A system for monitoring egg counts of the dengue disease mosquito (*Aedes aegypti*)

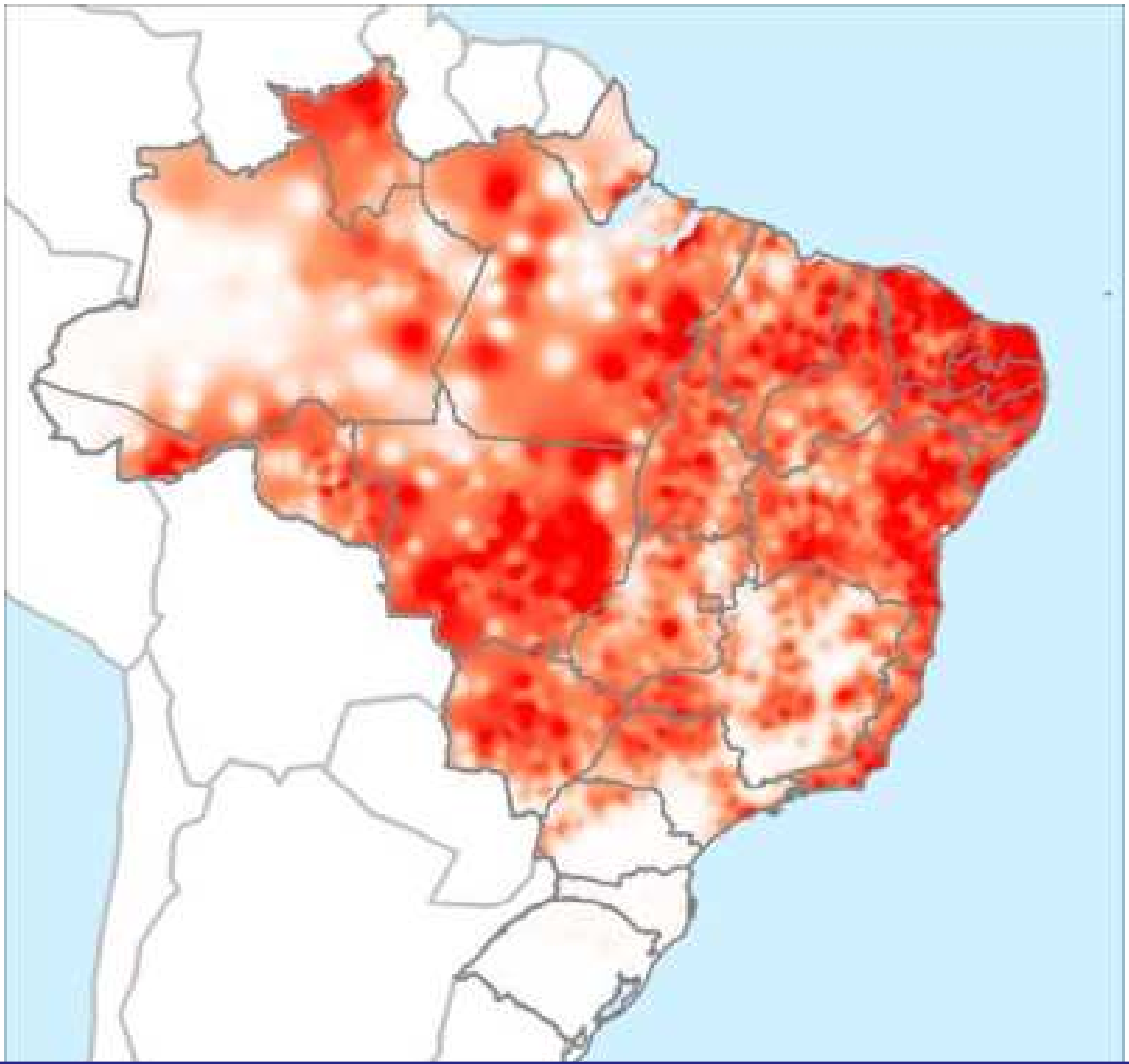
Wagner H. Bonat
Henrique S. Dallazuanna
Paulo J. Ribeiro Jr

Universidade Federal do Paraná, Brasil

Introduction - the context

- Epidemiological scenario: socio, cultural and environmental aspects
- Some important tropical diseases: dengue, leptospirose, tuberculose, etc.
- The **SAUDAVEL** project
- identification and modelling of risk factors: protection for endemic and epidemic periods
- developing resources for the practice of epidemiological and entomological surveillance
- Purpose here: illustrate a general framework covering biological, epidemiological, statistical, social and decision dimensions for surveillance and control

Dengue in Brazil

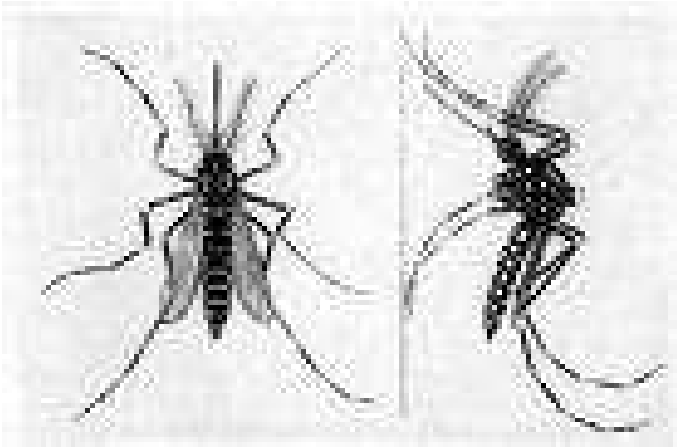


Dengue in Brasil

- currently is one of the most important diseases
- first sorotype detected in 1980
- explosive epidemics
- 500.000 cases/yr and 40 deaths/yr
- highly affected by demographic density and entomological control
- virus transmited (only) by *Aedes* mosquitos.

