Spatial Analysis

OIE Collaboration project

Workshop

OIE Collaborating Centre for epidemiology, training and control of emerging avian diseases

OIE Headquarters

Overview

- Where to retrieve spatial data
  - Sources of spatial data
- Type of data in spatial epidemiology
  - Points, polygons, lines or raster data
- Basic overview of spatial analysis

Spatial data in public health

- Management of territorial data
  - Location of farms
  - Landscape ecology
- Spatial aspects in disease surveillance
- Management of disease outbreaks and control measures
- Research
  - Spatial statistics
  - Risk Maps

Sources for animal and epidemiological data

- Specific research investigations
  - High quality but usually limited spatial extension
- Surveillance data
  - Field data
  - Laboratory data
  - Etc.
- National\International reporting systems
  - Animal Disease Notification System
  - WAHIS

Sources for spatial data

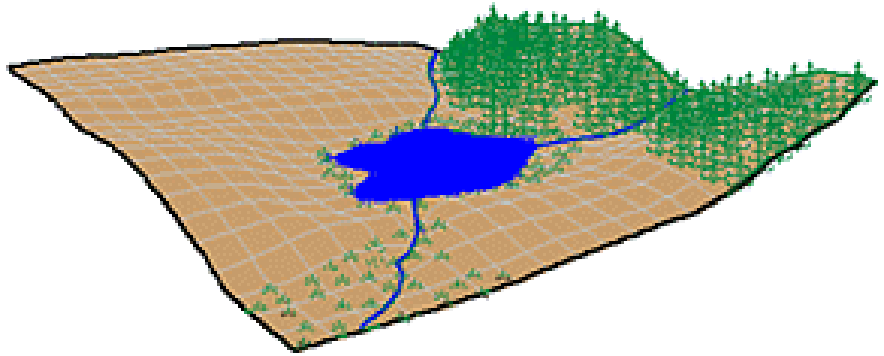
- Field data
  - GPS
- Surveillance data
  - Animal Registry
  - GIS
- Environmental data
  - Remote sensing

Spatial data formats

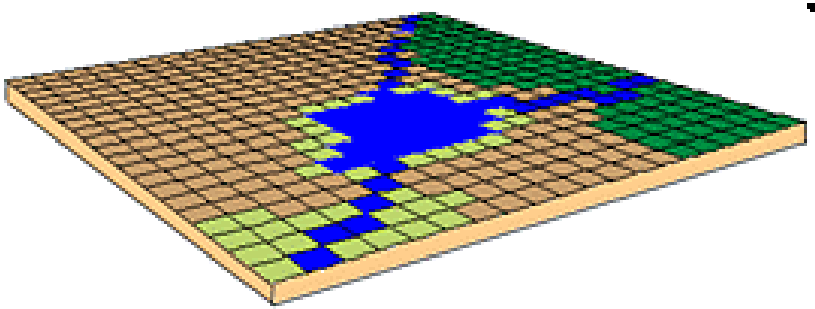
- Vector
  - Different feature types:
    - Points (Precise geographic coordinates)
    - Lines (artificial or natural features as roads and rivers)
    - Polygon (administrative boundaries, surveillance zones)
  - Aggregated vs Non-aggregated data
- Raster
  - Environmental data
  - Land Cover
  - Geostatistical elaborations (spatial interpolations)

