Spatial Modeling of grub density of the may beetle (forest cockchafer: *Melolontha hippocastani)* in the 'Hessischen Ried' area

Matthias Schmidt und Rainer Hurling

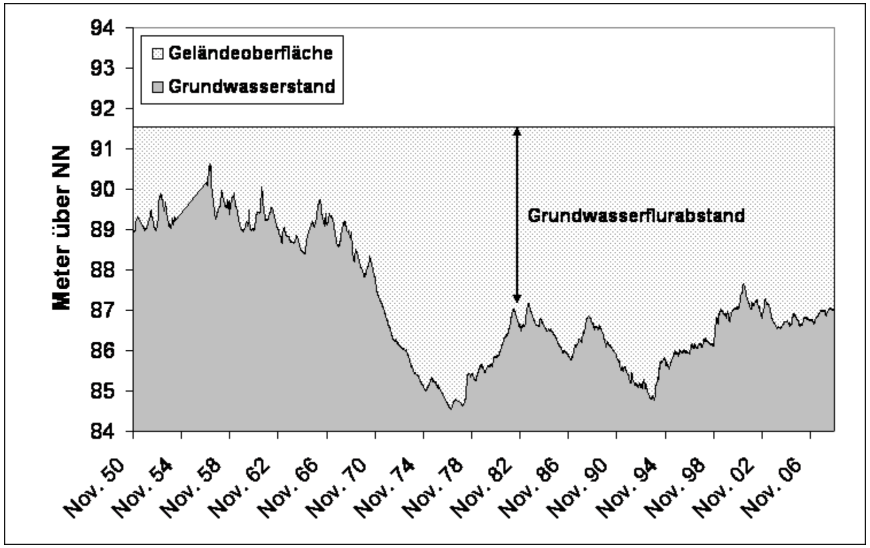
Motivation

• Since several decades the forests in the ‘Hessisch Ried’ area are subject to high abiotic and biotic stress factors. Subsequently in some areas the stands show an extensive decay.

• Temperature$_{veg}$ = 16,6 °C, Precipitation$_{veg}$ = 371 mm, Index of aridity = 38.4

• **Extensive drawdown of ground water**, fragmentation of forests, emission of pollutants (e.g. kerosin by the nearby Frankfurt airport, nitrogen emission)

• Outbreaks of **may beetle**, gipsy moth, bark beetle



Motivation

Motivation

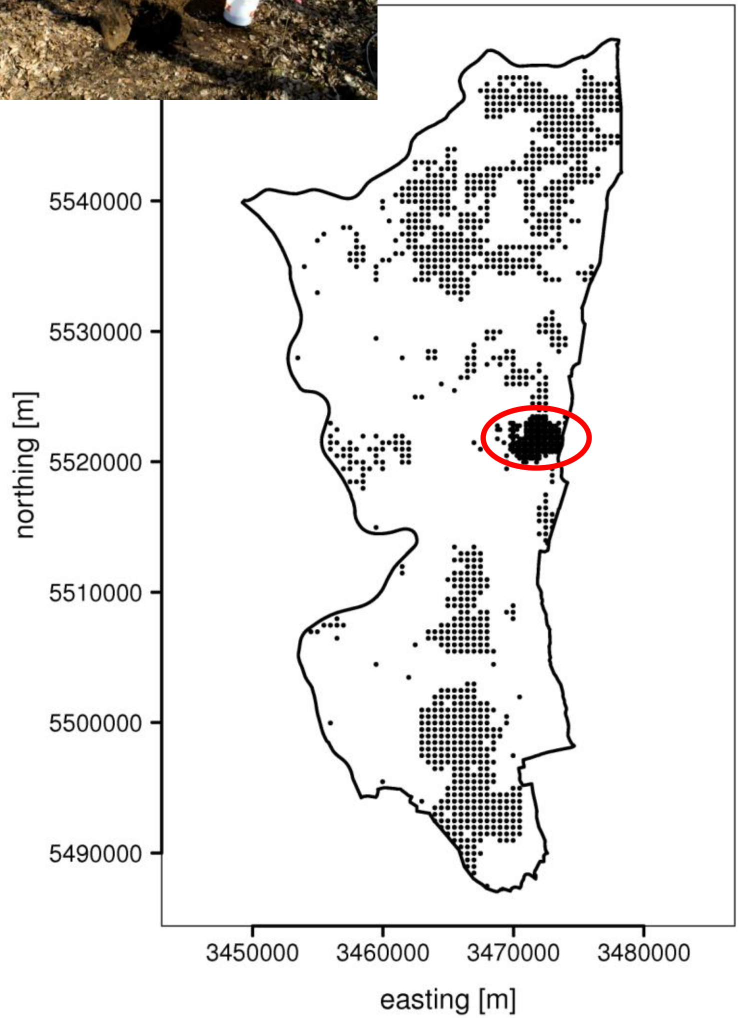
• **Biology**: The may beetle lives usually 3 years as grub below ground (including moulting), at the end of summer of the third year it cocoons and the adult beetle hatches out from the ground in May of the 4 years. An extensively ‘mature grub’ takes place, followed by copulation and death of the males, several egg depositions in the ground and death of the females. There are several tribes in the Rhine valley that follow their own cycle.

