W O R K R E P O R T

Institute for World Forestry

Temporal development of crown condition of *Picea abies*

two-step approach using statistical and geostatistical methods

by

VOLKER MUES and RICHARD FISCHER

Federal Research Centre
for Forestry and Forest Products

Bundesforschungsanstalt für Forst- und Holzwirtschaft Hamburg
(Federal Research Centre for Forestry and Forest Products)
Address: Leuschnerstr. 91, D-21031 Hamburg, Germany
Postal address: P.O. Box: 80 02 09, D-21002 Hamburg, Germany

Phone: +40 / 73962-101
Fax: +40 / 73962-480
E-mail: weltforst@holz.uni-hamburg.de
Internet: http://www.bfafh.de

Institute for World Forestry

Temporal development of crown condition of *Picea abies*

two-step approach using statistical and geostatistical methods

Work report of the Institute for World Forestry 2003 / 9

This work report was published in December 2001.

Hamburg, September 2003

Contents

Summary

The aim of the presented Internal Report 2001 is to analyse the temporal development of defoliation of *Picea abies* in Europe by means of the transnational Level I data set. Additional external deposition and meteorological data are used to describe the influence of single stress factors.

Evaluations are conducted using multivariate linear models in a two step analysis. Step 1 evaluates variables that were available in annual or higher time resolution in relation to annual mean plot defoliation for the years 1994 to 2000. The regression coefficient for the interaction term 'year*plotid' is interpreted as plot-wise time trend. It is presented in maps and after a geostatistical analysis detecting spatial autocorrelation interpolated by the geostatistical method “kriging”. Rooted mean squared errors (RMSE) are calculated for all plots. They give a sensitive measure for the discontinuity of defoliation development and serve as a conservative estimate for the accuracy of the linear models offering a tool for quality control when interpreting the time trends.

Results of step 1 show significant influences of insects, age and country as well as their interaction terms on annual mean defoliation. Time trends were only significant for plots and countries, which shows that for *Picea abies* there is no significant mean European wide trend. Fungi gave implausible results, whereas summer precipitation showed insignificant but mostly plausible effects. Larger regions with deteriorating defoliation are observed in southern Sweden and Finland, in Estonia as well as in alpine regions of Switzerland, Austria, and Slovenia. Improving crown condition is observed in northern Scandinavia, Lithuania, southern Poland, and in Slovakia.

The analysis of the RMSE reveals single plots with extreme defoliation values, these could however only be partly explained by information available from the Level I data base. The interpretation of time trends at these plots has to be conducted with care.

Step 2 of the analysis is the explanation of plot-wise time trends and RMSE values by predictors of low temporal resolution. Preliminary results show that trends of defoliation are correlated with the difference in SO_x deposition between 1998 and 1997, with water availability and base saturation. The promising results for sulphur deposition underline the necessity to include annual deposition values into step 1 of the analysis in the future. Site characteristics may show closer correlation when combined with meteorological data of higher time resolution.

After discussion of the results in an editorial group the evaluations are foreseen to be extended and applied to other main tree species for a presentation in the Forest Condition Report of UN/ECE and EU in 2002.

1 Introduction

Under the UNECE Convention on Long-range Transboundary Air Pollution the International Cooperative Programme on the Assessment and Monitoring of Air Pollution Effects on Forests (ICP Forests) is operated under the Lead of Germany with a participation of 39 countries. The Programme Co-ordinating Centre (PCC) of the ICP Forests is hosted by the Federal Research Centre for Forestry and Forest Products in Germany. The crown condition survey is conducted on a large-scale transnational gridnet (Level I) of the ICP Forests, which was established in 1986. It is conducted in close co-operation with the European Union's "Scheme on the Protection of Forests against Atmospheric Pollution". The survey aims to assess the spatial and temporal variation of forest condition in relation to natural and anthropogenic factors, particularly air pollution.

The spatial distribution of crown condition can be expressed by the medium-term mean defoliation. This derived variable was introduced already in the Technical Report 2001 (Lorenz et al. 2001), where the possibility of a preliminary adjustment for methodologically caused variation was analysed. As a result of this study the preliminarily adjusted defoliation (PAD) for the six main tree species in Europe was calculated. It expresses the deviation of defoliation from the respective age and country specific mean defoliation. It is a measure for the level of defoliation at a given point or region, preliminarily reduced by methodologically caused variation.

The aim of the analysis in the present report is to describe the temporal development of defoliation and its correlation with environmental factors. The temporal development of crown condition is described in a first step using predictor variables varying over time and those variables (age, country, and their interaction), which are useful to clean defoliation data from methodological differences (Lorenz et al., 2001). A serious drawback in this context is data availability. Only for a few potentially influencing factors time varying predictor variables were available. Many factors could only be described by time constant predictor variables in a second step. A two-step analysis was selected to use both types of predictor variables, time varying as well as time constant ones. Statistical and geostatistical methods are used to detect regions where a significant temporal development was observed and to indicate those variables and factors, which are expected to be responsible for this with high probability.

The presented work report is the unmodified version of the Internal Report 2001.

Richard Fischer is deputy Head of the Programme Co-ordinating Centre of ICP Forests which is hosted by the Institute for World Forestry. Dr. Volker Mues is scientist at the same institute.

2 Data and Methods

2.1 Data

In the 2001 Technical Forest Condition Report (Lorenz et al. 2001) spatial variation of crown condition was described as the so-called preliminarily adjusted defoliation (PAD). This parameter accounts for methodological differences between the countries. Results were presented for the six main tree species *Fagus sylvatica*, *Picea abies*, *Pinus pinaster*, *Pinus sylvestris*, *Quercus ilex*, and *Quercus robur et petraea*.

In general, for all main tree species spatio-temporal analyses of defoliation are planned in the Technical Report 2002. In this respect first evaluations focussing on *Picea abies* are presented in the present Internal Report 2001. The analyses focus on *Picea abies* because of mainly two reasons: (i) this tree species is observed at a very high number of plots so that even models with a higher number of degrees of freedom can be calculated and (ii) *Picea abies* shows no clear deviation from the linear correlation with age (Lorenz et al., 2001, pp. 56 f.). Furthermore the spatial distribution of *Picea abies* reaches from the north of Norway to Italy and from Spain to the east of Europe. The data from France and Italy were not used because of methodological changes from 1996 to 1997 in both countries, which would influence analysis of temporal development.

The study is based on data from all plots, for which continuous observations from 1994 to 2000 were available on average for more than 2 trees of *Picea abies*. Target variable of the first step of the analysis was annual mean defoliation with 7 observations for the years 1994 to 2000 for each survey plot.

2.1.1 Time-varying predictor variables

In a first step of the analysis (s. Figure 1, page 7) mainly variables of high temporal resolution with annual or monthly values for the whole observation period were used as predictors. These "time-varying" variables are:

- Precipitation (monthly mean precipitation; Global Precipitation Climatology Centre, GPCC; spatial resolution 1°). Summer precipitation was calculated as sum of precipitation from April to September
- mfungi (plot-wise mean occurrence of fungi, values range from 0 to 1 indicating 0% to 100% of plot trees are influenced by fungi; derived from the Level I database, T3)

- minsect (plot-wise mean occurrence of insects, values range from 0 to 1 indicating 0% to 100% of plot trees are influenced by insects; derived from the Level I database, T2)

mfungi and minsect are calculated from assessed values of the Level-I survey. They express the relative frequency of trees affected by insects or fungi, respectively. The precipitation values were calculated for each plot by a bilinear interpolation from a gridnet of the Global Precipitation Climatology Centre (GPCC) of low spatial resolution.