} **Remote Sensing**

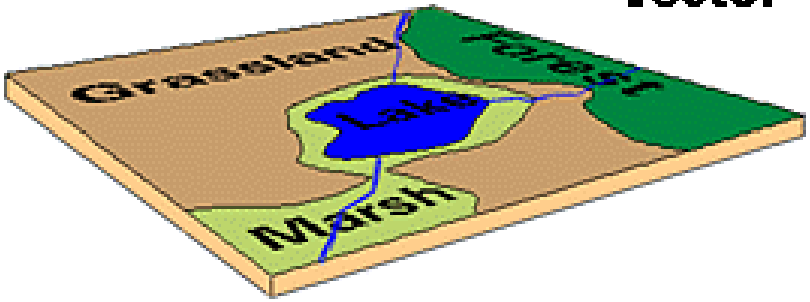
Real World



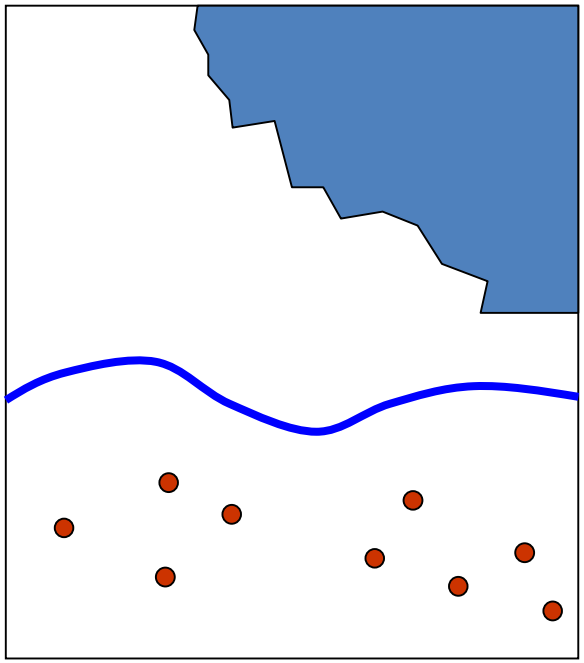
Raster / Image

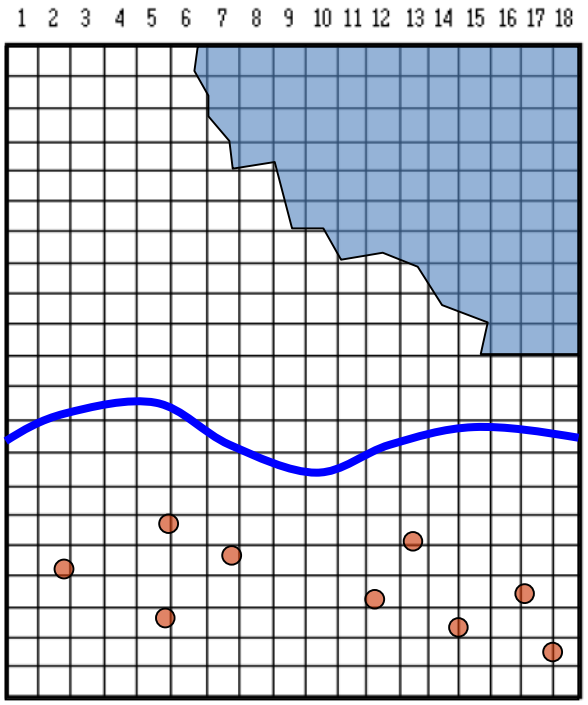


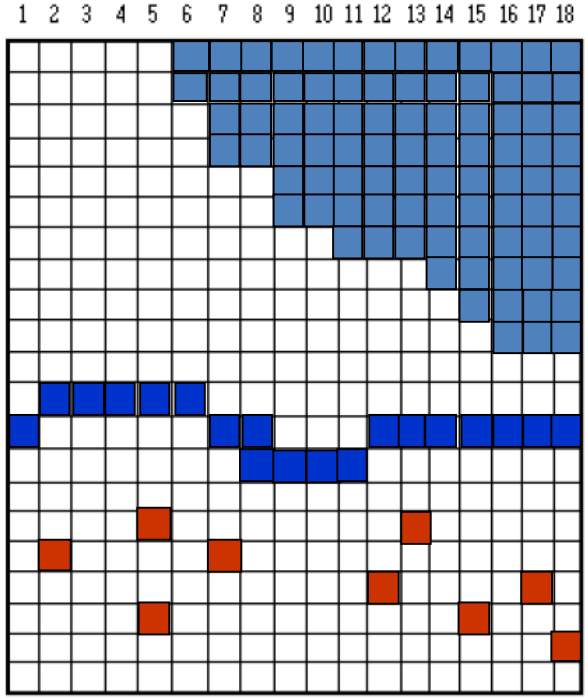
Vector

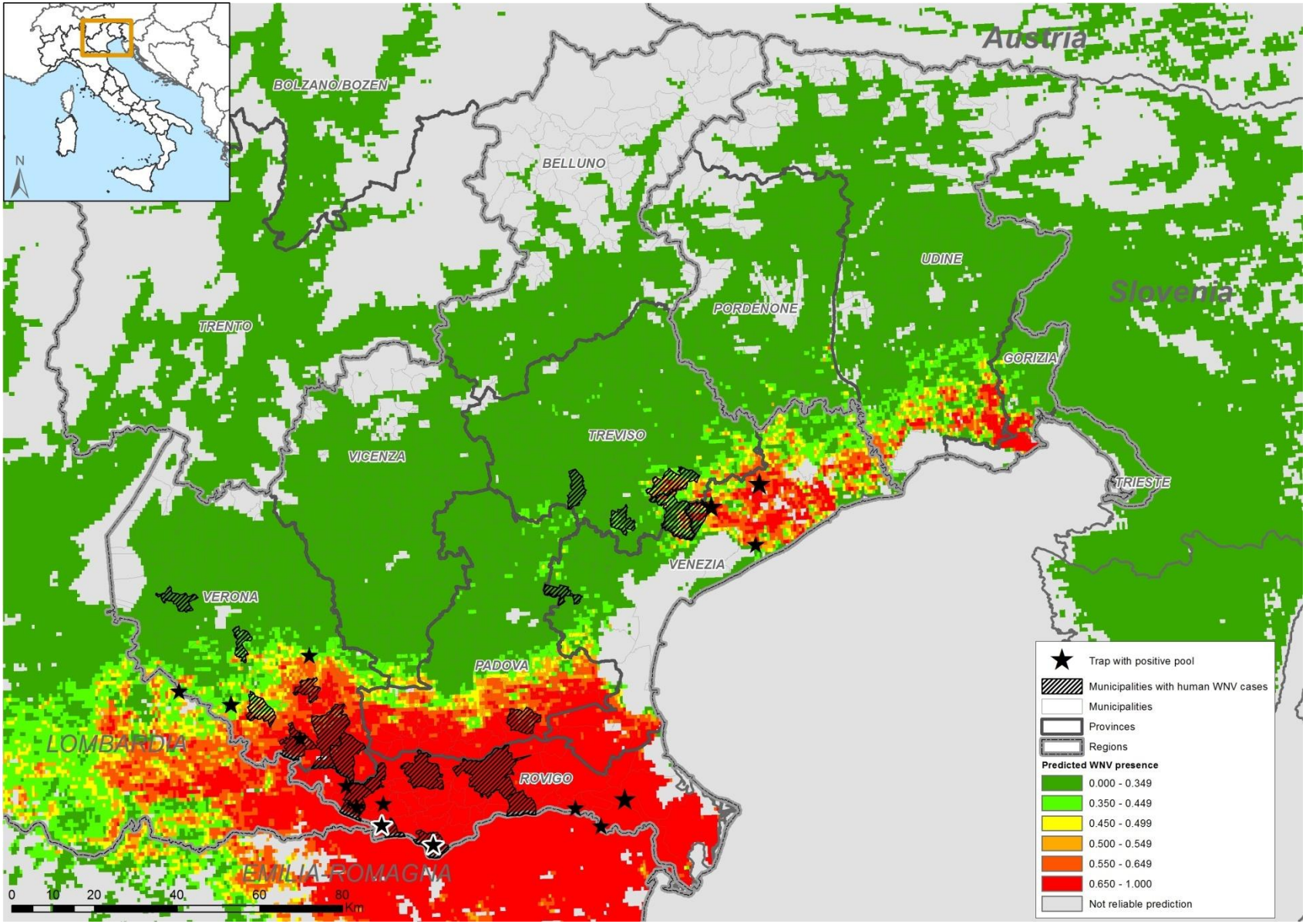


Raster




1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

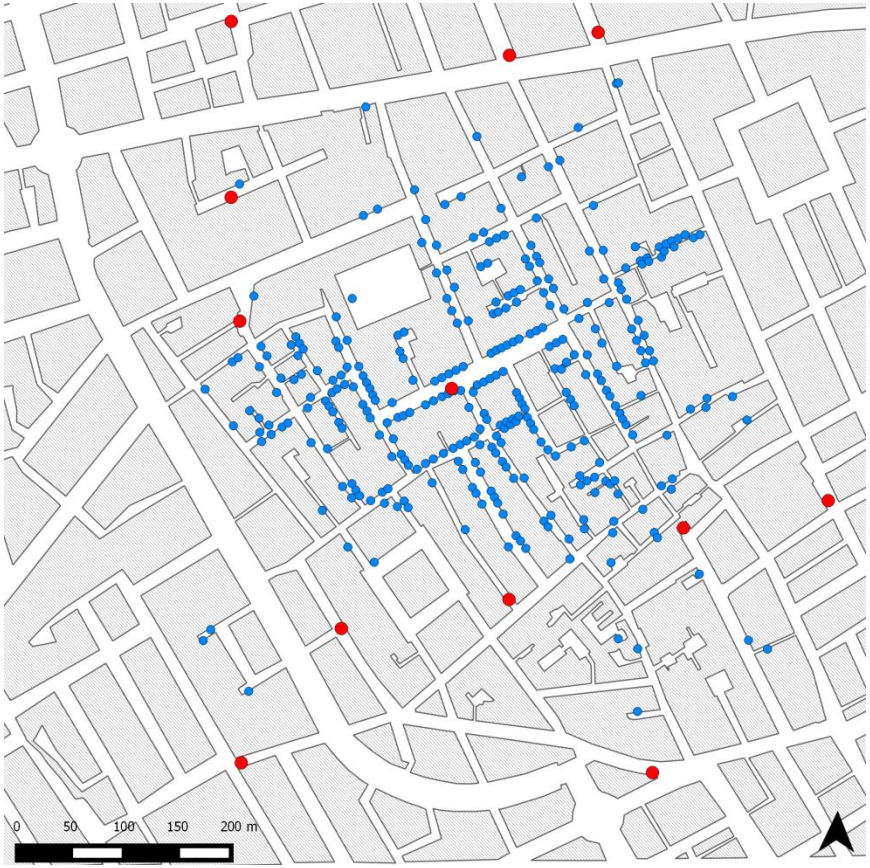

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18



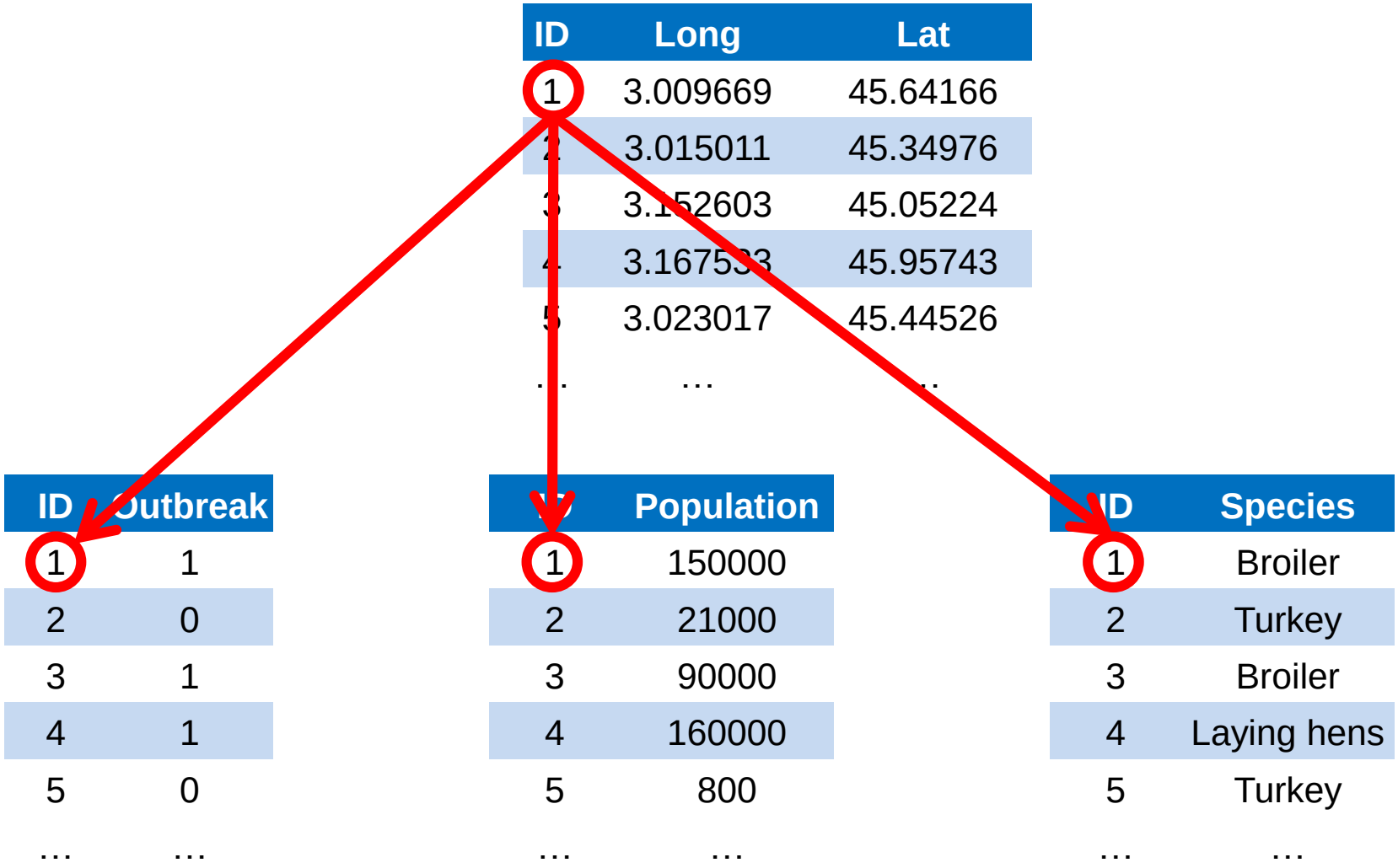
- Risk Maps for WNV in North-eastern Italy