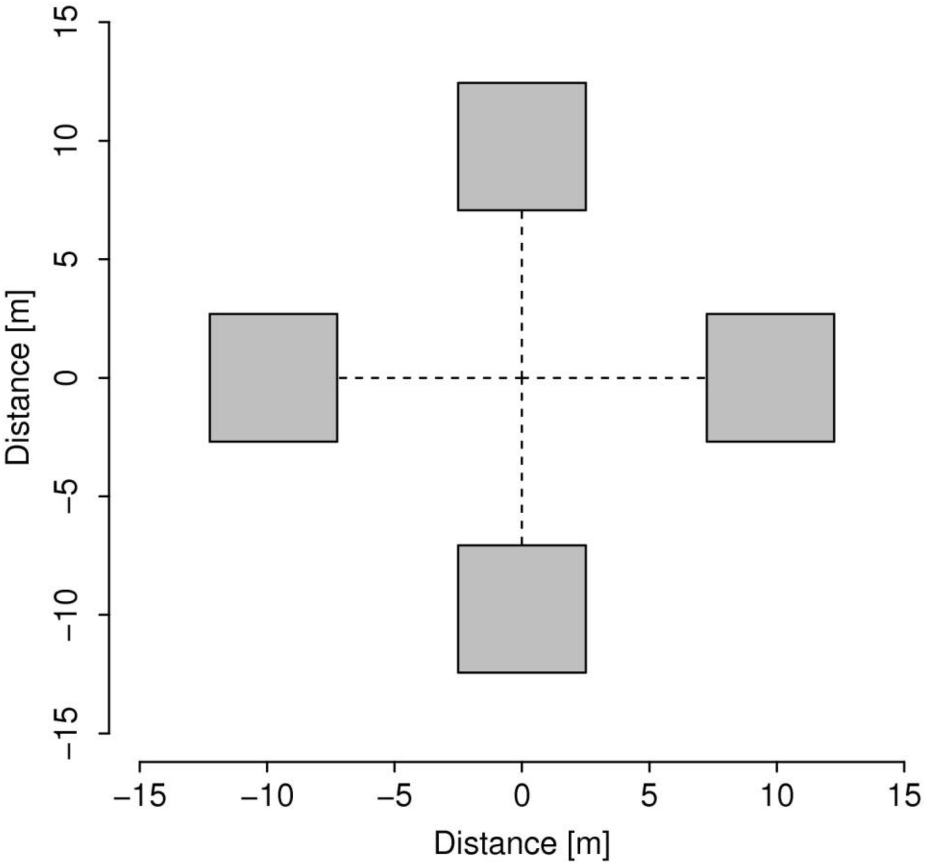
• **Spatial prognosis** of grub density (E_3) to support decision making in practical forestry concerning regeneration planning (planting, harrowing), (airborne) operations of forest protection applying chemical and biological pesticides against the adult beetle.

• **Tests of hypothesis:** Effect of distance to ground water table? Effect of soil type? Effect of forest stand type?…..on grub density. Quantifying the spatial pattern of grub density.

• **Method**: Generalized additive regression models (mgcv, gamlss)

Data base: grub sampling design



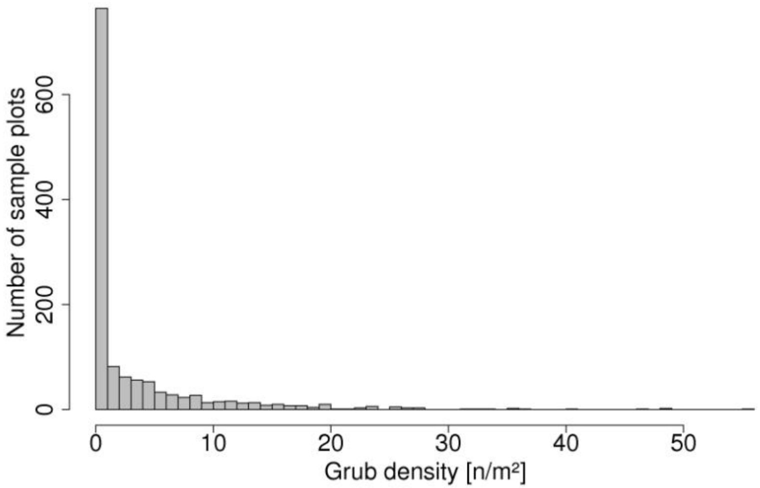


Data base: grub inventory 2009

Grubs E_3/m^2	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**>10**
Samples=1276	664	100	82	62	56	53	33	28	23	27	13	135
%	52	8	6	5	4	4	3	2	2	2	1	11

	Min.	**1^{st} Quartil**	**Median**	**Mean**	**3^{rd} Quartil**	**Max.**
Distance to GWT 2007 [m]	35.16	11.45	5.20	7.56	3.24	0.29
Clay thickness [%]	0.000	0.000	0.000	2.749	0.000	100.000

Number of samples with massive clay layer = 55



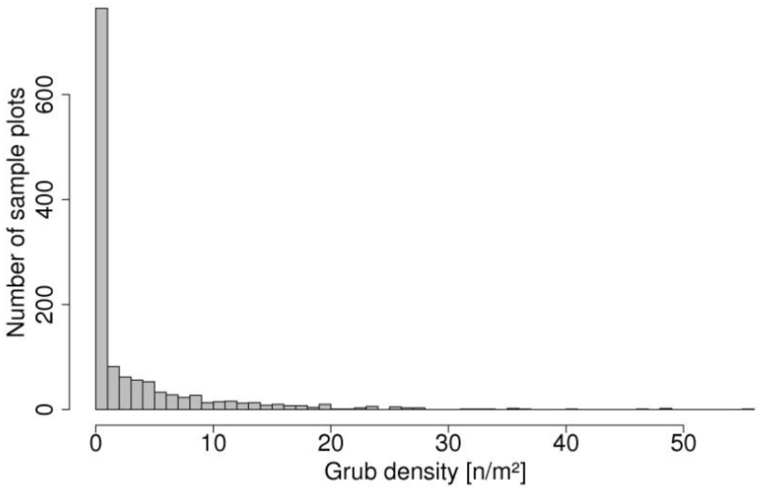
Data base: grub inventory 2009

• The grub densities are count data

• Count data usually show certain properties: integer values, positive data range, significant skewness to the right, high proportion of zeros

• Standard regression models are not appropriate

• Distribution assumption: Poisson, zero-inflated Poisson, negative binomial



Generalized additive models

$$g(\lambda_i)=\beta_0+f_1(DWT_i)+f_2(CTH_i)+f_3(east_i,north_i) \qquad [1]$$

GD_i		**Grub density at sample plot *i* [n/m²]**
		$GD_i \mid x_i \sim Poisson(\lambda_i)$ **with** $E(y_i \mid x_i) = \lambda_i$ **and** $Var(GD_i \mid x_i) = \lambda_i$; $GD_i = 0, 1, 2, \ldots.$
DWT_i	:	**Regionalized distance to GWT in October 2007 at sample plot *i* [m]**
CTH_i	:	**Regionalized clay thickness at sample plot *i* [%]**
$east_i$, $north_i$	:	**Easting and northing of sample plot *i* using the Gauß-Krüger-System referring to the 3. meridian**
f_1, f_2	:	**1-dimensional smoothing functions (penalized thin plate basis regressions splines)**
f_3	:	**2-dimensional smoothing function (penalized thin plate basis regressions spline)**

Distribution assumptions:

