Advanced GIS Virtual Training Course- Introduction to Spatial Cluster Analysis

Chris Compton and Art Subharat, EpiCentre, Massey University, New Zealand

August 2021

Overview

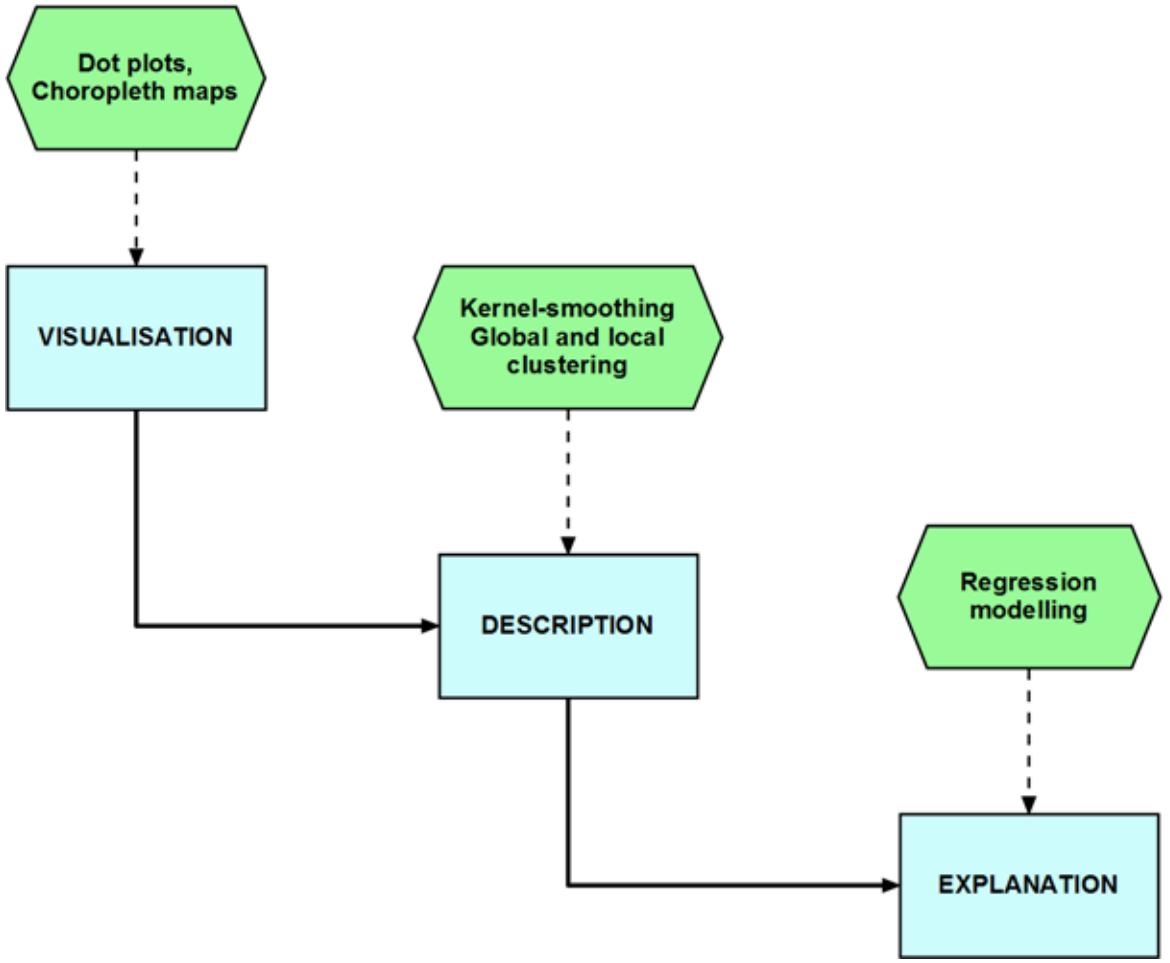
Session outline

- Day 03 - 04 review
- Introduction to spatial clustering
- Exercises on spatial clustering
- Wrap-up

Day 05 timetable

Times	Activities
10:00 - 10:20	Review Day 03 - 04
10:20 - 10:35	Spatial clustering presentation
10:25 - 11:15	Exercises- 3.1, 3.2 and extension
11:15 - 11:25	Group feedback to class
11:25 - 11:30	Wrap-up

‘Clustering’ in spatial analysis workflow



Resources

- Notes with exercises “IntroductionToClusterDetection.docx”
- Presentation pdf file “IntroductionToClusterDetection-Pres.pdf”
- Course discussion forum

Introduction

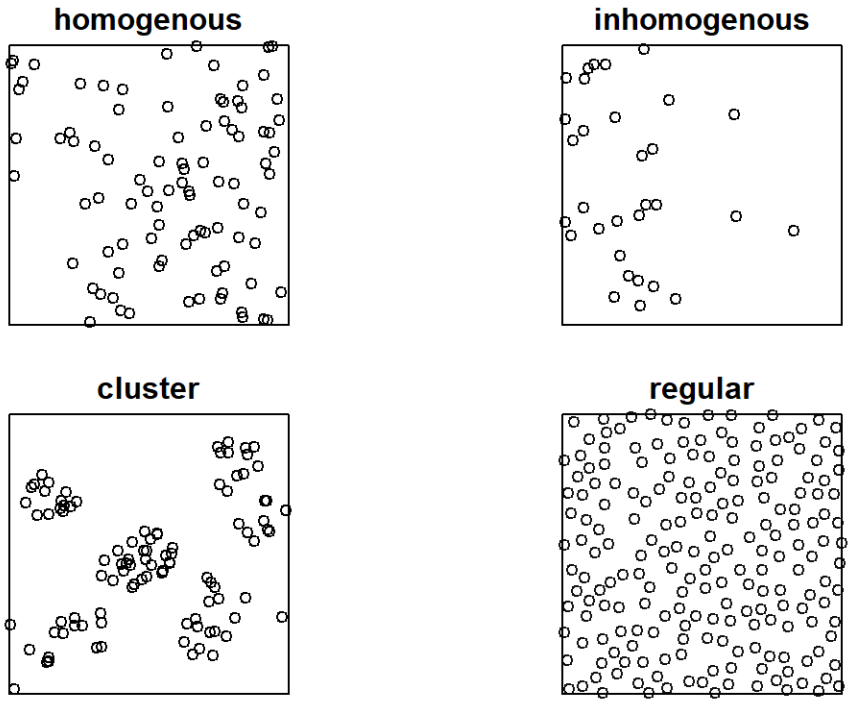
Why investigate clustering?