How to integrate animal and epidemiological data with space?

ID	Long	Lat	Cases
1	-0.140795	51.515913	1
2	-0.140379	51.512679	3
3	-0.140304	51.512809	1
4	-0.140501	51.516230	2
5	-0.140375	51.515406	1
6	-0.140454	51.516284	2
7	-0.140722	51.518504	1
8	-0.140403	51.517066	1
9	-0.140196	51.515577	1
10	-0.139714	51.511973	1
…	…	…	…



How to integrate animal and epidemiological data with space?



ID	Long	Lat
1	3.009669	45.64166
2	3.015011	45.34976
3	3.152603	45.05224
4	3.167533	45.95743
5	3.023017	45.44526
…	…	…

ID	Outbreak
1	1
2	0
3	1
4	1
5	0
…	…

ID	Population
1	150000
2	21000
3	90000
4	160000
5	800
…	…

ID	Species
1	Broiler
2	Turkey
3	Broiler
4	Laying hens
5	Turkey
…	…

How to integrate animal and epidemiological data with space?

ID	Long	Lat	Outbreak	Population	Species
1	3.009669	45.64166	1	150000	Broiler
2	3.015011	45.34976	0	21000	Turkey
3	3.152603	45.05224	1	90000	Broiler
4	3.167533	45.95743	1	160000	Laying hens
5	3.023017	45.44526	0	800	Turkey
…	…	…	…	…	…

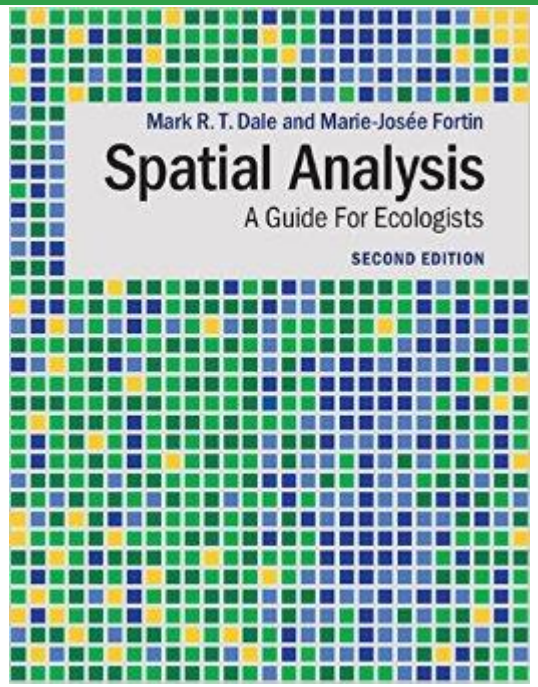
Spatial Analysis

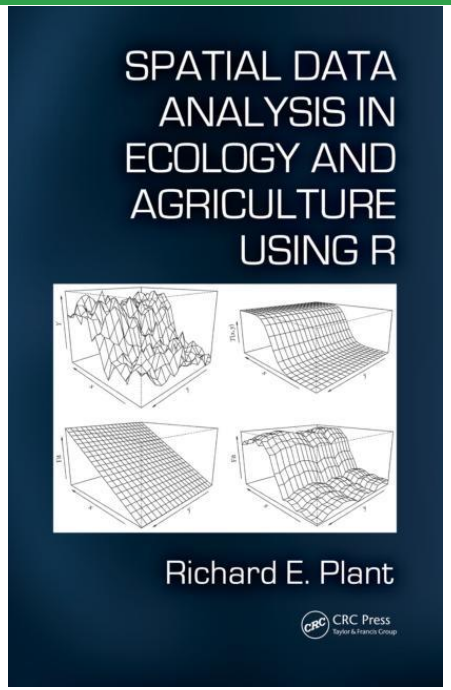
'Analytical techniques to determine the spatial distribution of a variable, the relationship betwee n the spatial distribution of variables, and the association of the variables of an area…
… It refers to the analysis of phenomena distributed in space and having physical dimensions' (Data West Research Agency)

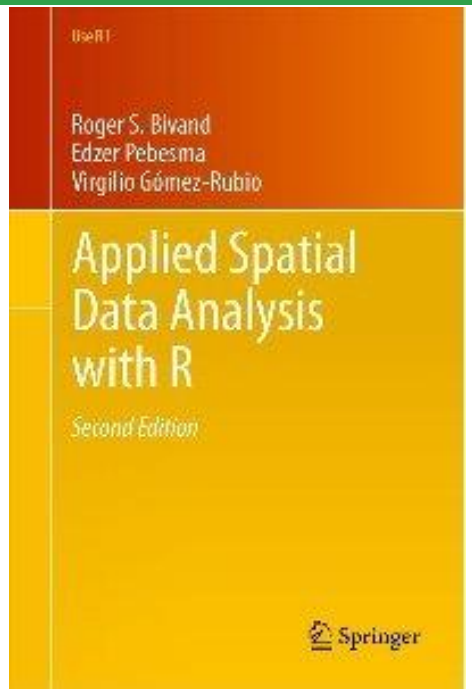
- Statistical/Mathematical methods accounting for geographical locations

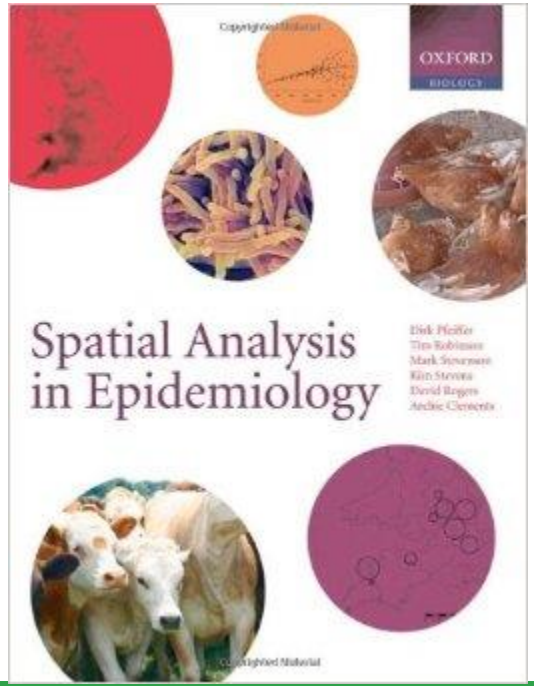
Spatial Analysis

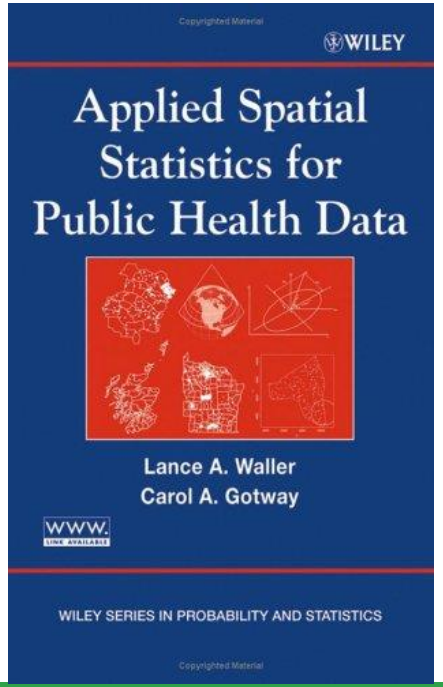
- Search for unexpected patterns of events
  - Random vs Clustered vs Dispersed
  - Globally vs Locally
  - 1st Order vs 2nd Order
  - Isotropic vs Anisotropic
  - Point Pattern Analysis vs Polygons
- Ecological analyses
  - Species distribution
  - Species abundance