Poisson,
zero inflated Poisson,
negative binomial

library(mgcv) Wood

library(gamlss) Rigby & Stasinopoulos

library(AER) Kleiber and Zeileis

Poisson distribution

	Model	AIC	Dispersion Parameter
$g(\lambda_i)= \beta_0+f_1(DWT_i)+f_2(CTH_i)$	[2.0]	10692.3	9.31***
$g(\lambda_i)= \beta_0+f_1(DWT_i)+f_2(CTH_i)+f_3(east_i,north_i)$,			
edf for $f_3(east_i,north_i)$= 97.505	[2.1]	5501.2	2.89***
edf for $f_3(east_i,north_i)$= 318.15	[2.2]	4749.9	1.90***

zero-inflated Poisson distribution

	Modell	AIC	Mixture Parameter
$g_1(\lambda_i)=\beta_{01}+f_{11}(DWT_i)+f_{21}(CTH_i)$	[3.0]	7075.9	$g_2(\omega_i) = \beta_{02}$
$g_1(\lambda_i)=\beta_{01}+f_{11}(DWT_i)+f_{21}(CTH_i)+f_{31}(east_i,north_i)$			
edf for $f_{31}(east_i,north_i)$ = 90.622	[3.1]	5038.9	$g_2(\omega_i)=\beta_{02}+f_{12}(DWT_i)$

negative binomial distribution

	Model	AIC	Dispersion Parameter
$g_1(\mu_i)=\beta_{01}+f_{11}(DWT_i)+f_{21}(CTH_i)$	[4.0]	5170.9	$g_2(1/\phi) =\beta_{02}$; ϕ= 0.323
$g_1(\mu_i)=\beta_{01}+f_{11}(DWT_i)+f_{21}(CTH_i)+f_{31}(east_i,north_i)$,			
edf for $f_3(east_i,north_i)$= 87.553	[4.1]	4192.9	$g_2(1/\phi) =\beta_{02}$; ϕ=1.551
edf for $f_3(east_i,north_i)$= 117.346	[4.2]	4161.3	$g_2(1/\phi) =\beta_{02}$; ϕ= 1.711
edf for $f_3(east_i,north_i)$= 125.152	[4.3]	4164.6	$g_2(1/\phi) =\beta_{02}$; ϕ= 1.734

Validation / quantile residuals

- Assumption: the observations $y_1, \ldots, y_n$ are independently distributed with probability density function (pdf) $f(y_i, \mu_i, \phi)$.
- $F(y_i, \mu_i, \phi)$ denotes the corresponding (cumulative) distribution function (cdf).

- the quantile residual r_i corresponding to observation y_i is defined as:

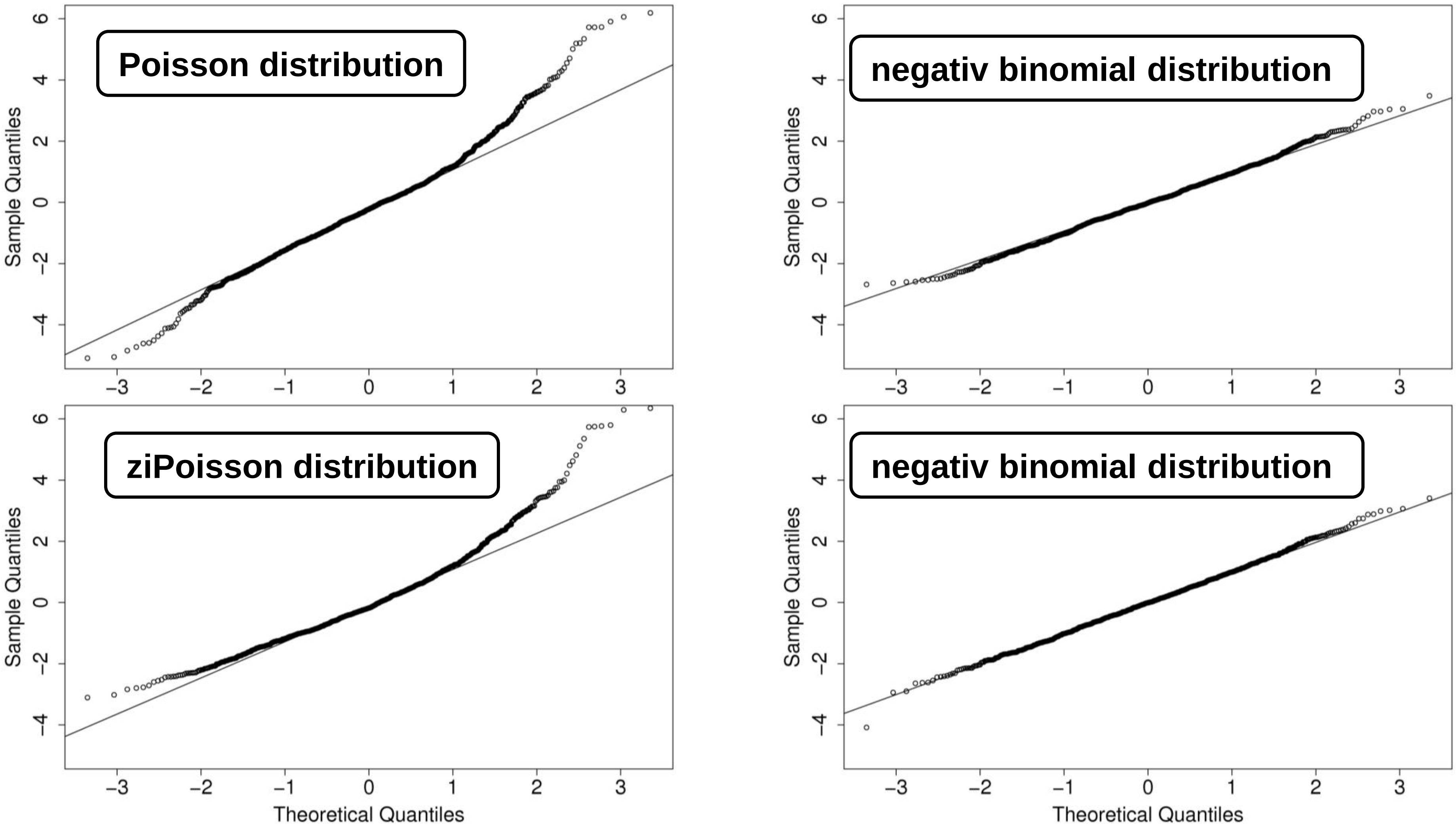
$$r_i = \Phi^{-1}\{ F(y_i; \hat{\mu}_i, \hat{\phi}) \}$$

with Φ and Φ^{-1} denoting the distribution and quantile function of the standard normal distribution.