Aedes aegypti

- suitable environmental conditions at tropical countries
- transmitions by adult mosquito females, daily habits
- transmition after 8 to 10 days after first contact and vertical transmission



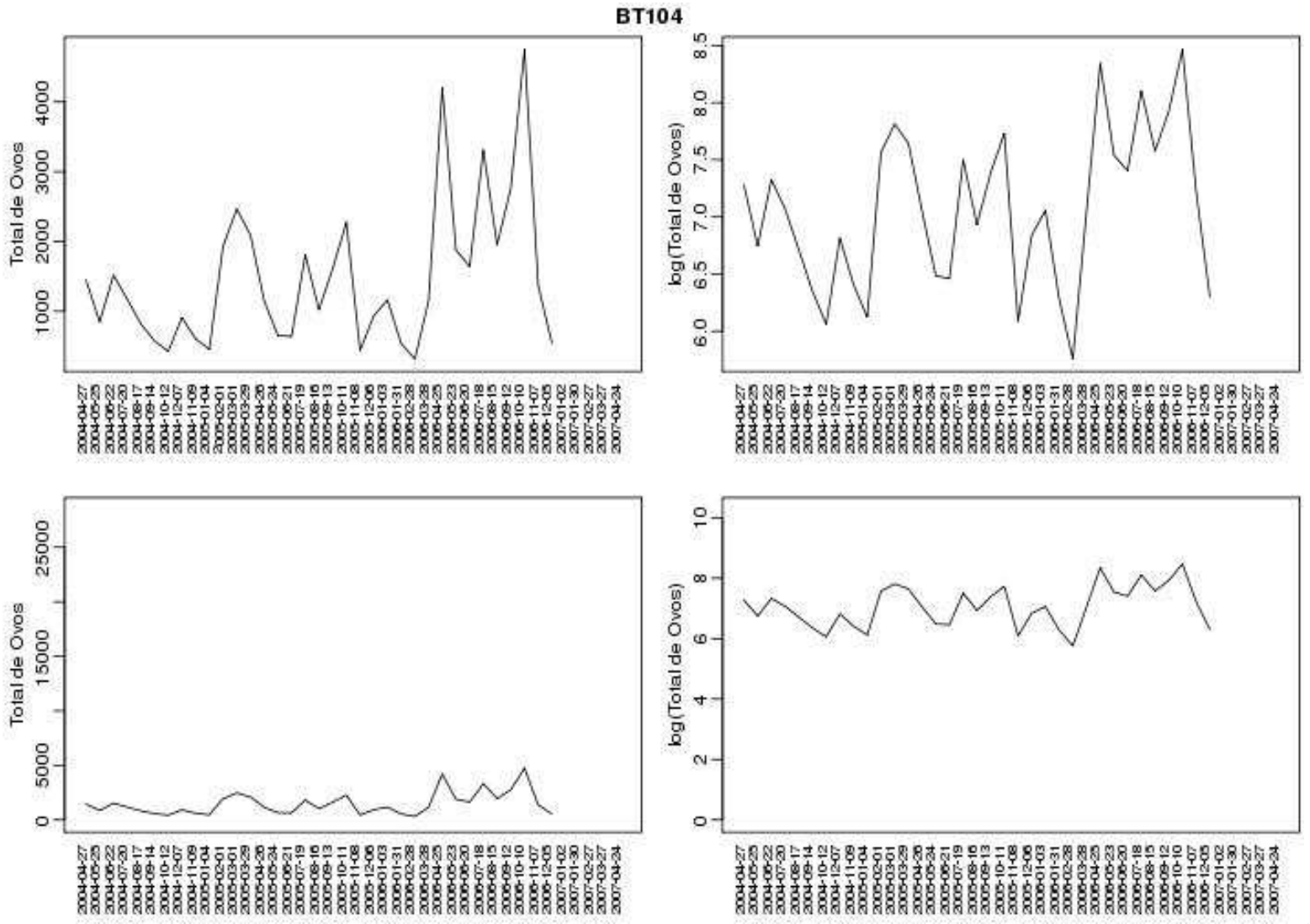
Entomologic surveillance

- monitoring of eggs instead of adults and/or larvae (ovitraps)
- density at urban area, spatial-temporal risk maps, infestation indexes
- monitoring of water deposits.
- domestic habits, short range
- *risk* = exposure X vulnerability
  - exposure: nature, environment, extent of events
  - vulnerability: social factors, adaptation, social inequalities

The Recife-SAUDAVEL experiment

- Desenvolvido em Recife/PE.
- start in 2004: 564 ovitraps at 6 neighbourhoods
- 7 days cycle for data colection with 25% at ech week
- May 2007: 19.068 data colections, 14.829.557 eggs counted.
- **local**: place type, backyard, water supply (type, quality, frequency), recipients (size and lid)
- environmental: temperature(s), umidity, rainfall
- weekly grouped, up to 12 weeks lag
- movie and slides

Her majesty, the data



Some issues

- relate egg counts with local and environmental covariates
- effects:
  - $Y \sim Amb + Loc$
  - $Y \sim Loc + Amb$
- spatial and/or temporal
- relations within and between neighborhoods
- relations at city level
- prediction, surveillance, alarm system, intervention

Modelo 1: No eggs $\sim$ Environment

- $\sim \beta_0 + \beta_1 X_t + \beta_2 X_{t-1} + \beta_2 X_{t-2} + \ldots$
- alternatives to non/semi-parametrics, transfer functions, RW, CAR, etc
- factorial analysis
- negative binomial

Tabela: Model 1

Effects	Estimates	Std. errors
Intercepto	1.818	1.174
TEMP.1	-0.150	0.049
TEMP.3	0.257	0.038
UMI.2	0.0382	0.008

Fators - max. temp.

loads	Factor 1	Factor 2	Factor 3
Semana 1	**0.9808**	-0.1799	0.0117
Semana 2	**0.9152**	-0.1335	0.1008
Semana 3	**0.7789**	-0.0726	0.2388
Semana 4	**0.6652**	-0.0263	0.3462
Semana 5	0.4625	0.0197	**0.5299**
Semana 6	0.3216	0.1194	**0.5867**
Semana 7	0.2220	0.2333	**0.5752**
Semana 8	0.1359	0.4048	**0.4842**
Semana 9	0.0795	**0.6043**	0.3287
Semana 10	0.0559	**0.7245**	0.2030
Semana 11	0.0098	**0.8607**	0.0706
Semana 12	-0.0950	**0.8775**	0.0837

