Spatial spread of the brown rat resistance to rodenticides in Flanders

ISEC, Montpellier, July 1, 2014

Ivy Jansen and Kristof Baert

ivy.jansen@inbo.be

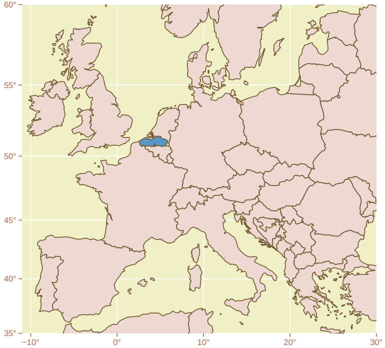
inbo

Research Institute
for Nature and Forest

Overview

Introduction

- Monitor the brown rat resistance to rodenticides in Flanders (by means of genetic mutations)

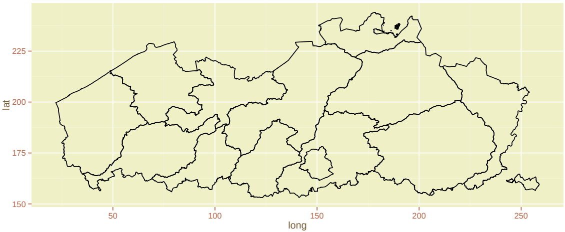


- Two screening periods
  - 2003 – 2005
  - 2006 – 2010

Introduction

Aims of the study

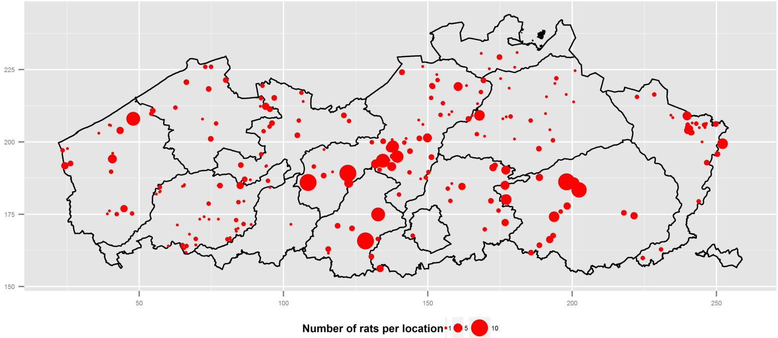
- Determine degree of resistance against rodenticides in Flanders
- Are there differences between the 12 river basins?



- Is the resistance increasing over time?

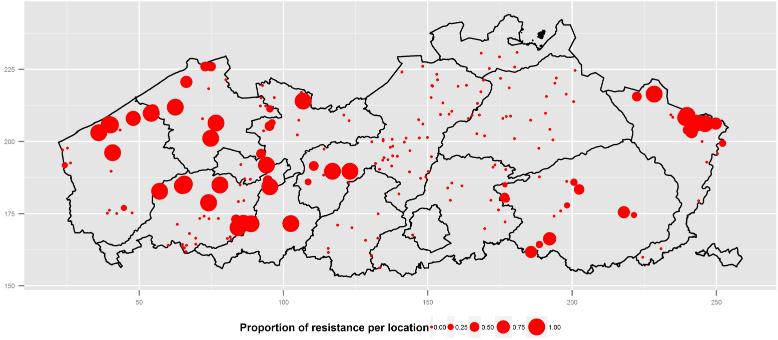
Exploratory Analysis Screening I (2003 – 2005)

- 581 rats on 238 locations
- Different number of rats per location (aim was 3)



Exploratory Analysis Screening I (2003 – 2005)

