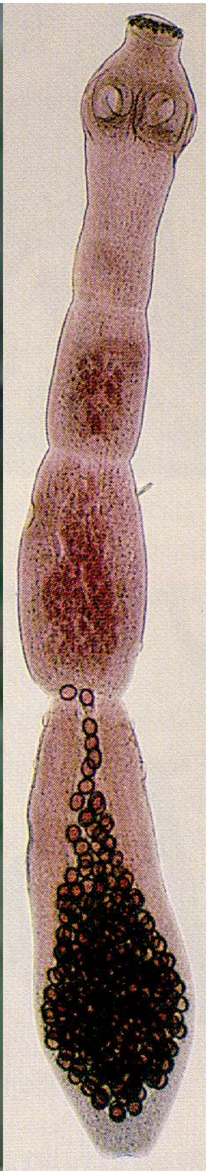
Bayesian modeling of factors potentially influencing the distribution of *Echinococcus multilocularis* in foxes

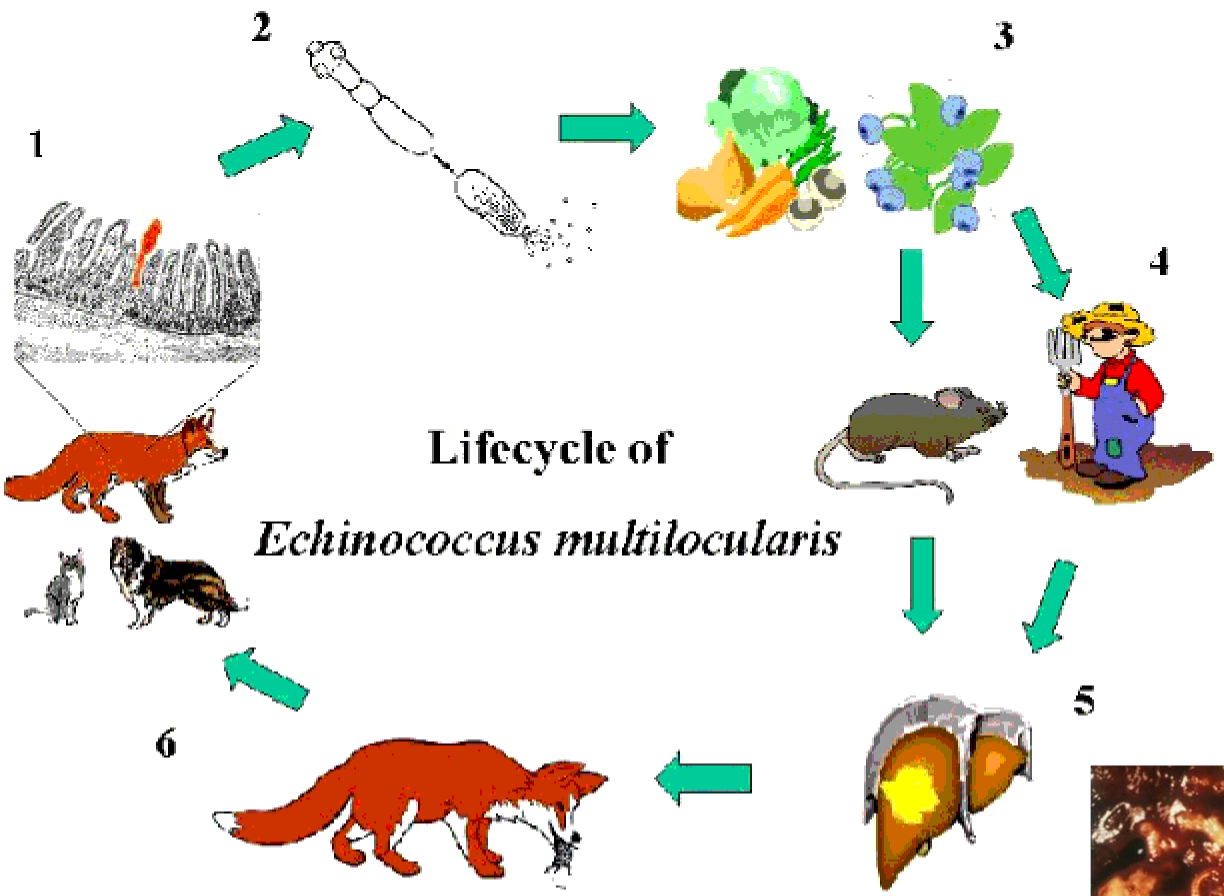
Staubach, C.[1], Schwarz, S.[1], Sutor, A.[1], Hoffmann, L.[2], Tackmann, K.[1], Schmid, V.[3], Conraths, F.J.[1]

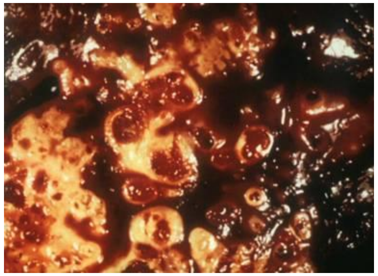
[1]Friedrich-Loeffler-Institut, Federal Research Institute for Animal Health, Seestraße 55, 16868 Wusterhausen, Germany

[2]Thuringian State Authority for Food Safety and Consumer Protection, Tennstedter Str. 8/9, 99947 Bad Langensalza, Germany

[3]Department of Statistics, Ludwig-Maximilians-University, Ludwigstr. 33, 80539 München, Germany







http://www.cdfound.to.it/html/ae1.htm

Alveolar echincoccosis

- Alveolar echinococcosis in humans is caused by the larval stage of *Echinococcus multilocularis*

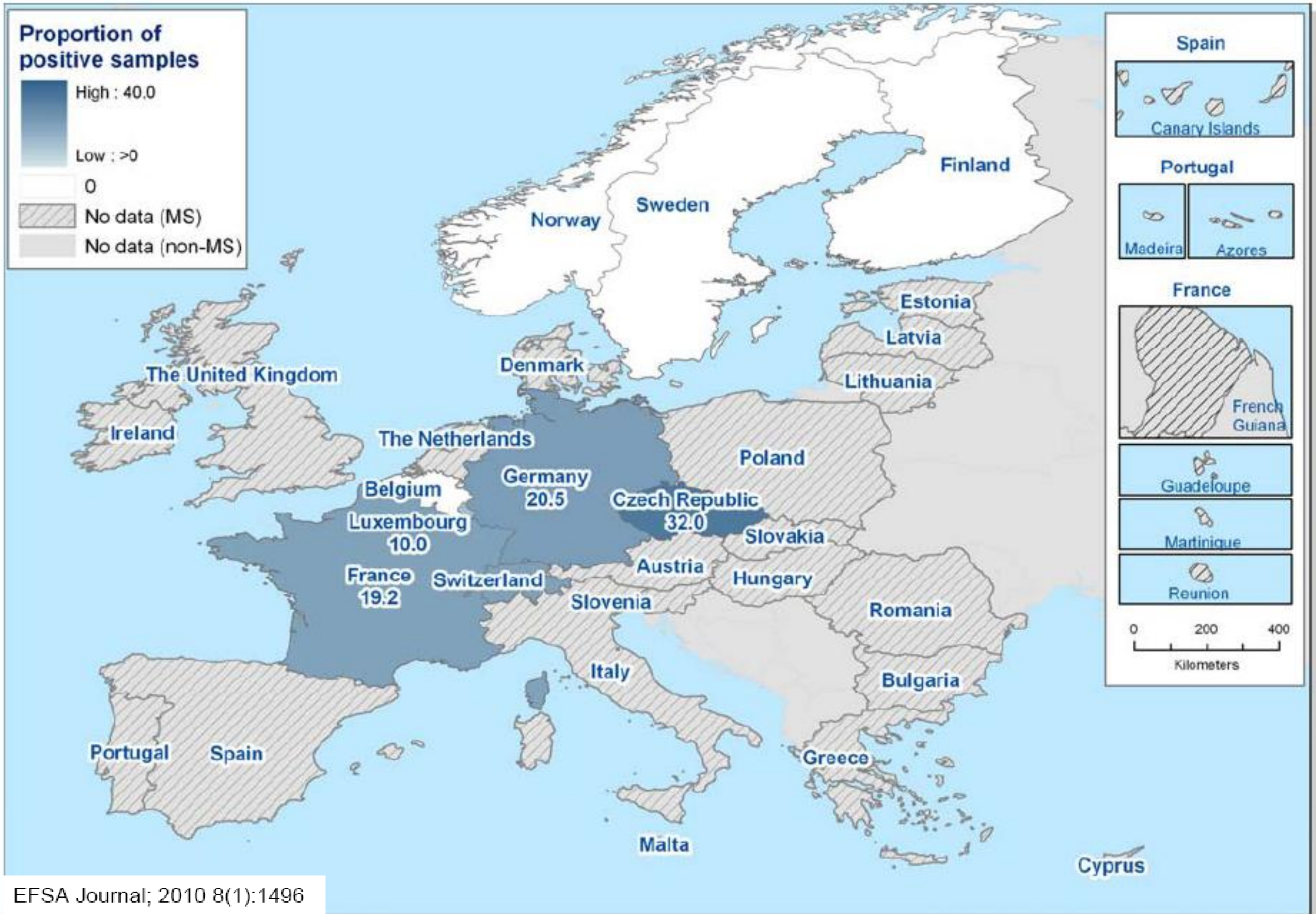
- Obligate 2-host parasitic cycle is predominantly sylvatic