Model 2: No eggs $\sim$ Environment + Local

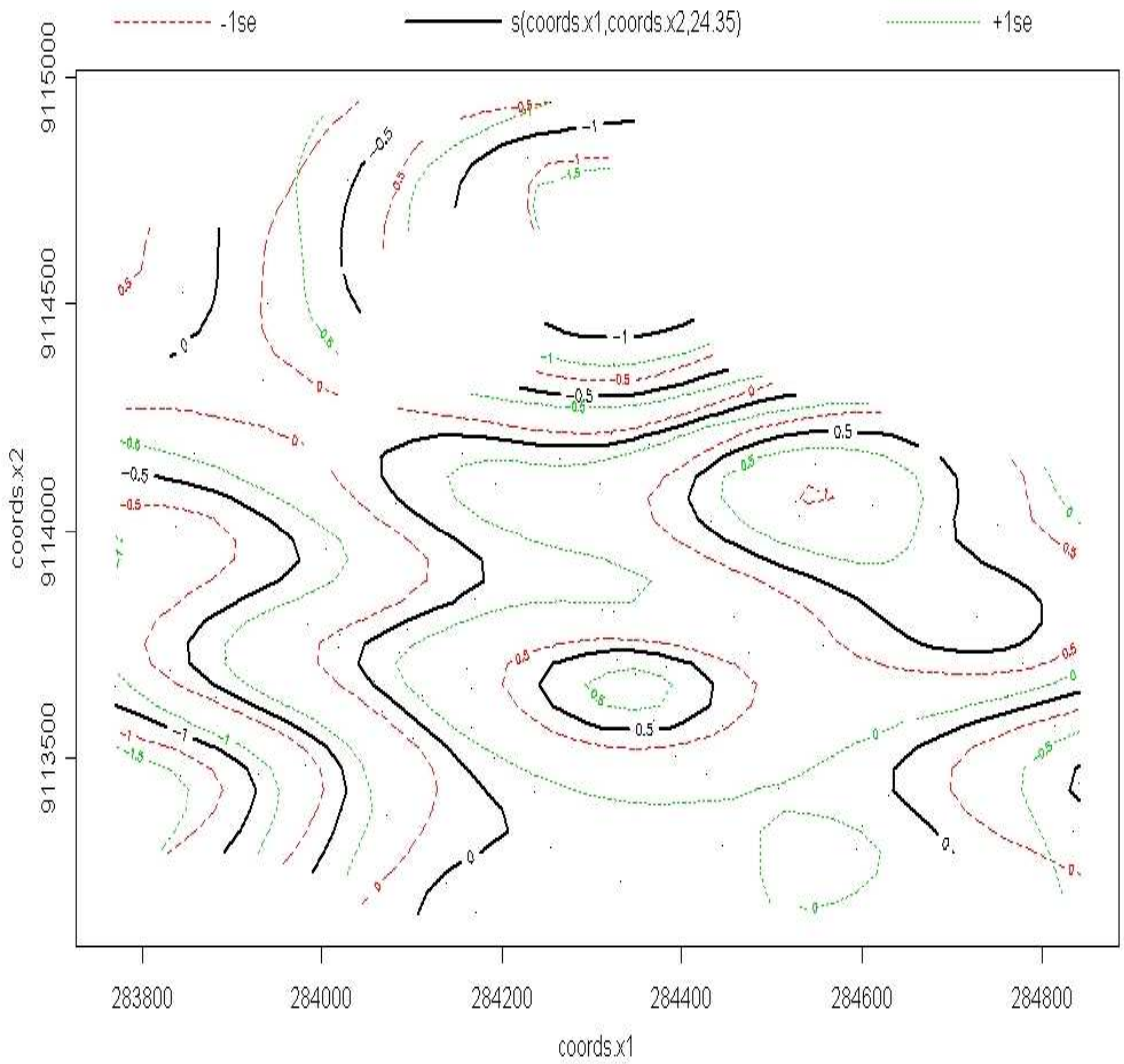
effects	Estimates	Std. erros
Intercepto	1.55	1.099
TEMP.1	-0.160	0.046
TEMP.3	0.268	0.036
UMI.2	0.0395	0.007
RES.PEQ.SEM	0.574	0.103
RES.PEQ.COM	0.178	0.06
FREQ ABAST	0.420	0.088
AGUA REDE	-0.292	0.098