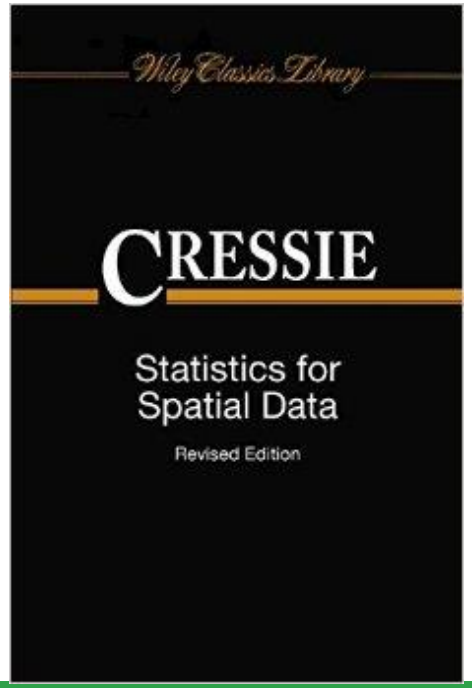

Mark R. T. Dale and Marie-Josée Fortin
Spatial Analysis
A Guide For Ecologists
SECOND EDITION


SPATIAL DATA ANALYSIS IN ECOLOGY AND AGRICULTURE USING R
Richard E. Plant
CRC Press


Use R!
Roger S. Bivand
Edzer Pebesma
Virgilio Gómez-Rubio
Applied Spatial Data Analysis with R
Second Edition
Springer


OXFORD
Spatial Analysis in Epidemiology


WILEY
Applied Spatial Statistics for Public Health Data
Lance A. Waller
Carol A. Gotway
WILEY SERIES IN PROBABILITY AND STATISTICS


