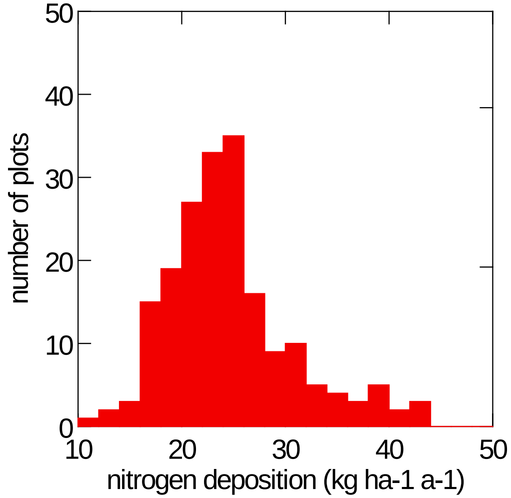
Epidemiological data analysis in forest research

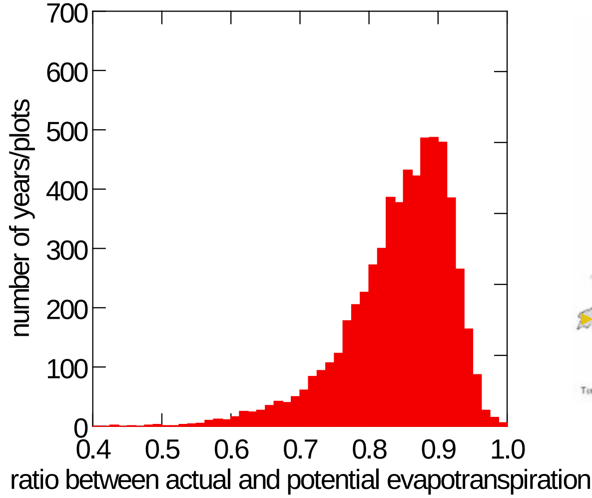
Sabine Braun[1)] and Christian Schindler[2)]
[1)]Institute for Applied Plant Biology, Schönenbuch, Switzerland
[2)] Swiss TPH, University of Basel, Switzerland

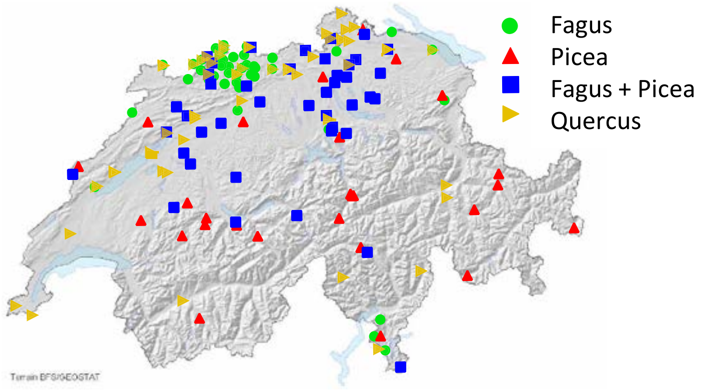
Epidemiological analysis in forest research

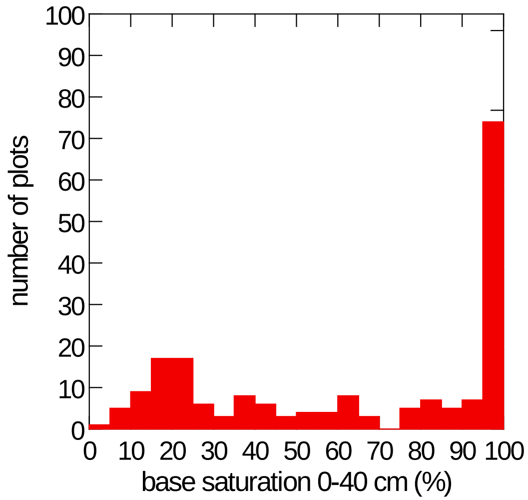
- Limitations according to DeVries et al. 2014:
  - high variability of forests being examined (non-homogeneous experimental material)
  - unbalanced and virtually biased allocation of plots to air pollution levels (non-random allocation of treatment, absence of replication)
  - effects of confounding factors and covariates
  - absence of a true control
  - potential inadequacy of sampling numbers for proper statistical analysis and interpretation
  - co-linearity between multi data
  - constraints in data availability
  - nature of the available response indicators