Model 3: No eggs $\sim$ Env. + Local + Space + Time

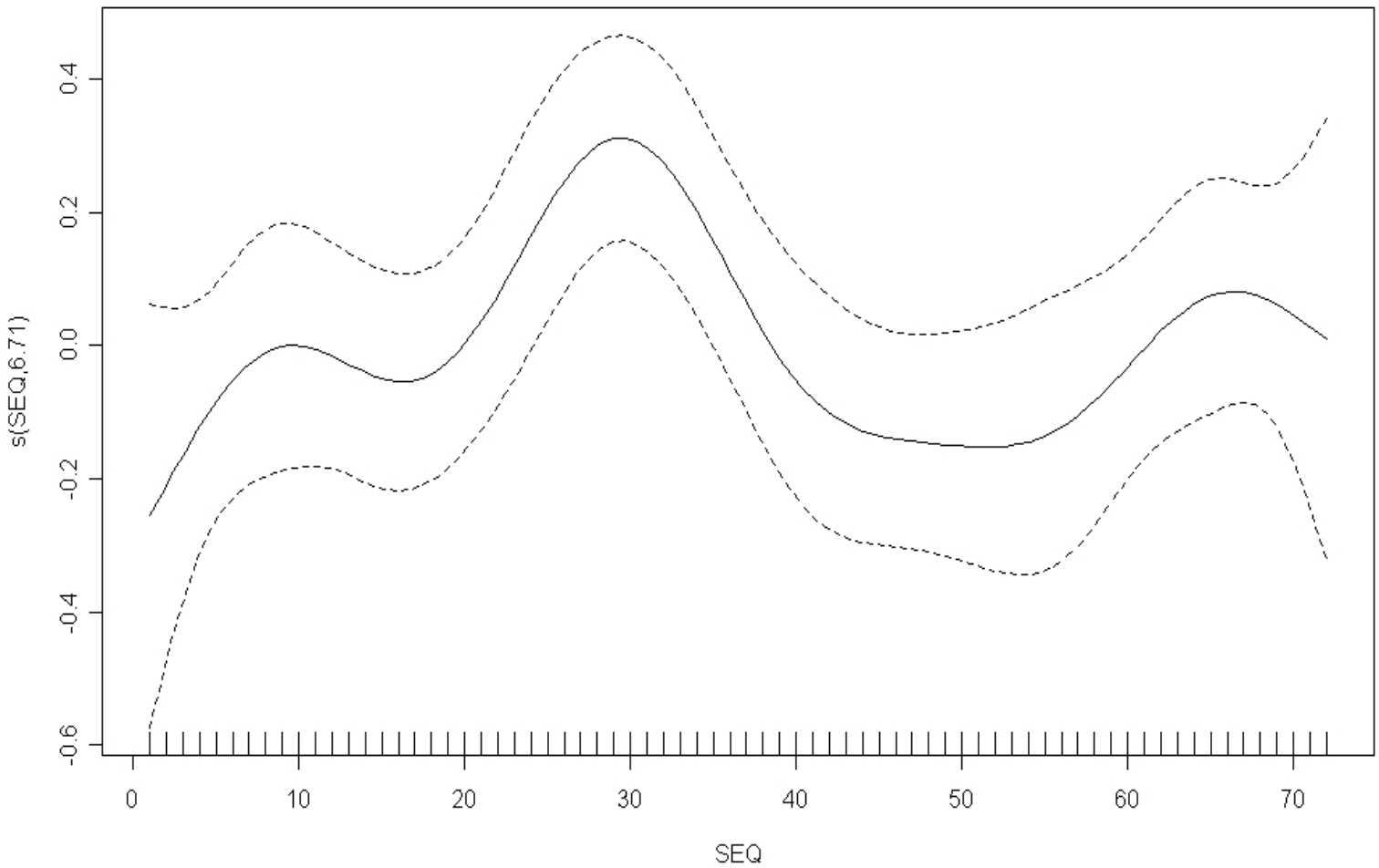
Effects	Estimates	Std Errors
Intercepto	1.209	0.980
TEMP.1	-0.161	0.041
TEMP.3	0.284	0.032
UMI.2	0.0378	0.006
RES.PEQ.SEM	0.426	0.104
Smooth functions	EDF	F
Coordinates	24.346	10.61
Time	6.71	3.39

- $Y \sim$ Negative Binomial
- GAM model
  $g(E(Y)) = \beta_0 + \sum_a \beta_a x_a + \sum_l \beta_l z_l + s(c_x, c_y) + s(t)$