Additionally, a variable expressing the flow of time was introduced as a transformation of the year of observation. As the variation of the original variable (range: 2000-1994+1=7, standard deviation: 2) is very low compared with the mean (1997) its explaining power is also rather low. This difficulty can be solved by deriving the variable year1994 as

- year1994 = year – 1994 . (1)

The regression coefficient for year1994 can be interpreted in the same way as that of year would be. It is the mean difference of defoliation from one year to the previous year.

Time trend of defoliation can be modelled by several factors[1)]:

- year1994 (time trend of defoliation in Europe)
- year1994*country (country-wise mean deviation from European time trend of defoliation; methodological and/or real time trends of defoliation per country)
- year1994*plotid[2)] (plot specific time trend, which can not be explained by other compartments of time trend)

Table 1: descriptive statistics for time varying variables

		1994	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**
minsect	mean	0.0137	0.0093	0.0079	0.0092	0.0091	0.0079	0.0105
	std dev	0.0874	0.0672	0.0526	0.0625	0.0617	0.0466	0.0563
	min	0	0	0	0	0	0	0
	max	1.00	0.92	0.80	1.00	0.98	0.94	0.77
mfungi	mean	0.0264	0.0158	0.0288	0.0323	0.0272	0.0273	0.0265
	std dev	0.1210	0.0837	0.1316	0.1362	0.1148	0.1145	0.1092
	min	0	0	0	0	0	0	0
	max	1.00	1.00	1.00	1.00	1.00	1.00	1.00
precipitation	mean	415	447	413	417	469	415	408
summer	std dev	116	125	144	121	94	149	110
	min	187	210	202	176	234	180	171
[mm]	max	799	809	771	765	793	932	823
precipitation	mean	759	787	693	727	815	782	812
year	std dev	188	213	200	187	169	240	236
	min	380	481	394	372	476	452	279
[mm]	max	2086	2071	1412	1942	2059	2214	2065

The descriptive statistics of the time-varying variables (s. Table 1) show limitations of the indices minsect and mfungi. Their distributions are both right skewed in all years of the observed period (low mean relative to centre of range, mean–std.dev.

1) Additionally the interaction year1994*age was tested but was not significant in any model/combination of predictor variables

2) Plotid is used as categorical (class) variable in this study

negative!). For most years 90% of the plots had the value 0%. In 1995 and 1996 even the 95% quantile is 0%. The precipitation variables indicate interesting differences between the years: Whereas 1998 was a year with high precipitation, e.g. for 1996 low precipitation with a high standard deviation was calculated for the *Picea abies* plots.

2.1.2 Time-constant predictor variables

Besides the time-varying variables there are time-constant variables with only one or two available values in the evaluation period. These are:

- e.g. water availability and soil type from Level I crown condition assessment
- e.g. pH, base saturation, and cation exchange capacity (CEC) from Level I soil survey
- meteorological data: long term (1961-1990) mean monthly air temperature and mean monthly sum of precipitation
- Depositions are available from the EMEP Eulerian model for the years 1997 and 1998 (spatial resolution 50 x 50 km); further indices can be calculated (e.g. differences of depositions)

These variables can be used directly or indirectly by developing derived explanatory variables for a statistical analysis of the spatio-temporal development of defoliation. An example for transformation, in this case a grouping of a classification variable, is shown for soil type (grouping s. Table 2). 57 FAO soil types occurring at the evaluated *Picea abies* plots were grouped into 9 groups. The absolute frequencies of these groups (Table 2) are sufficiently high to calculate meaningful statistics e.g. of the respective mean defoliation.

Table 2: frequency distribution of soilgroup

soilgroup	frequency	percent
1 Podzols	285	29.66
2 Cambisols	268	27.89
3 Leptosols	73	7.60
4 Arenosols	70	7.28
5 Regosols	52	5.41
6 Luvisols	60	6.24
7 Histosols	44	4.58
8 Gleysols	39	4.06
9 Other	70	7.28
missing	86	

In order to estimate the risk of soil water deficiencies the 57 FAO soil types were classified into two groups. Soil draught group 1 comprises Regosols, Leptosols, Arenosols (without gleyic Arenosols), Vertisols, vertic Luvisols. They are regarded as soil types with low water holding capacity and thus an elevated risk of water stress in periods with low precipitation. Group 2 comprises all the other soil types, where increased soil dependent drought risk (sdr) is not expected.

Whereas the time varying parts of the models of step1 are to be tested with respect to their contribution to variance explanation there are the time-constant variables age, country, their interaction, and plotid. They are used to explain methodological differences of the age-effect (Lorenz et al., 2001) or to enable the quantification of that part of variance of annual mean defoliation, which can be explained by the interpretation of plots as subplots of the study area (compare 2.2). The explanation of defoliation by plotid combines the explanation by all time constant variables.

2.2 Analyses

In a preliminary investigation auto-correlative structures for defoliation were detected only for very few plots (PROC AUTOREG, PROC ARIMA, SAS 8.1 online documentation). This is probably due to the shortness of the studied time series and high temporal variance of defoliation values. Thus, auto-correlative effects were not modelled directly.

Instead, a split-plot analysis (Diggle et al., 1994; Hendriks et al., 2000) was calculated. The split plot analysis takes into account that the (7) observations of a single plot are expected to react more similar to an incoming factor than those of different plots. The test, if a predictor variable is contributing significantly to a model's explanation of the target variable's variance. Split-plot test for significance are more conservative compared to normal ANOVA tests which are conducted under the assumption of statistically independent observations.

Following the character of the predictors, which can be divided into time-varying and time-constant variables a two step analysis was conducted (Figure 1). In the first step the time-varying predictor variables (year1994, precipitation, minsects, mfungi) and time-constant variables (age, country, their interaction, and plotid), which describe mainly methodological differences, were used to explain annual mean plot defoliation. The resulting plot-wise regression coefficient of the interaction term year1994*plotid can be interpreted as the time trend for each plot. It quantifies the plot specific mean change in defoliation from one year to the following. As an important result of the models, its spatial variation is presented in maps. In Figure 2 it is graphically presented as the slope of the regression line. In models, which additionally include year1994 and the interaction term year1994*country the plot-wise regression coefficient (for year1994*plotid) does not describe those parts of the defoliation trend, which can be explained by European or country-wise mean trends.

The plot-wise calculated RMSE value is the rooted mean of the squared vertical differences between the annual mean defoliation values and the regression line. Extreme deviations from the regression line (years 1995 and 1996 in Figure 2) lead to significant high values of the plot-wise RMSE. It is a sensitive measure for discontinuities of annual defoliation values.