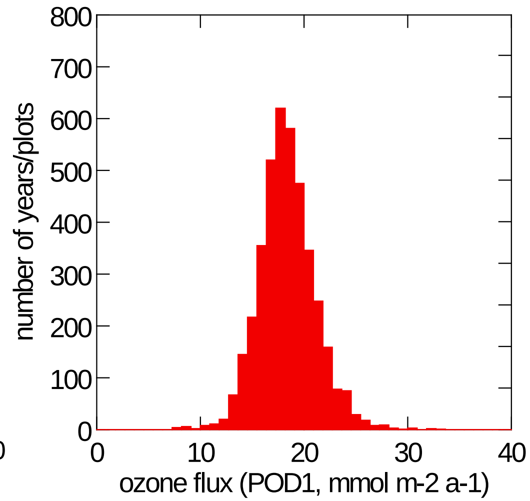
Gradients of predictor variables



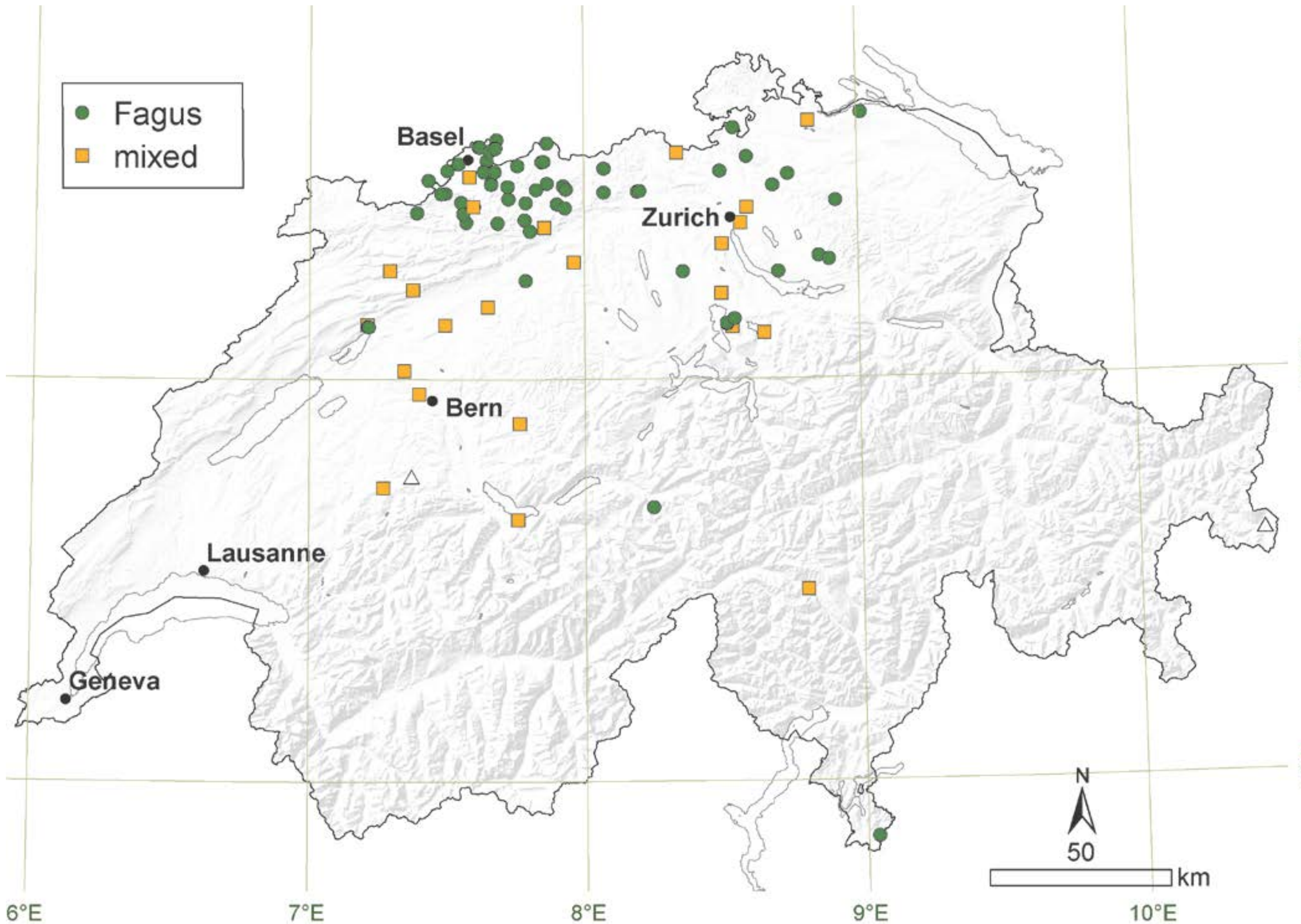








Beech observation plots included in the ozone study



Get an overview of the raw data

	A	B	C	D	E	F	G	H
1	site	tree	BHD91	BHD95	BHD98	BHD02	BHD06	BHD10
4619	57.0	156		63.1	64.2	66.9	68.6	70.6
4620	57.0	157		76.4	77.5	79.2	80.3	81.3
4621	57.0	158	59.5	60.7	61.6	62.5	63.5	64.4
4622	57.0	159		54.6	56.1	57.2	59.5	61.2
4623	57.0	161		57.5	58.6	59.4	60.5	61.8
4624	57.0	162	48.6	49.5	50.4	51.6	52.6	53.8
4625	57.0	163	61.8	63.4	64.3	65.4	66.7	67.8
4626	57.0	164		55.7	56.7	57.7	59.3	60.4
4627	57.0	165		50.7	51.5	52.5	53.1	53.9
4628	57.0	166	62.1	63.6	64.6	65.9	67.0	68.6
4629	57.0	167		75.2	76.5	78.4	80.1	81.8
4630	57.0	168	62.2	63.9	64.8	65.7	66.8	69.1
4631	57.0	169	47.0	48.4	49.4	50.6	51.9	53.4
4632	57.0	170	52.2	53.2	53.6	54.0	54.7	55.4
4633	57.0	171	56.9	58.5	59.9	61.2	62.7	
4634	57.0	172	38.8	39.7	40.1	40.6	41.0	
4635	57.0	173	46.8	48.9	50.4	52.3	54.4	56.3
4636	57.0	174	48.1	49.5	50.2	51.0	51.9	52.5
4637	57.0	175	45.3	46.5	47.1	47.9	48.6	49.5
4638	57.0	176	60.1	61.6	63.0	64.7	66.3	67.8
4639	57.0	177	43.7	45.2	45.8	46.5	47.0	47.7
4640	57.0	178		42.1	42.7	43.3	44.0	44.5
4641	57.0	179	60.3	62.5	64.0	65.3	66.5	67.2
4642	57.0	180	49.2	51.1	52.4	53.8	54.8	55.7
4643	57.0	181		51.8	53.0	54.6	56.1	57.5
4644	57.0	182		72.7	73.5	75.9	77.3	78.9
4645	57.0	183	35.2	36.7	38.0	38.8	39.6	40.5
4646	57.0	186	35.7	38.3	40.2	42.4	44.6	47.1
4647	57.0	187		41.4	42.3	43.8	44.5	45.3
4648	57.0	188		53.1	54.0	54.8	55.9	56.9

Data preparation

- Remove trees with implausible data (e.g. measurement errors, mismatches of trees)
- Calculate growth
  – Difference between diameters of subsequent years
  – Calculate as basal area increment (smaller age trends)
  – In the ozone study, relative data were analyzed (deviation from the average over the observation period: for growth on the basis of individual trees, for site factors on the basis of site average). This removes effects of all site factors which are constant over the observation period (e.g. N deposition, altitude), and also tree effects (e.g. social position).
- Transform data into a sequential file with one observation per tree and year
- Add covariates to the file

Regression models with grouped datasets

Use mixed regression (PROC MIXED in SAS, lme or lmer in R)

- For model comparison, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) are available
- As the goodness of fit always increases when the number of parameters is increased, both criteria penalize for the number of parameters (p=number of parameters without intercept, n=sample size):
  Penalty in AIC = 2*p
  Penalty in BIC = p*ln(n)

Selection of covariates

- Select explanatory variables according theoretical considerations prior to conducting regression analyses. Optimize the regression for minimum AIC and minimum BIC. As a rule of thumb, the number of parameters in the regression should be ≤ 1/10 of the number of sites.
- Explore the data set using factor analysis. Strongly intercorrelated variables may be replaced by a common factor from the factor analysis or by the variable which most strongly correlates with the factor.
- Backward selection starting from a large model is better than model building by forward selection. However, test similar parameters in separate equations and compare the goodness of fit.

Selection of covariates

The following predictors have been used successfully:

- Ozone
- Drought
- Age
- Altitude
- N Deposition

Calculation of drought

- Hydrological model Wasim-ETH
- Meteorological input data interpolated on a daily basis by Meteotest (temperature, humidity, irradiation, precipitation, wind speed)
- Local soil and stand characteristics

Land use table of Wasim-ETH

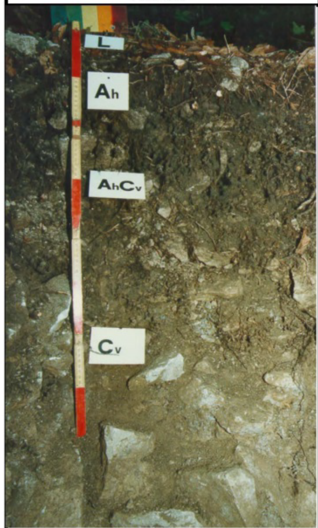
```
[multilayer_landuse]
1 # count of multilayer landuses
1  mixed_forest  { Landuse_Layers = 1, 2, 3;                                #
                 k_extinct = 0.3 ;   # extinction coefficient for reducing radiation
                 LAI_scale = 20 ;    # Scaling factor for calculating the aerodynamic resistencies of layer 2..n dependent on the
cumulated leaf area index
                                              }
[landuse_table]
3                   # number of following land use codes, per row one use
#Co name of the
#de Landuse type
#-- --------------
1 beech  {method   = VariableDayCount;
        RootDistr        = 1.0;       # parameter for root density distribution
        TReduWet         = 0.95;      # relative theta value for beginning water stress
        LimitReduWet     = 0.5;       # minimum relative reduction factor of real transpiration at saturation
        HReduDry         = 3.5;              # hydraulic head (suction) for beginning dryness stress
        IntercepCap      = 0.6;              # optional: specific thickness of the water layer on the leafes in mm.
JulDays          = 15    74    115   125   135   166    196    227    285    310    319   ; #
Albedo           = 0.17  0.17  0.17  0.17  0.17  0.17   0.17   0.17   0.17   0.17   0.17   ;  #
rsc              = 100   100   95    75    65    65     65     65     65     85     100   ;    # leaf surface resistance in s/m
rs_interception  = 10    10    10    10    10    10     10     10     10     10     10     ;
rs_evaporation   = 500   500   500   500   500   500    500    500    500    500    500   ;    # SOIL surface resistance in s/m
LAI              = 1     1     1     3     5     5      5      5      4      1      1      ;   # Leaf Area Index (1/1)
Z0               = 2.0   2.0   2.0   2.0   2.0   2.0    2.0    2.0    2.0    2.0    2.0    ;   # Roughness length in m
VCF              = 0.3   0.3   0.3   0.5   0.6   0.6    0.6    0.6    0.6    0.5    0.3    ;   # Vegetation covered fraction
RootDepth        = 1.0   1.0   1.0   1.0   1.0   1.0    1.0    1.0    1.0    1.0    1.0    ;   # Root depth in m
AltDep           = 0.025 0.025 0.025 0.025 0.025 0.025 -0.025 -0.025 -0.025 -0.025 -0.025 ;    # altitude correction for JulDays
                        }
```

Soil table of Wasim-ETH

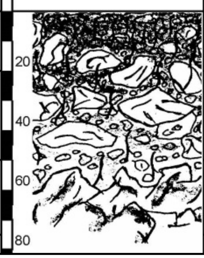
```
[soil_table]
1                  # number of following entries
#co- name of the
#de  soil profile
#-- ---------------
1 profil_31bu           {method = MultipleHorizons;
PMacroThresh            = 1.0 ; # precipitation capacity thresholding macropore runoff in mm per hour
MacroCapacity           = 3   ; # capacity of the macropores in mm per hour
CapacityRedu            = 0.8 ; # reduction of the macropore capacity with depth (reduction ratio/meter)
MacroDepth              = 0.5 ; # maximum depth of the macropores
horizon     =           1          2          3          4          5;
Name        =           Ah         Al         BvAl       BvBt       Bt;
 ksat         =             2.31e-5    8.1e-6    3.18e-7    3.18e-7    3.18e-7 ;
 k_recession  =                  1          1          1          1          1 ;
 theta_sat    =               0.46       0.46       0.45      0.266      0.266 ;
 theta_res    =              0.034      0.034      0.067      0.048      0.048 ;
 alpha        =                1.6        1.6          2        0.8        0.8 ;
 Par_n        =               1.37       1.37       1.41       1.09       1.09 ;
 Par_tau      =                0.5        0.5        0.5        0.5        0.5 ;
 thickness    =               0.03       0.07        0.3        0.1        0.5 ;
 layers       =                  1          1          1          1         48 ;
                              }

# C&P      soil_types  = U         U          Lu         T          T           ;
```

Soil profiles

Profilfoto	Standort-Daten		
	Politische Gemeinde	Bennwil	
	Standortbezeichnung	17 Bennwil	
	Koordinaten (1:25'000)	625 900	249 350
	Datum der Aufnahme	23.08.1984	
	Geologie		
	Neigung 21°	Höhe	665
	Exposition	W	
	Lage im Relief	Hang	
	Verbale Kurzbeschreibung		
	Bodenform	Rendzina	
	Humusform	L-Mull	
	Durchwurzelungstiefe (cm)	50	mittel
	nutzbare Feldkapazität des Wurzelraums (mm)	48	sehr gering
	Feldkapazität des Wurzelraums (mm)	68	sehr gering
	Vernässung (Stufe 0-6)	0	nicht staunass
	Luftkapazität (Stufe 1-5)	1	sehr gering
	Basensättigung 0-40 cm (%)	100	hoch
	pH(CaCl2) 0-40 cm	7.09	sehr schwach alkalisch
Lage: zw. Bu 30 (5m südl.), 36 (10m), 35 (5m)	pflanzensoziologische Einheit	12*, 12a	

Profil-Aufnahme



Horizont Tiefe	Bez. 62	Farbe Munsell	pH CaCl2	Kalk Mü. 94	Bd.art Mü. 80	Gefüge Mü. 125	Poren GPV /LK /nFK /FK	Dichte Mü. 126	Humus org. S%	Skelett Vol%	Wurz. Mü. 130	BS %	BCAL
0-20	Ah	10YR 2/1	6.9	8.2	Ul3	kru 1,4 sub-pol 2 3,4	59/ 11/ 37/ 51	1.5	28.4	70	5	100	63
20-50	AhCv	10YR 3/2	7.3	13.2	Uls	sub-pol 2-3,4	54/ 11/ 31/ 43	1.5	11.5	80	3.5	100	.
50-75	Cv	10YR 4/3-7/6	7.7	50.8	Us	ein 1-z.T. koh	50/ 15/ 26/ 35	1.5		95	0	100	.

