(0,n.store),rho=rep(0,n.store),
    U.0=matrix(0,n,n.store),U.t=array(0,c(n,T,n.store)))
beta <- rep(0,n.store)
theta <- matrix(0,length(theta0),n.store)
rho <- rep(0,n.store)
U.0 <- matrix(0,n,n.store)
U.t <- array(0,c(n,T,n.store))


# SDs for proposal distributions
sd.beta <- 0.01
sd.theta <- 0.01
sd.rho <- 0.005
sd.U.0 <- 80
sd.U.t <- 2


# Prior probabilities
lprior.beta0 <- dnorm(beta0,mean=c,sd=sqrt(h),log=TRUE)
lprior.theta0 <- dmvnorm(theta0,mean=q,sigma=diag(rep(g,3)),log=TRUE)
lprior.rho0 <- dunif(rho0,min=-1,max=1,log=TRUE)
lprior.U.00 <- dlogis(U.00,location=m,scale=s,log=TRUE)
lprior.U.t0 <- matrix(0,n,T)
lprior.U.t0[,1] <- dlogis(U.t0[,1],location=rho0*U.00+beta0,scale=1,log=TRUE)
for (t in 2:T) {
    lprior.U.t0[,t] <- dlogis(U.t0[,t],location=rho0*U.t0[,t-1]+
        beta0*t,scale=1,log=TRUE)
}
lprior.U.01 <- lprior.U.00; lprior.U.t1 <- lprior.U.t0


# Metropolis-Hastings loop
for (j in 1:n.rep) {
    # Store only every thin'th iteration
store <- (floor(j/thin)==j/thin)
j.store <- j/thin


# Beta proposal
```

```
beta1 <- rnorm(1,mean=beta0,sd=sd.beta)
# New priors if beta updated (beta, U.t)
lprior.beta1 <- dnorm(beta1,mean=c,sd=sqrt(h),log=TRUE)
lprior.U.t1[,1] <- dlogis(U.t0[,1],location=rho0*U.00+beta1,
    scale=1,log=TRUE)
for (t in 2:T) {
    lprior.U.t1[,t] <- dlogis(U.t0[,t],location=rho0*U.t0[,t-1]+beta1*t,
        scale=1,log=TRUE)
}
# Accept/reject
alpha.test <- (lprior.beta1+sum(lprior.U.t1))-
    (lprior.beta0+sum(lprior.U.t0))
if (alpha.test>log(runif(1))) {
lprior.beta0 <- lprior.beta1; lprior.U.t0 <- lprior.U.t1
n.accept$beta <- n.accept$beta + 1
beta0 <- beta1
}
if (store) {
    beta[j.store] <- beta0
    alpha$beta[j.store] <- alpha.test
}


# Theta proposal
theta1 <- rmvnorm(1,mean=theta0,sigma=diag(rep(sd.theta^2,3)))
# New likelihood if theta updated (NB ordering restriction
# is checked here rather than when calculating new prior probabilities)
if (theta1[1]<theta1[2] && theta1[2]<theta1[3]) {
y.check <- findInterval(U.t0,vec=theta1)
like.is.1 <- identical(y,y.check)
} else
like.is.1 <- FALSE
# New priors if theta updated (theta only)
lprior.theta1 <- dmvnorm(theta1,mean=q,sigma=diag(rep(g,3)),log=TRUE)
# Accept/reject
alpha.test <- lprior.theta1-lprior.theta0
if (alpha.test>log(runif(1)) && like.is.1) {
lprior.theta0 <- lprior.theta1
n.accept$theta <- n.accept$theta + 1
theta0 <- theta1
}
if (store) {
    theta[,j.store] <- theta0
    alpha$theta[j.store] <- alpha.test+like.is.1
}


# Rho proposal
rho1 <- rnorm(1,mean=rho0,sd=sd.rho)
# New priors if rho updated (rho, U.t)
lprior.rho1 <- dunif(rho1,min=-1,max=1,log=TRUE)
lprior.U.t1[,1] <- dlogis(U.t0[,1],location=rho1*U.00+beta0,
    scale=1,log=TRUE)
for (t in 2:T) {
    lprior.U.t1[,t] <- dlogis(U.t0[,t],location=rho1*U.t0[,t-1]+beta0*t,
        scale=1,log=TRUE)
}
# Accept/reject
```