  definitive host: red fox (*Vulpes vulpes*) in Central Europe

  intermediate host: different rodent species

- Alveolar echinococcosis is considered the most dangerous autochthonous parasitic zoonosis in central Europe

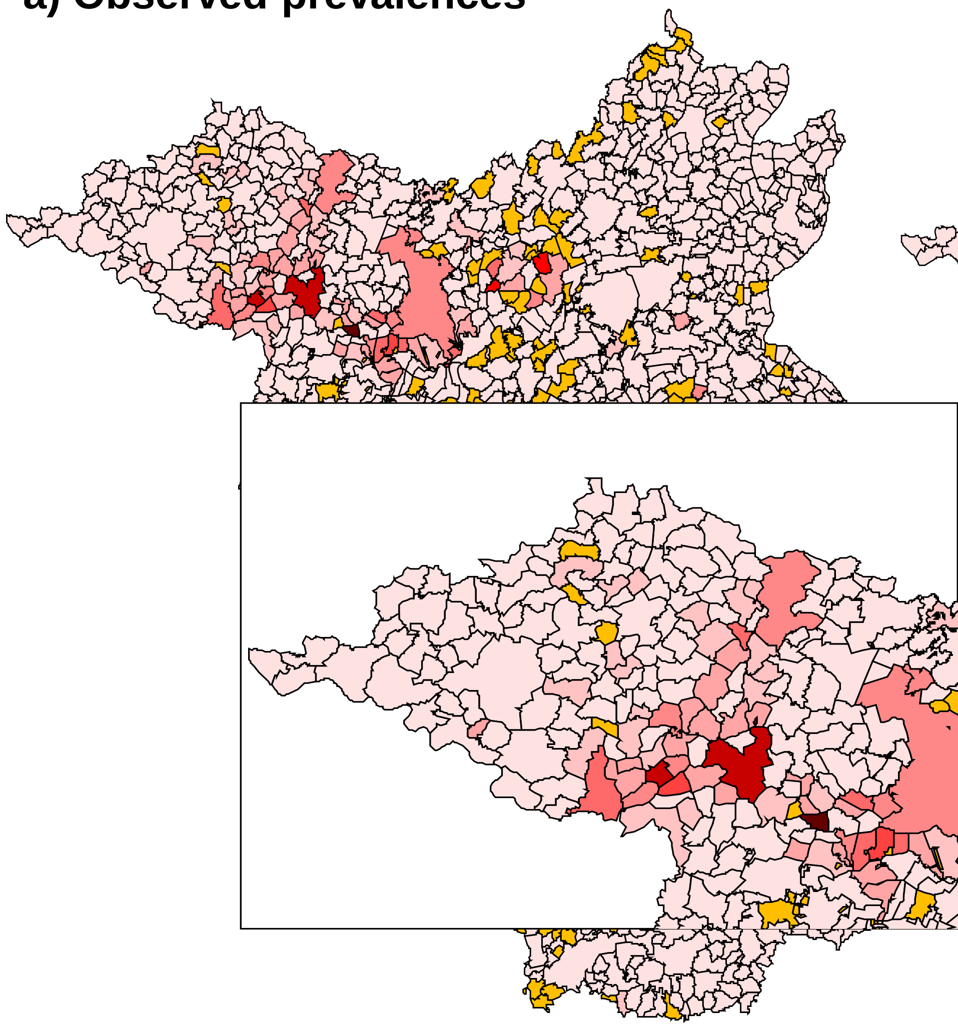
Figure EH4. Findings of* Echinococcus multilocularis *in foxes, 2008