Bennwil
plant-available soil water 46 mm
very low

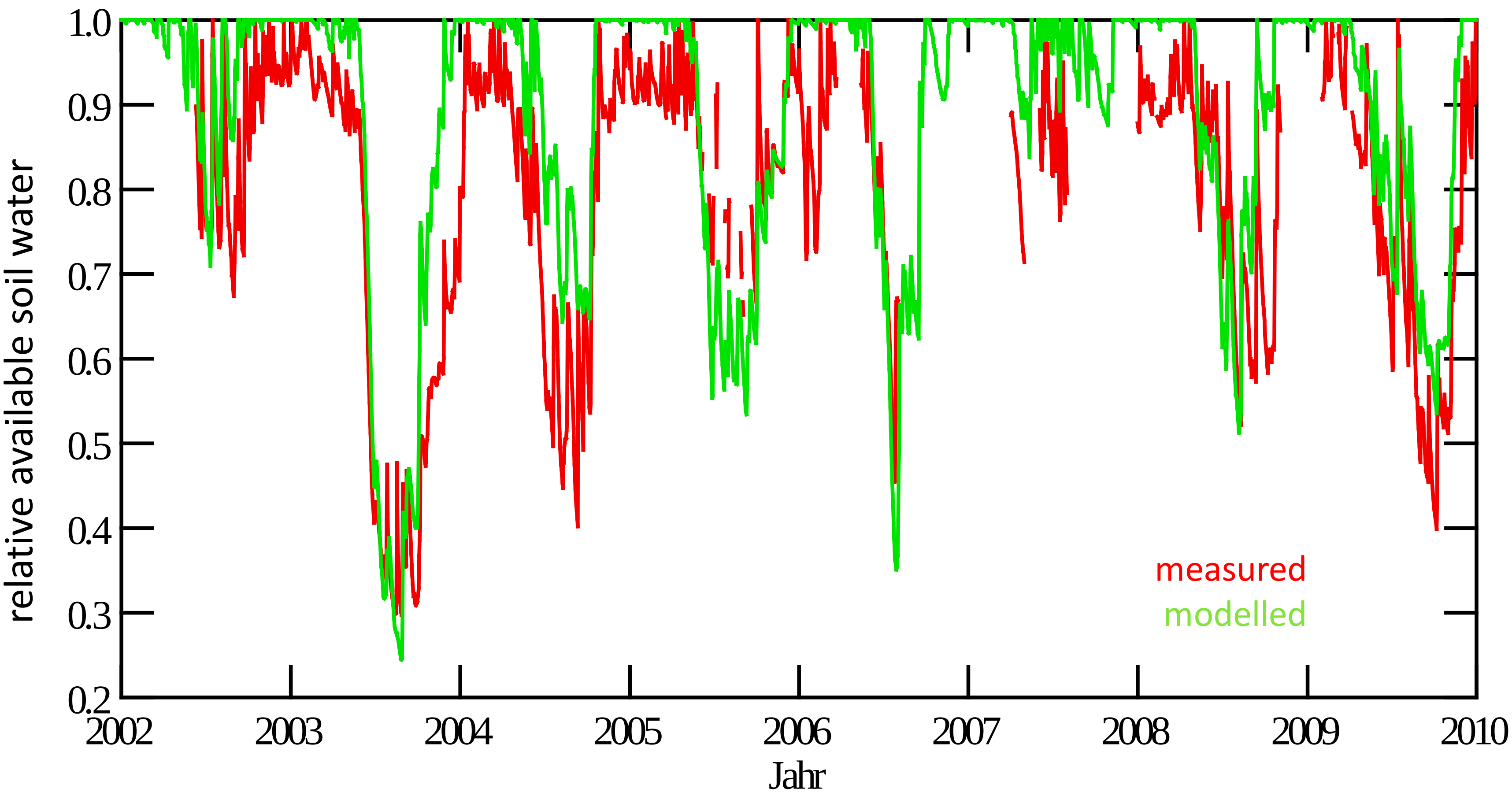
Profilfoto	Standort-Daten		
	Politische Gemeinde	Aarwangen	
	Standortbezeichnung	77 Aarwangen	
	Koordinaten (1:25'000)	626220/232686	(GPS)
	Datum der Aufnahme	15.03.00	
	Geologie		
	Neigung 9°	Höhe	470
	Exposition	NW	
	Lage im Relief		
	Verbale Kurzbeschreibung		
	Bodenform	Braunerde	
	Humusform	L-Mull	
	Durchwurzelungstiefe (cm)	100	tiefgründig
	nutzbare Feldkapazität des Wurzelraums (mm)	220	sehr hoch
	Feldkapazität des Wurzelraums (mm)	378	mittel
	Vernässung (Stufe 0-6)	1	sehr schwach staunass
	Luftkapazität (Stufe 1-5)	3	mittel
	Basensättigung 0-40 cm (%)	21	tief
Lage:	pH(CaCl2) 0-40 cm	4.05	stark sauer
zw. Bu 17, 13, 18, 36 (6.4m, 9.8m, 5.1m, 9.8m)	pflanzensoziologische Einheit	7aaB	

Profil-Aufnahme

Horizont Tiefe	Bez. 62	Farbe Munsell	pH CaCl2	Kalk Mü. 94	Bd.art Mü. 80	Gefüge Mü. 125	Poren GPV /LK /nFK /FK	Dichte Mü. 126	Humus org. S%	Skelett Vol%	Wurz. Mü. 130	BS %	BCAL
0-10	Ah	10YR 3/3	3.95	0	Ls2	pol2-3	69/ 12/ 37/ 57	1	7.9	0	5	42	7.5
10-25	AhAl	10YR 3/4	4.13	0	Ut4	pol3	55/ 13/ 28/ 42	2	4.4	0	3	16	4.4
25-55	Al	10YR 4/4	4.02	0	Ut4	pol3	45/ 8/ 22/ 37	3	1.8	0	2	14	4.9
55-100	Bv	10YR 5/4	4.11	0	Ls2	pol4	43/ 10/ 17/ 33	3	1.1	0	1	44	10.8
10[illegible]											0	65	9.4

Aarwangen
plant-available soil water 220 mm
very high

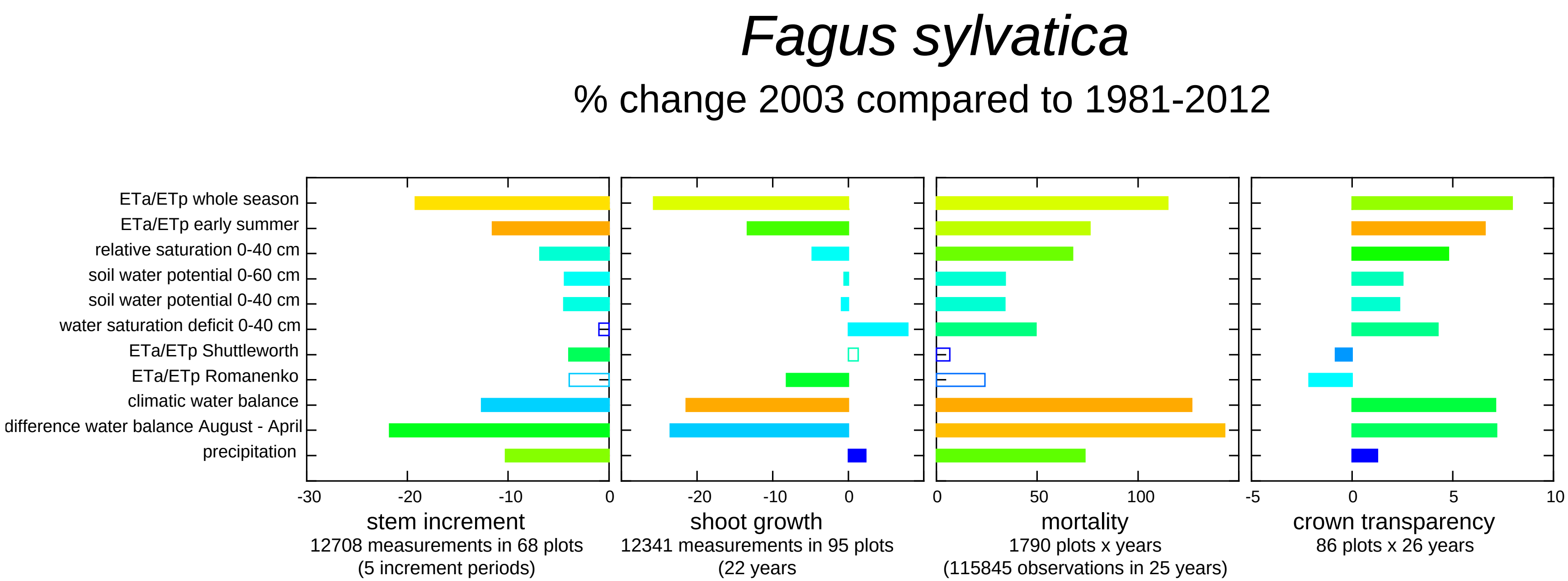
Validation against measured data
example Brislach 20 cm