$r_i = \sim N(0,1)$, if $f(y_i, \mu_i, \phi)$ is indeed the correct model for the observations

Dunn and Smyth, 1996

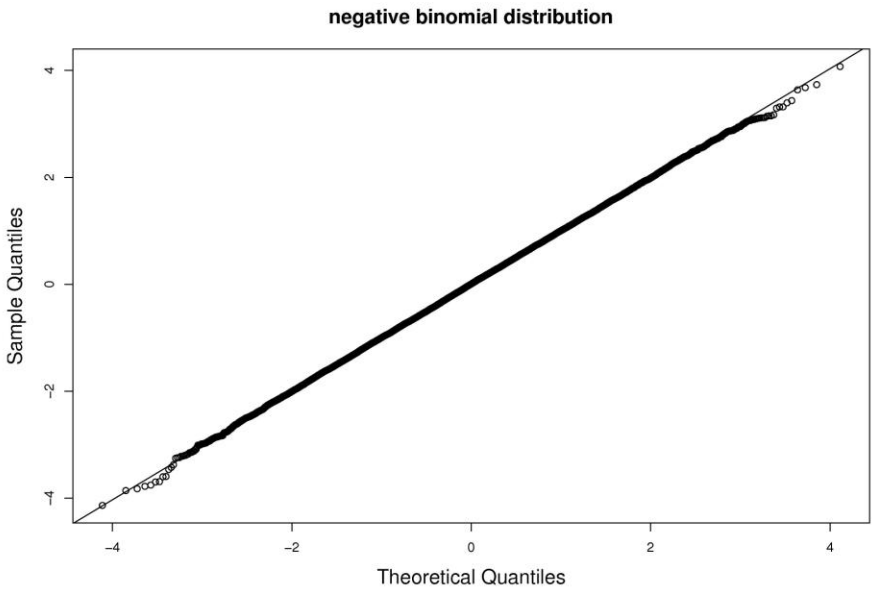
Validation quantile residuals



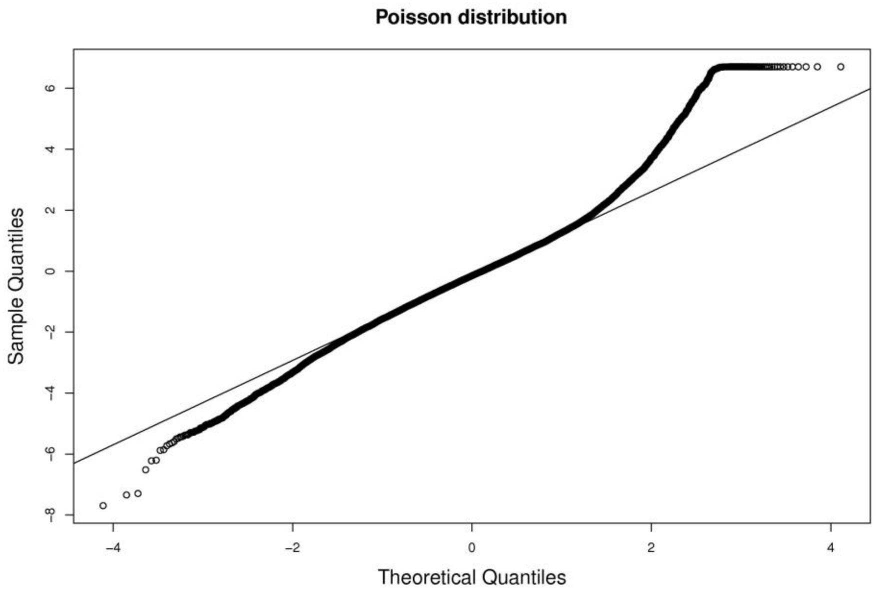
Validation quantile residuals

• Draw a random sample from the fitted negative binomial model of 10 times the original sample size applying the original predictors and fitting a Poisson and a negative binomial model to the data.