Wiley Classics Library
CRESSIE
Statistics for Spatial Data
Revised Edition

Spatial statistics software

SaTScan™

Software for the spatial, temporal, and space-time scan statistics

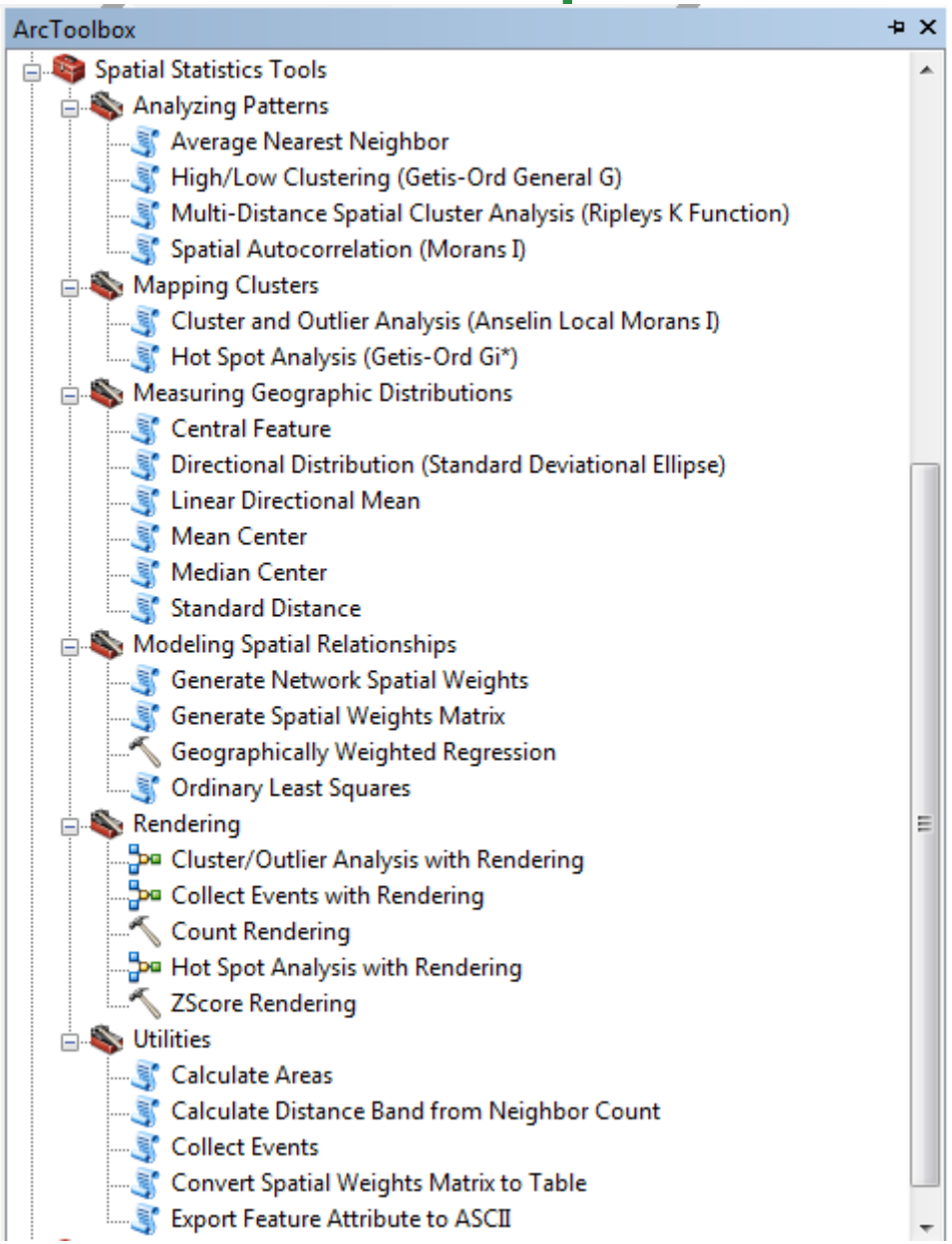
GeoDa™

ESRI ArcGIS™

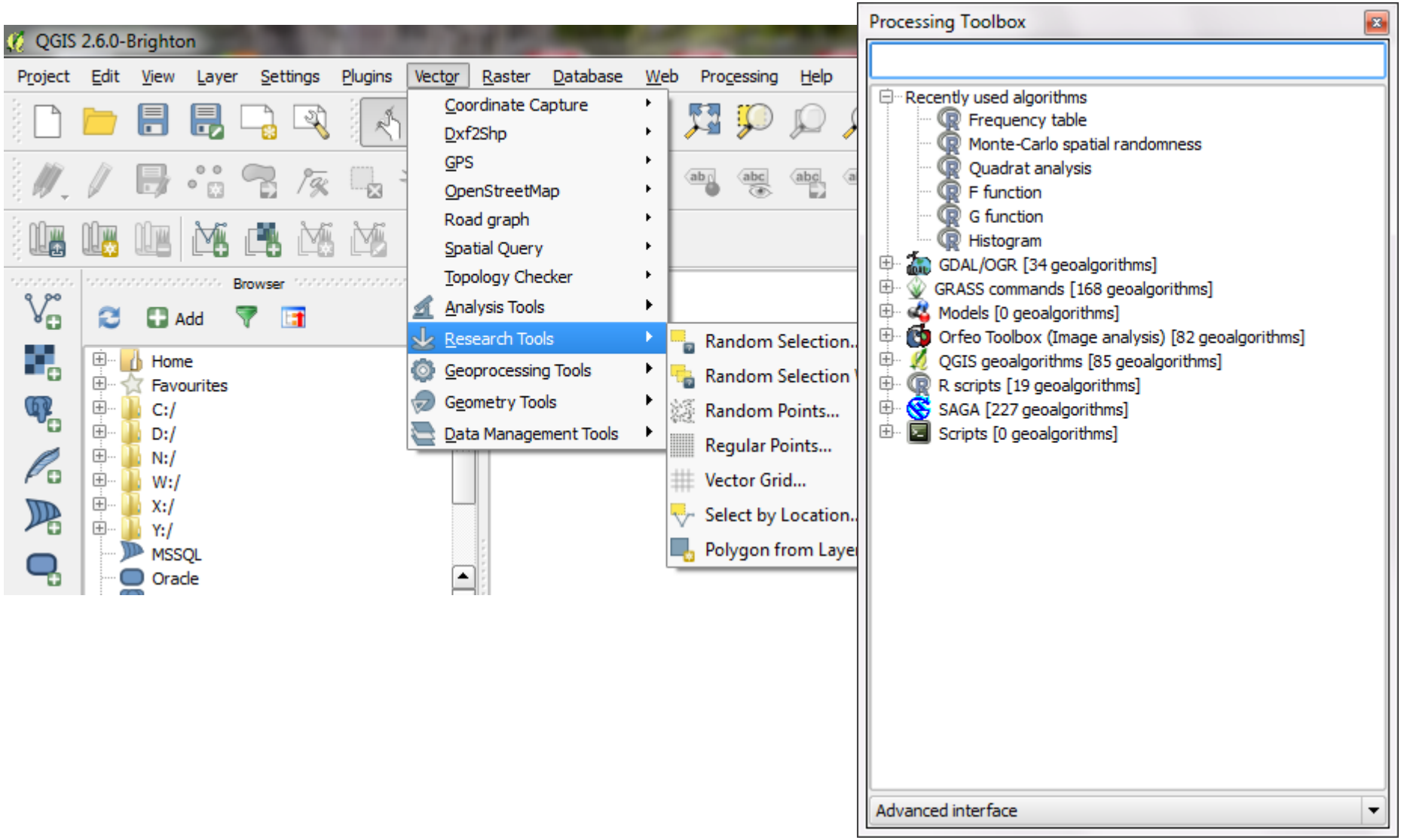
QGIS Trademark

R

ESRI – ArcInfo & Spatial Statistics

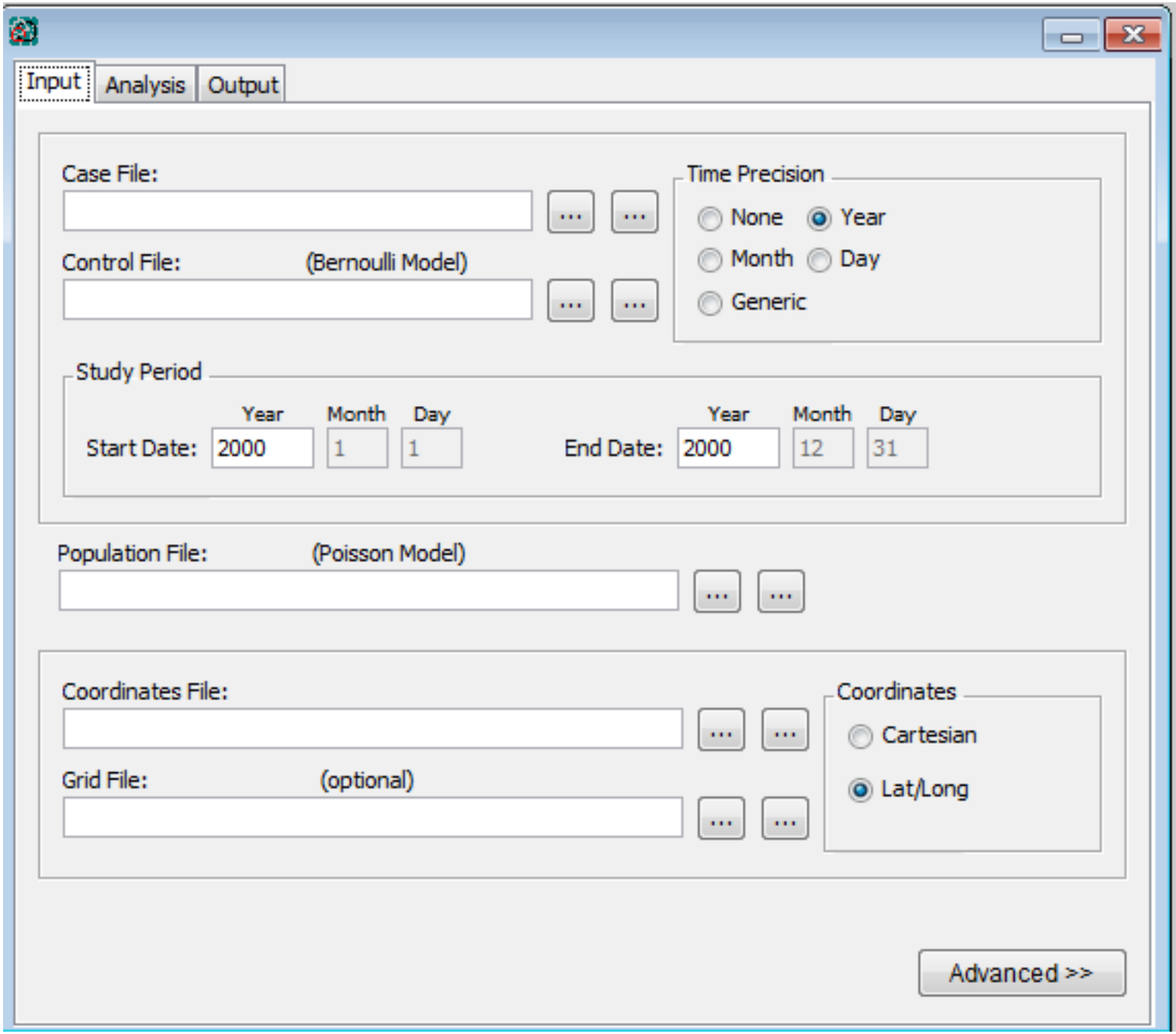


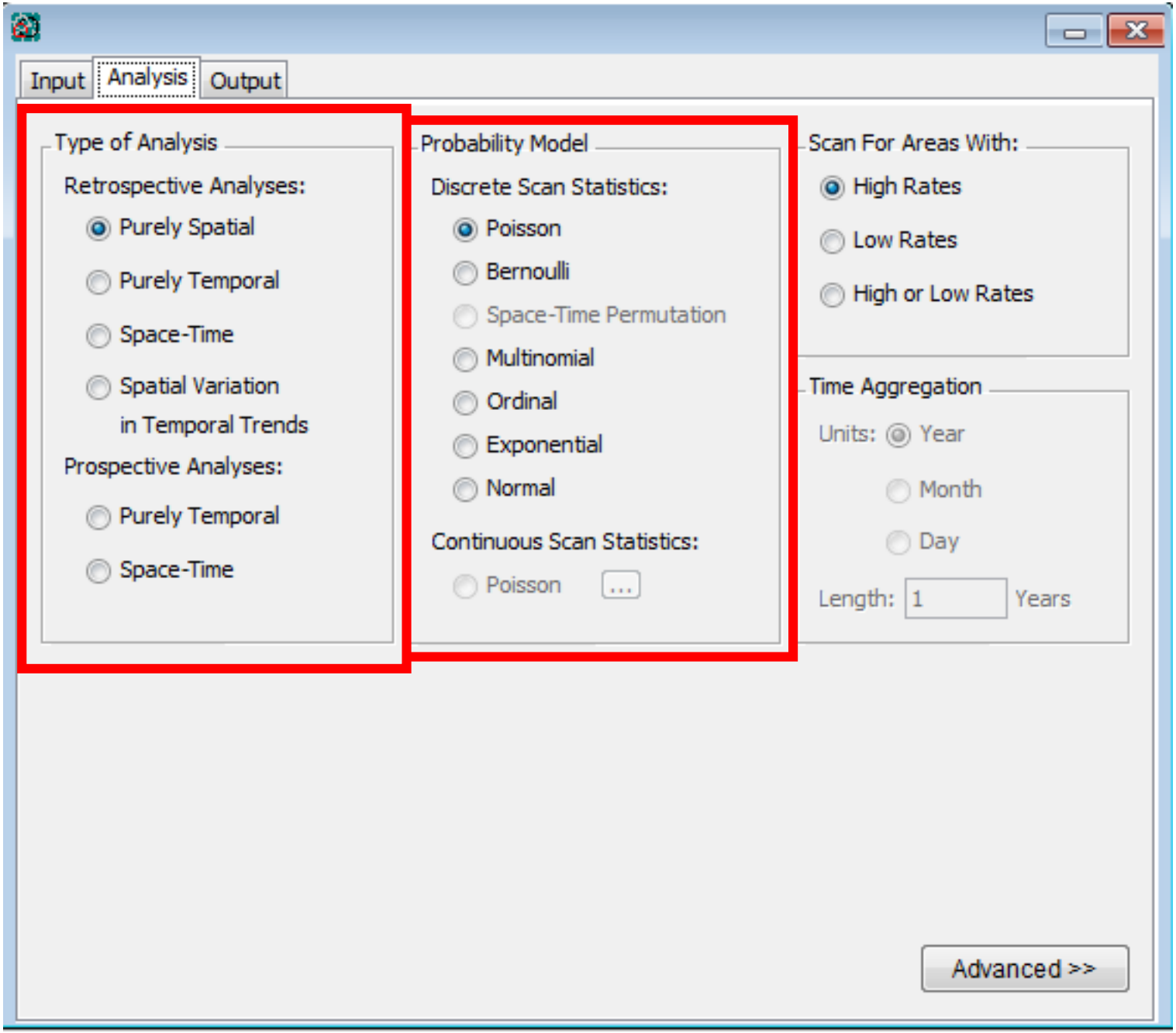
QGIS

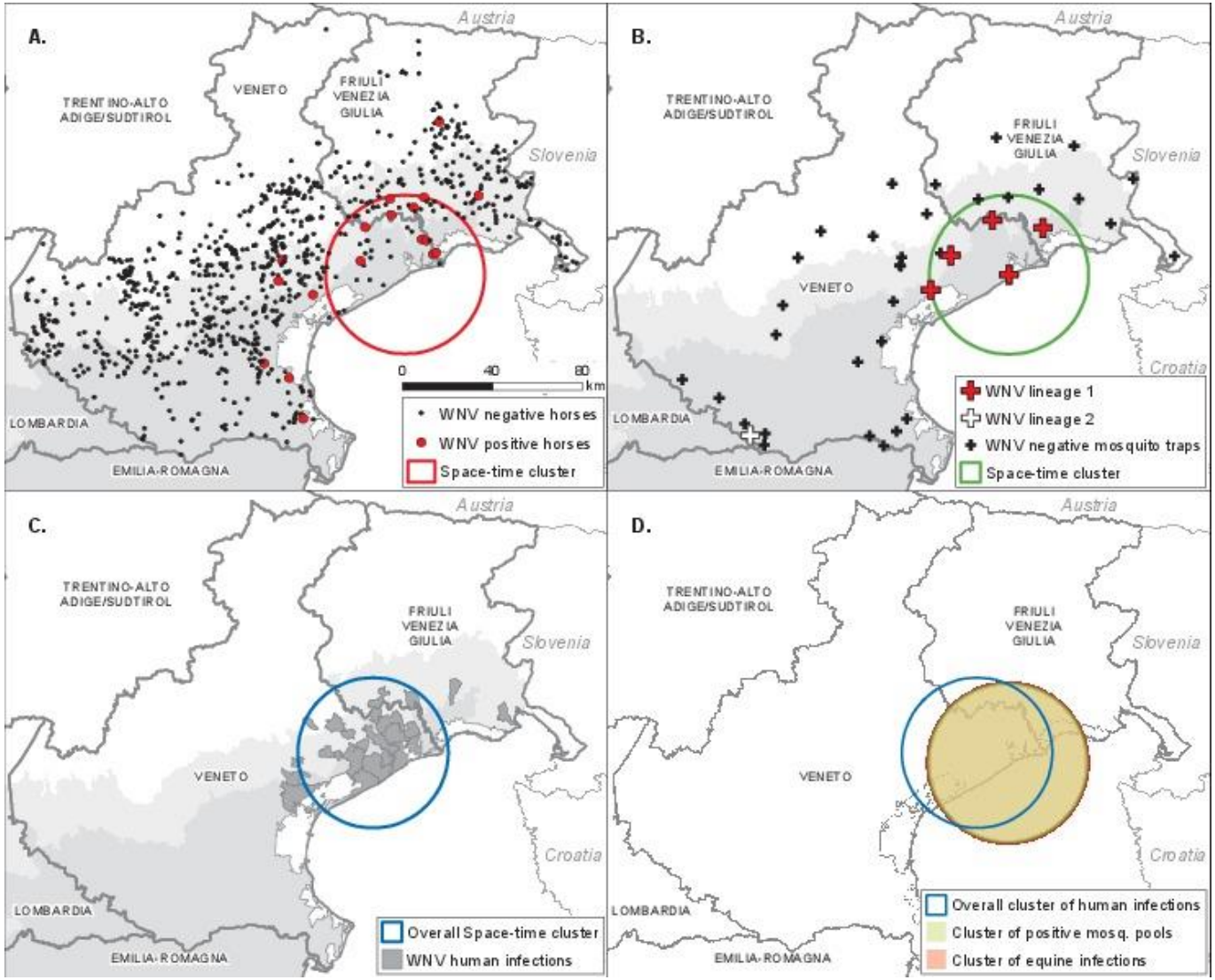

QGIS 2.6.0-Brighton
Project
Edit
View
Layer
Settings
Plugins
Vector
Raster
Database
Web
Processing
Help
Coordinate Capture
Dxf2Shp
GPS
OpenStreetMap
Road graph
Spatial Query
Topology Checker
Analysis Tools
Research Tools
Geoprocessing Tools
Geometry Tools
Data Management Tools
Random Selection...
Random Selection
Random Points...
Regular Points...
Vector Grid...
Select by Location...
Polygon from Layer
Browser
Add
Home
Favourites
C:/
D:/
N:/
W:/
X:/
Y:/
MSSQL
Oracle
Processing Toolbox
Recently used algorithms
Frequency table
Monte-Carlo spatial randomness
Quadrat analysis
F function
G function
Histogram
GDAL/OGR [34 geoalgorithms]
GRASS commands [168 geoalgorithms]
Models [0 geoalgorithms]
Orfeo Toolbox (Image analysis) [82 geoalgorithms]
QGIS geoalgorithms [85 geoalgorithms]