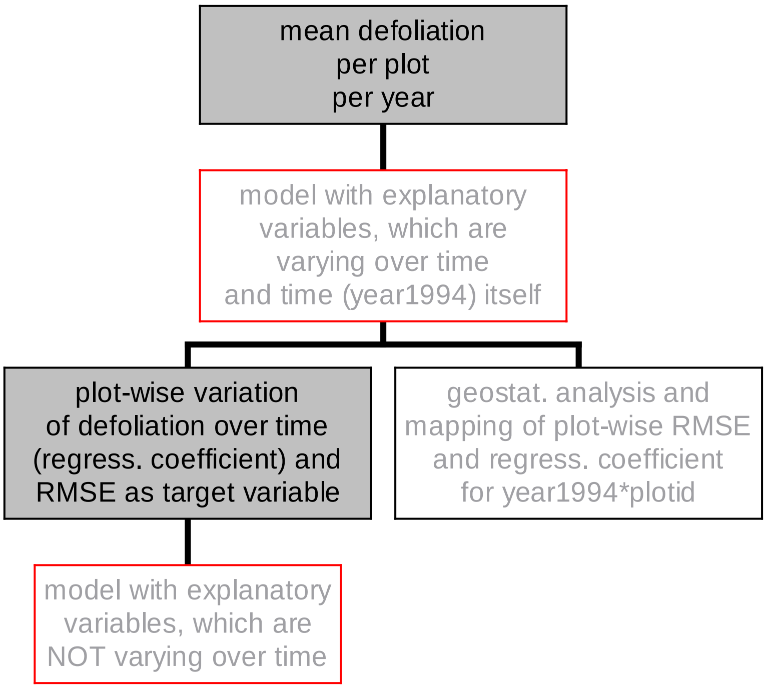


Figure 1: two-step analysis of temporal development of defoliation

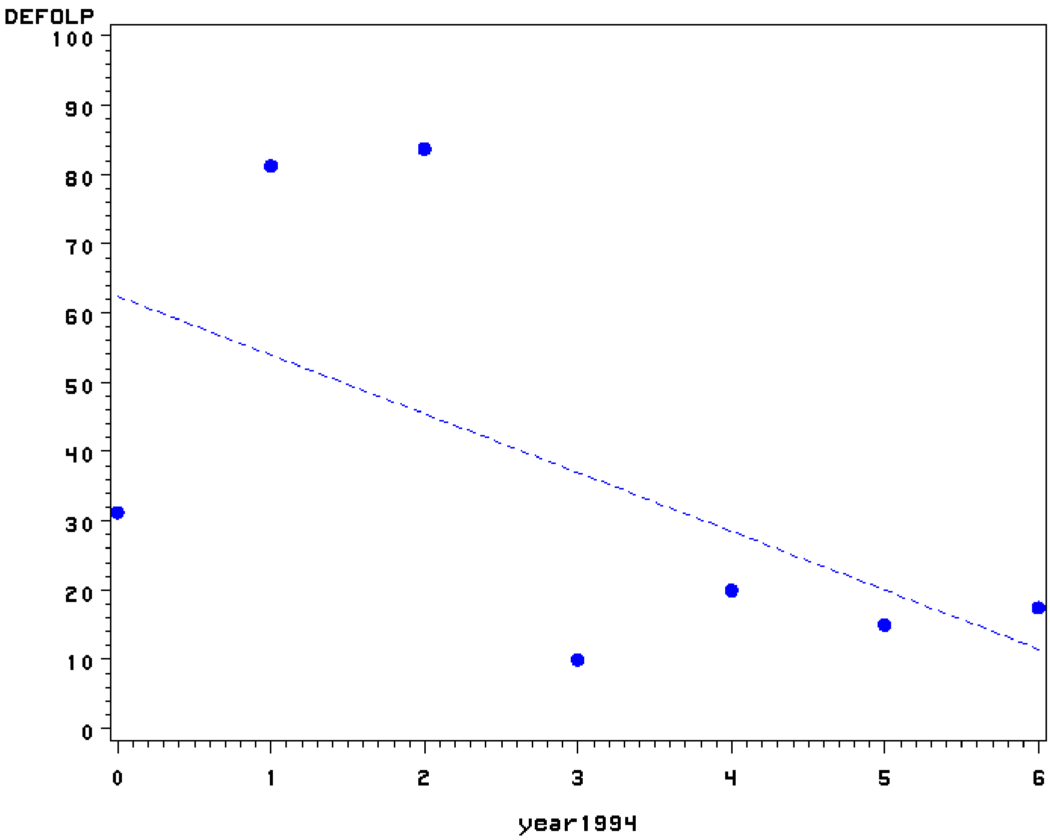


Figure 2: mean defoliation at plot 560 in Lithuania over time (year1994) with linear regression showing heavy discontinuity

Linearity of the relations between defoliation and its predictors was assumed. For the relation between defoliation and age this assumption could be supported for Spruce by Lorenz et al. (2001). Generally, deviations from the linear assumption of the models can be identified by high values of the rooted mean squared error (RMSE). The RMSE is known as usual statistic quantifying accuracy of regression or ANOVA models. Here it is calculated for each plot to enable a geo-referenced representation of model accuracy.

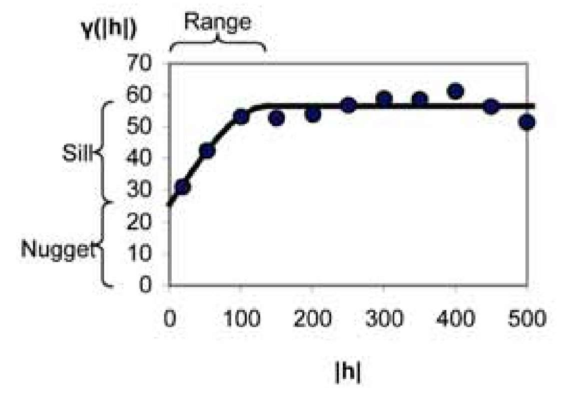
High RMSE values are caused by discontinuous temporal development of defoliation or simply by a high variance of defoliation. Such discontinuities are additionally detected by plot-wise scatter plots of defoliation over time (e.g. Figure 2). The reasons for these discontinuities can be manifold. They may be caused by extreme weather conditions, insect pests, fungi infections, damages by timber harvest or other forest management action. In most cases the Level I data set only allows for a limited interpretation of these discontinuities. Nevertheless, high RSME values indicating a discontinuous development demand for a cautious interpretation of the time trends at the respective plots. In this sense they are an important tool for quality control and are presented in maps.

In the second step the plot-wise regression coefficient of year1994*plotid is used as target variable ("derived variable", Diggle et al., 1994) for a second model in which its correlation with time constant variables (e.g. deposition level, long time mean precipitation, soil and stand parameters) is used.

2.2.1 geostatistics

The fundamental assumption of geostatistics is, that a regionalised variable may consist of a deterministic, a correlative and a random component (Ripley, 1981; see also Schall 1999). The deterministic component, the "drift", can be described e.g. by regression or covariance models (step 1 in Figure 1). The correlative component expresses, that points located close together show smaller differences concerning the regionalised variable than points with a large spatial distance. Because this is a spatial correlation of values of *one* variable, it is called spatial (intra-variable) autocorrelation. This component can be used, to calculate weights for an interpolation by the data themselves instead of those subjectively chosen, like e.g. inverse squared distance weighted interpolations. The random part is that part of the target variable's variance, which is neither determined by predictor variables nor influenced by the target variable itself in terms of (spatial) autocorrelation.

The spatial autocorrelation of the regionalised variable can be described by an empirical semivariogram which expresses the dissimilarity increasing with distance h between (sample) points x_i and $x_i + h$ (Fig. 2.4.5.2-1). Each point in the empirical semivariogram is calculated using equation (5) for the particular distance or class (lag) of distance h. The semi-variance is the mean squared difference between i pairs of values of the regionalised variable from i pairs of points/locations within the spatial distance h.



$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i) - z(x_i + h)]^2 \quad (2)$$

$N(h)$ – number of point pairs with distance h
$z(x_i)$ – regionalised variable at sample point x_i

Figure 3: experimental semivariogram of average dissimilarities over spatial distance |h| [m] and a modelled spherical semivariogram: nugget: 25.5 sill: 31.0 range: 136 km.