Drought indicators tested

- Ratio between actual and potential evapotranspiration calculated with Wasim (Penman equation), averaged either for the whole season or for early summer
- Ratio between actual and potential evapotranspiration, calculated after Romanenko or Shuttleworth
- Soil water potential averaged over 40 or 60 cm
- Soil moisture deficit in the uppermost 40 cm
- Relative soil water (AWB) in the uppermost 40 cm
- Water balance
- Precipitation

Selection of optimum drought indicator



Multivariable models including other significant predictors depending on the parameter

The colors are continuous between orange and blue and represent the range of the AIC (orange: best, blue: worst model)

Empty columns: regression not significant

Output of lme in R for the beech dataset

```
Linear mixed-effects model fit by REML
 Data: stbugrp3
      AIC      BIC   logLik
  10606.9 10673.97 -5294.45

Random effects:
 Formula: ~1 | SITE
        (Intercept)  Residual
StdDev:   0.0426545 0.3649011

Fixed effects: GZUWREL ~ POD1REL + ETAETPSPRING + as.factor(YEAR)
                          Value  Std.Error    DF    t-value p-value
(Intercept)           0.0508870 0.15450531 12667   0.329354  0.7419
POD1REL              -0.2842555 0.06549411 12667  -4.340169  0.0000
ETAETPSPRING          1.5215673 0.14436378 12667  10.539813  0.0000
as.factor(YEAR)1998   0.0653975 0.01670516 12667   3.914811  0.0001
as.factor(YEAR)2002  -0.1894897 0.01772971 12667 -10.687698  0.0000
as.factor(YEAR)2006   0.0963410 0.02310224 12667   4.170202  0.0000
as.factor(YEAR)2010  -0.0329751 0.02143736 12667  -1.538205  0.1240
```

Grouping factor = site

Factor analysis

"Variance" Explained by Rotated Components:	1	2	3
	9.361	7.968	3.356
Percent of Total Variance Explained	1	2	3
	39.006	33.199	13.982

Factor loadings

	irradiation				air temperature						ETa/ETp averaged over various periods of the season											other soil moisture indicators		
Factor	STRAREL	STRAB1RL	STRAB2RL	STRAF1RL	LTEMPREL	LTEMPB1RL	LTEMPB2RL	LTEMPB3RL	LTEMPB4RL	LTEMPB10RL	ETVERHREL	ETVB1RL	ETVB2RL	ETVB3RL	ETVB4RL	ETVB5RL	ETVB6RL	ETVB7RL	ETVB8RL	ETVB9RL	ETVB10RL	RS40REL	SMD40REL	PF60REL
1	0.576	0.235	0.186	0.005	0.142	0.654	-0.165	0.611	0.779	0.724	-0.957	-0.954	-0.099	-0.453	-0.792	0.012	0.013	-0.290	-0.690	-0.908	-0.934	-0.893	0.803	0.770
2	-0.684	-0.825	-0.886	-0.635	0.140	0.022	-0.644	-0.314	0.031	0.108	0.057	0.039	0.971	0.842	0.517	0.930	0.950	0.926	0.619	0.213	0.183	0.206	-0.072	-0.154
3	0.078	0.341	-0.225	-0.484	0.929	0.740	0.694	0.445	0.519	0.638	-0.077	-0.103	0.024	0.199	0.137	-0.106	-0.121	0.078	0.191	-0.076	-0.025	-0.124	0.081	0.108

Use of factors from factor analysis as covariate

```
 Data: stbugrp3
       AIC      BIC    logLik
  10597.69 10679.66 -5287.847

Random effects:
 Formula: ~1 | SITE
        (Intercept)  Residual
StdDev:  0.04148319 0.3644801

Fixed effects: GZUWREL ~ POD1REL + FACTOR.1. + FACTOR.2. + FACTOR.3.
 + as.factor(YEAR)
                         Value  Std.Error    DF    t-value p-value
(Intercept)          1.5220019 0.07395083 12665  20.581268  0.0000
POD1REL             -0.3246860 0.06611709 12665  -4.910772  0.0000
FACTOR.1.           -0.1091736 0.00990636 12665 -11.020554  0.0000
FACTOR.2.            0.0028045 0.00859626 12665   0.326252  0.7442
FACTOR.3.           -0.0150367 0.01008694 12665  -1.490712  0.1361
as.factor(YEAR)1998  0.0988068 0.02198910 12665   4.493444  0.0000
as.factor(YEAR)2002 -0.1321246 0.02217762 12665  -5.957562  0.0000
as.factor(YEAR)2006  0.2573075 0.03207253 12665   8.022677  0.0000
as.factor(YEAR)2010  0.0352308 0.03089130 12665   1.140477  0.2541
```

Grouping of dataset

- 5 increment measurement periods (YEAR)
- 62 plots
- 60 trees per plot

Several groupings are possible:

- year
- site
- tree within site

Test different groupings
Dependent variable: basal area growth, relative

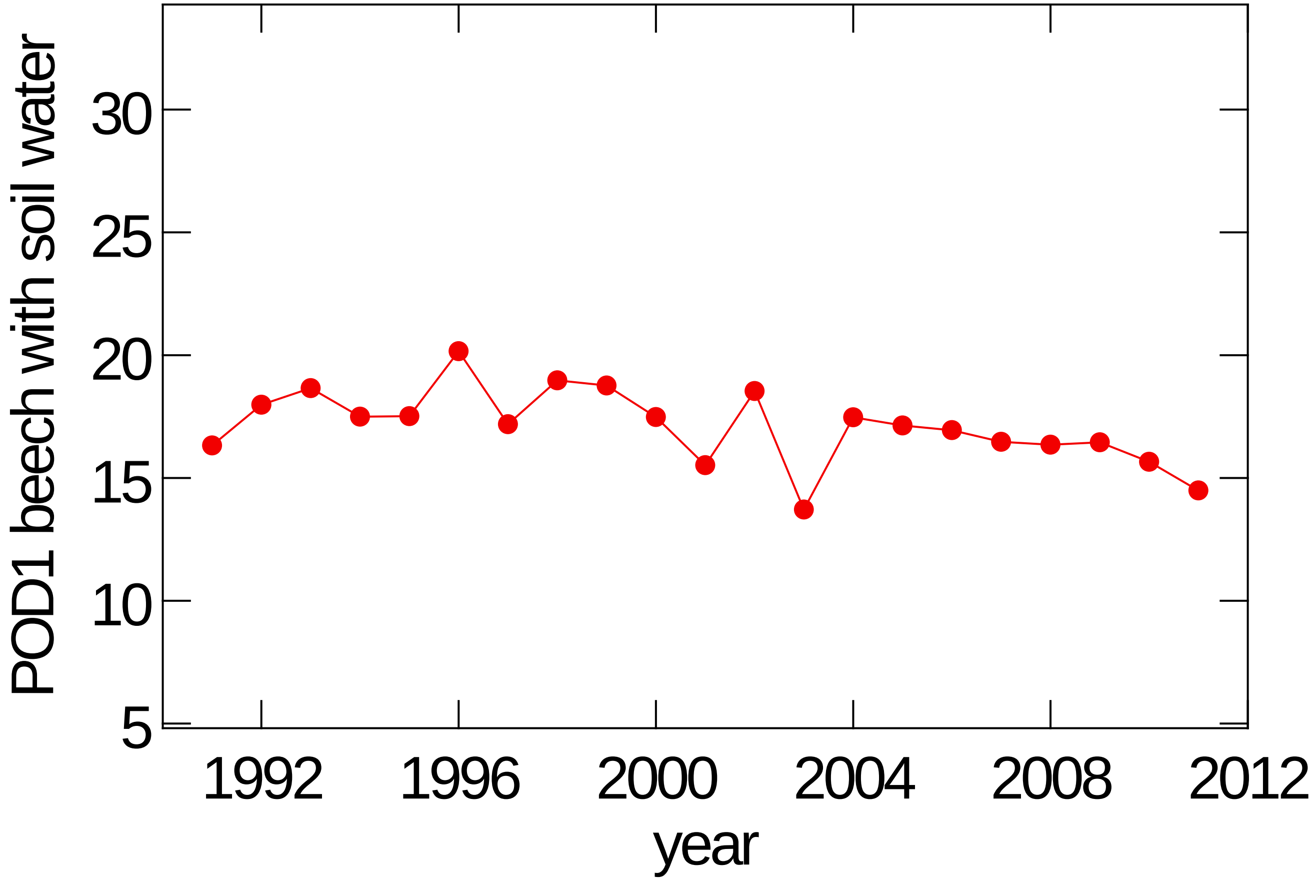
Group	number of groups	AIC	parameter	coeff	SE	t-value	p-value
site	68	11203	(Intercept)	1.846	0.076	24.181	0.000
			POD1	0.299	0.043	6.988	0.000
			ETa/ETp	-0.886	0.094	-9.379	0.000
tree within site	3487	11269	(Intercept)	1.830	0.076	23.977	0.000
			POD1	0.288	0.042	6.798	0.000
			ETa/ETp	-0.862	0.094	-9.128	0.000
year	5	10674	(Intercept)	-0.068	0.168	-0.407	0.684
			POD1	-0.331	0.065	-5.096	0.000
			ETa/ETp	1.670	0.168	9.964	0.000
site, with year as factor variable	68	10620	(Intercept)	-0.019	0.169	-0.111	0.912
			POD1	-0.366	0.068	-5.395	0.000
			ETa/ETp	1.678	0.170	9.881	0.000
			as.factor(JAHR)1998	0.059	0.017	3.548	0.000
			as.factor(JAHR)2002	-0.204	0.018	-11.494	0.000
			as.factor(JAHR)2006	0.114	0.024	4.682	0.000
			as.factor(JAHR)2010	-0.053	0.021	-2.490	0.013