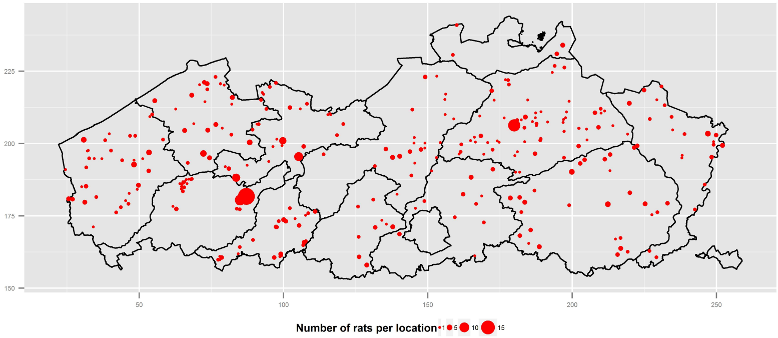
- Large differences in proportion of resistance between locations and river basins



Exploratory Analysis Screening II (2006 – 2010)

Same conclusions as in screening I

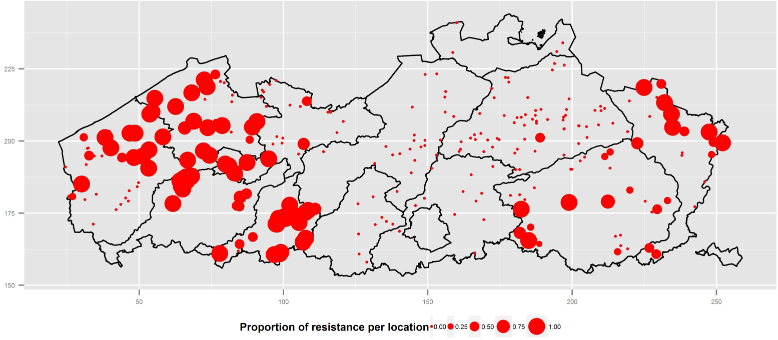
- 680 rats on 320 locations
- Different number of rats per location (aim was 3)



Exploratory Analysis Screening II (2006 – 2010)

Same conclusions as in screening I

- Large differences in proportion of resistance between locations and river basins



Statistical Analysis – Logistic Regression (Screening I)

Significant difference in resistance between river basins?

$$\text{logit}(\pi_{ij}) = \text{Basin}_i$$

1. Ignore clustering of rats within location
   - Highly significant difference between river basins ($D = 120.66$, $p < 0.001$)
   - Coefficients:

```
                       Estimate Std. Error z value Pr(>|z|)
BekkenBeneden_Schelde   -3.1850     0.4166  -7.645 2.09e-14 ***
BekkenBoven_Schelde     18.5661  4612.2020   0.004 0.996788
BekkenBrugse_Polders    -0.4308     0.3563  -1.209 0.226597
BekkenDemer             -0.9520     0.2510  -3.793 0.000149 ***
BekkenDender           -18.5661  1058.1117  -0.018 0.986001
BekkenDijle            -18.5661  1058.1117  -0.018 0.986001
BekkenGentse_kanalen    -2.0149     0.5323  -3.785 0.000153 ***
BekkenIjzer             -0.5108     0.2981  -1.713 0.086646 .
BekkenLeie              -0.8910     0.2969  -3.001 0.002690 **
BekkenMaas              -0.5534     0.2617  -2.115 0.034455 *
BekkenMarke            -18.5661  2465.3257  -0.008 0.993991
BekkenNete             -18.5661  1135.4456  -0.016 0.986954
```

   - Estimates on the boundary of parameter space (all rats sensitive or resistant)

Statistical Analysis – Logistic Regression (Screening II)

Significant difference in resistance between river basins?

$$\text{logit}(\pi_{ij}) = \text{Basin}_i$$

1. Ignore clustering of rats within location
   - Highly significant difference between river basins ($D = 270.27$, $p < 0.001$)
   - Coefficients:

```
                   Estimate Std. Error z value Pr(>|z|)
RBBenedenschelde  -18.5661  1006.4650  -0.018 0.985282
RBBovenschelde      1.2192     0.3157   3.862 0.000112 ***
RBBrugse Polder     0.3930     0.2491   1.578 0.114560
RBDemer            -1.1987     0.2617  -4.580 4.65e-06 ***
RBDender          -18.5661  1537.4007  -0.012 0.990365
RBDijle           -18.5661  1232.6628  -0.015 0.987983
RBGentse kanalen   -0.7563     0.2528  -2.992 0.002769 **
RBIjzer            -0.3947     0.2324  -1.698 0.089439 .
RBLeie              0.9163     0.2789   3.286 0.001018 **
RBMaas Antwerpen  -18.5661  1966.6495  -0.009 0.992468
RBMaas Limburg     -0.3429     0.2788  -1.230 0.218619
RBNete             -4.6913     1.0046  -4.670 3.01e-06 ***
```