Three parameters are usually used to describe the shape of the variogram: nugget, sill and range. The nugget is the semi-variance, which is observed for the distance h = 0. It can be interpreted as the random component of the regionalised variable. Mainly two conditions lead to a nugget value greater than zero:

- The underlying measurement grid net has a too low density, so that the spatial structure/autocorrelation could not be detected completely.
- The underlying spatial structures are covered by inaccuracies of data assessment or other reasons of "noise".

The sill is quantifying the auto-correlative component of the regionalised variable. The range is the distance, in which spatial autocorrelation is observed. The closer a plot is lying to an estimation (target) point x_i, the lower is the particular value of the semivariogram $\gamma(h)$ and the higher is – in general – the (kriging-) weight of this plot for the interpolation (kriging) of the regionalised variable at any estimation point $z^*(x_i)$.

The result of modelling the empirical semivariogram is the so-called theoretical semi-variogram. A popular function of the theoretical semivariogram is the spherical model, defined by the following equation:

$$\gamma(h) = \begin{cases} \text{nugget} & ; 0 = h \\ \text{nugget} + \text{sill} \cdot \left\{ 1{,}5 \left(\frac{h}{\text{range}} \right) - 0{,}5 \left(\frac{h}{\text{range}} \right)^3 \right\} & ; 0 < h < \text{range} \\ \text{nugget} + \text{sill} & ; h \geq \text{range} \end{cases} \quad (3)$$

Spherical semivariograms are used in the present study due to their good interpretability. General introductions to applied geostatistics are given by e.g. Ripley (1981).

Only for those points a value of the regionalised variable was estimated, for which at least 12 Level I plot values are available in a radius of 800 km and for which at least 4 plot values are available within a radius of 100 km. The latter precondition was defined in order to reduce the area of extrapolation beyond the sample area. For the calculation of the kriging values however plots within the 800 km radius were used.

Another application of semivariography is the detection of outliers. Plots with residuals, which do not fit to the spatial pattern lead to very high variogram values and can be identified e.g. by calculating so called variogram clouds. This is of main interest when conducting the geostatistical analysis of the residuals of the first step of analysis (Figure 1).

3 Results

According to the two-step analysis (Figure 1) the results are presented in two parts. Whereas the resulting models of step one describe the time-trend by time-varying predictor variables, the results of step 2 show the possibility to explain regression parameters of step 1 as derived variable by time-constant predictor variables.

3.1 Step 1

The first part of the statistical analysis resulted in a number of multivariate models explaining mean plot defoliation for *Picea abies* as target variable. Their corresponding values for R^2 and RMSE as well as their predictor variables are presented in Table 3. Each model was calculated with plotid as predictor variable (line a) as well as without plotid (line b). Split plot analyses (signifcance tests) could only be conducted for the full models (first line), whereas the R^2 values of the second line give a more realistic picture about the explaining potential of the included predictor variables. A reference model (ref.) is included into the table in order to show how much variance of defoliation can be explained without using any time varying predictor variable.

Table 3: results of analysis of annual mean defoliation (step 1 in Figure 1)

model		R^2	RMSE	mfungi	minsect	prec. summer	y	inter. y*c	a	c	inter. a*c	inter. y*p	p
I	a	90.34	4.88	o	x	x	x	x	x	x	x	x	x
	b	49.44	9.48	o	x	x	x	x	x	x	x		
II	a	90.34	4.88		x	x	x	x	x	x	x	x	x
	b	49.39	9.48		x	o	x	x	x	x	x		
III	a	90.35	4.88		x		x	x	x	x	x	x	x
	b	49.41	9.48		x		x	x	x	x	x		
IV	a	90.35	4.88		x				x	x	x	**x**	x
	b	48.61	9.54		x				x	x	x		
ref.	a	90.27	4.90						x	x	x	x	x
	b	48.03	9.60						x	x	x		

x – included
o – implausible coefficient
x – significant (split-plot analysis)

y – year1994
c – country
a – age

inter. – interaction
p – plotid

In the first model the index for fungi had an implausible negative regression coefficient (the higher the index the less defoliation) and was therefore excluded in the following models. Model II shows a plausible negative regression coefficient for precipitation in summer only as long as the plot-wise parameters year1994*plotid and plotid are included in the model. Model III only includes predictor variables that show plausible regression coefficients and which are mostly significant (split-plot analysis). There is no significance for an European trend of defoliation but there are significantly different trends in some countries (compare Table 4). Because it is still not possible to determine whether these differences between countries are caused by inconsistencies in the assessment or by real changes in defoliation, the further analysis of step 2 (Figure 1) were conducted with the plot-wise time trend, calculated with model IV a. The regression coefficient for minsects is 8.77% for this model.

Table 4: descriptive statistics for plot-wise time trends (regression coefficients for year1994*plotid; model IVa, Table 3)

	n plots	mean	std dev	min	max
			[% defol. / year]		
all plots	1047	0.15	1.62	-9.5	9.4
Austria	66	0.41	1.02	-4.6	3.0
Belgium	7	-1.32	3.88	-9.5	2.1
Denmark	8	-0.61	2.43	-4.4	3.3
Finland	148	0.29	1.40	-6.6	5.3
Germany	180	0.04	1.47	-8.1	4.9
Ireland	3	1.95	2.04	0.5	4.3
Luxembourg	2	0.51	1.99	-0.9	1.9
Spain	1	0.93	.	0.9	0.9
Sweden	162	0.57	1.51	-5.7	9.4
United Kingdom	11	0.35	1.18	-2.8	1.6
Bulgaria	3	1.00	0.82	0.1	1.6
Croatia	1	2.05	.	2.1	2.1
Czech Republic	83	0.13	1.40	-3.2	3.6
Estonia	31	1.37	1.67	-1.7	5.6
Lativia	38	0.47	1.48	-3.3	4.4
Lithuania	24	-0.44	1.99	-7.8	2.5
Norway	130	-0.22	1.71	-5.8	5.4
Poland	26	-1.82	1.77	-6.0	1.1
Romania	33	0.20	1.59	-1.7	4.8
Slovak Republic	44	-1.06	1.47	-5.2	0.9
Slovenia	21	1.31	1.47	-1.3	4.9
Switzerland	25	0.40	1.33	-2.3	3.6

The means of plot-wise time trends show marked differences between the countries. From the countries with more than 10 plots Estonia shows a mean annual worsening of 1.37% defoliation, which equals a mean deterioration of 8.22% defoliation in the evaluation period. Highest improvements were reported for the Slovak Republic and especially Poland.

3.1.1 Spatial distribution of plot-wise trend

The regression coefficient for year1994*plotid represents the plot-wise linear time trend in the years from 1994 to 2000. The spatial distribution underlines comparatively large variations throughout Europe (Figure 5). A few plots in southern Scandinavia, Latvia and Estonia show an annual deterioration between 5% and 10% defoliation, which gives a mean deterioration between 30% and 60% in the 6 years from 1994 to 2000. Clusters of improving plots are located in south and middle Norway as well as in Poland and Slovakia. Border effects seem to appear between Czech Republic and Poland but are not confirmed by a closer view, which shows an increasing improvement of crown condition going from the north of the Czech Republic to the south of Poland.

The spatial interpolation (kriging) of the regression coefficients (plot-wise time trends) levels out the large variations and thus results in a lower maximum and a higher minimum for the 32*32 km grid (Figure 6) compared with the plot values (Figure 5). On the other hand the clusters of plots with similar trends become more evident. Large areas with deteriorating crown condition of *Picea abies* are located around the Baltic sea (southern Finland, south-eastern Sweden, Estonia and Latvia). Also in alpine areas of Switzerland, Austria and Slovenia a deterioration prevails. Intra-country gradients appear in Romania, Czech Republic, Norway and Finland. Predominantly improving trends are found in Slovakia, Poland, and Lithuania.