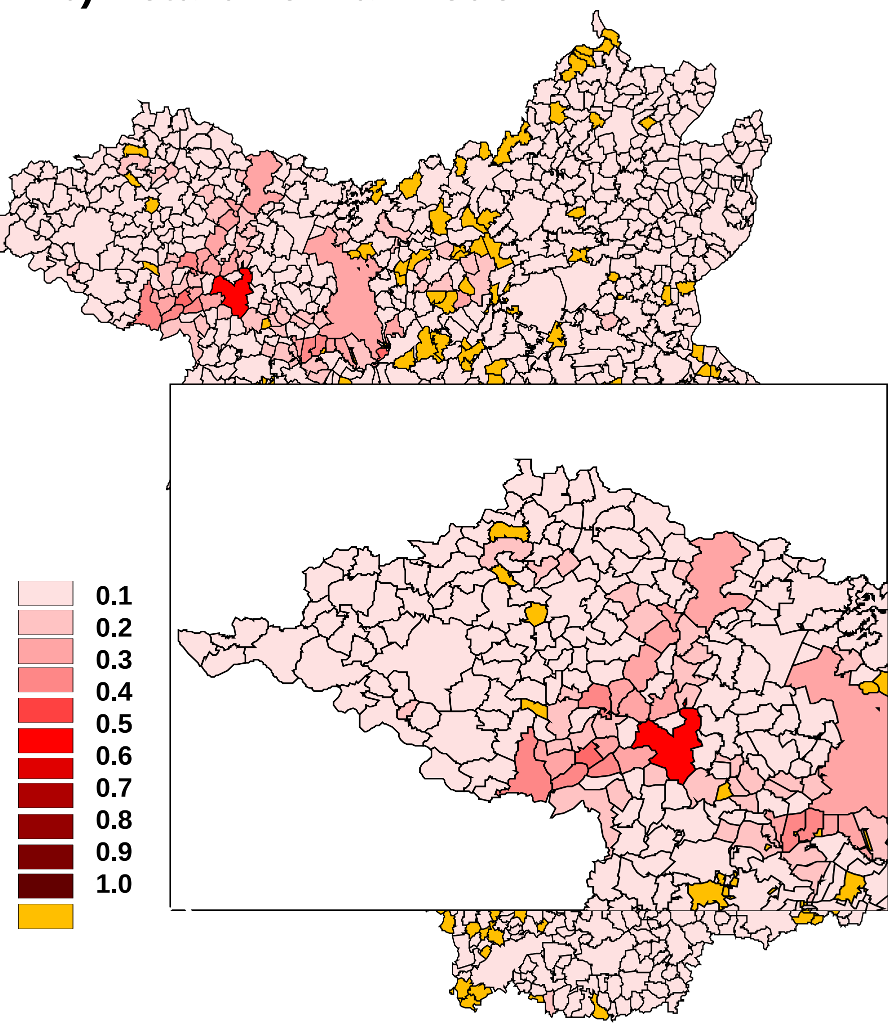
EFSA Journal; 2010 8(1):1496

Period prevalence of *E. multilocularis* 1991-1995

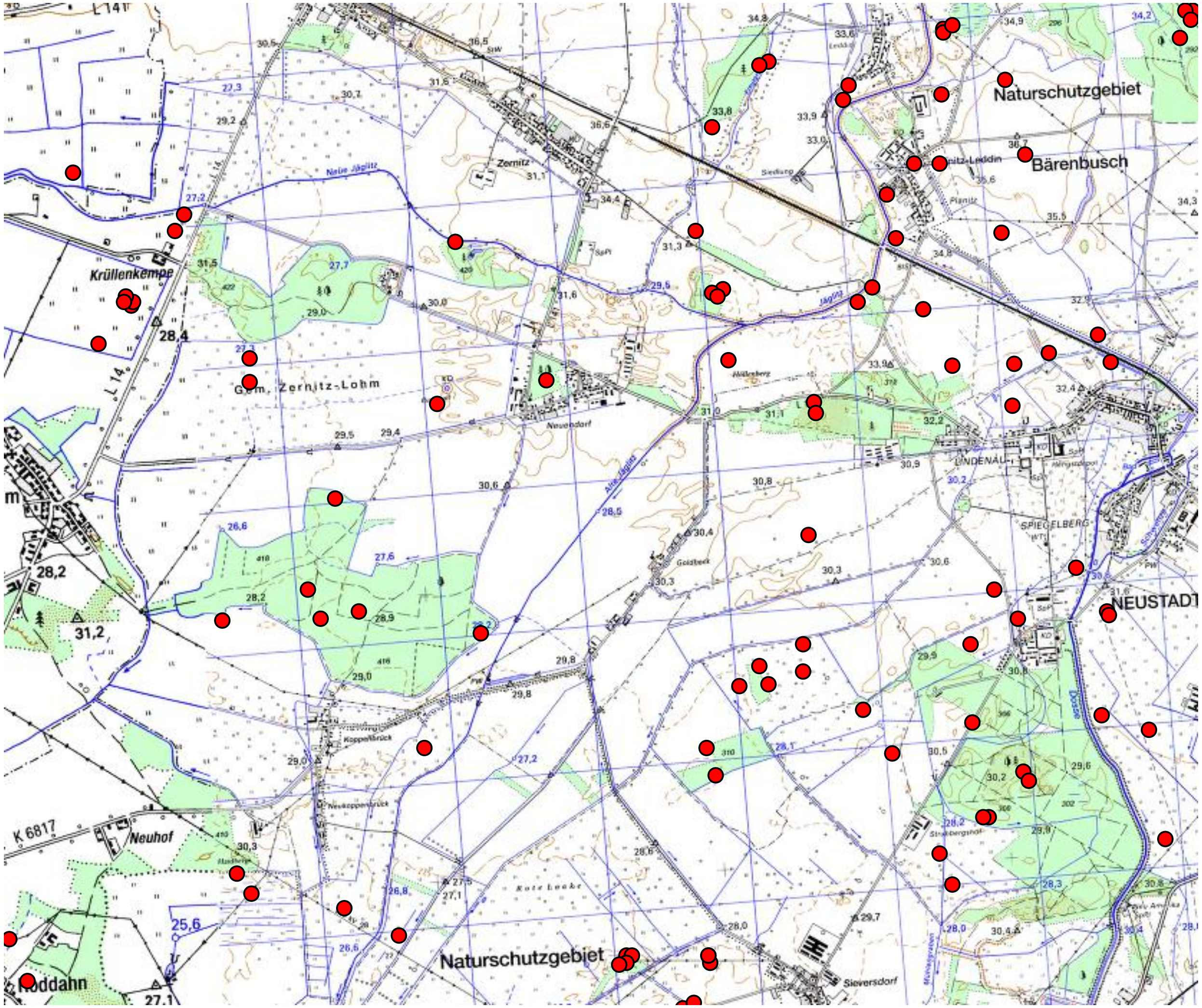
a) Observed prevalences

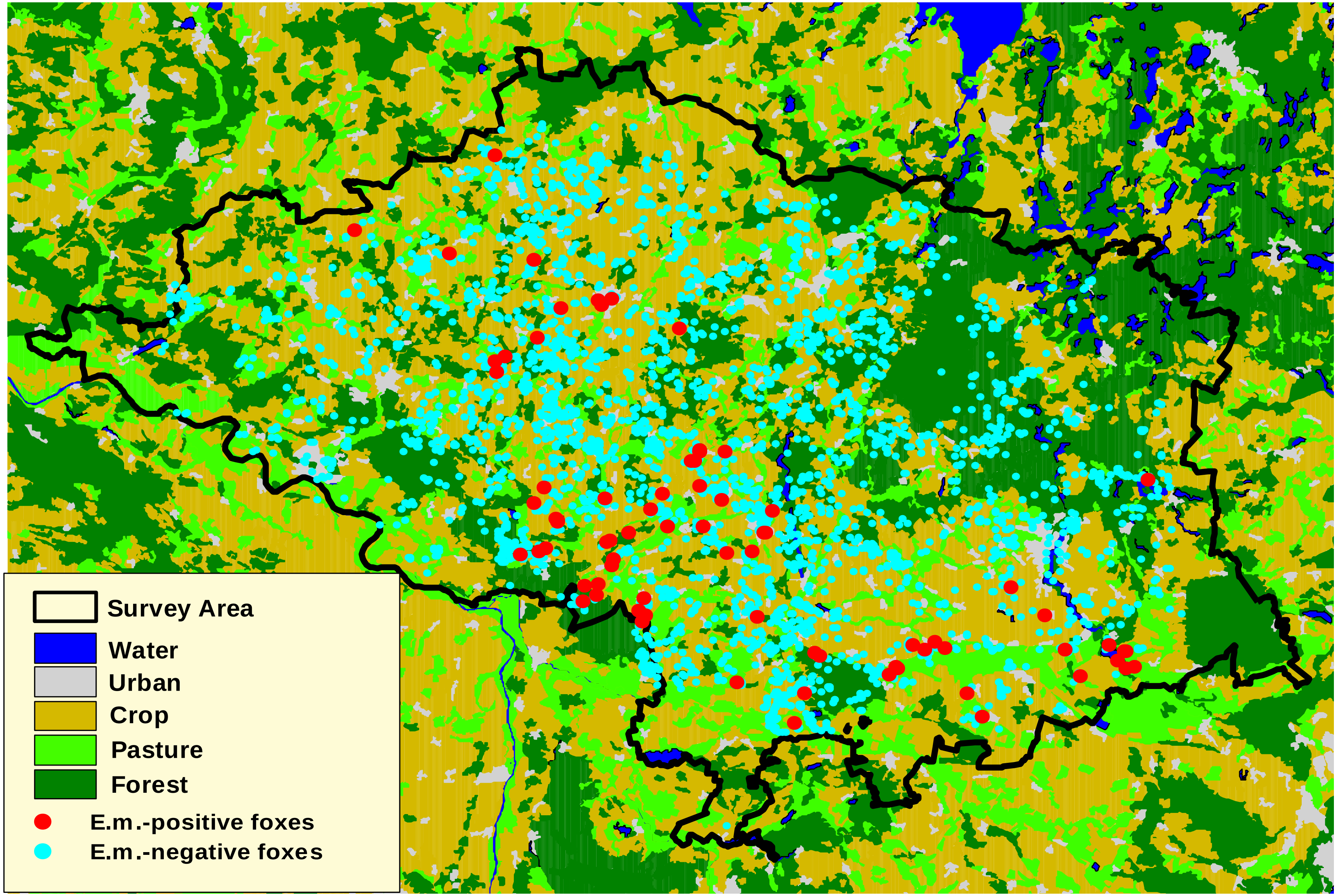


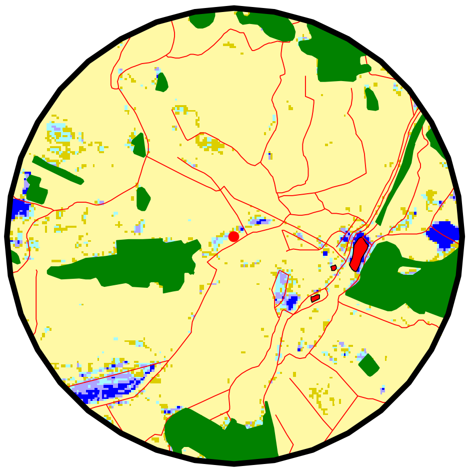
