modern statistical methods for spatio-temporal data – understanding patterns in space and time

Janine Illian

University of St Andrews,UK and NTNU Trondheim, Norway

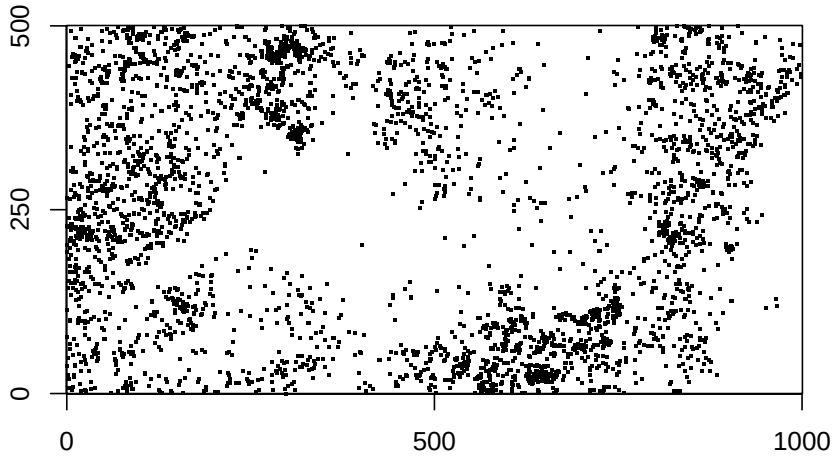
April 2015

spatial patterns

What do muskoxen, the rainforest, crimes and terrorism have in common?

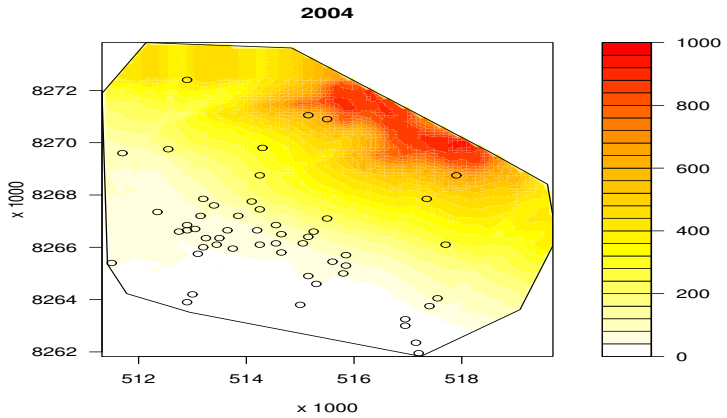
spatial patterns

What do muskoxen, the rainforest, crimes and terrorism have in common?
They all form **spatial patterns**...



spatial patterns

What do muskoxen, the rainforest, crimes and terrorism have in common?
They all form **spatial patterns**...

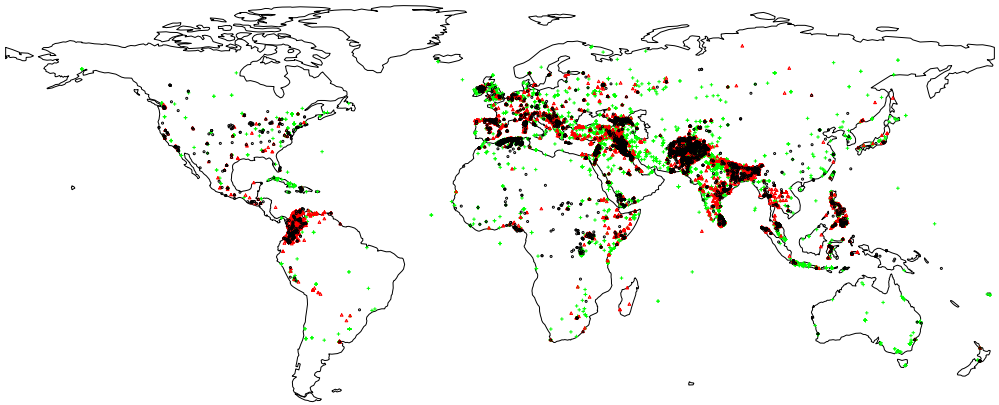


spatial patterns

What do muskoxen, the rainforest, crimes and terrorism have in common?
They all form **spatial patterns**...

spatial patterns

What do muskoxen, the rainforest, crimes and terrorism have in common?
They all form **spatial patterns**...



making sense of spatial patterns

aim

spatial pattern analysis:
understanding underlying mechanisms by analysing the (abstract) spatial pattern

spatial data I

spatial statistics...