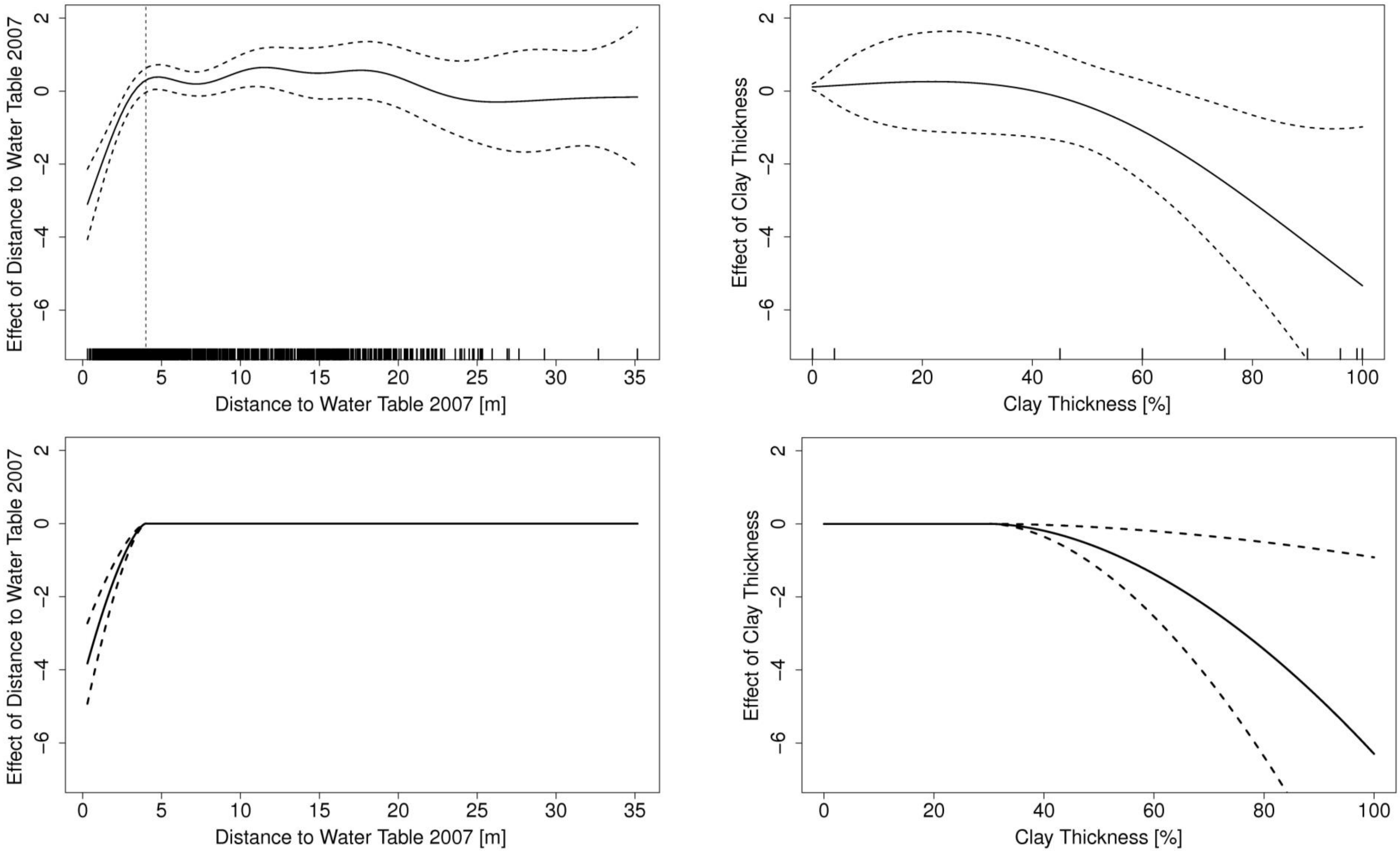
negativ binomial distribution



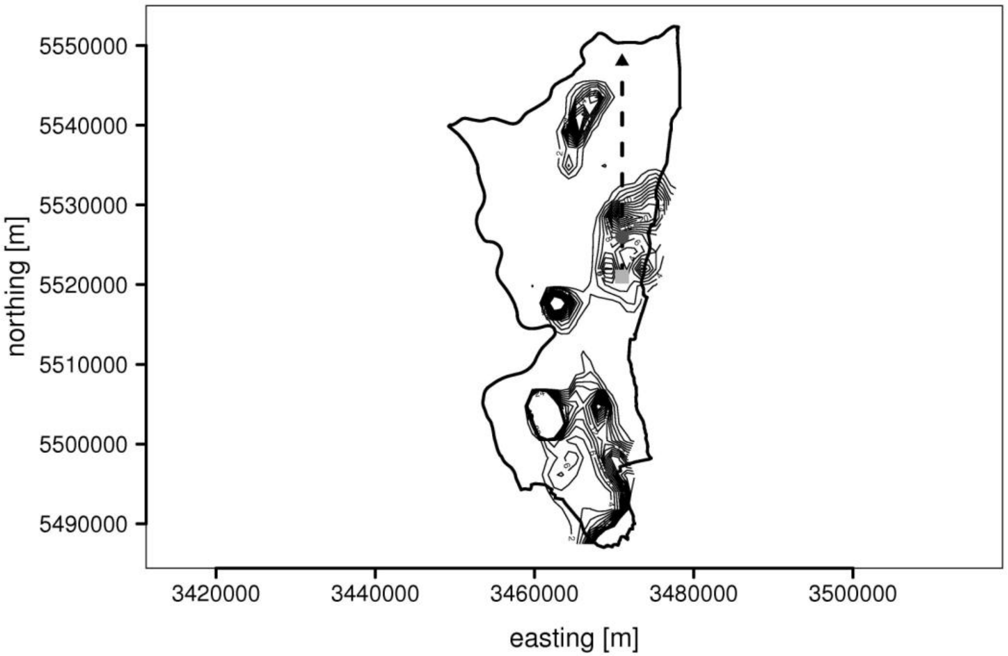
Poisson distribution

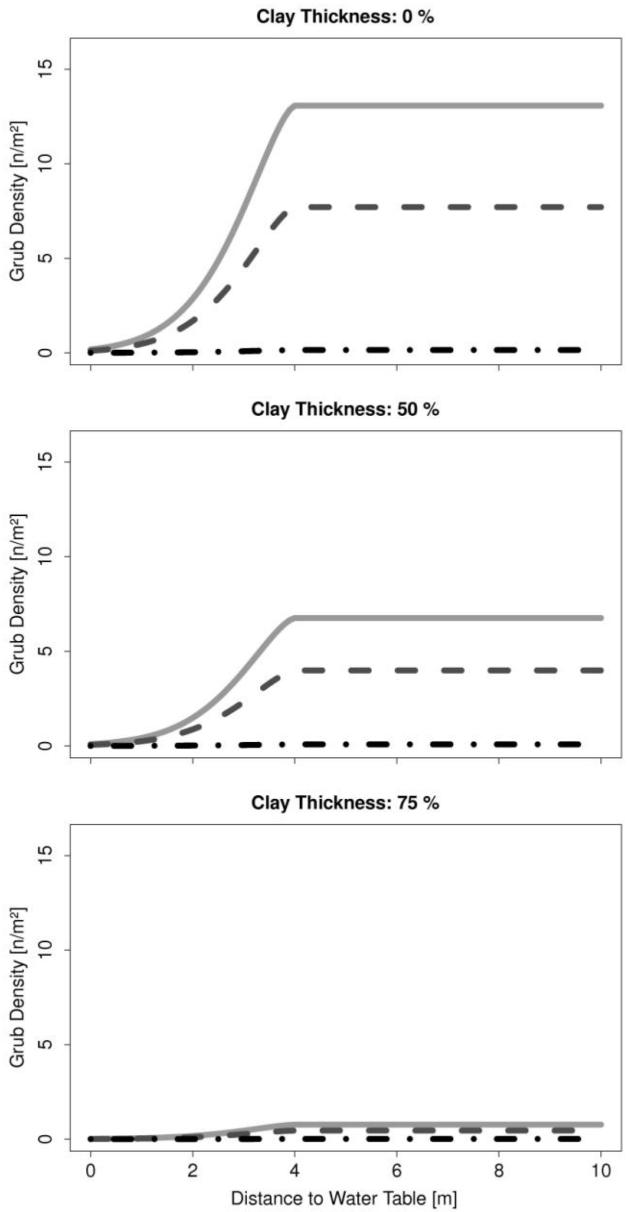


Results model effects

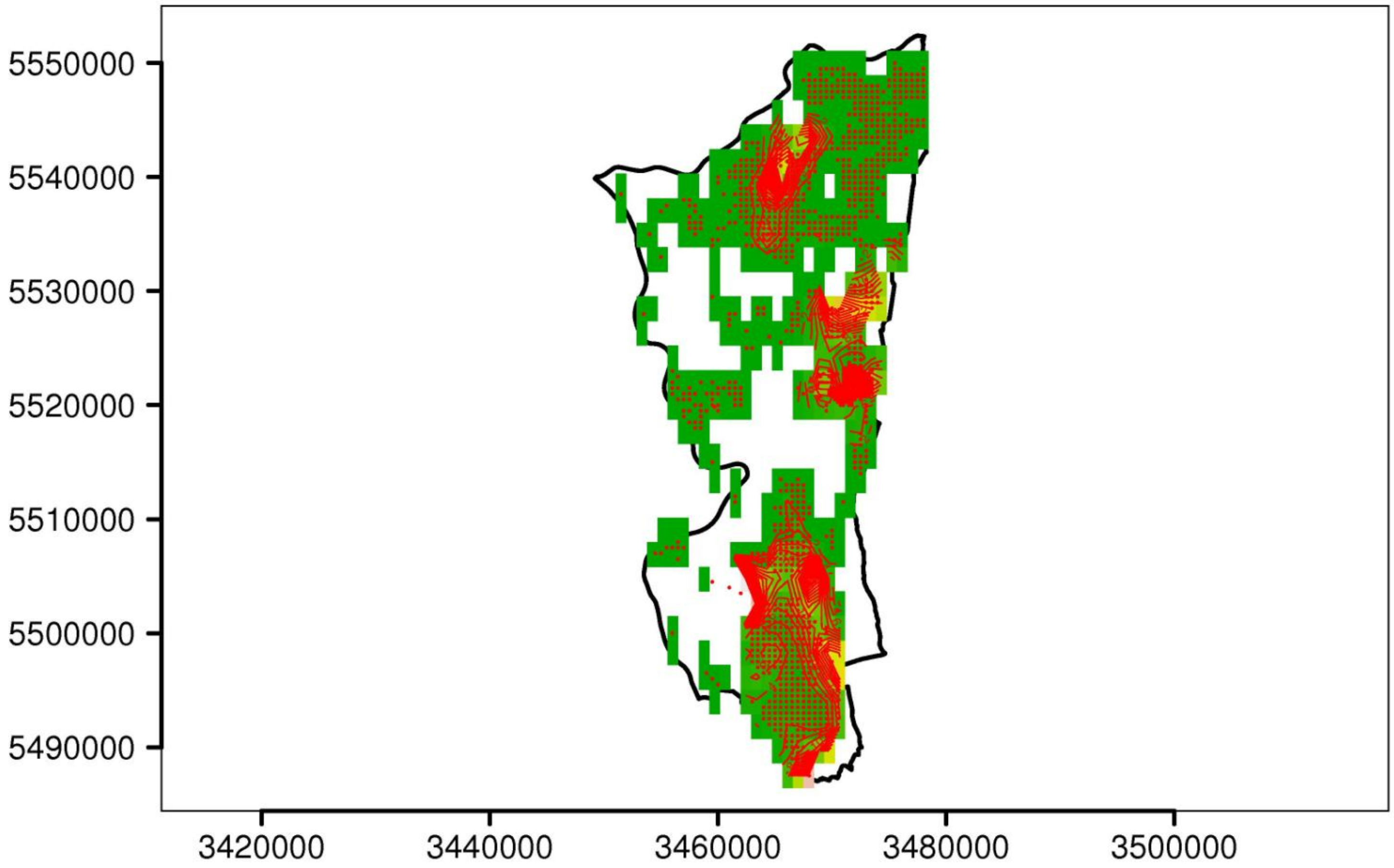


Model predictions: Effect of distance to GWT





Effect of the spatial trend function



Conclusions

• Generalized additive regression models (~negative binomial distribution) are a well suited approach for modeling count data of may beetle grubs .

• The most important model effect on grub density results from **geographic location.**