Test different groupings

Group	number of groups	AIC	parameter	coeff	SE	t-value	p-value
site	68	11203	(Intercept)	1.846	0.076	24.181	0.000
			POD1	0.299	0.043	6.988	0.000
			ETa/ETp	-0.886	0.094	-9.379	0.000
site, with year as factor variable	68	10620	(Intercept)	-0.019	0.169	-0.111	0.912
			POD1	-0.366	0.068	-5.395	0.000
			ETa/ETp	1.678	0.170	9.881	0.000
			as.factor(JAHR)1998	0.059	0.017	3.548	0.000
			as.factor(JAHR)2002	-0.204	0.018	-11.494	0.000
			as.factor(JAHR)2006	0.114	0.024	4.682	0.000
			as.factor(JAHR)2010	-0.053	0.021	-2.490	0.013

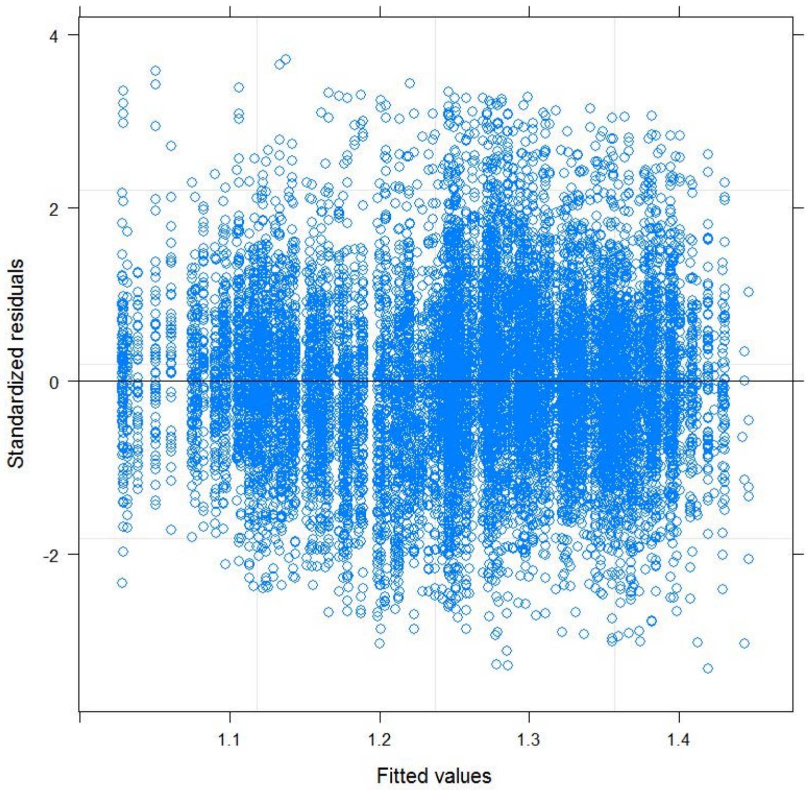
This suggests that the variable year is confounding

Variation of POD1 with time

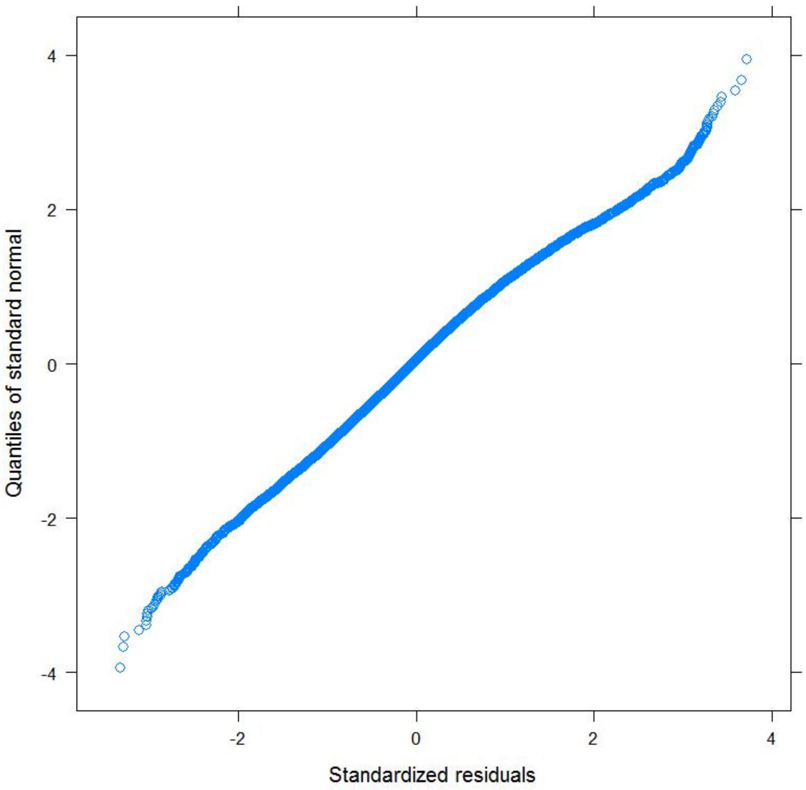


Model validation: residual plots

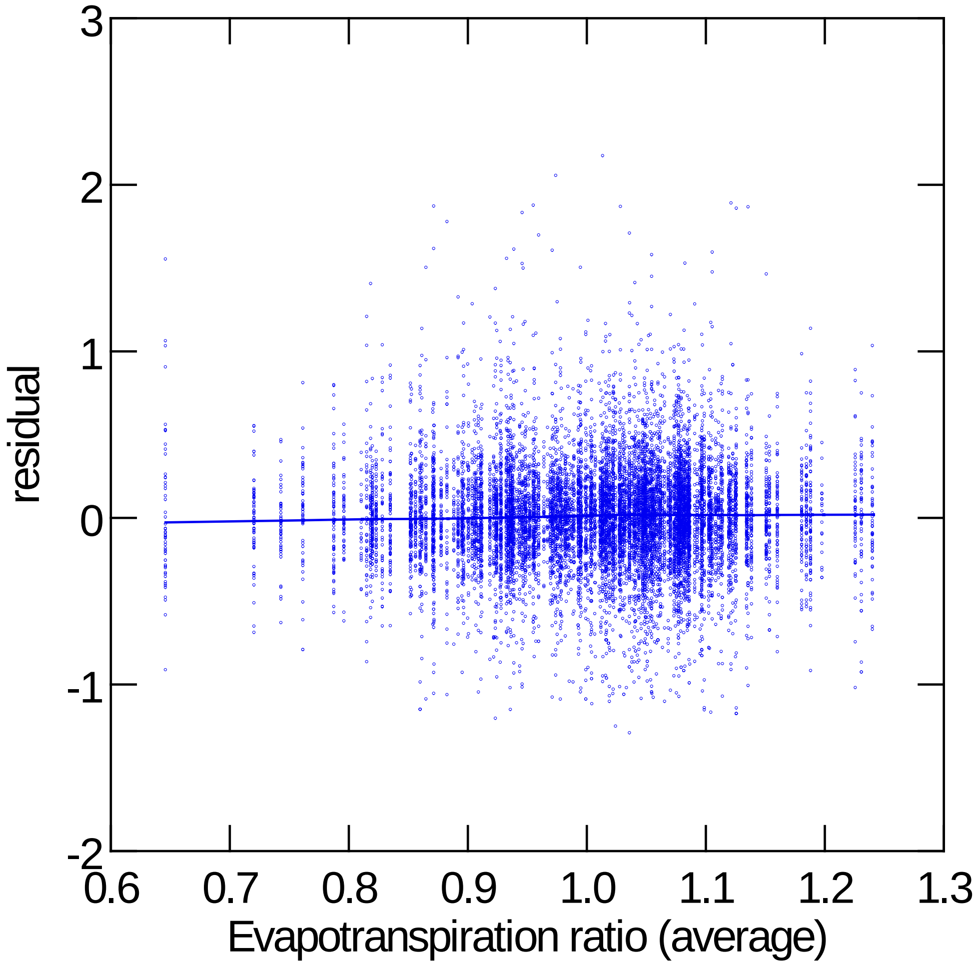
plot(model1)