R scripts [19 geoalgorithms]
SAGA [227 geoalgorithms]
Scripts [0 geoalgorithms]
Advanced interface

SaTScan

- Easily performs [only] Scan Statistics analyses
  - Wizard to import data in the software
    - Cases
    - Population
    - Controls
    - Coordinates
    - Likely foci (points aroudn which clustering is expected)
  - Possibility to directly export shapefiles (ESRI) or kml files (Google Earth)
    - Need to use lon/lat coordinate systems


Input
Analysis
Output
Case File:
Control File:
(Bernoulli Model)
Time Precision
None
Year
Month
Day
Generic
Study Period
Year
Month
Day
Start Date:
2000
1
1
End Date:
2000
12
31
Population File:
(Poisson Model)
Coordinates File:
Grid File:
(optional)
Coordinates
Cartesian
Lat/Long
Advanced >>


Input
Analysis
Output
Type of Analysis
Retrospective Analyses:
Purely Spatial
Purely Temporal
Space-Time
Spatial Variation in Temporal Trends
Prospective Analyses:
Purely Temporal
Space-Time
Probability Model
Discrete Scan Statistics:
Poisson
Bernoulli
Space-Time Permutation
Multinomial
Ordinal
Exponential
Normal
Continuous Scan Statistics:
Poisson
...
Scan For Areas With:
High Rates
Low Rates
High or Low Rates
Time Aggregation
Units: Year
Month
Day
Length: 1 Years
Advanced >>