The empirical and the modelled theoretical semivariogram for plot-wise time trend/regression coefficient is presented in Figure 4. The decrease of spatial auto-correlation is slower between the first and second range at 97.2km and 520km, respectively. This can indicate the occurrence of at least two unexplained factors influencing the temporal development of defoliation.

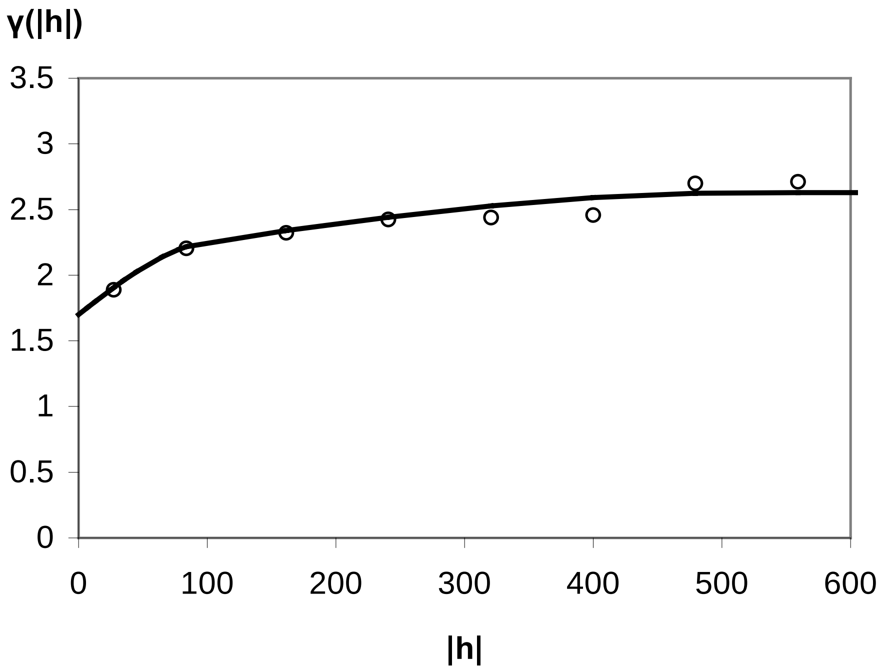


Figure 4: empirical variogram (dots) and modelled nested spherical variogram (line) for plot-wise time trend; |h| = distance in km; nugget: 1.7, sill1: 0.4, range1: 97.2 km, sill2: 0.53, range2: 520 km.