```
alpha.test <- (lprior.rho1+sum(lprior.U.t1))-
    (lprior.rho0+sum(lprior.U.t0))
if (alpha.test>log(runif(1))) {
lprior.rho0 <- lprior.rho1; lprior.U.t0 <- lprior.U.t1
n.accept$rho <- n.accept$rho + 1
rho0 <- rho1
}
if (store) {
rho[j.store] <- rho0
alpha$rho[j.store] <- alpha.test
}


# Updating one obs (tree) at a time
for (i in 1:n) {
    # U.i0 proposal
    U.i01 <- rnorm(1,mean=U.00[i],sd=sd.U.0)
    # New priors if U.i0 updated (U.i0, U.i1)
    lprior.U.01[i] <- dlogis(U.i01,location=m,scale=s,log=TRUE)
    lprior.U.t1[i,1] <- dlogis(U.t0[i,1],location=rho0*U.i01+beta0,
        scale=1,log=TRUE)
    # Accept/reject
    alpha.test <- (lprior.U.01[i]+lprior.U.t1[i,1])-
        (lprior.U.00[i]+lprior.U.t0[i,1])
    if (alpha.test>log(runif(1))) {
    lprior.U.00[i] <- lprior.U.01[i]
    lprior.U.t0[i,1] <- lprior.U.t1[i,1]
n.accept$U.0[i] <- n.accept$U.0[i] + 1
U.00[i] <- U.i01
}
if (store) {
        U.0[i,j.store] <- U.00[i]
        alpha$U.0[i,j.store] <- alpha.test
}
}


for (t in 1:T) {
        for (i in 1:n) {
        # U.it proposal
U.it1 <- rnorm(1,mean=U.t0[i,t],sd=sd.U.t)
# New likelihood if U.it updated (current values of i,t only)
y.check <- findInterval(U.it1,vec=theta0)
like.is.1 <- identical(y[n*(t-1)+i],y.check)
# New priors if U.it updated (U.it, U.i(t+1))
if (t>1) {
        lprior.U.t1[i,t] <- dlogis(U.it1,location=rho0*U.t0[i,t-1]
            +beta0*t,scale=1,log=TRUE)
} else {
lprior.U.t1[i,t] <- dlogis(U.it1,location=rho0*U.00[i]
    +beta0*t,scale=1,log=TRUE)
}
if (t<T) {
        lprior.U.t1[i,t+1] <- dlogis(U.t0[i,t+1],location=rho0*U.it1
            +beta0*(t+1),scale=1,log=TRUE)
}
# Accept/reject
```

```
if (t<T) {
    alpha.test <- (lprior.U.t1[i,t]+lprior.U.t1[i,t+1])-
        (lprior.U.t0[i,t]+lprior.U.t0[i,t+1])
} else {
alpha.test <- lprior.U.t1[i,t]-lprior.U.t0[i,t]
}
if (alpha.test>log(runif(1)) && like.is.1) {
    lprior.U.t0[i,t] <- lprior.U.t1[i,t]
if (t<T) {
    lprior.U.t0[i,t+1] <- lprior.U.t1[i,t+1]
}
n.accept$U.t[i,t] <- n.accept$U.t[i,t] + 1
U.t0[i,t] <- U.it1
}
if (store) {
    U.t[i,t,j.store] <- U.t0[i,t]
alpha$U.t[i,t,j.store] <- alpha.test+like.is.1
}
    }
}
}