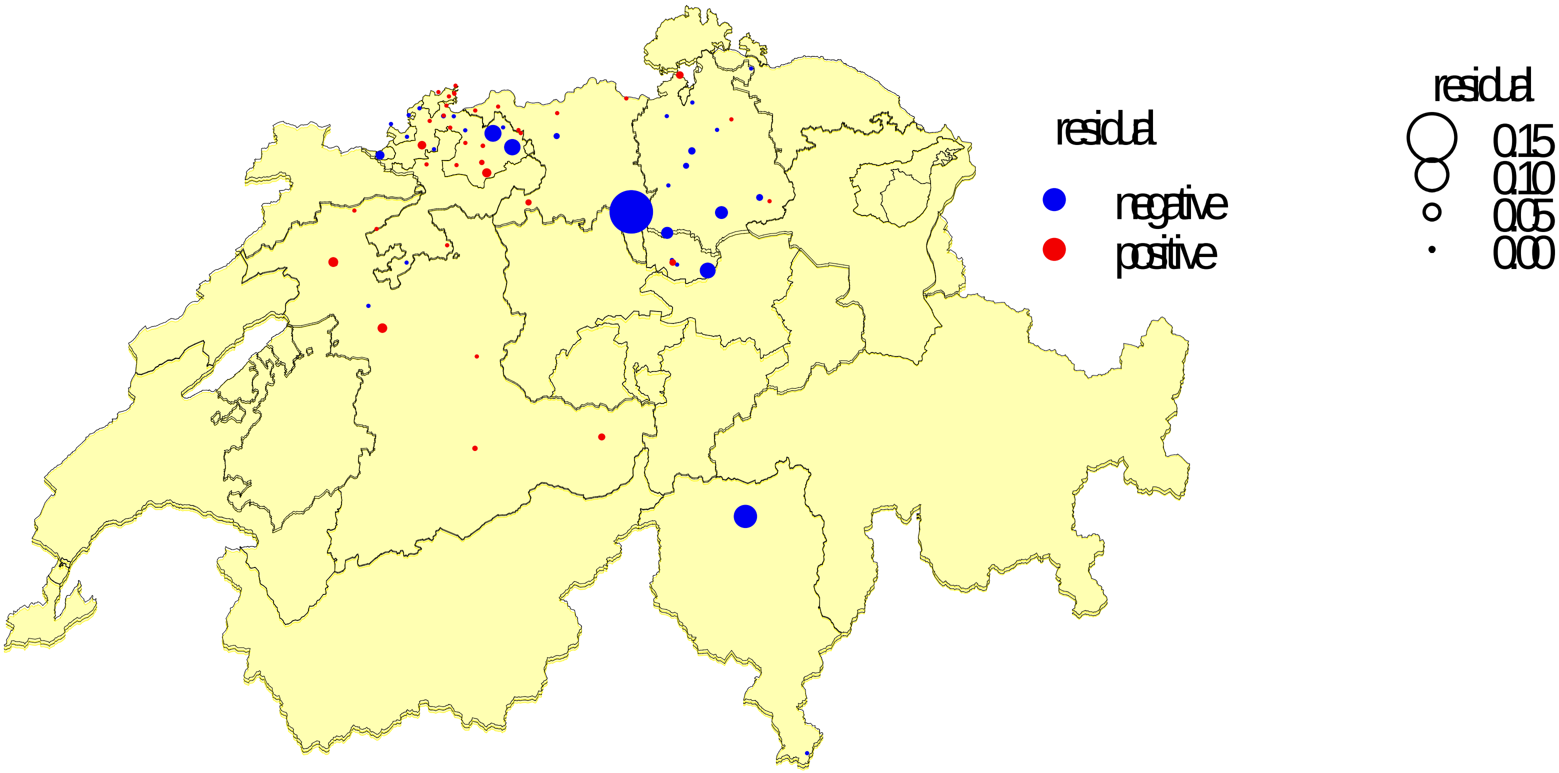
qqnorm(model1)



Model validation
Residuals vs. evapotranspiration ratio

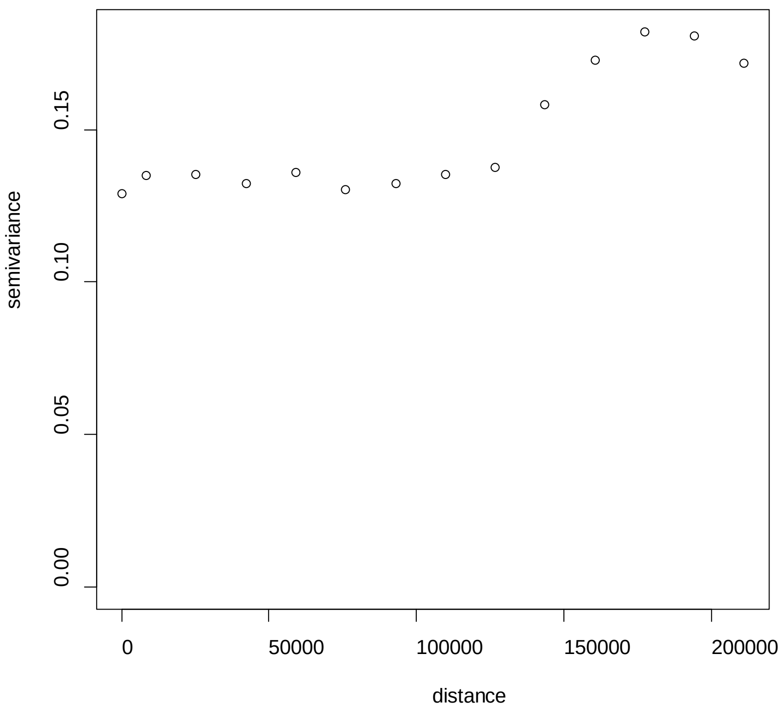


Model validation
Geographical distribution of the residuals

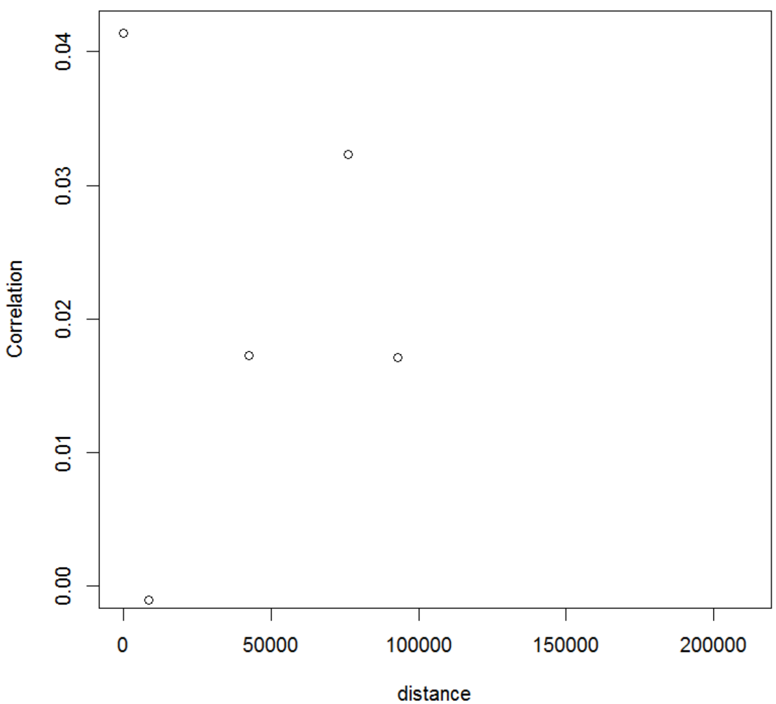


Spatial autocorrelation (library geoR)

```
resid<-residuals(st1)
coo<-
cbind(ozonstzuwbual14$KW,ozonstzuwbual14$KN)
vg<-variog(coords=coo,data=resid)
plot(vg)
```