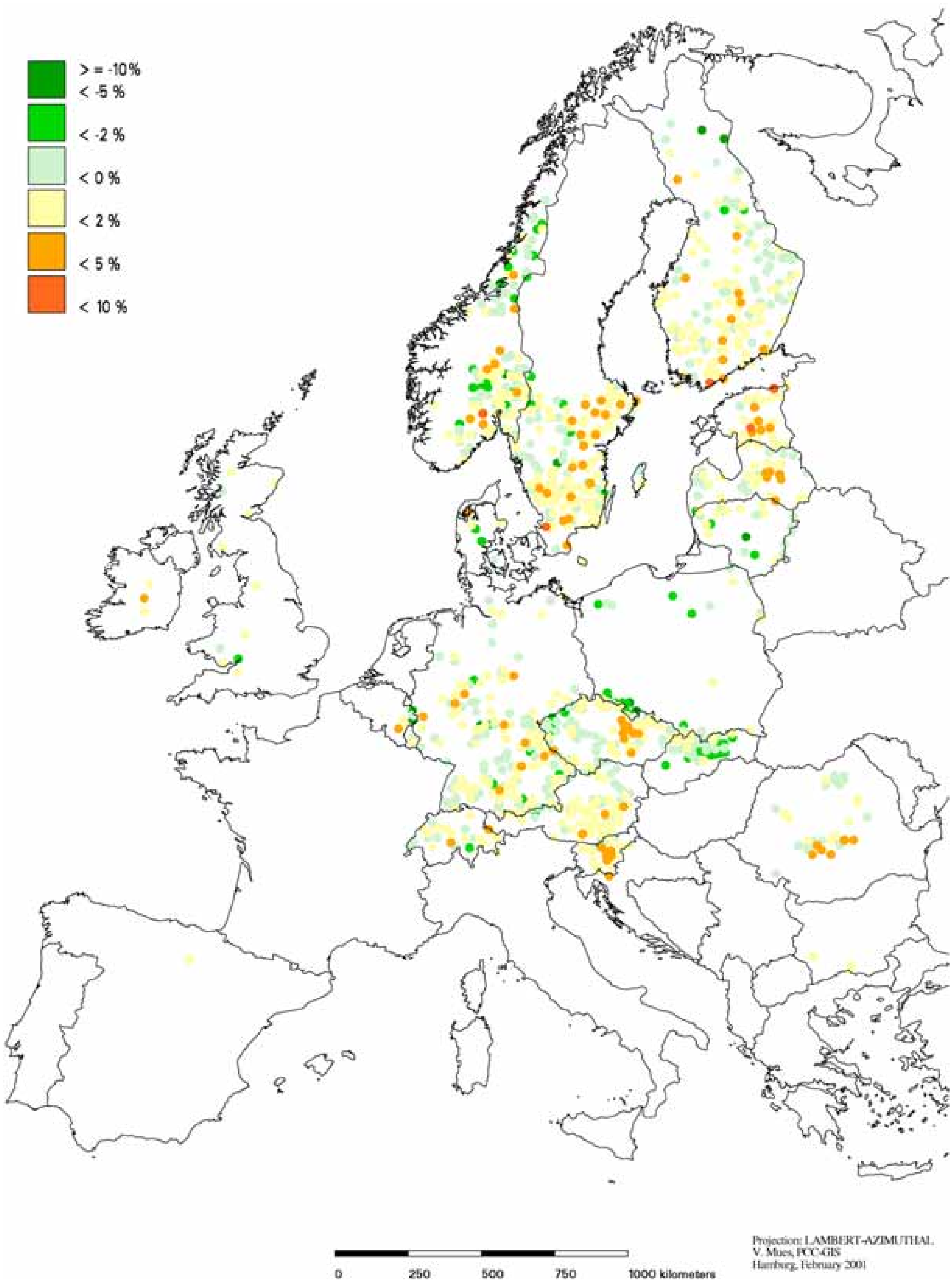


Figure 5: plot-wise time trend of defoliation for the years1994 – 2000

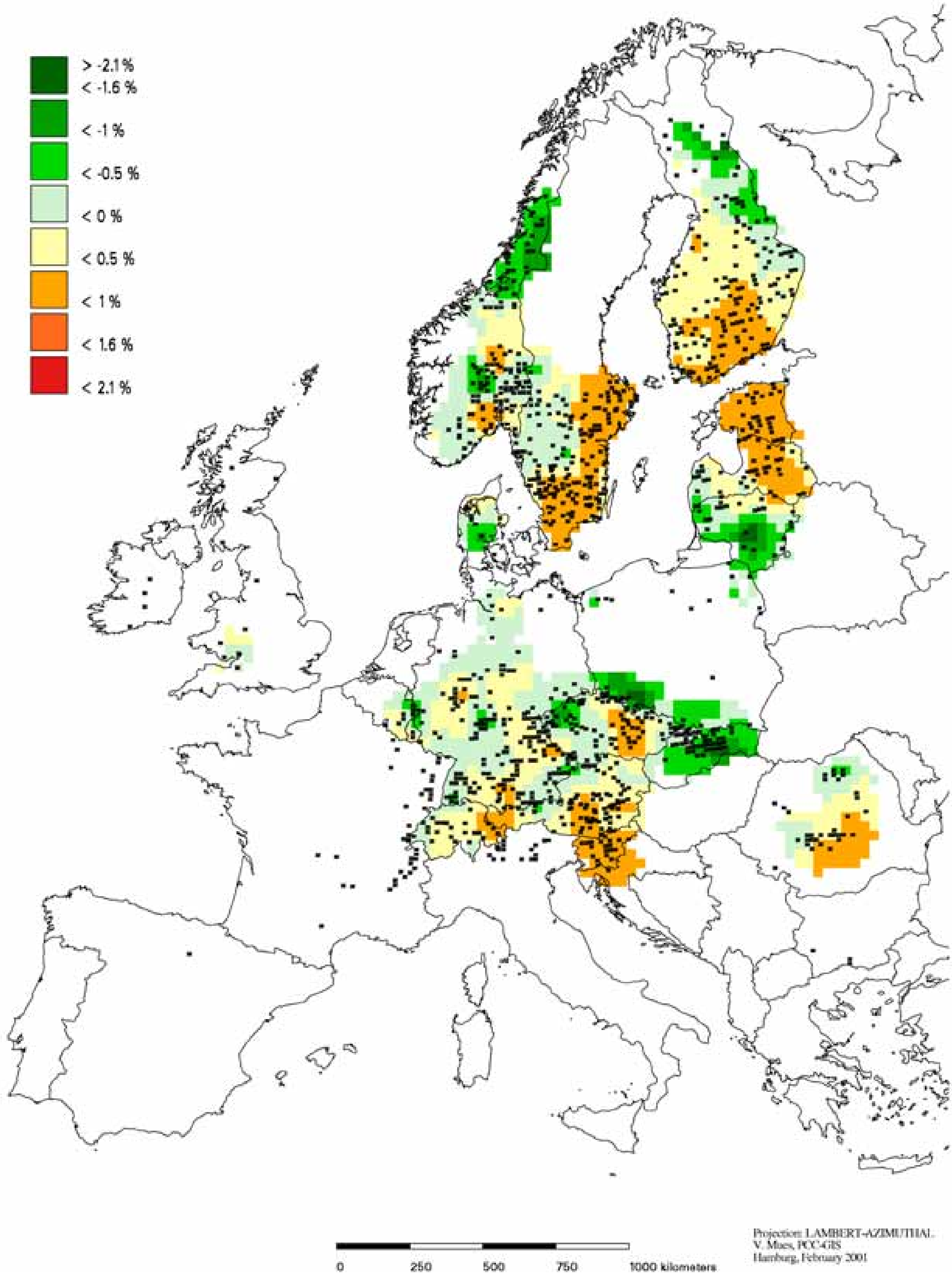


Figure 6: kriged plot-wise time trend of defoliation for the years1994 – 2000

3.1.2 Spatial distribution of plot-wise RMSE

The plot-wise rooted mean squared error is a measure for the accuracy of the linear model explaining the time trends. It gives valuable background information for an interpretation of the modelled time trends and is particularly sensitive to discontinuities. Results show RMSE values of < 10% for the large majority of the plots.

Only a few plots are characterised by high errors. The map of plot-wise values (Figure 9) shows namely two red plots of very high RMSE-values: A scatter plot of the survey-plot in Lithuania (Figure 2) shows very high values for the years 1995 and 1996. In both years the value for minsect is with 0.75 very high (75% of the trees are assessed as influenced by insects). For the other years the respective value always is 0.

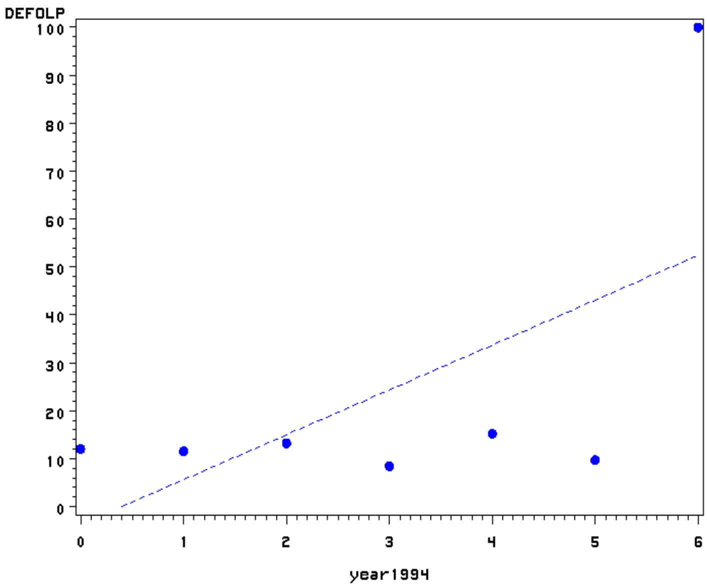


Figure 7: scatter plot of annual mean defoliation for plot 1716 in Sweden

The Swedish plot 1716 also shows heavy discontinuities. The scatter plot of this survey-plot is showing an un-explained discontinuity from 1999 to 2000 (Figure 7). The defoliation of 100% in 2000 can yet not be explained by the data.

The semivariogram analysis shows a hole effect at 300km (Figure 8). In this distance from the two extreme values the density of plots is lower compared to the rest of the evaluated area (low density in Poland, Baltic sea; see Figure 9). Therefore, for these distances less pairs of observations with partici-pation of one of the two extreme values could be used for the calculation of the empirical semivariogram.

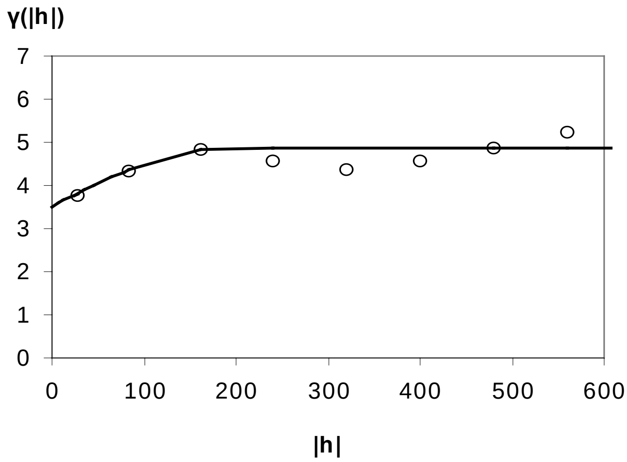


Figure 8: Empirical variogram (dots) and modelled spherical variogram (line) for plot-wise RMSE; |h| = distance in km; nugget: 3.51, sill: 1.35, range: 186.3km.

The map of kriged RMSE values (Figure 10) shows only for the region in Lithuania a very high discontinuity. This is mainly but not only due to plot 560 and its southern neighbour plot.