b) Beta-binomial model

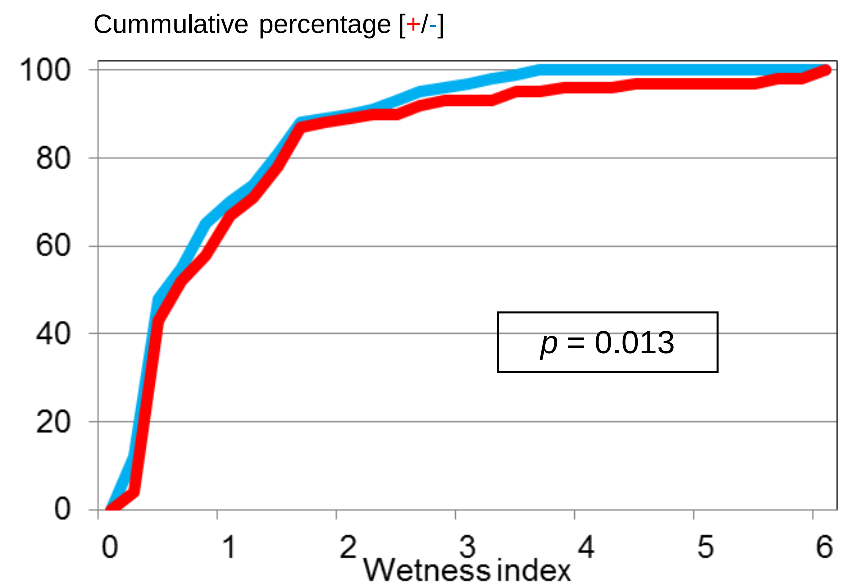


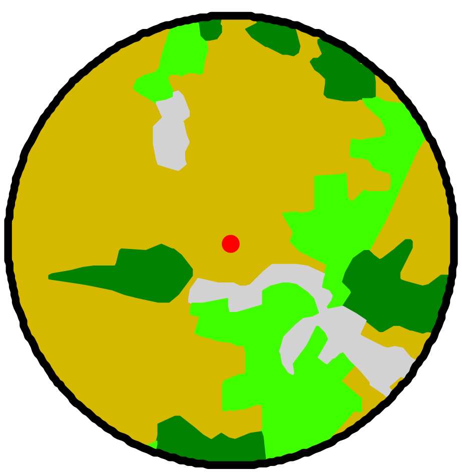
Digitised positions of shot foxes

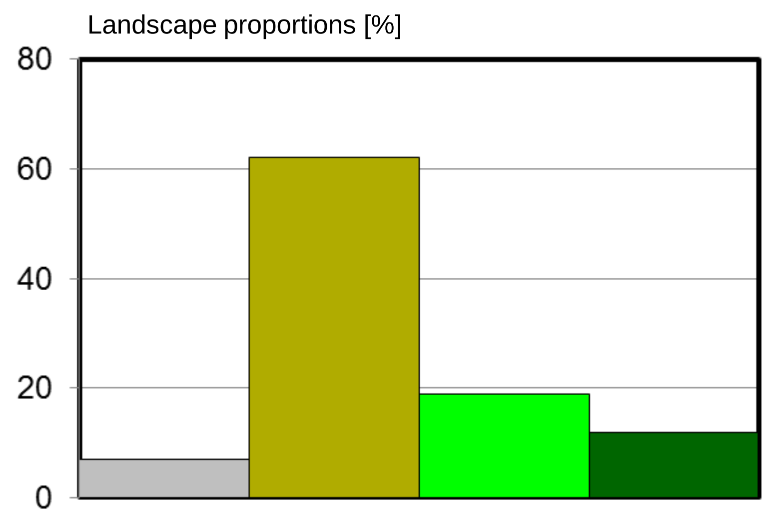



Survey Area
Water
Urban
Crop
Pasture
Forest
E.m.-positive foxes
E.m.-negative foxes