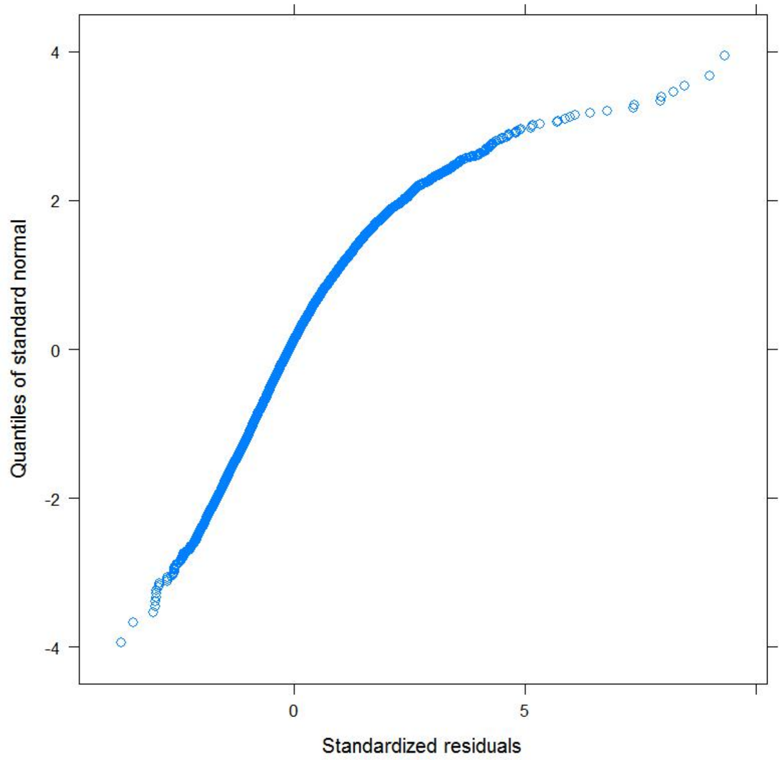


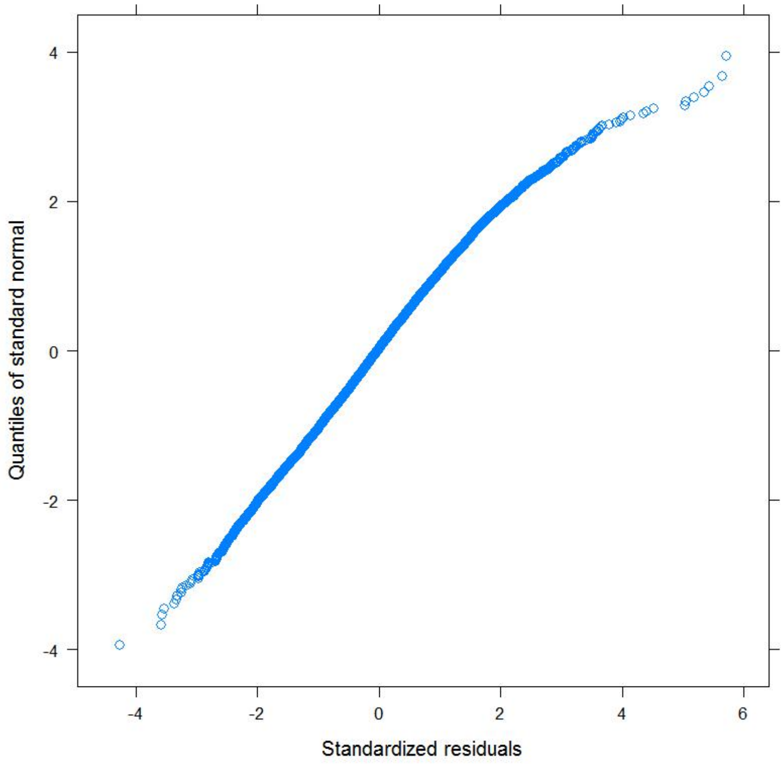

```
varr<-var(resid)
vg_<-vg
vg_$v<-(1-vg$v/varr)
plot(vg_,ylab='Correlation')
```



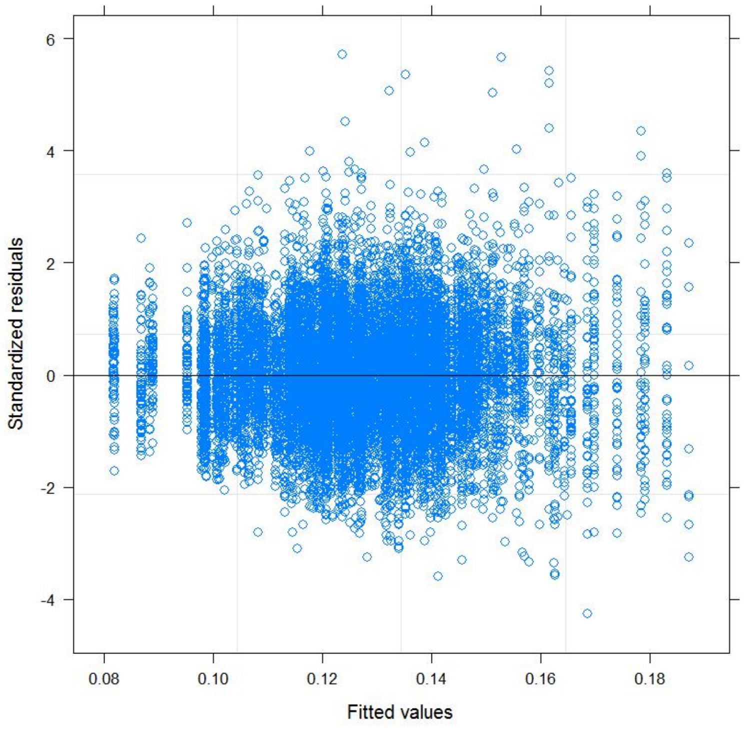
Normal distribution
Dependent variable: basal area growth

after square root transformation:

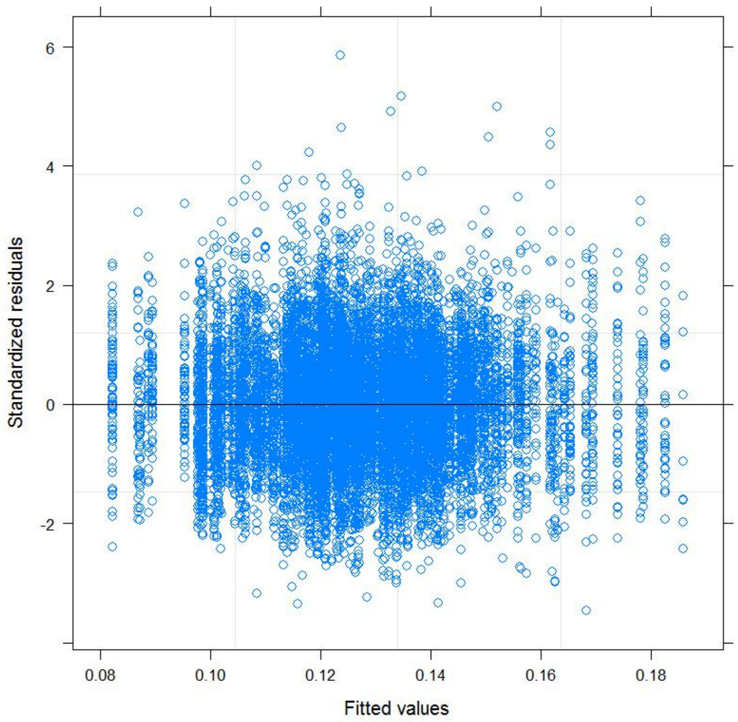




Heteroscedasticity




```
st3a <- update(st3,weights=varPower())
anova(st3,st3a)
```




```
Model df        AIC       BIC    logLik   Test  L.Ratio p-value
st3      1 10 -55084.13 -55009.61 27552.07
st3a     2 11 -55344.45 -55262.48 27683.22 1 vs 2 262.3176  <.0001
```

Extraction of a quantitative relationship from a multivariate output

Dependent variable: GZUWREL				
	Value	Std.Error	t-value	p-value
(Intercept)	0.043	0.155	0.274	0.784
POD1REL	-0.287	0.066	-4.374	0.000
ETa/ETp early summer	1.534	0.146	10.532	0.000
as.factor(JAHR)1998	0.065	0.017	3.892	0.000
as.factor(JAHR)2002	-0.190	0.018	-10.719	0.000
as.factor(JAHR)2006	0.096	0.023	4.155	0.000
as.factor(JAHR)2010	-0.039	0.021	-1.808	0.071

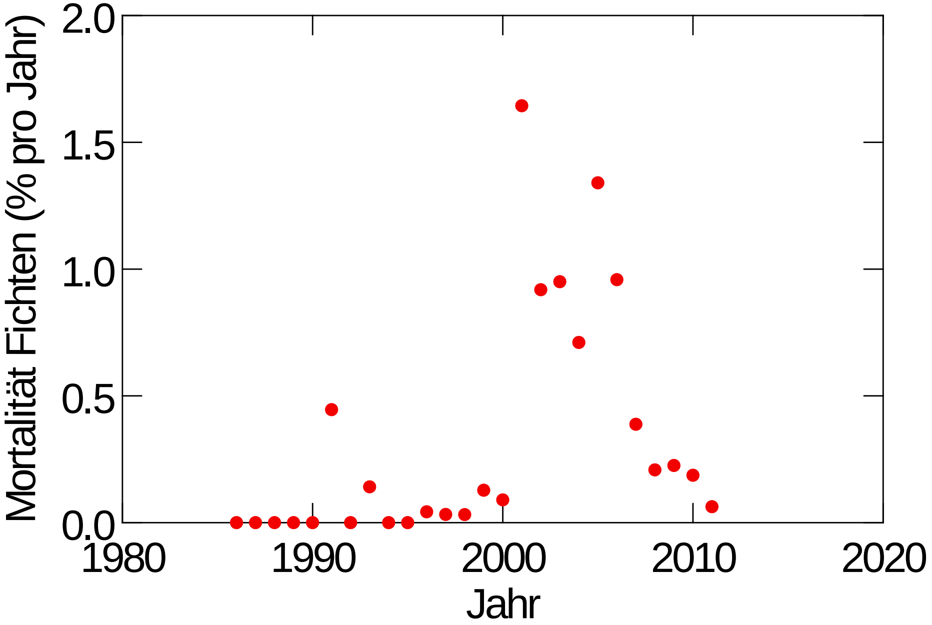
n Factors: Define n-1 binary variables which get the value 1 when $factor_{1..(n-1)}$ is true otherwise 0. In the above example JAHR1=1 when YEAR=1998, JAHR2=1 when YEAR=2002 etc.