• The 'causal' predictors **distance to GWT** and **clay thickness** show significant effects on grub density also. If distance to GWT exceeds more than 4 m and **clay thickness** is less than 40% no effects are observed.

• For validation purposes **quantile residuals** should be used.

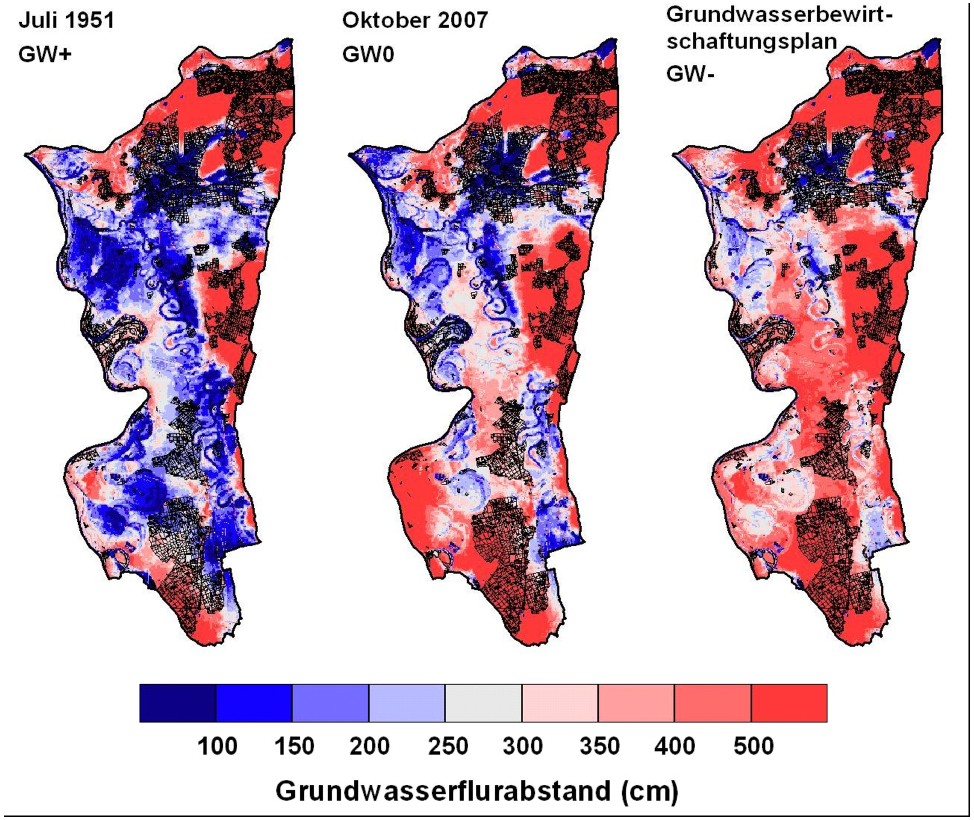
• In 2013 the next grub inventory will take place on the same grid and the model will be parameterized again. The potential changes in the spatial pattern of grub density over time and the effects of applied pesticides will be tested.