- Risk Maps for refining WNV surveillance

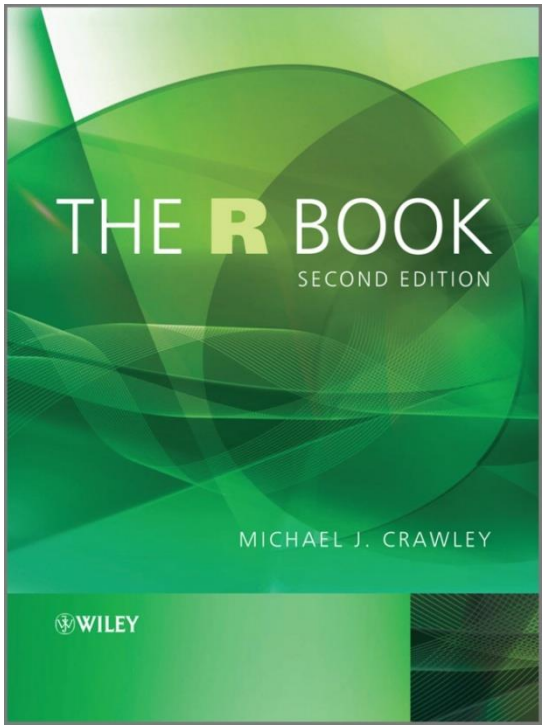
What's  ?

- A computer language and environment for statistical computing and graphics

- Related to the computing language S, which has been developed into a commercial product (S-PLUS)