E(Y) = 0.043 – 0.287*POD1REL + 1.534*mean(Eta/Etp) + 0.065*mean(year98) – 0.19*mean(year02) + 0.096*mean(year06) – 0.039*mean(year10)
where yearxx = indicator variable for observations from period xx and mean(yearxx) = relative frequency of observations from period xx.

Second example
mortality of Norway spruce with logistic regression

- Dependent variable: the number of dead and of surviving trees by site
- Grouped logistic regression
- Package lme4 in R

Development of mortality of Norway spruce in observation plots

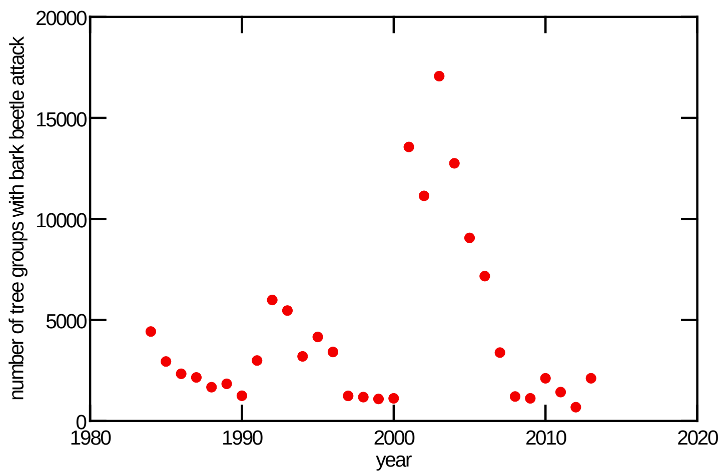



```
mortfi3 <-glmer(cbind(GEST,NGEST) ~
KK14+ETVERH*KNDEP21+ANZHERDE+(1|STNRNEU),
family=binomial,data=mortalitatfi)
```

Possible covariates for the time series

Introducing a group for the years 2001-2006 : has no explanatory value
Number of windthrown trees: Is itself correlated with drought, nitrogen deposition and base saturation

Number of tree groups with bark beetle attack (national database, average for CH)



With or without interaction?

	AIC	BIC		
	1435	1465		
	Estimate	SE	zvalue	pvalue
(Intercept)	-3.193	1.043	-3.061	0.002
foliar K (≤2.8, >2.8 mg g^{-1})	-0.786	0.289	-2.72	0.007
ETa/ETp	-4.073	0.951	-4.281	0.000
N-Deposition (≤20, >20 kg N ha^{-1} yr^{-1})	-0.727	0.445	-1.634	0.102
tree groups with bark beetle (*1000)	0.128	0.011	11.999	0.000

	AIC	BIC		
	1429	1464		
	Estimate	SE	zvalue	pvalue
(Intercept)	-14.310	4.910	-2.915	0.004
foliar K (≤2.8, >2.8 mg g^{-1})	-0.835	0.290	-2.883	0.004
ETa/ETp	8.626	5.514	1.564	0.118
N-Deposition (≤20, >20 kg N ha^{-1} yr^{-1})	10.970	4.978	2.205	0.027
tree groups with bark beetle (*1000)	0.126	0.011	11.841	0.000
ETa/ETp*Ndep	-13.320	5.603	-2.377	0.017

With or without bark beetle frequency?

	AIC	BIC		
	1429	1464		
	Estimate	SE	zvalue	pvalue
(Intercept)	-14.310	4.910	-2.915	0.004
foliar K (≤2.8, >2.8 mg g^{-1})	-0.835	0.290	-2.883	0.004
ETa/ETp	8.626	5.514	1.564	0.118
N-Deposition (≤20, >20 kg N ha^{-1} yr^{-1})	10.970	4.978	2.205	0.027
tree groups with bark beetle (*1000)	0.126	0.011	11.841	0.000
ETa/ETp*Ndep	-13.320	5.603	-2.377	0.017

ETa/ETp	Ndep	drought effec	N effect
0.852	0		
0.735	0	0.367	
0.852	1		0.689
0.735	1	1.726	3.242

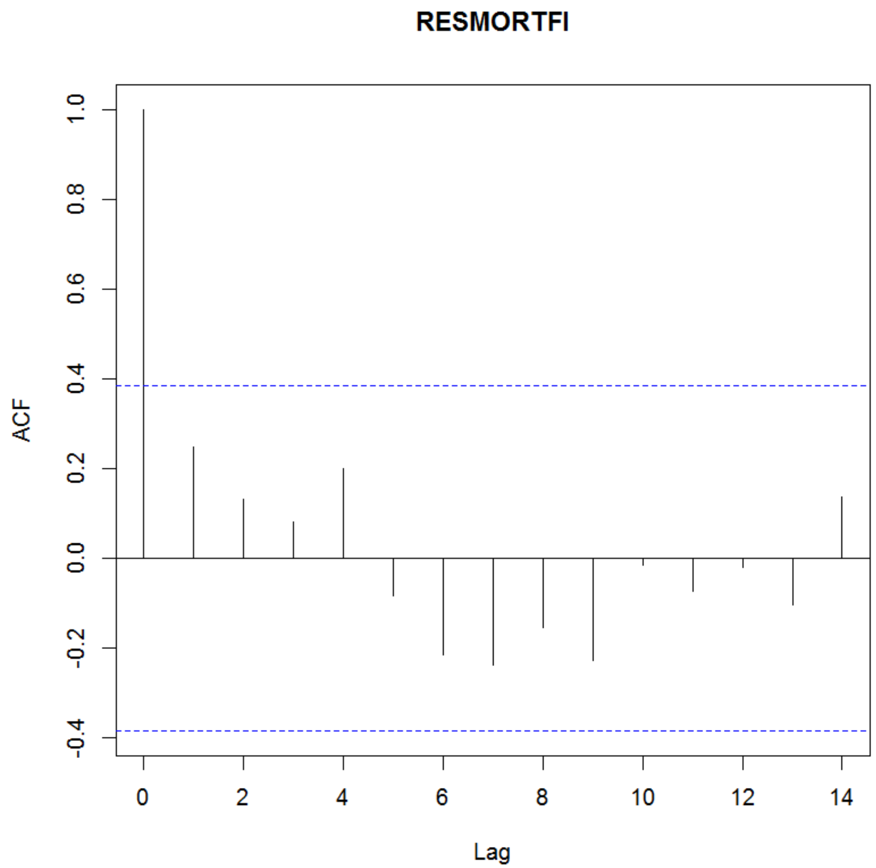
	AIC	BIC		
	1570	1600		
	Estimate	SE	zvalue	pvalue
(Intercept)	-13.564	5.370	-2.526	0.012
K (≤2.8, >2.8)	-1.605	0.258	-6.214	0.000
ETa/ETp	9.902	6.031	1.642	0.101
N-Deposition (≤20, >20)	14.311	5.432	2.635	0.008
ETa/ETp*Ndep	-17.726	6.112	-2.9	0.004

ETa/ETp	Ndep	drought effec	N effect
0.852	0		
0.735	0	0.316	
0.852	1		0.457
0.735	1	2.483	3.588

The number of bark beetles is an intermediate end point and should not be removed

Residual analysis for autocorrelation

Residuals of the model averaged per year -> resmortfijr
acf (resmortfijr)



Epidemiological analysis in forest research

Limitations according to DeVries et al. 2014:

- high variability of forests being examined (non-homogeneous experimental material)
- unbalanced and virtually biased allocation of plots to air pollution levels
- effects of confounding factors and covariates
- absence of a true control

Answers

- high variability can often be explained by measurable factors which can be included in a statistical model
- unbalanced data sets can be handled using statistical models
- confounding factors can be examined and controlled in statistical models
- A «true control» is not needed if there are varying levels of the exposure.
- Addition experiments in the field do not have a better control than epidemiology especially when the dose-response curve is not linear

Epidemiological analysis in forest research

Limitations according to DeVries et al. 2014 (contd.):

- potential inadequacy of sampling numbers for proper statistical analysis and interpretation
- co-linearity between multi data
- constraints in data availability
- nature of the available response indicators

Answers

We do have:

- large datasets with a number of replicates per factor analyzed
- pollution maps with high spatial resolution
- all data in one database
- various quantitative response indicators

Epidemiological analysis in forest research

- Epidemiology in forest research needs not only high quality input data of a dependent variable but also high quality pollution maps (both in time and space) or local pollution data for a large number of plots
- Epidemiology cannot prove causality but it can provide strong indications for causality. Causality can be fully established only by experiments.
- Epidemiology can disentangle and quantify the contributions of different predictor variables to an overall effect (e.g. growth)
- According to a rule of thumb, in epidemiological analysis 10x more sites are needed than the number of explanatory/confounding variables considered
- Epidemiological findings can generate hypotheses deserving further study in experiments
- The results from epidemiology are essential for modelling growth under different climate change scenarios