• An additional inventory of pest-damages should be used to quantify the relation between E_3/m^3 and corresponding damages.

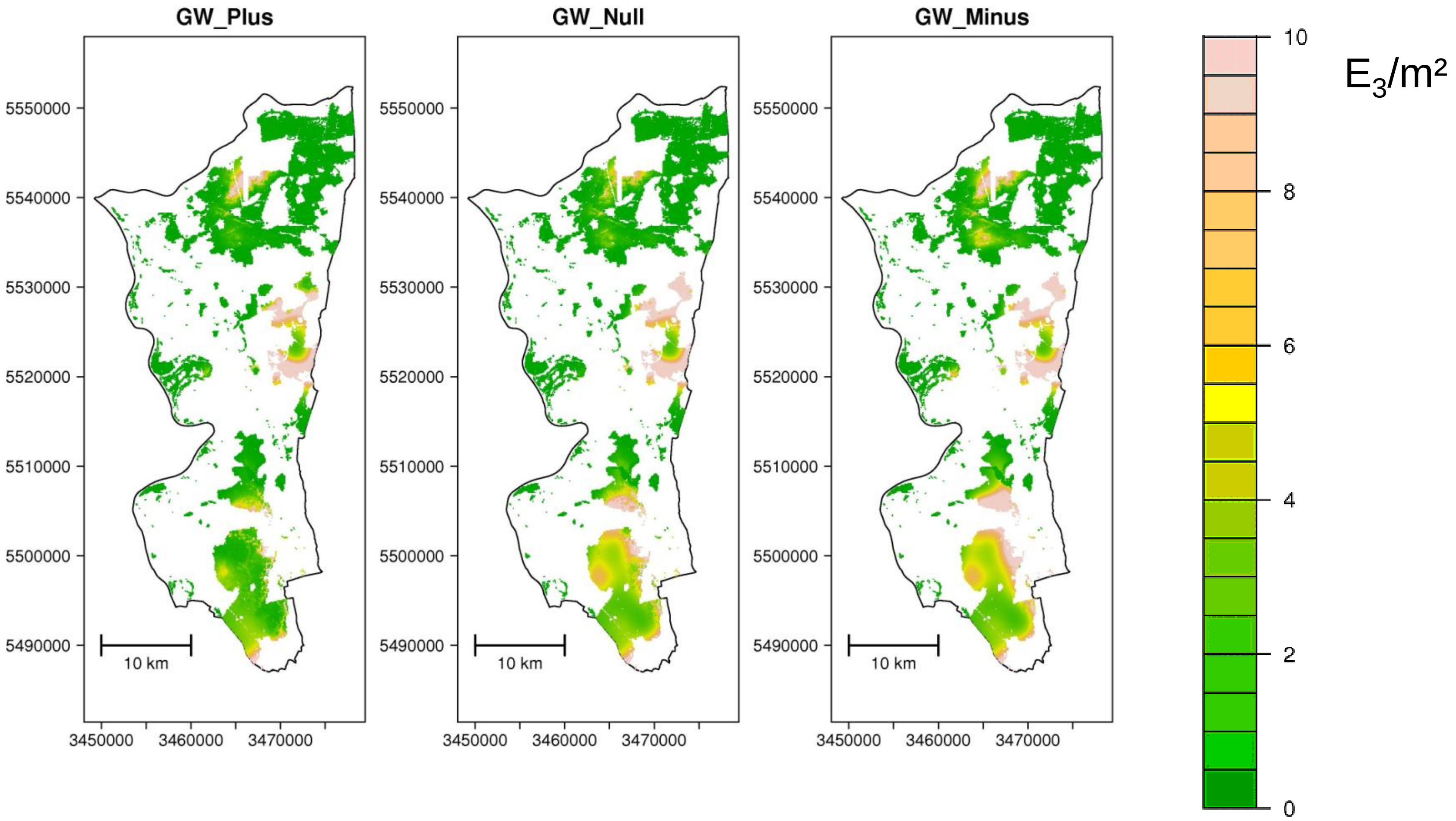
Chemical and biological operations of forest protection both require a good information base!

Busch (1865)

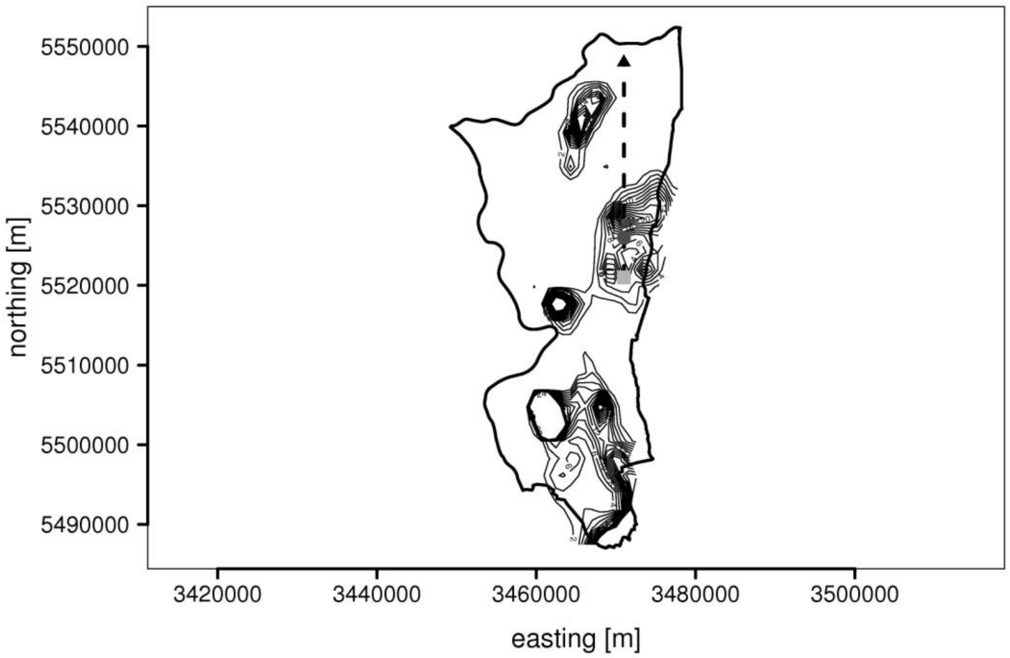
Simulation der Effekte unterschiedlicher Grundwasserbewirtschaftungspläne auf die Engerlingsdichte