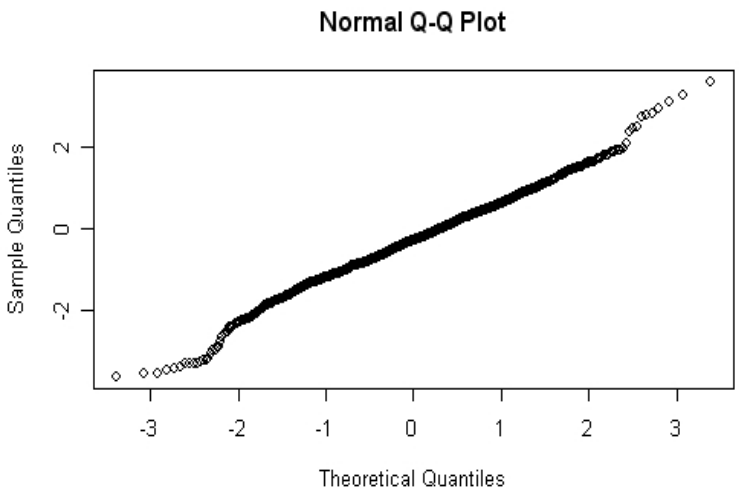
Spacial effect

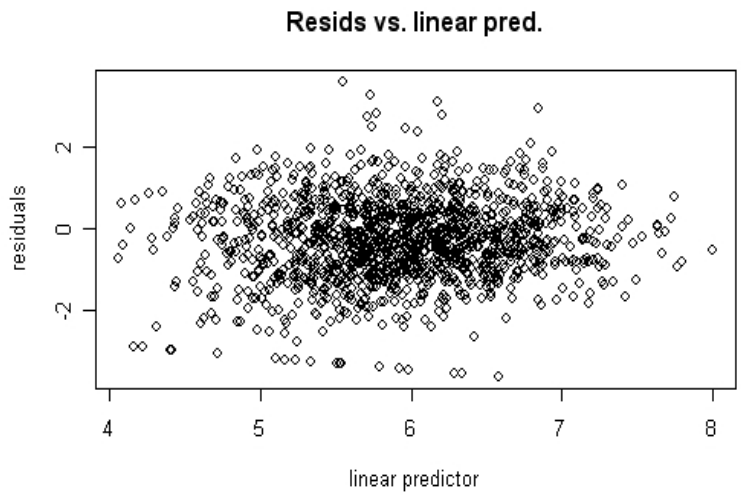


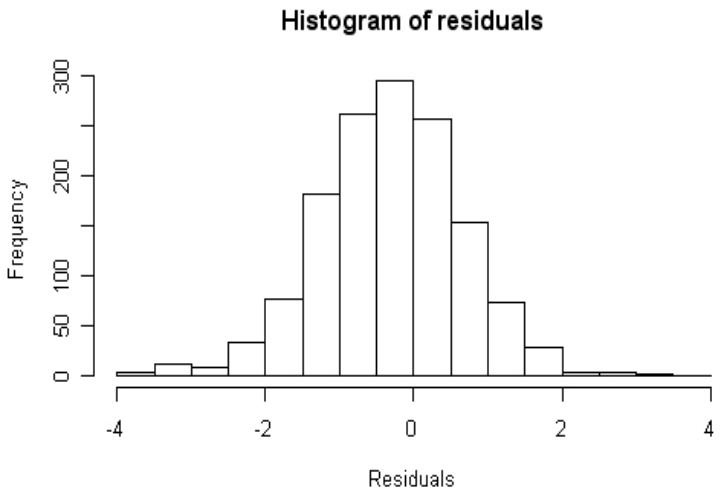
Temporal effect

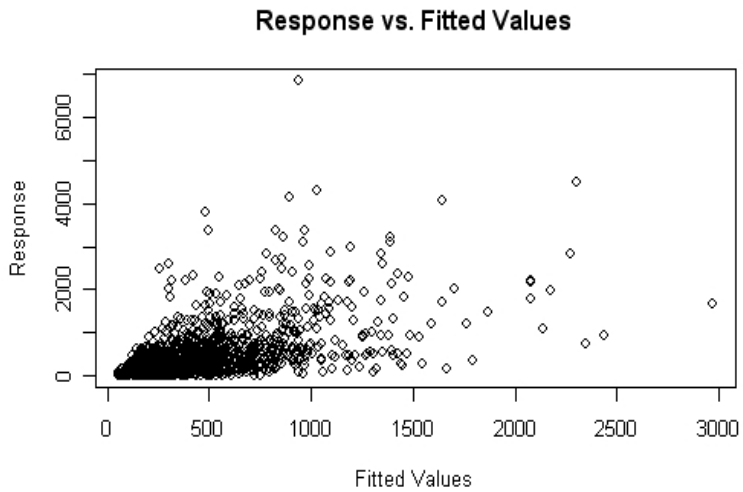


Residuals

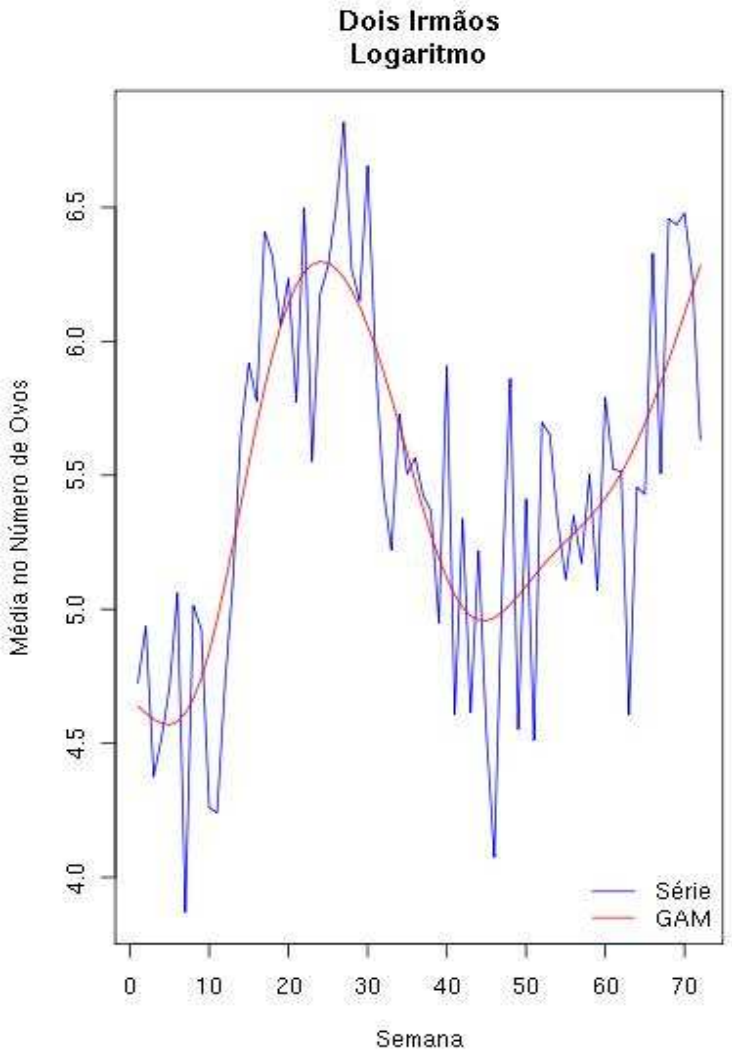

Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles