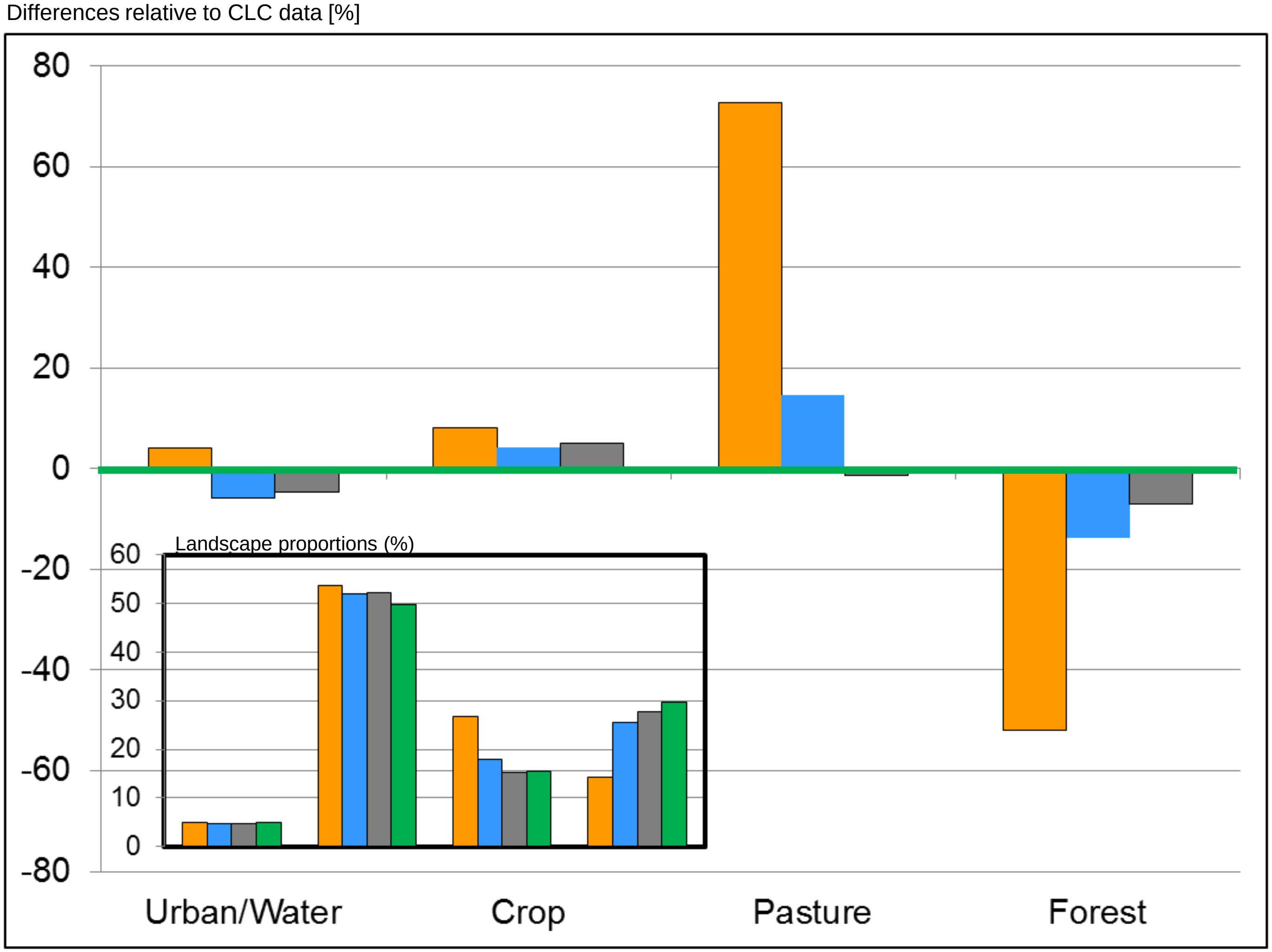

Cummulative percentage [+/-]
100
80
60
40
20
0
0
1
2
3
4
5
6
Wetness index
p = 0.013

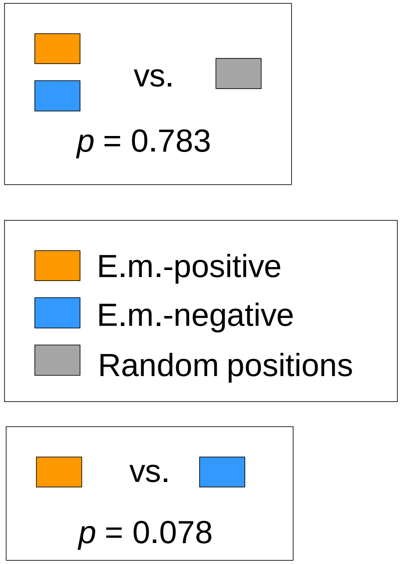



Landscape proportions [%]
80
60
40
20
0

Landscape proportions surrounding each shot fox an relative differences between E.m.-positive and –negative foxes (n = 3,521)

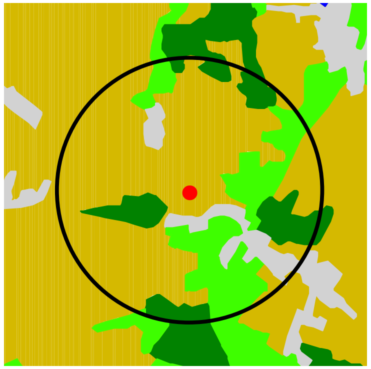
Estimation based on buffer zones of 2.5 km radius

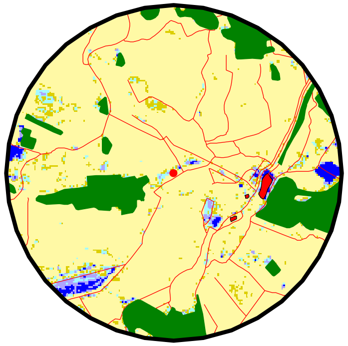


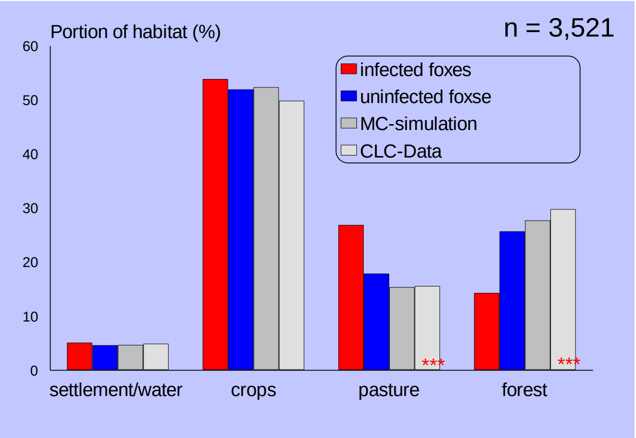


Modeling of factors potentially influencing the distribution

- Hypothesis
  - The oncosphere infectivity can be effectively reduced by elevated temperature and dessication
  - Microclimate and habitat can be suspected as potential risk factors
- Data
  - landscape composition per spatial unit was derived from a high-resolution land-survey vector database and supplemented by a digital elevation model
  - 42,861 foxes were sampled in two Federal States of Germany