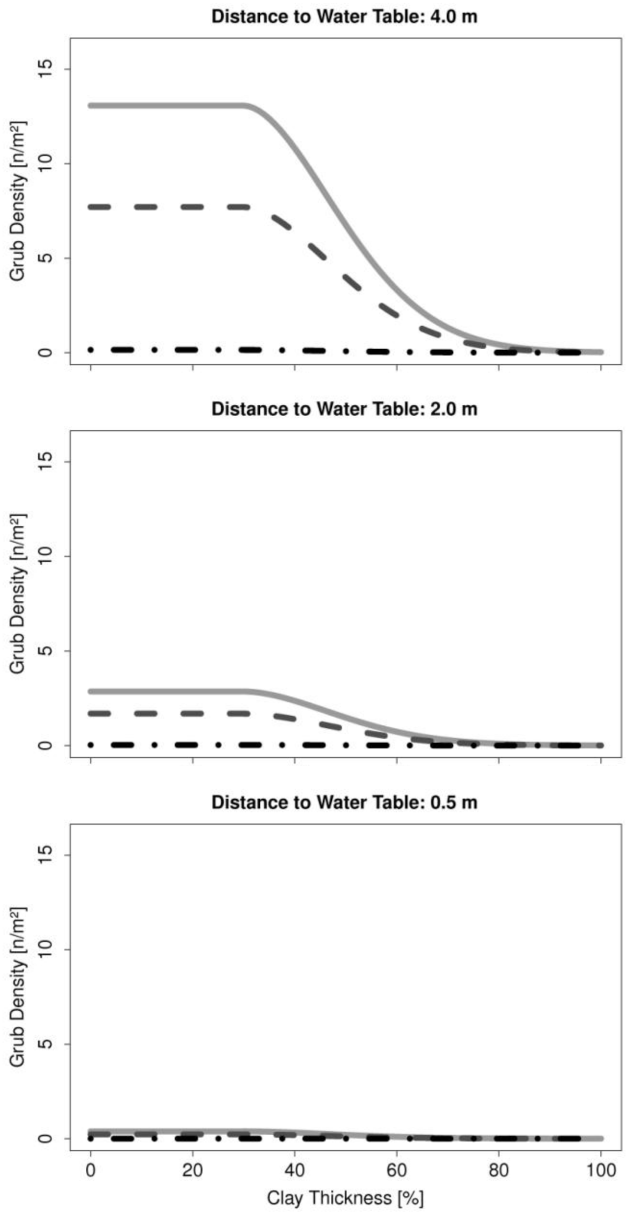


Simulation der Effekte unterschiedlicher Grundwasserbewirtschaftungspläne auf die Engerlingsdichte



Model predictions: Effect of clay thickness





Poisson distribution

$$g(\lambda_i) = \eta_i = x'_i\beta = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_k x_{ik} \qquad [1]$$

or

$$\lambda_i = \exp(\eta_i) = \exp(x_i\beta) = \exp(\beta_0) \cdot \exp(\beta_1 x_{i1}) \cdot \ldots \cdot \exp(\beta_k x_{ik})$$

mit $g(.)$: Verknüpfungsfunktion: natürlicher Logarithmus

und $y_i|x_i \sim Poisson(\lambda_i)$ mit $E(y_i|x_i) = \lambda_i$ and Var $(y_i|x_i) = \lambda_i$; $y_i = 0, 1, 2, \ldots$

Wahrscheinlichkeitsdichtefunktion der Poissonverteilung:

$$f(y;\lambda) = \frac{\lambda^y e^{-\lambda}}{y!} \qquad y = 0, 1, 2, \ldots..$$

zero inflated Poisson distribution

$g_1(\lambda_i) = \eta_{1i} = x'_i\beta_1 = \beta_{01} + \beta_{11}x_{i1} + \ldots + \beta_{k1}x_{ik}$ **[3a]**

mit $g_1(.)$ **: Verknüpfungsfunktion: natürlicher Logarithmus**

$g_2(\omega_i) = \eta_{2i} = x'_i\beta_2 = \beta_{02} + \beta_{12}x_{i1} + \ldots + \beta_{k2}x_{ik}$ **[3b]**

und $g_2(.)$ **: Verknüpfungsfunktion: logistisch**

und

$y_i|x_i \sim ZIP(\lambda_i, \omega_i)$ **mit** $E(y_i|x_i) = (1-\omega_i)\lambda_i$ **und** $Var(y_i|x_i) = \lambda_i(1-\omega_i)(1+\lambda_i\omega_i)$ $\quad y_i = 0, 1, 2, \ldots.$

Wahrscheinlichkeitsdichtefunktion der Zero-inflated Poissonverteilung:

$$f(y;\omega,\lambda) = \begin{cases} \omega + (1-\omega)\,e^{-\lambda} & , y = 0 \\ (1-\omega)\,\dfrac{\lambda^y}{y!}e^{-y} & , y = 1,2,\ldots. \end{cases}$$

negativ binomial distribution

$g_1(\mu_i) = \eta_{1i} = x'_i\beta_1 = \beta_{01} + \beta_{11}x_{i1} + \ldots + \beta_{k1}x_{ik}$ [2a]

mit $g_1(.)$: Verknüpfungsfunktion: natürlicher Logarithmus

und $y_i|x_i \sim Poisson(\mu_i)$ with $\mu_i \sim Gamma$(shape=ϕ, scale=$1/\phi$) =

$Gamma$(Mittelwert=1, Varianz=1/ϕ)

und $y_i|x_i \sim NegBin(\mu_i, \phi)$ mit $E(y_i|x_i)=\mu_i$ und $Var(y_i|x_i) = \mu_i + \mu_i^2/\phi$; $y_i = 0, 1, 2, \ldots.$

Wahrscheinlichkeitsdichtefunktion der negativ Binomialverteilung:

$$f(y, \theta, \mu) = \frac{\Gamma(\phi + y)}{\Gamma(\phi)\, y!} \frac{\mu^y \phi^\phi}{(\mu + \phi)^{\phi + y}}$$ **, mit Γ Gammafunktion.** $y = 0, 1, 2, \ldots.$

$g_2(1/\phi_i) = \eta_{2i} = x'_i\beta_2 = \beta_{02} + \beta_{12}x_{i1} + \ldots + \beta_{k2}x_{ik}$ [2b]

mit $g_2(.)$: Verknüpfungdfunktion: natürlicher Logarithmus