-3
-2
-1
0
1
2
3
-2
0
2

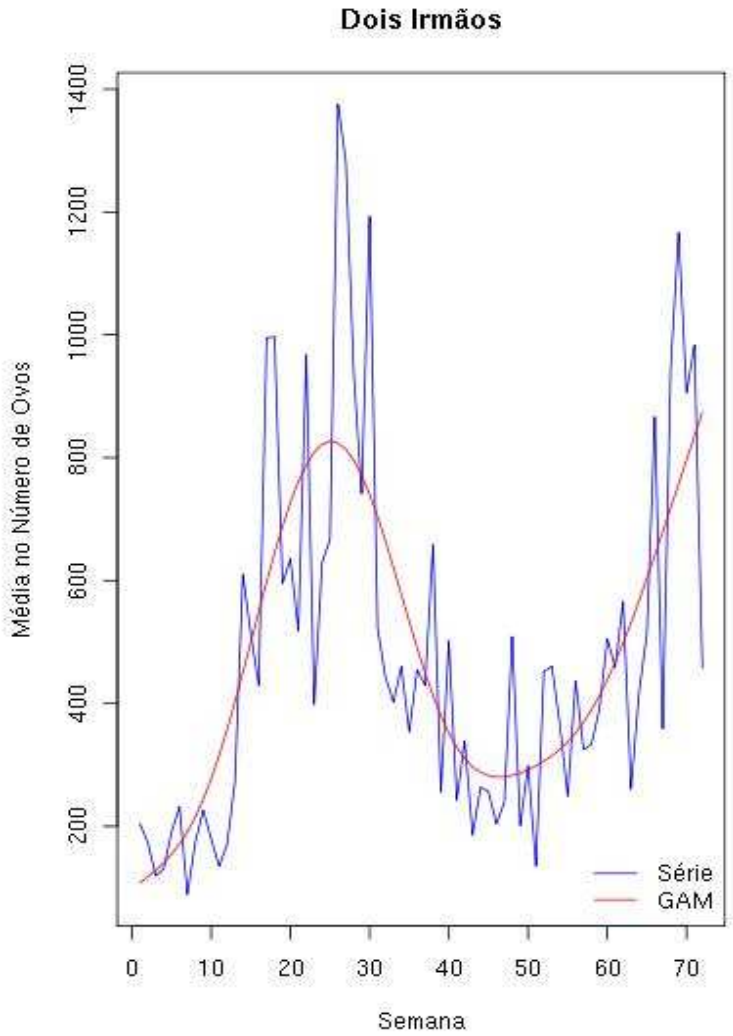

Resids vs. linear pred.
residuals
linear predictor
4
5
6
7
8
-2
0
2


Histogram of residuals
Frequency
Residuals
0
50
100
200
300
-4
-2
0
2
4


Response vs. Fitted Values
Response
Fitted Values
0
2000
4000
6000
0
500
1000
1500
2000
2500
3000