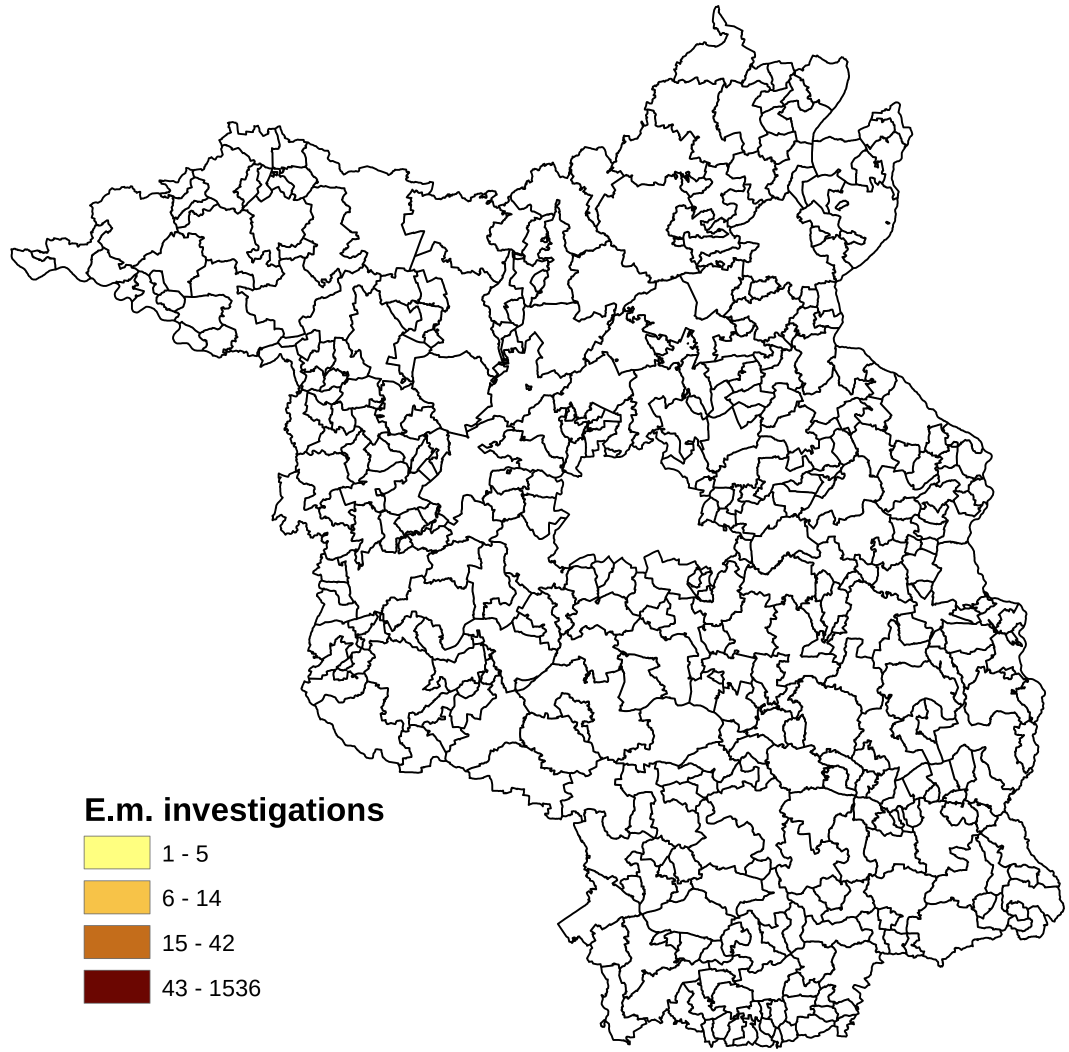
Staubach et al. 2001

Using Hierarchical Bayesian spatial models on municipality level to estimate the parameters of the potential influencing environmental factors

E.m. Data Brandenburg

14,422 foxes
2000 – 2012
- Location (municipality)

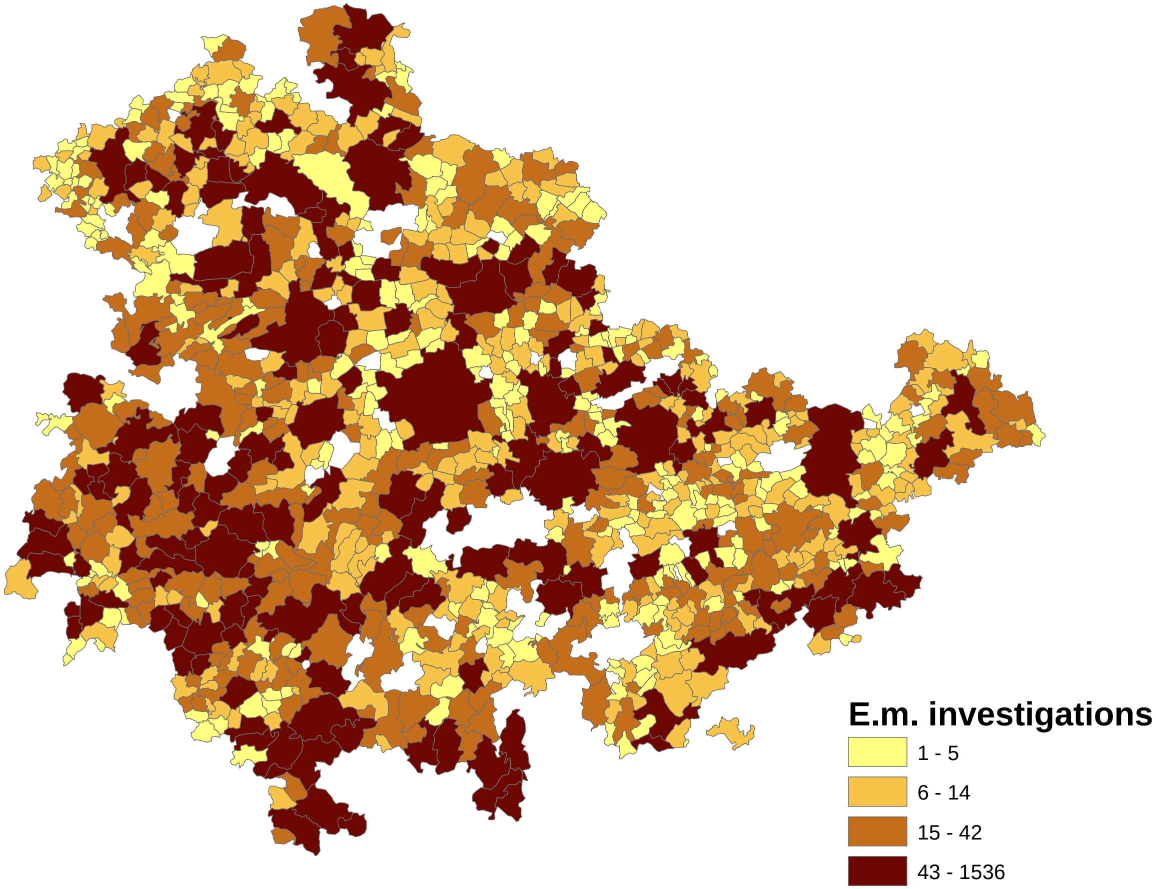
1,819 foxes positive for *E. multilocularis*
- Intestinal scraping technique