   - Estimates on the boundary of parameter space (all rats sensitive or resistant)

Statistical Analysis – Logistic Regression (Screening I)

Significant difference in resistance between river basins?

$$\text{logit}(\pi_{ij}) = \text{Basin}_i$$

2. Reduce to presence or absence of resistance per location
   - Highly significant difference between river basins ($D = 75.575$, $p < 0.001$)
   - Coefficients:

```
                        Estimate Std. Error z value Pr(>|z|)
BekkenBeneden_Schelde    -2.5257     0.5196  -4.861 1.17e-06 ***
BekkenBoven_Schelde      18.5661  4612.2020   0.004   0.9968
BekkenBrugse_Polders      0.2877     0.5401   0.533   0.5943
BekkenDemer               0.6061     0.5075   1.194   0.2324
BekkenDender            -18.5661  2174.2129  -0.009   0.9932
BekkenDijle             -18.5661  1809.0545  -0.010   0.9918
BekkenGentse_kanalen     -1.0986     0.5774  -1.903   0.0571 .
BekkenIjzer               0.1054     0.4595   0.229   0.8186
BekkenLeie               -0.8473     0.3450  -2.456   0.0141 *
BekkenMaas               -0.1431     0.3789  -0.378   0.7057
BekkenMarke             -18.5661  2917.0127  -0.006   0.9949
BekkenNete              -18.5661  1423.3564  -0.013   0.9896
```

   - Estimates on the boundary of parameter space (all rats sensitive or resistant)

Statistical Analysis – Logistic Regression (Screening II)

Significant difference in resistance between river basins?

$$\text{logit}(\pi_{ij}) = \text{Basin}_i$$

2. Reduce to presence or absence of resistance per location
   - Highly significant difference between river basins ($D = 176.73$, $p < 0.001$)
   - Coefficients:

```
                     Estimate Std. Error z value Pr(>|z|)
RBBenedenschelde  -18.56607 1331.42803  -0.014 0.988874
RBBovenschelde      1.99243    0.61546   3.237 0.001207 **
RBBrugse Polder     0.47000    0.40311   1.166 0.243641
RBDemer            -0.26826    0.36844  -0.728 0.466545
RBDender          -18.56607 2306.10099  -0.008 0.993576
RBDijle           -18.56607 1882.92358  -0.010 0.992133
RBGentse kanalen   -0.61310    0.34437  -1.780 0.075015 .
RBIjzer            -0.05407    0.32892  -0.164 0.869433
RBLeie              2.52573    0.73485   3.437 0.000588 ***
RBMaas Antwerpen  -18.56607 3261.31930  -0.006 0.995458
RBMaas Limburg      0.57536    0.41667   1.381 0.167318
RBNete             -4.15888    1.00778  -4.127 3.68e-05 ***
```

   - Estimates on the boundary of parameter space (all rats sensitive or resistant)

Statistical Analysis – Mixed Models (Screening I)

Include random location effect to take into account clustering of rats

$$\text{logit}(\pi_{ijk}) = \text{Basin}_i + b_{ij} \quad \text{with} \quad b_{ij} \sim N(0, \sigma_b^2)$$