- Surveillance
  - Are patterns of disease incidence changing?
- Response to cluster alerts
  - Are further responses needed?
- Investigate disease aetiology
  - Generate hypotheses for further studies/control

Key question in spatial clustering analyses

- Are differences in occurrence of disease
  - 'Global' or large-scale?
    - Overall trend in variation in density of disease ("inhomogeneous")
    - e.g. due vaiation in large-scale geographic/climatic determinants of disease
    - **1st order process**
  - Local scale?
    - Subjects in proximity interact with one-another ('dependence')
    - Due to
      - Infectious nature of disease causing local spread
      - Alikeness of other local factors
      - **2nd order process**

Spatial patterns



Methods

Methods to evaluate clustering (autocorrelation)

- Global methods (for autocorrelation)
  - Point data
    - Ripley's *K*-function difference
  - Areal data
    - Global Moran's *I*
- Localised non-focused methods
  - Point data
    - Spatial scan statistic (areal data also)

Clustering of point data

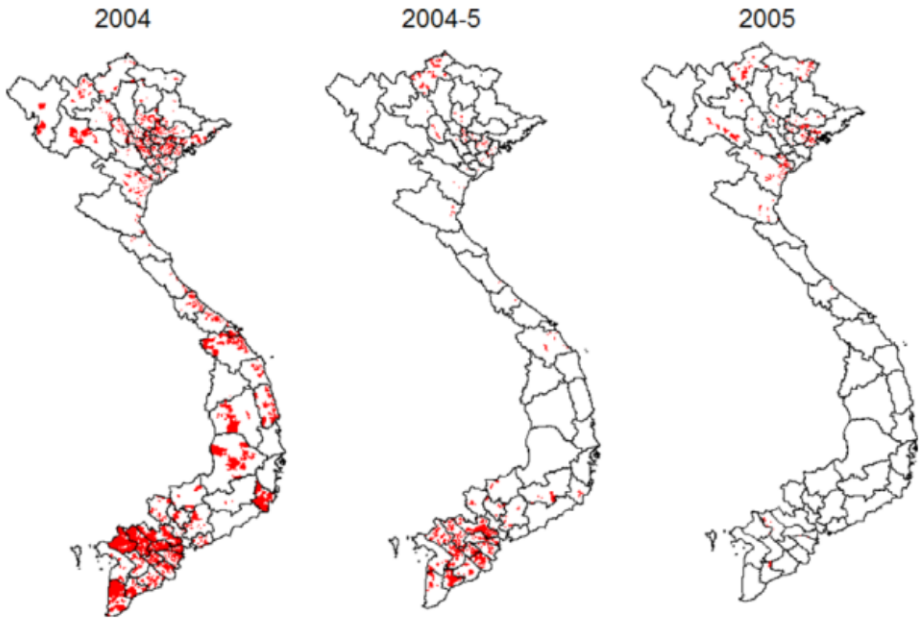
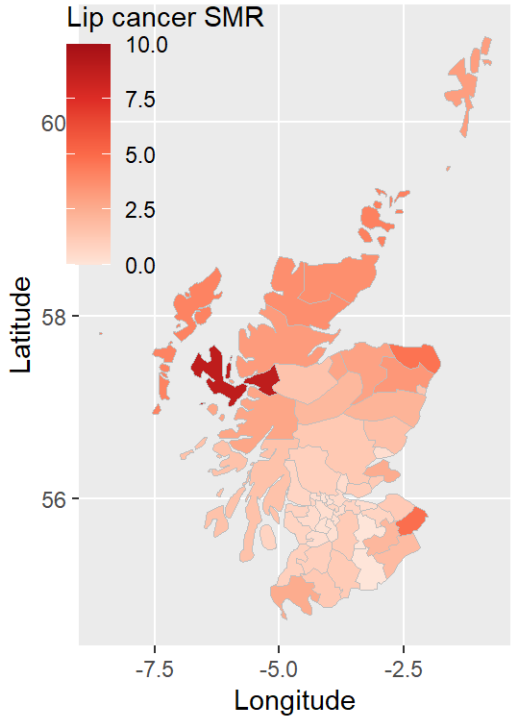


Fig. 2. Spatial pattern of affected communes for each of the three epidemic periods for HPAI outbreaks in Vietnam (province boundaries also shown).

from Pfeiffer et al. 2007 The Vet. J.

Clustering of areal data

- Lip cancer SMR Scotland



Exercises

Exercises

- Exercises 3.1 & 3.2 from “IntroductionToClusterDetection.docx”
- Repeat above analysis for Mandalay/Nay Phi Taw data set
  - Download from Stream - Cluster analysis of spatial data - Files of teaching material and data sets for cluster analysis -

Group feedback

Wrap-up

- Cluster analysis needs to be informed by the data and biology
  - Follow ‘visualise-describe-explain’ steps
- Be aware of limitations of data
  - Same problems with bias as with other epidemiologic studies