Predicting from the model

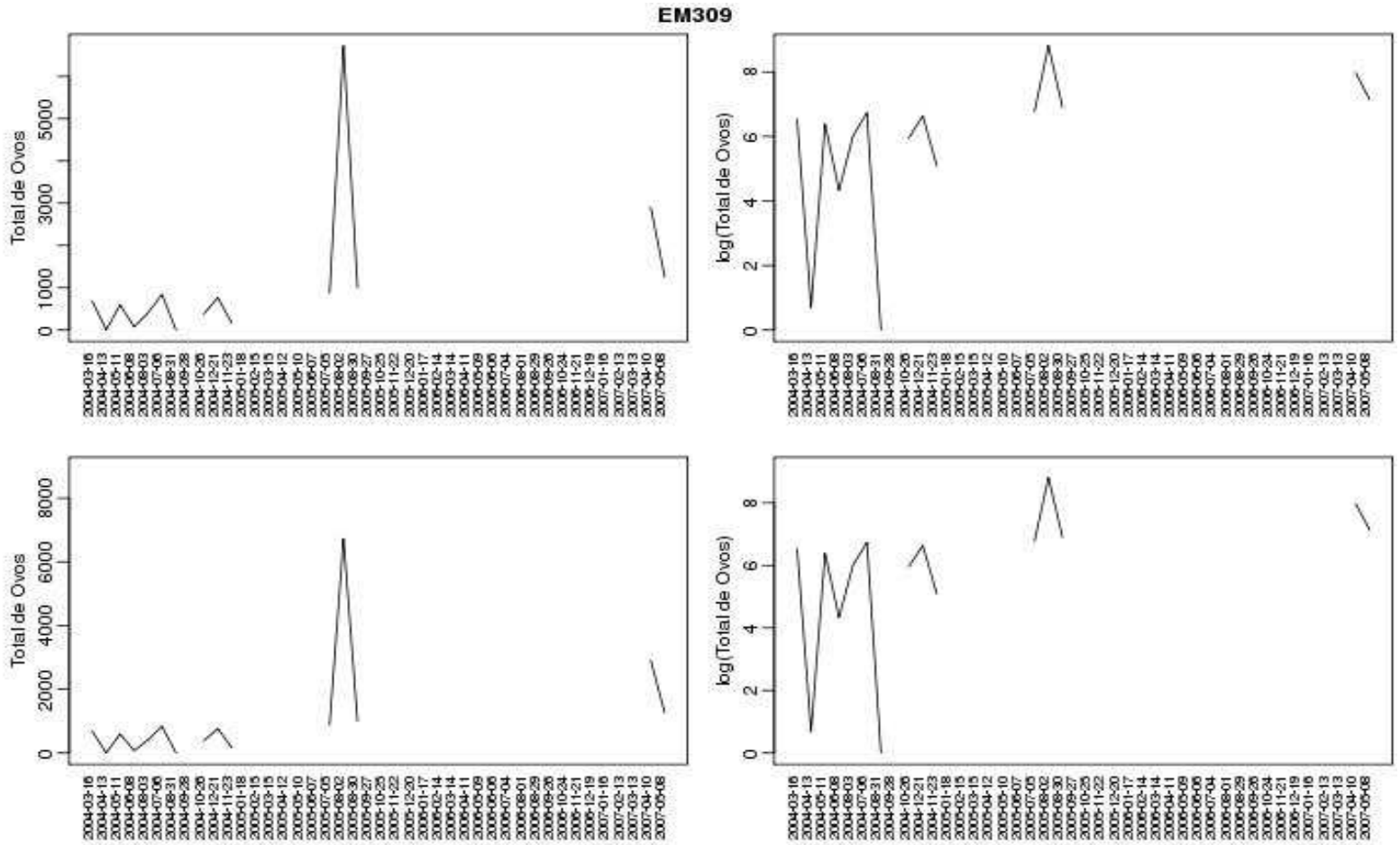




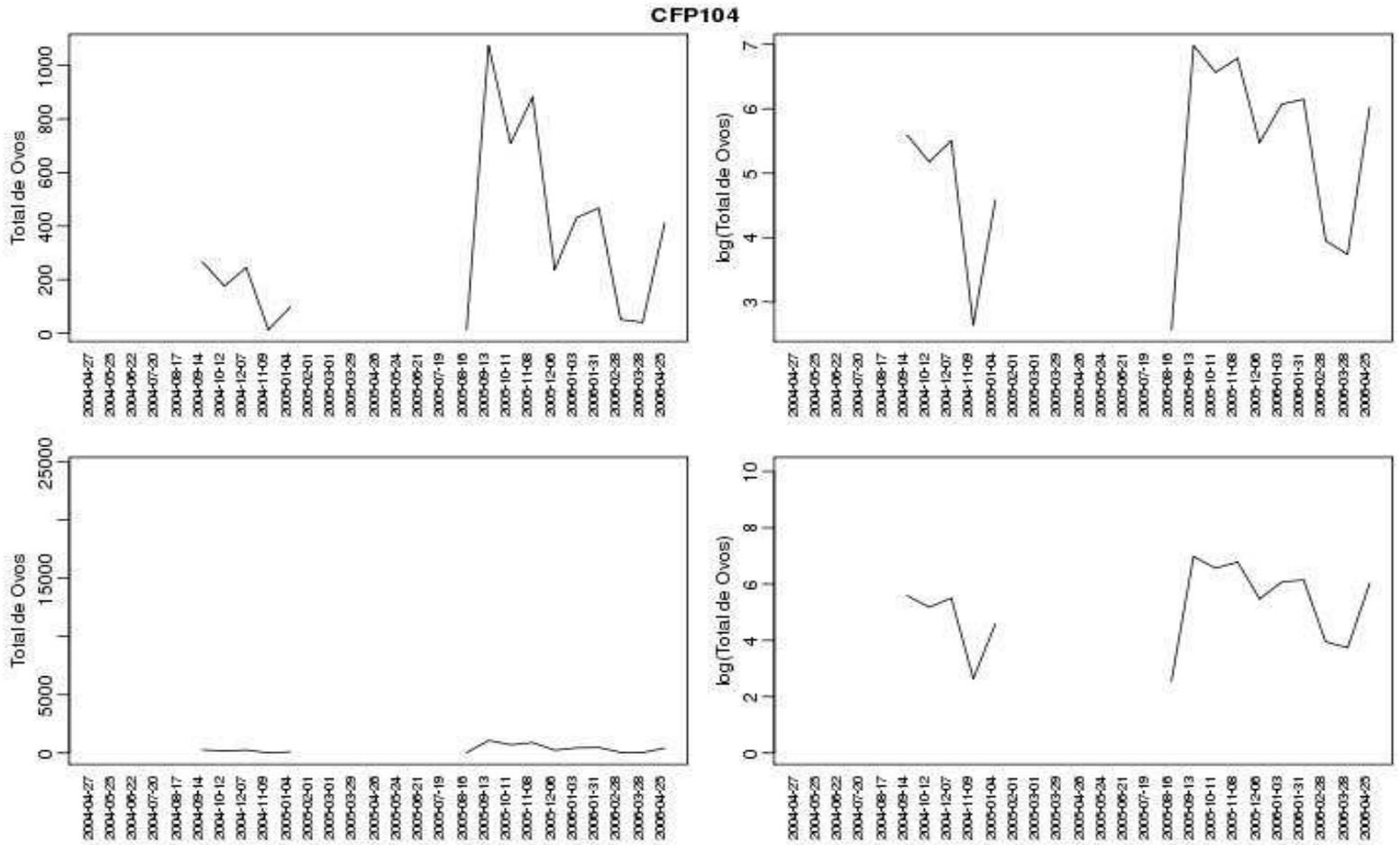
Limitations

- Field conditions for the experiment, limited control of sources of variation
- *screening* / exploratory
- reliability of the data: missingness, peaks, outlines, local outliers