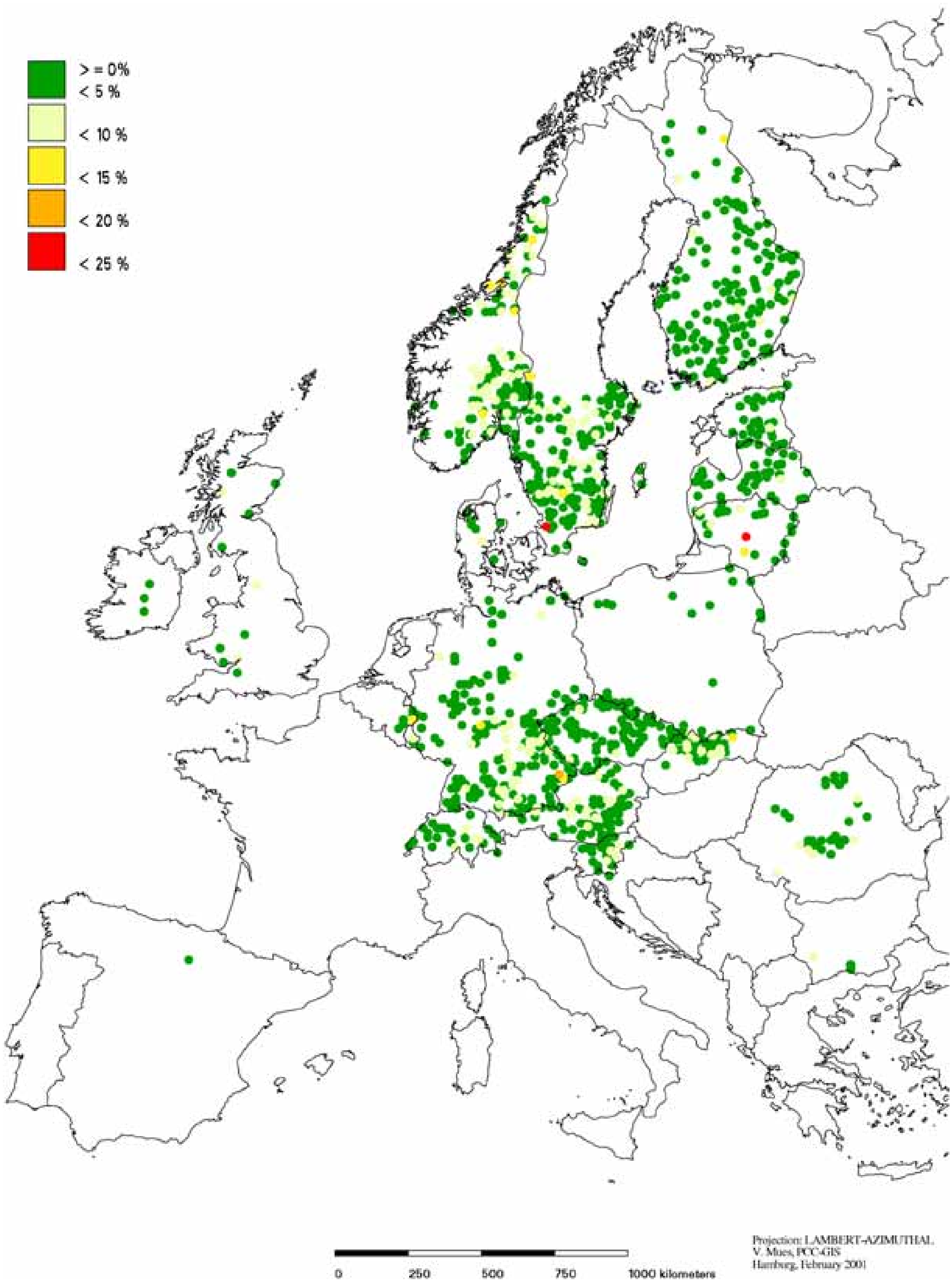


Figure 9: plot-wise rooted mean squared error (RMSE) of model for time trend of defoliation for the years 1994 – 2000

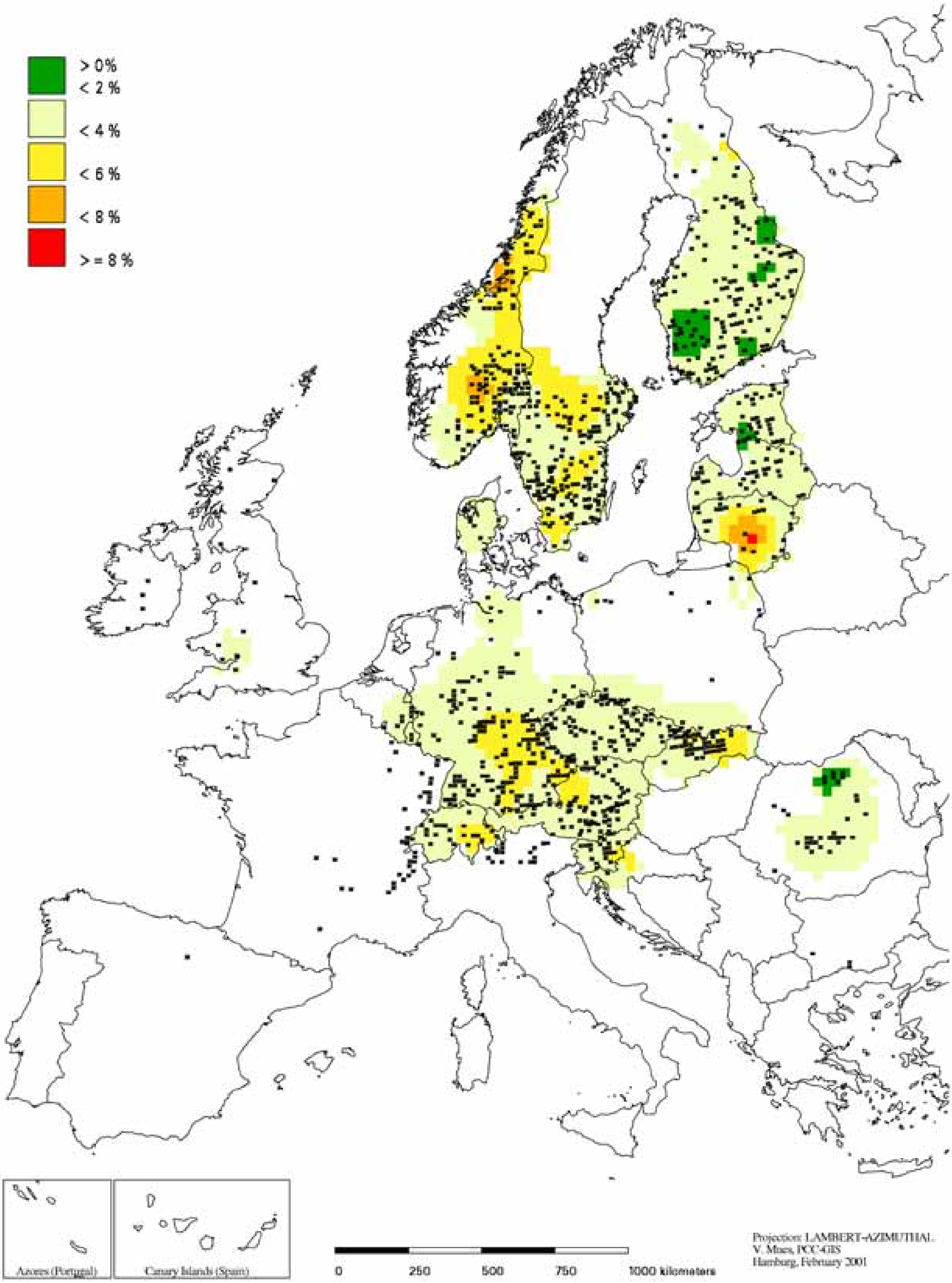


Figure 10: kriged plot-wise rooted mean squared error (RMSE) of model for time trend of defoliation for the years 1994 – 2000

3.2 Step 2

Two parameters derived from the analysis of step 1 are used as target variables for step2: (i) the plot-wise time trend (or regression coefficient) and the plot-wise rooted means of squared error (RMSE). Both of them were presented above (Figure 5 and Figure 9 respectively).