R remains free

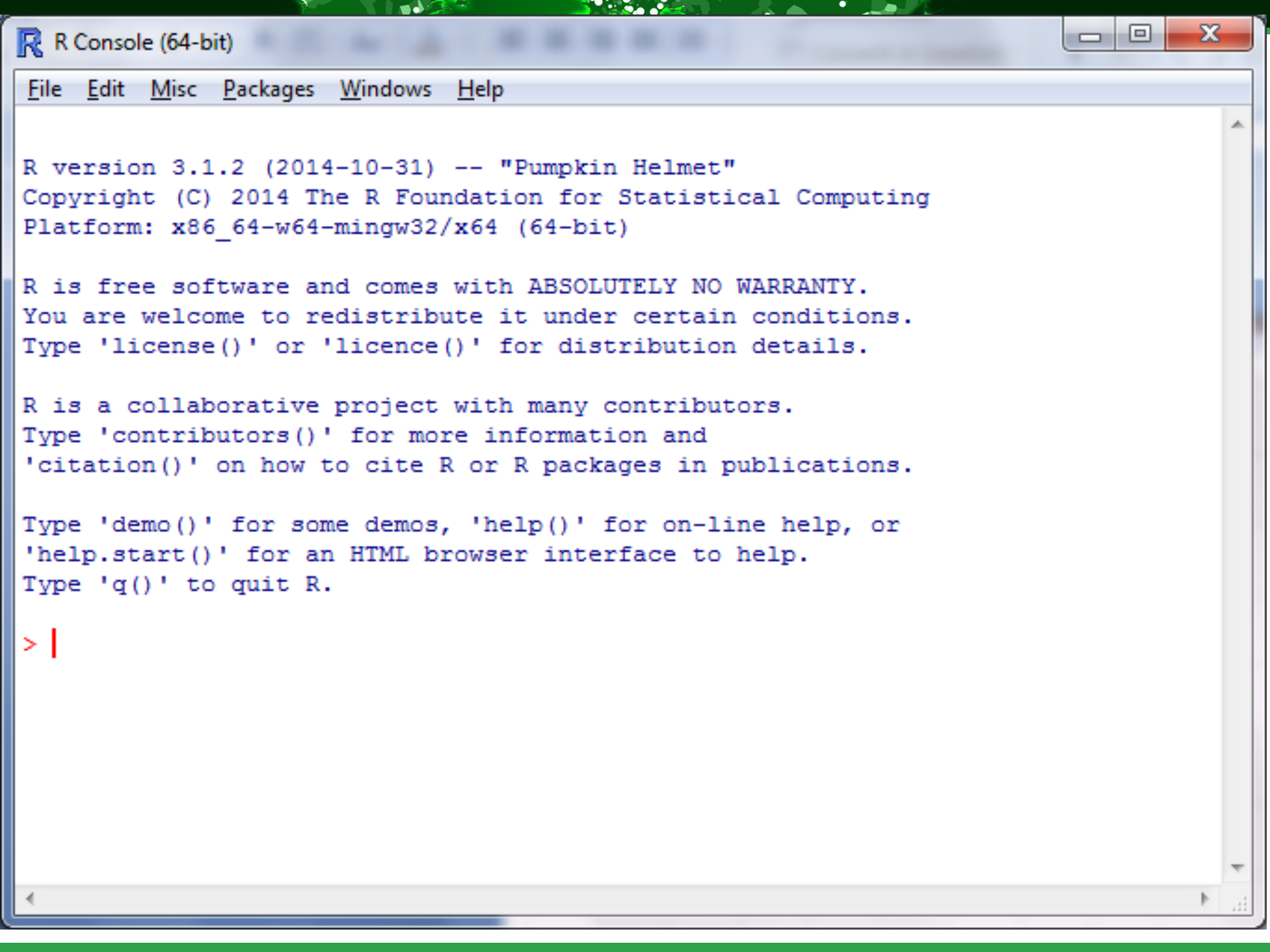
R – Pro's

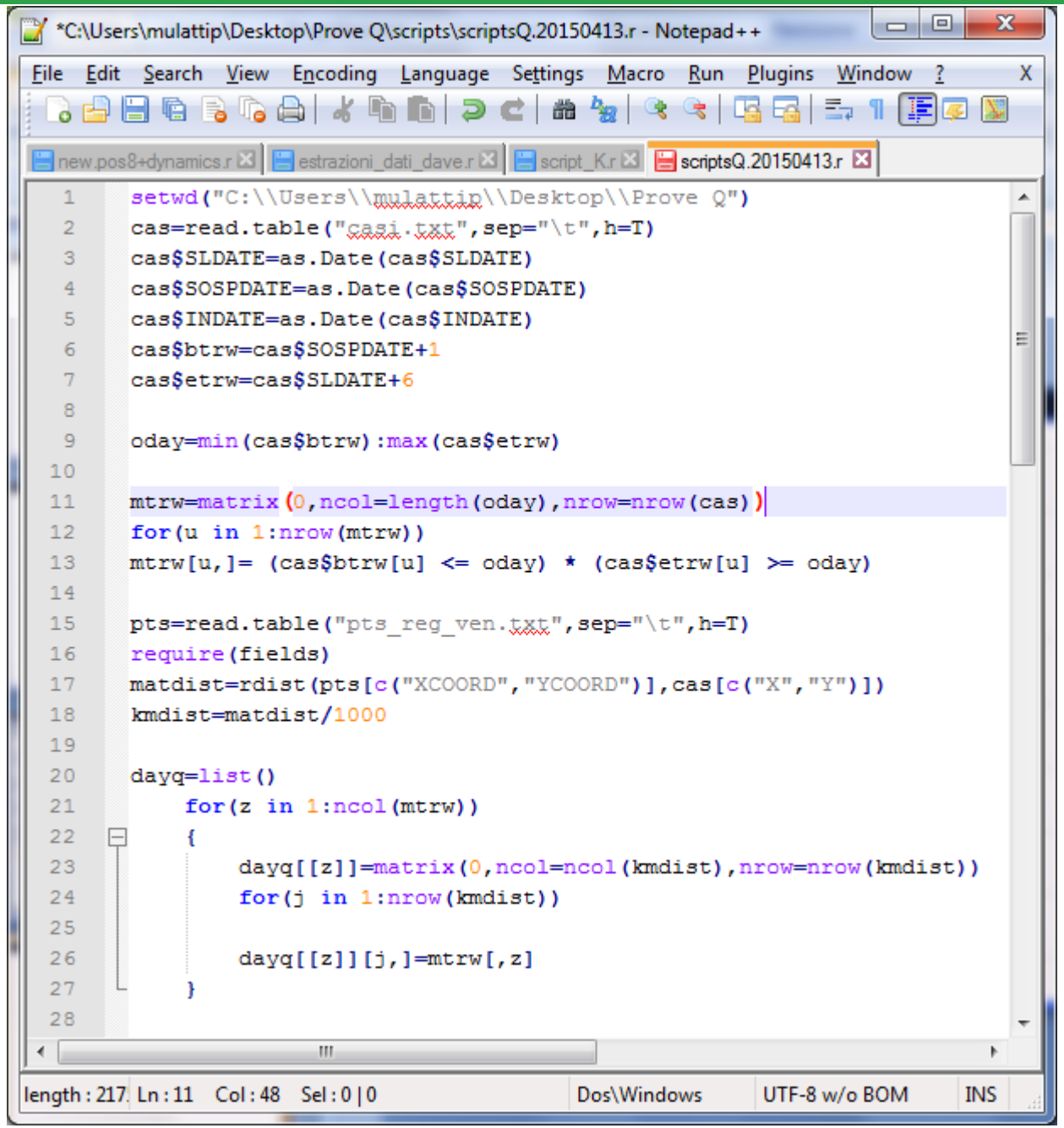


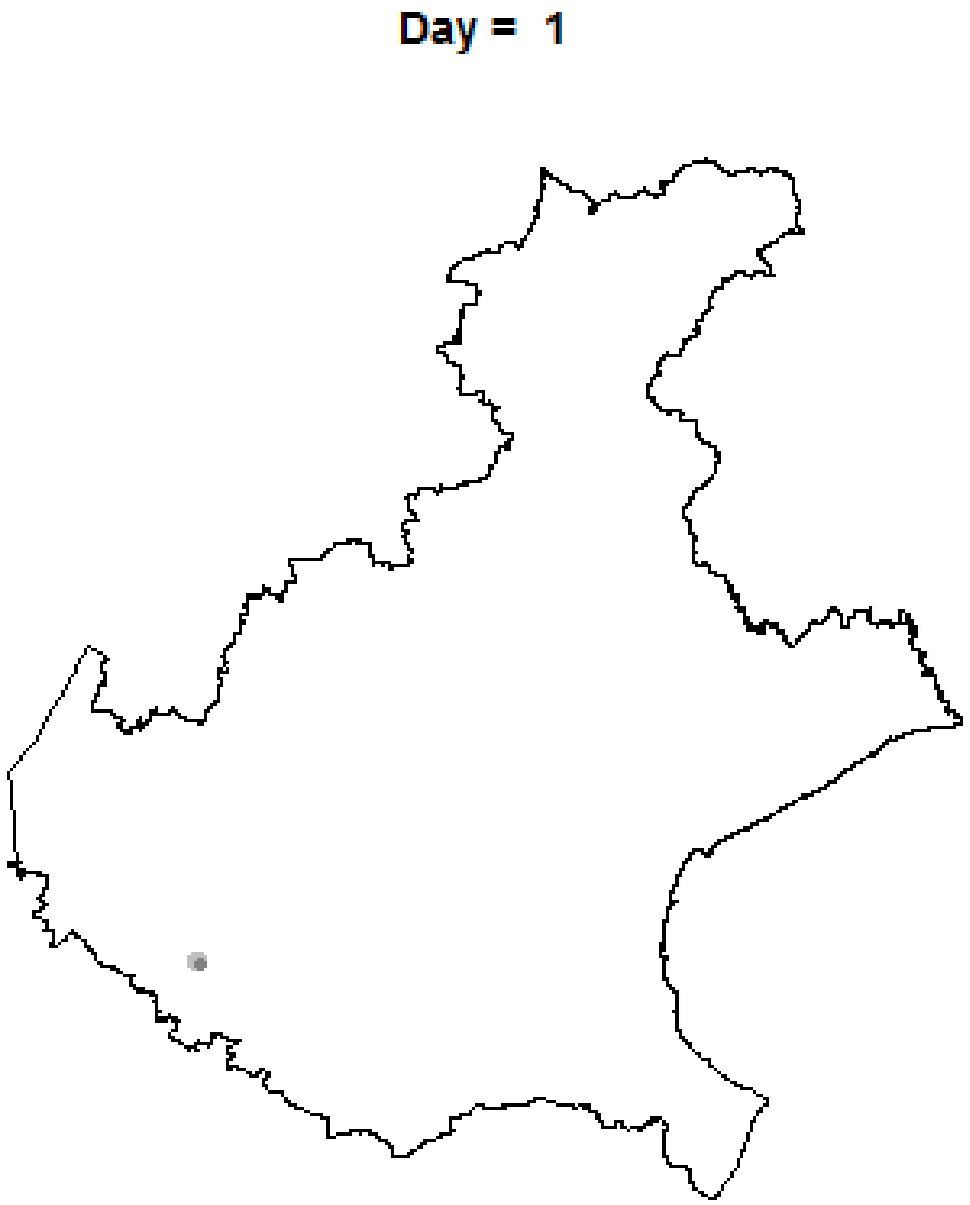
- Highly flexible and versatile
  - Data exploration and visualisation
  - Data manipulation
  - Modeling
- Well supported and documented
- Has been officially recognised in the academic sector
- It is free!

– Con's

- Hard to get into
  - Steep learning curve
- Not completely user-friendly
  - Plenty of help can be found on the web
  - Help is "embedded" for each function
  - Quite good error reporting (but it needs to get used to)


R Console (64-bit)
File Edit Misc Packages Windows Help
R version 3.1.2 (2014-10-31) -- "Pumpkin Helmet"
Copyright (C) 2014 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
>


*C:\Users\mulattip\Desktop\Prove Q\scripts\scriptsQ.20150413.r - Notepad++
File Edit Search View Encoding Language Settings Macro Run Plugins Window ?
new.pos8+dynamics.r
estrazioni_dati_dave.r
script_K.r
scriptsQ.20150413.r
1 setwd("C:\\Users\\mulattip\\Desktop\\Prove Q")
2 cas=read.table("casi.txt",sep="\t",h=T)
3 cas$SLDATE=as.Date(cas$SLDATE)
4 cas$SOSPDATE=as.Date(cas$SOSPDATE)
5 cas$INDATE=as.Date(cas$INDATE)
6 cas$btrw=cas$SOSPDATE+1
7 cas$etrw=cas$SLDATE+6
8
9 oday=min(cas$btrw):max(cas$etrw)
10
11 mtrw=matrix(0,ncol=length(oday),nrow=nrow(cas))
12 for(u in 1:nrow(mtrw))
13 mtrw[u,]= (cas$btrw[u] <= oday) * (cas$etrw[u] >= oday)
14
15 pts=read.table("pts_reg_ven.txt",sep="\t",h=T)
16 require(fields)
17 matdist=rdist(pts[c("XCOORD","YCOORD")],cas[c("X","Y")])
18 kmdist=matdist/1000
19
20 dayq=list()
21 for(z in 1:ncol(mtrw))
22 {
23 dayq[[z]]=matrix(0,ncol=ncol(kmdist),nrow=nrow(kmdist))
24 for(j in 1:nrow(kmdist))
25
26 dayq[[z]][j,]=mtrw[,z]
27 }
28
length : 217 Ln : 11 Col : 48 Sel : 0 | 0
Dos\Windows
UTF-8 w/o BOM
INS



Example

- LPAI H9N2 in turkey farms in Veneto region
- 143 Infected Farms (2012-2013)

- Risk of “Neighbourhood Infection”

$$Q_{ikt} = \sum_{j=1;j\neq i}^{N-1} \eta_{ijkt} c_{jt} w_{ij},$$

Mulatti et al, Pre Vet Med 95 (2010) 267-274

Spatial Analysis

- Candidate requirement
  - Knowledge of statistics:
    - Statistical distributions
    - Statitical models (GLM, GLMM)
  - Basic programming abilities
    - Previous (limited) exposure to programming
    - R (desiderable)
  - Basic knowledge of GIS methods
    - Not strictly required, but helpful

Thanks for your attention!

Scan Statistics

- Place circular (purely spatial) or cylindrical (space-time) windows on the study area
- Calculate the “excess” of events within each scanning window
- Based on comparison between observed vs expected events
- Account for the population density
- Statistica signigicance obtained through simulations