A critical view of the data



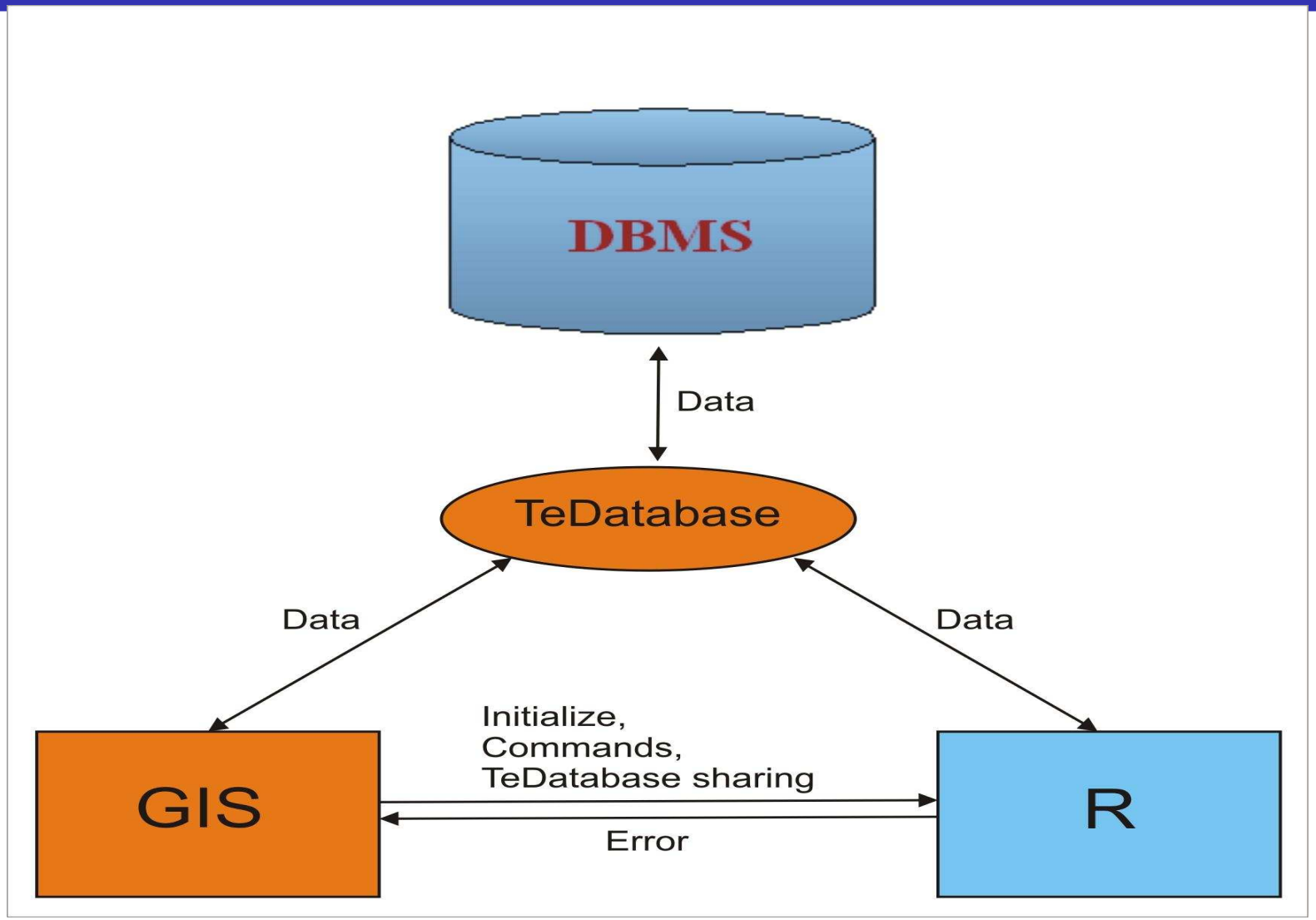
critical view of the data



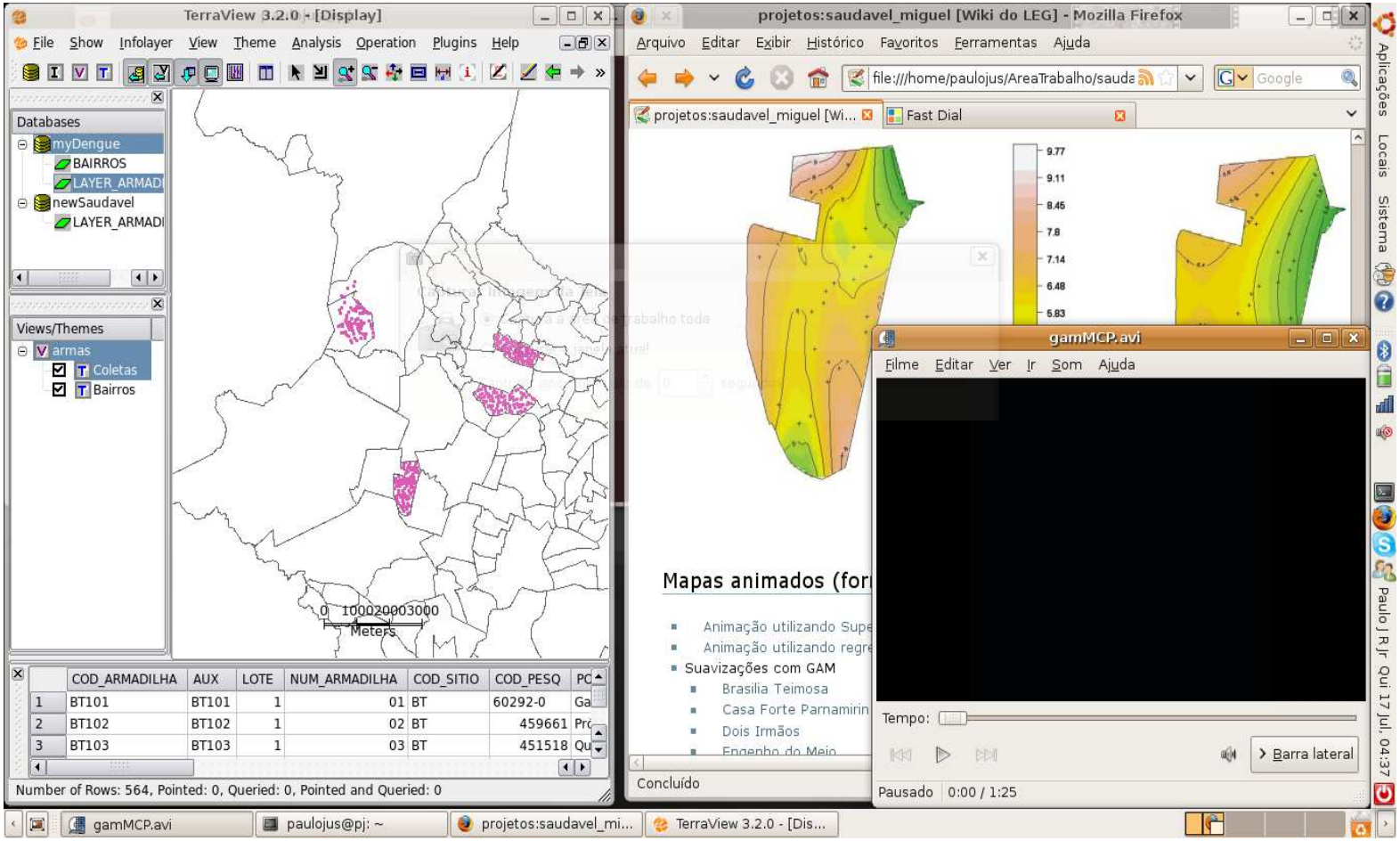
Technology transfer

- Implemeting methods:
  - environment for prototyping methods and analysis
  - R, packages and programing languages (fortran, C, C++, jags);
  - taylored/customised packages (RDengue, . . .)
- Computational structure
  - Linux (cross-plataform);
  - DBMS: (MySQL, postgres, etc);
  - GIS: Terralib e TerraView
  - **aRT project:** API R-Terralib;
  - automatic reports in various formats (Sweave/html): GIS, DBMS, WEB
- 100% **free software**, code and data will be made available
- rational usage of (scarse) human resources
- field implementation: SMCP project – live CD's

aRT


DBMS
Data
TeDatabase
Data
Data
Initialize,
Commands,
TeDatabase sharing
GIS
R
Error

Integrated environments



Acknowledgements

- Field and Lab teams at Health Secretary Recife/CVA and CPqAM/Fiocruz.
- The SAUDAVEL network.