- Highly significant difference between river basins ($\chi^2 = 66.442$, $p < 0.001$)
- Fixed effects:

```
                        Estimate Std. Error z value Pr(>|z|)
BekkenBeneden_Schelde   -4.2638     0.6954  -6.131 8.72e-10 ***
BekkenBoven_Schelde     18.5661  7604.2378   0.002 0.998052
BekkenBrugse_Polders    -0.6193     0.6048  -1.024 0.305852
BekkenDemer             -1.2693     0.4941  -2.569 0.010199 *
BekkenDender           -18.5661  1744.5421  -0.011 0.991509
BekkenDijle            -18.5661  1744.5420  -0.011 0.991509
BekkenGentse_kanalen    -2.6959     0.7890  -3.417 0.000633 ***
BekkenIjzer             -0.8985     0.5389  -1.667 0.095433 .
BekkenLeie              -1.4740     0.4463  -3.303 0.000957 ***
BekkenMaas              -1.0057     0.4499  -2.235 0.025394 *
BekkenMarke            -18.5661  4064.6399  -0.005 0.996356
BekkenNete             -18.5665  1872.4378  -0.010 0.992089
```

- Variability between locations $\sigma_b^2 = 2.6388$

```
Random effects:
 Groups  Name        Variance Std.Dev.
 locatie (Intercept) 2.6388   1.6244
Number of obs: 581, groups: locatie, 235
```

Statistical Analysis – Mixed Models (Screening II)

Include random location effect to take into account clustering of rats

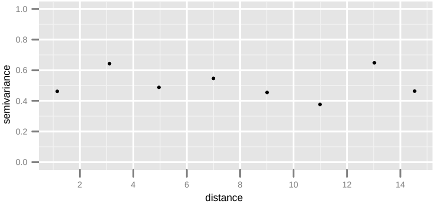
$$\text{logit}(\pi_{ijk}) = \text{Basin}_i + b_{ij} \quad \text{with} \quad b_{ij} \sim N(0, \sigma_b^2)$$