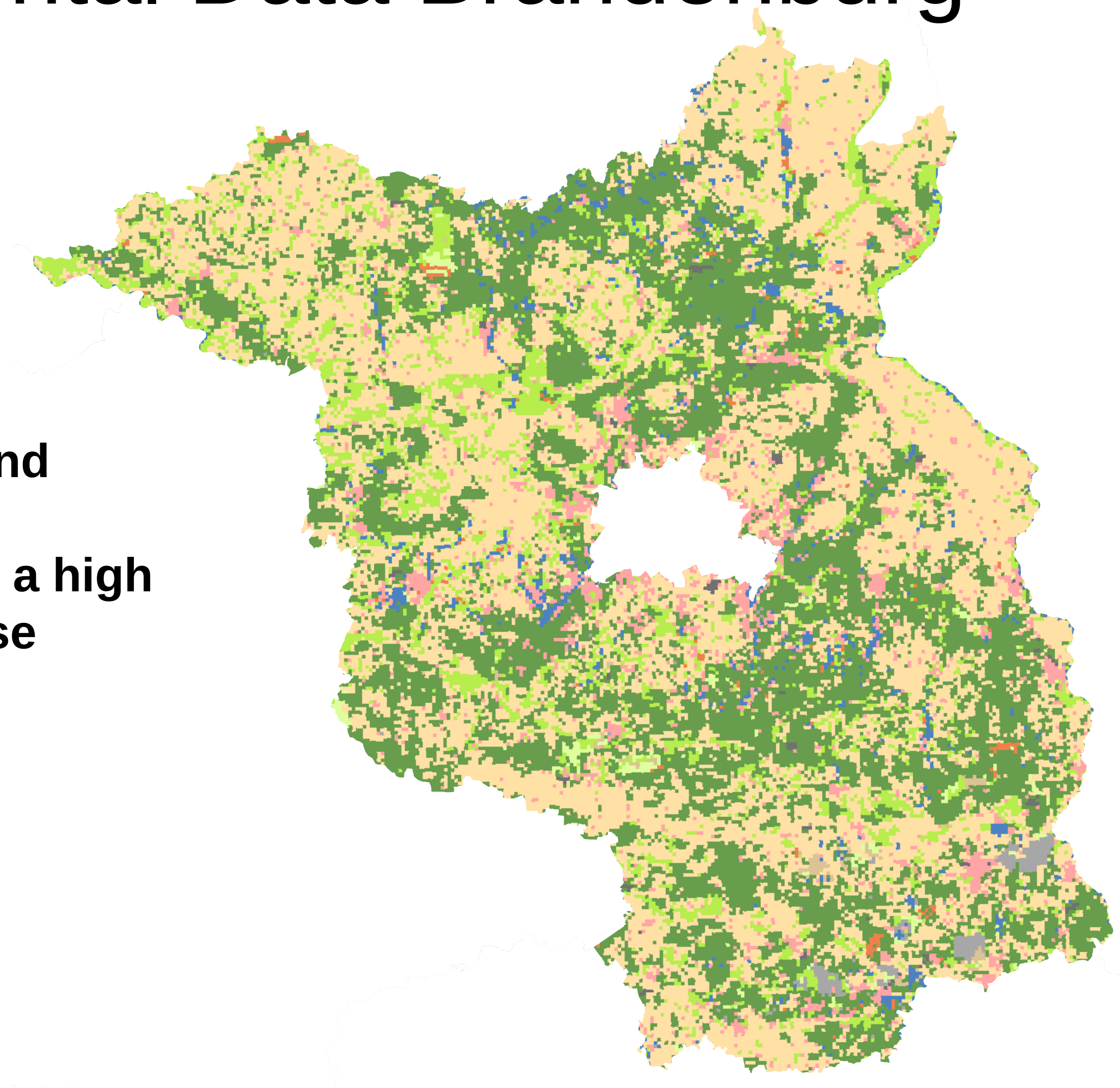
E.m. Data Thuringia

24,224 foxes
1990 – 2009
- Location (municipality)

6,302 foxes positive for *E. multilocularis*
- Intestinal scraping technique



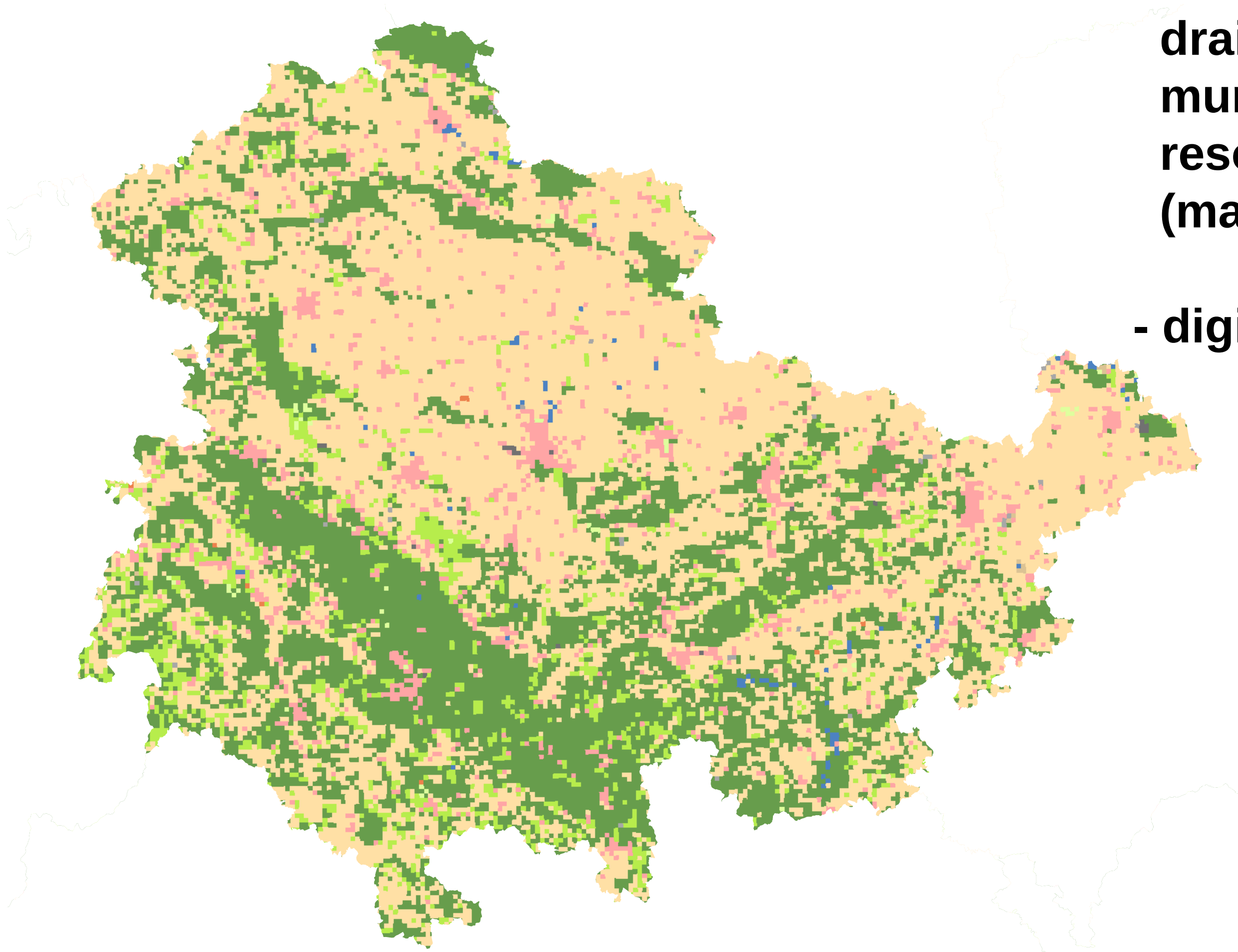
Environmental Data Brandenburg



- **landscape composition and drainage rivers per municipality derived from a high resolution survey database (map scale 1: 5,000)**