# Save results
results <- list(beta=beta,theta=theta,rho=rho,U.0=U.0,U.t=U.t)
```

Bibliography

Daniel Adler, Duncan Murdoch, and others. *rgl: 3D visualization device system (OpenGL)*, 2014. URL `http://CRAN.R-project.org/package=rgl`. R package version 0.94.1143.

J.H. Albert and S. Chib. Sequential ordinal modeling with applications to survival data. *Biometrics*, 57:829–836, 2001.

N.H. Augustin, M. Musio, K. von Wilpert, E. Kublin, S.N. Wood, and M. Schumacher. Modelling spatiotemporal forest health monitoring data. *Journal of the American Statistical Association*, 104(487):899–911, 2009.

C. Belitz, A. Brezger, T. Kneib, S. Lang, and N. Umlauf. *BayesX: Software for Bayesian Inference in Structured Additive Regression Models. Version 2.1.*, 2012.

N.E. Breslow and D.G. Clayton. Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*, 88:9–25, 1993.

J. Eichhorn, P. Roskams, M. Ferretti, V. Mues, A. Szepesi, and D. Durrant. *Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests*, chapter Visual Assessment of Crown Condition and Damaging Agents. UNECE ICP Forests Programme Co-ordinating Centre, Hamburg, 2010.

L. Fahrmeir and S. Lang. Bayesian inference for generalized additive mixed models based on Markov random field priors. *Journal of the Royal Statistical Society C (Applied Statistics)*, 50:201–220, 2001.

F.E. Harrell. *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*, chapter Ordinal Logistic Regression, pages 335–336. Springer-Verlag, New York, 2001.

M.D. Higgs and J.M. Ver Hoef. Discretized and aggregated: Modeling dive depth of harbor seals from ordered categorical data with temporal autocorrelation. *Biometrics*, 68:965–974, 2012.

X. Lin and D. Zhang. Inference in generalized additive mixed models by using smoothing splines. *Journal of the Royal Statistical Society Series B*, 61:381–400, 1999.

M. Lorenz, G. Becher, V. Mues, R. Fischer, R. Becker, V. Calatayud, N. Dise, G.H.M. Krause, M. Sanz, and E. Ulrich. Forest condition in europe: 2005 technical report of icp forests and futmon. Technical report, ICP Forests, Hamburg, 2005.

A. Michel and W. Seidling. Forest condition in europe: 2014 technical report of icp forests. Technical report, BFW Austrian Research Centre for Forests, Vienna, 2014.

M. Mund, W.L. Kutsch, C. Wirth, T. Kahl, A. Knohl, M.V. Skomarkova, and E.-D. Schulze. The influence of climate and fructification on the inter-annual variability of stem growth and net primary productivity in an old-growth, mixed beech forest. *Tree Physiology*, 30(6): 689–704, 2010.

R. Ospina and S.L.P. Ferrari. Inflated beta distributions. *Statistical Papers*, 51:111–126, 2010.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2014. URL `http://www.R-project.org/`.

R.A. Rigby and D.M. Stasinopoulos. Generalized additive models for location, scale and shape (with discussion). *Applied Statistics*, 54:507–554, 2005.

B.W. Silverman. Some aspects of the spline smoothing approach to nonparametric regression curve fitting. *Journal of the Royal Statistical Society, Series B*, 47:1–52, 1985.

D. Stegmueller. Modelling dynamic preferences: A bayesian robust dynamic latent ordered probit model. *Political Analysis*, pages 1–20, 2013.

Nikolaus Umlauf, Thomas Kneib, Stefan Lang, and Achim Zeileis. *R2BayesX: Estimate Structured Additive Regression Models with BayesX*, 2013. URL `http://CRAN.R-project.org/package=R2BayesX`. R package version 0.3-1.

W. N. Venables and B. D. Ripley. *Modern Applied Statistics with S*. Springer, New York, fourth edition, 2002. URL `http://www.stats.ox.ac.uk/pub/MASS4`. ISBN 0-387-95457-0.

X. Wang and K.M. Kockelman. Bayesian inference for ordered response data with a dynamic spatial-ordered probit model. *Journal of Regional Science*, 49:877–913, 2009.

S.N. Wood. Low-rank scale-invariant tensor product smooths for generalized additive mixed models. *Biometrics*, 62(4):1025–1036, 2006a.

S.N. Wood. *Generalized Additive Models. An Introduction with R*. Chapman & Hall/CRC, Boca Raton, 2006b.

S.N. Wood. Fast stable restricted maximum likelihood and marginal likelihood estimation of semi-parametric generalized linear models. *Journal of the Royal Statistical Society Series B*, 73:3–36, 2011.

S.N. Wood, M.V. Bravington, and S.L. Hedley. Soap film smoothing. *Journal of the Royal Statistical Society Series B*, 70(5):931–955, 2008.