- Highly significant difference between river basins ($\chi^2 = 521.43$, $p < 0.001$)
- ```
  Fixed effects:
                     Estimate Std. Error z value Pr(>|z|)
  RBBenedenschelde  -18.5661  1659.3905  -0.011 0.991073
  RBBovenschelde      2.0136     0.5740   3.508 0.000451 ***
  RBBrugse Polder     0.3605     0.4997   0.721 0.470706
  RBDemer            -1.8957     0.4930  -3.845 0.000121 ***
  RBDender          -18.5661  2534.7520  -0.007 0.994156
  RBDijle           -18.5661  2032.3259  -0.009 0.992711
  RBGentse kanalen   -1.3245     0.4760  -2.783 0.005389 **
  RBIjzer            -0.6962     0.4365  -1.595 0.110728
  RBLeie              2.1734     0.6240   3.483 0.000496 ***
  RBMaas Antwerpen  -18.5661  3242.4624  -0.006 0.995431
  RBMaas Limburg     -0.4780     0.5089  -0.939 0.347563
  RBNete             -5.8783     1.6556  -3.550 0.000385 ***
  ```
- Variability between locations $\sigma_b^2 = 3.6025$
  ```
  Random effects:
   Groups  Name        Variance Std.Dev.
   locatie (Intercept) 3.6025   1.898
  Number of obs: 680, groups: locatie, 320
  ```
  ```

Statistical Analysis – Spatial Analysis (Screening I)

Additional spatial correlation between locations?

- Study variogram of the residuals within 15km range
  - No extra spatial correlation necessary

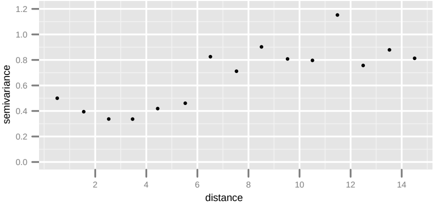


- Isotropy (equal variance in all directions) fulfilled

Statistical Analysis – Spatial Analysis (Screening II)

Additional spatial correlation between locations?

- Study variogram of the residuals within 15km range
  - Smaller variability between closeby locations ($<$ 6km)



  - Take into account this correlation in further analyses
  - Aim for a good spatial spread in future monitoring

- Isotropy (equal variance in all directions) fulfilled

Statistical Analysis – Evolution over time

- Has not been analysed, since data collection in both screening periods was not comparable.

	Screening I	lower	upper	Screening II	lower	upper
Beneden_Schelde	0.04	0.02	0.09	0.00		
Boven_Schelde	1.00			0.77	0.65	0.86
Brugse_Polders	0.39	0.24	0.57	0.60	0.48	0.71
Demer	0.28	0.19	0.39	0.23	0.15	0.33
Dender	0.00			0.00		
Dijle	0.00			0.00		
Gentse_kanalen	0.12	0.04	0.27	0.32	0.22	0.44
Ijzer	0.38	0.25	0.52	0.40	0.30	0.52
Leie	0.29	0.19	0.42	0.71	0.59	0.81
Maas	0.37	0.26	0.49	0.00		
Marke	0.00			0.42	0.29	0.55
Nete	0.00			0.01	0.00	0.06

- A follow-up monitoring program, with enough rats and locations per river basin, will be designed to answer all questions more accurately in the future.

Conclusions

- Ignore clustering $\Rightarrow$ WRONG !!
- Necessary to take into account correlation between rats from the same location
  - either aggregate to presence/absence of resistance
    $\Rightarrow$ loss of information/observations
  - or add a random location effect
    $\Rightarrow$ variability between locations can be estimated

Conclusions

- Ignore clustering $\Rightarrow$ WRONG !!
- Necessary to take into account correlation between rats from the same location
  - either aggregate to presence/absence of resistance
    $\Rightarrow$ loss of information/observations
  - or add a random location effect
    $\Rightarrow$ variability between locations can be estimated
- Highly significant differences in proportion of resistance between river basins
- River basins with only sensitive or resistant rats have parameter estimates on the boundary of the parameter space (but do not influence other estimates)
- Accuracy of estimates depends on number of rats per river basin (for some basins very small, < 20)

Future Monitoring

- Find an optimal monitoring design
  1. Either choose for uncorrelated data, so only 1 rat per location
  2. Or opt for a clustered design with more rats per location
- Sample size calculations
  - Calculate sample size within each river basin
  - Depends on current proportion $\overline{P}$ and the assumed increase Δ in proportion of resistance over 1 monitoring period

Future Monitoring

- Find an optimal monitoring design
  1. Either choose for uncorrelated data, so only 1 rat per location
  2. Or opt for a clustered design with more rats per location
- Sample size calculations
  - Calculate sample size within each river basin
  - Depends on current proportion $\overline{P}$ and the assumed increase Δ in proportion of resistance over 1 monitoring period

Example

- Assume $\alpha = 0.10$, $1 - \beta = 0.80$, $\Delta = 10\%$
- *n* = number of locations
- *n* ranges from 54 (when $\overline{P} = 0$ or $\overline{P} = 1$) to 236 (when $\overline{P} \sim 0.5$)

Future Monitoring

Influence of choices

- $\Delta = 5\% \Rightarrow n \times 2$
  $\Delta = 1\% \Rightarrow n \times 10$
- $\alpha = 0.05, 1 - \beta = 0.90 \Rightarrow n \times 2$
- More rats per location

$$K_T = n(K_L + SK_R)$$

  - Total cost $= K_T$
  - Selection cost for a location $= K_L$ (transportation, traps, ...)
  - Sample cost for a rat $= K_R$ (genetic analysis, ...)
  - Number of locations $= n$
  - Number of rats per location $= S$

Future Monitoring

Provisional design: 100 locations / river basin / 3y

- $\alpha = 0.05, 1-\beta = 0.80 \Rightarrow \Delta$ within $7\% - 20\%$
- $\alpha = 0.05, 1-\beta = 0.90 \Rightarrow \Delta$ within $8\% - 22\%$
- $\alpha = 0.10, 1-\beta = 0.80 \Rightarrow \Delta$ within $6\% - 17\%$
- $\alpha = 0.10, 1-\beta = 0.90 \Rightarrow \Delta$ within $7\% - 20\%$