- **digital elevation model**

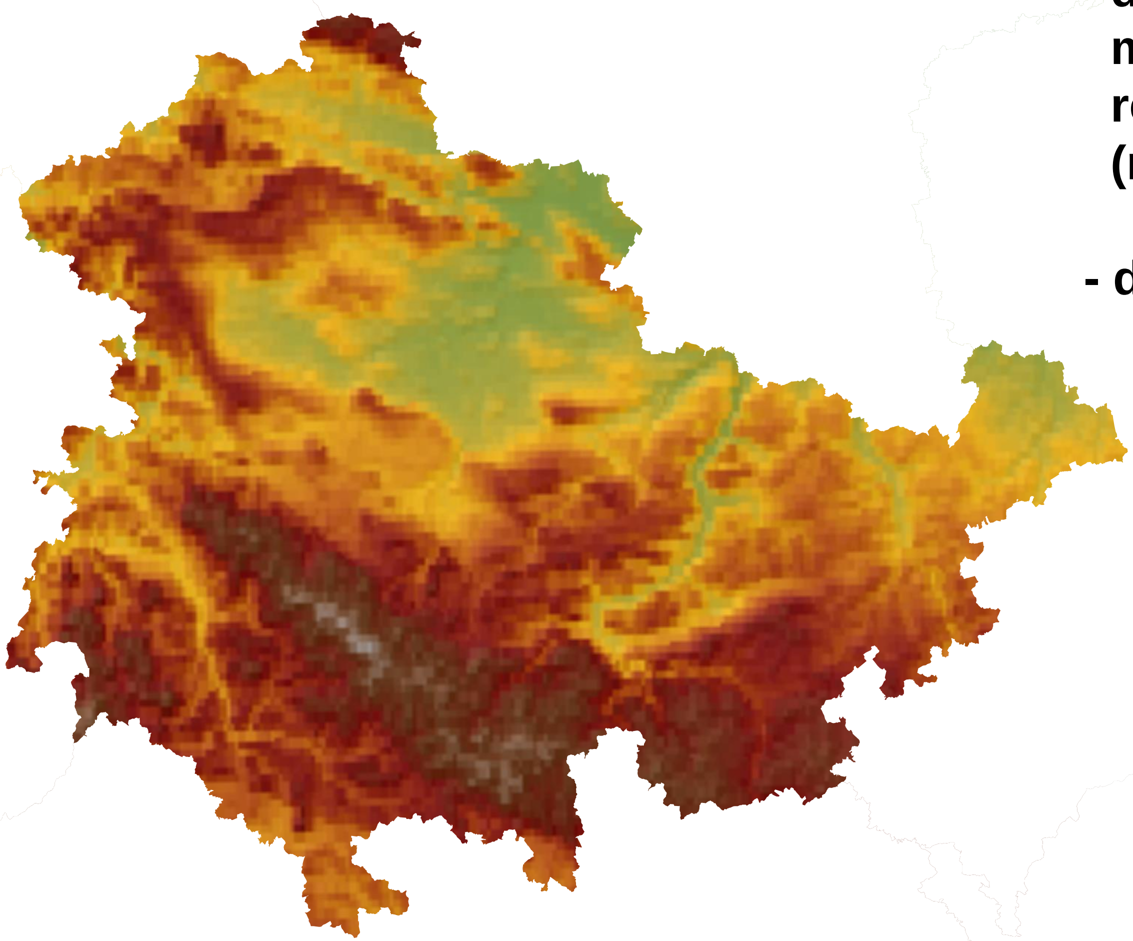
Environmental Data Thuringia



- **landscape composition and drainage rivers per municipality derived from a high resolution survey database (map scale 1: 5,000)**

- **digital elevation model**

Elevation Data Thuringia



- landscape composition and drainage rivers per municipality derived from a high resolution survey database (map scale 1: 5,000)

- digital elevation model

Full Bayesian Model

1. Each individual *i* has an unknown probability π_i that the animal is positive or negative depending on area the animal lives and one or more covariates describing the environmental conditions
2. Parameter π_i is modeled using a logistic model with random intercept and structured effect per municipality
3. The structured effect displays spatial heterogeneity using a Gaussian Markov Random field (GMRF)
4. We choose uninformative prior distributions [$\tau \sim Ga(0.001, 0.001)$] for the unknown variance parameters

Implementation of the Bayesian Model

- The joint distribution of the full Bayesian model is analytically intractable
- Therefore, we used a Markov Chain Monte Carlo algorithm (MCMC) to estimate the parameters of the model
  - Burn-in: 1,000 iterations
  - Sample: every 50th iteration of the next 50,000 iterations
  - The freely available software BayesX 2.0.1 was used for inference (http://www.stat.uni-muenchen.de/~bayesx/bayesx.html)
- And we used INLA (Integrated Nested Laplace Approximation) a new approach for Bayesian inference available as R package (www.r-inla.org)
  - very fast
  - accurate approximation to the marginal posterior density for the hyperparameters

Model Comparison

- Deviance
  - posterior mean of the deviance
  - deviance is defined as - 2 * log(likelihood)
- Deviance Information Criterion (DIC)
  - generalisation of Akaike's Information Criterion (AIC)
  - the model with the smallest DIC is estimated to be the model that would best predict a replicate dataset which has the same structure as that currently observed
- Effective number of parameters (pD)
  - pD is the posterior mean of the deviance minus the deviance of the posterior means
  - penalizing the deviance in the DIC
  - if the DIC between different models is comparable the model with the smallest pD should be the most parsimonious model

Model comparison and parameter estimates regarding *E.m.* in foxes in Brandenburg

Model	Deviance	DIC	pD	Estimate (95 % BCI)	Mean/St.Dev.
Spatial	9537.97	9809.30	135.66	1.726 (1.239; 2.382)	
Urban	9549.81	9808.06	129.13	-4.109 (-6.252; -2.063)	3.805
Crop	9549.74	9800.96	125.61	1.877 (1.164; 2.611)	5.066
Pasture	9548.41	9809.76	130.68	2.541 (0.970; 4.168)	3.190
Forest	9546.96	9801.93	127.49	-1.742 (-2.431; -0.960)	4.739
Drainage	9540.06	9807.41	133.68	264.665 (8.494; 508.946)	2.050
Elevation	9536.87	9809.19	136.16	0.0025 (-0.005; 0.009)	0.665
Crop + Drainage	9551.48	9801.21	124.86	1.806 (1.035; 2.511) 177.31 (-99.81; 420.03)	
Forest + Drainage	9545.96	9803.25	128.64	-1.693 (-2.471; -0.934) 38.75 (-234.58; 303.27)	

DIC = Deviance Information Criterion; pD = Effective number of parameters; BCI = Bayesian Credible Interval

Model comparison and parameter estimates regarding *E.m.* in foxes in Thuringia

Model	Deviance	DIC	pD	Estimate (95 % BCI)	Mean/St.Dev.
Spatial	24688.22	25294.90	303.34	0.767 (0.602; 0.957)	
Urban	24689.15	25298.58	304.71	-0.104 (-0.104; 0.850)	0.210
Crop	24687.59	25294.67	303.54	0.375 (0.057; 0.715)	2.243
Pasture	24711.36	25289.42	289.03	1.785 (1.139; 2.484)	5.315
Forest	24687.75	25286.34	299.29	-0.654 (-0.985; -0.316)	3.872
Drainage	24687.67	25298.32	305.33	72.153 (-172.92; 313.92)	0.581
Elevation	24690.19	25293.12	301.46	-0.0011 (-0.0020; -0.0003)	2.639
Pasture + Elevation	24715.99	25287.69	285.85	1.7184 (1.066; 2.397) -0.0011 (-0.0019; -0.0002)	
Forest + Elevation	24688.53	25288.01	299.73	-0.588841 (-0.939; -0.234) -0.0004 (-0.0014; 0.0006)	

DIC = Deviance Information Criterion; pD = Effective number of parameters; BCI = Bayesian Credible Interval

Comparison of GoF parameters using BayesX and INLA (*E.m.* in foxes in Brandenburg)

	BayesX		INLA	
Model	**DIC**	**pD**	**DIC**	**pD**
Spatial	9809.30	135.66	9809.95	134.01
Urban	9808.06	129.13	9807.36	127.09
Crop	9800.96	125.61	9801.91	124.25
Pasture	9809.76	130.68	9810.34	127.83
Forest	9801.93	127.49	9803.07	126.35
Drainage	9807.41	133.68	9809.87	133.86
Elevation	9809.19	136.16	9810.24	134.70
Crop + Drainage	9801.21	124.86	9801.90	124.16
Forest + Drainage	9803.25	128.64	9803.11	126.38

Comparison of GoF parameters using BayesX and INLA (*E.m.* in foxes in Thuringia)

	BayesX		INLA	
Model	**DIC**	**pD**	**DIC**	**pD**
Spatial	25294.90	303.34	25295.90	297.92
Urban	25298.58	304.71	25297.07	298.64
Crop	25294.67	303.54	25293.87	297.91
Pasture	25289.42	289.03	25287.24	283.33
Forest	25286.34	299.29	25286.04	294.02
Drainage	25298.32	305.33	25295.94	297.97
Elevation	25293.12	301.46	25293.10	295.80
Pasture + Elevation	25287.69	285.85	25285.62	280.52
Forest + Elevation	25288.01	299.73	25286.56	294.46

Summary

- A hierarchical Bayesian model was established to analyse potential factors influencing the distribution of *E. multilocularis* infections in foxes using data on municipality level
- The study confirmed results of a previous publication, which utilized exact locations and micro-habitat data of foxes in a much smaller region and using lower resolution data
- The statistical inference using INLA reduces the calculation time dramatically (MCMC approx. 2.5-3.0 hours to below 1 minute using INLA)
- Furthermore, INLA provides easily accessible Goodness of fit measures (e.g. local.DIC, CPO, Marginal Predictive Likelihood) not here demonstrated
- Nevertheless, an easy method for the selection of variables and the best model choice is still missing

Thank you
for listening!