Until now the results for both are very preliminary but the following results seem to be significant for the two derived variables:

3.2.1 plot-wise time trend

There are three factors influencing the time trend:

1. deposition (diffSOx = depSOx98 – depSOx97)
2. water availability (WATER)
3. tree nutrition (base saturation 10 – 30 cm)

All these factors until now can be described only by time-constant variables. Nevertheless, especially the highly significant correlation between plot-wise time trend and diffSOx emphasises the necessity of including time-varying information about the deposition situation into the analysis of step 1.

The R^2 values of all reasonable models tested yet were below 5%.

3.2.2 plot-wise RMSE

Mainly two variables were detected, which are influencing the plot-wise RMSE. The positive correlation of long-time mean precipitation in summer and that of base saturation (10 – 30 cm) with the RMSE is plausible because discontinuities should be more probable in regions where in normal years a good nutrition and water availability is given. Extreme weather conditions should lead to more exposed discontinuity.

Nevertheless, the strong inter-correlation between some of the predictor variables has to be analysed more intensively.

Also for the plot-wise RMSE all the R^2 values of all reasonable models tested yet were below 5%.

4 Interpretation of results

The two-step approach for an analysis of temporal development of annual mean defoliation leads to a good explanation of the variance. The part of variance, which is explained by methodological differences (reference model b in Table 3, page 11, R2: 48.0%), is very high. The additional explanation with plotid can increase the R^2 to 84.2% (not in Table 3). The further inclusion of its interaction with time (year1994) leads to an R^2 of 90.3% (reference model a in Table 3).

The inclusion of plotid in a model of defoliation means the description of the plot-specific combination of time-constant stand and site properties. Its interaction with year1994 (year1994*plotid) describes the mean plot-wise time trend. A further improvement of the model is only possible by including time-varying variables, which can explain deviations from this mean time trend. A significant improvement was until now only possible by the inclusion of the variable minsect (model IV in Table 3) describing the mean occurrence of insects on trees of the respective plot in the respective year. The contribution of minsect to the model was significantly but R^2 could be increased only by 0.08%.

The models built during step1 of the analyses show that for the defoliation data a meaningful model can not be built without an explanation of methodologically caused differences between the data values. These variables explain a big part of defoliation's variance. An additional big part can be explained by plotid and the plot-wise time trend of defoliation. It is not astonishing that further significant influences like e.g. minsect will only lead to comparatively low increases of R^2. Nevertheless, their importance should not be underestimated (see also Hendriks et al., 2000).

The predictor variables minsect and mfungi additionally are limited in their explanation power by their binary definition (occurrence of absence of insects or fungi, respectively). Transformed into percentages of damaged trees per plot, the variables are metric on the one hand, but do not quantify the degree of defoliation by the abiotic damage. This is the reason for an unsatisfactory description of extreme damage as it was e.g. observed for plot 560 in Lithuania (Figure 2, page 7). The regression coefficient of minsect (8.77%), which was calculated as mean influence of minsect over all plots, can not fully explain the discontinuous increase in defoliation of more than 50%. The example from Lithuania demonstrates, that a much higher increase of defoliation can be associated with the occurrence of insects.

The right skewed distribution of both variables, minsect and mfungi, was to be expected. Nevertheless there is a danger of confounding effects as only a few plots with extreme values might be the basis for high correlations, which are not caused by some kind of cause-effect relationship. Plausibility tests of the resulting regression coefficients, thus, are a very important task.

In step 1 minsect was the only time-varying variable, which could contribute to the model significantly with a plausible regression coefficient. The other variables, like e.g. annual summer precipitation did not contribute as clearly. Nevertheless it is expected, that transformations of annual precipitation (e.g. its relation to the long-term mean value, relations to threshold values, or interactions with variables, which express soil water availability) will correlate more clearly with defoliation. This expectation is supported by the results of step 2 concerning the further analysis of the plot-wise time trends (regression coefficients). Only those time-constant variables seem to be considerably correlated with the plot-wise time trend, that are expressing factors, which do vary over time in reality. Especially the correlation between plot-wise time trend and diffSOx emphasises the necessity of including time-varying deposition information into the analysis of step 1.

Perhaps more promising than the further analysis of the plot-wise time trend is the analysis of the coefficients for plotid. As explained above, this variable describes the plot-specific combination of time-constant stand and site properties. Thus, it is logical to analyse, whether its values can be explained by time-constant variables, which describe stand and site properties.

Calculation and analysis of the plot-wise RMSE gives information about deviation from the assumed linearity in the sense of discontinuity. This information is valuable as on plots with high RMSE values the time trends have to be interpreted with care and a further analysing seems to be a promising task. The RMSE is a conservative estimator of model accuracy. In contrast e.g. to the mean absolute error, high deviations from the model are weighted over-proportional. That is exactly the property, which is needed to detect discontinuities.

On the other hand values of about 5% should not be taken as a mean model error. Thus, perhaps the kriged maps of RMSE, where the plot-wise values are smoothed, show values, which are a little more realistic in the sense of model accuracy.

The assumption of time consistency of assessments can be analysed only if future ICCs allow an exact quantification of this task. For assessments in the past a clear deviation between real mean trend of defoliation and methodologically caused trend (time in-consistence) of defoliation will not be possible. Today it is not expected that a time in-consistence of significant quantity occurred in the participating countries. The high variability of time trends within countries is an additional statistical document for the assumption of time consistency.

5 Recommendations

The mulitvariate statistical approach of step 1 has delivered conclusive results by the inclusion of predictor variables changing over time. The presented statistical models may still be improved if additional variables are included. In this respect, annual deposition values will be of specific interest (EMEP, EDACS).

Further improvement is expected by inclusion of interaction terms between meteorology and stand and soil properties as predicting variables.

The applied split-plot analysis into account the correlation between plot specific defoliation values. Temporal autocorrelation will be examined in future approaches, if longer time series are available. The evaluation of time series beginning in the eighties might reveal a more significant influence of sulphur depositions on defoliation, as the deposition values were higher at these times. On the other hand the database is not as comprehensive for the first years of the crown condition assessments.

In future evaluations all variables expressing influences changing over time are recommended to be included in step 1 of the analysis. Step 2 on the other hand might deliver more conclusive results when focussing on the target variable ‘plotid’. As this variable describes plot-specific combinations of time-constant stand and site properties, predictor variables should be chosen which also express stand and site properties.

6 References

Diggle, P.J., Liang, K.-Y., Zeger, S.L. 1994: Analysis of Longitudinal Data. Clarendon Press Oxford.

Hendriks, C.M.A., Olsthoorn, A.F.M., Klap, J.M., Goedhart, P.W., Oude Voshaar, J.H., Blecker, A., de Vries, F., van der Salm, C., Voogd, J.C.H., de Vries, W., Wijdeven, S.M.J. 2000: Relationships between crown condition and its determining factors in The Netherlands for the period 1984 to 1994. Alterra-rapport 161, Wageningen.

Lorenz, M., 1993. Die Europäische Waldzustandserfassung. Z. Ökologie u. Naturschutz 2. 245-251.

Lorenz, M., Seidling, W., Mues, V., Becher, G., Fischer, R., 2001. Forest condition in Europe: 2001 Technical Report. United Nations Economic commission for Europe, European Commission (eds.), Geneva, Brussels. 103 p + Annexes.

Ripley, B. D., 1981: Spatial Statistics. Wiley & Sons, New York.

Schall, P., 1999: Proposal for extrapolation of relationships identified for Level II plots with available data at Level I plots. Internal paper, unpubl.