- analyses data collected "in space"

spatial data I

spatial statistics...

- analyses data collected "in space"
- **not** this kind of space:

spatial data I

spatial statistics...

- analyses data collected "in space"
- **not** this kind of space:

spatial data I

spatial statistics...

- analyses data collected "in space"
- **not** this kind of space:

- but data for which the **sampling location is known and "relevant"**

spatial data II

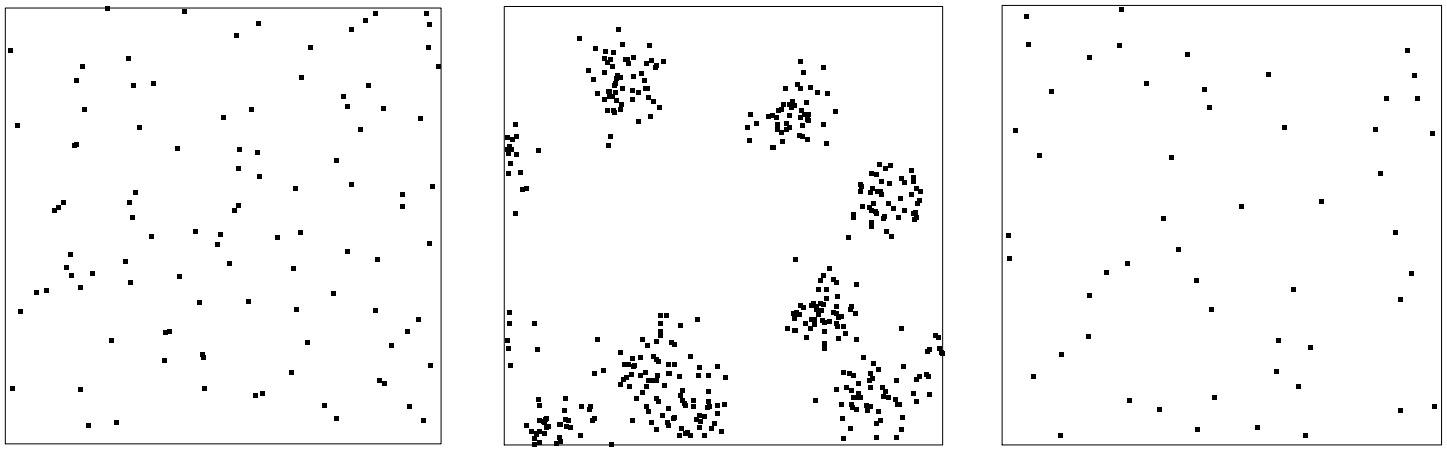
well, what do we mean by **relevant**?

two scenarios:

1. spatial structure in the data is of interest;
   aim is to **describe and analyse** the spatial structure
2. spatial structure in the data is a "nuisance";
   aim is to account for the spatial structure or **spatial autocorrelation**

We typically have to deal with both issues...

spatial structure...?



random, clustered and regular patterns

making sense of spatial patterns of crime

aim

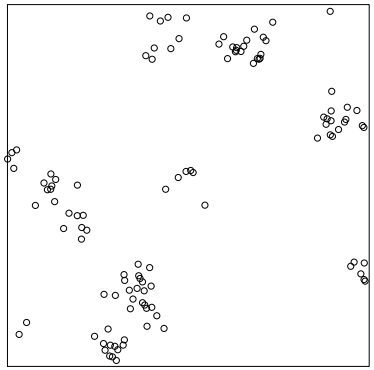
spatial statistics understanding spatial structures in patterns of crime by identifying clusters ("hotspots")

clustering revisited

Finding clusters – why do we need statistics – can't we just see them?

Let's see...

clustered pattern (artificial data)

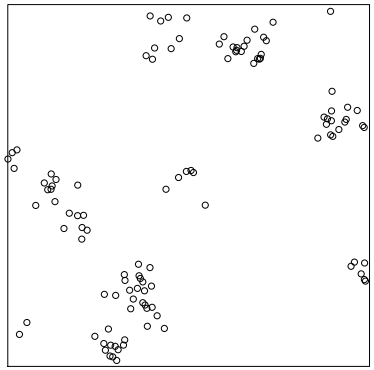


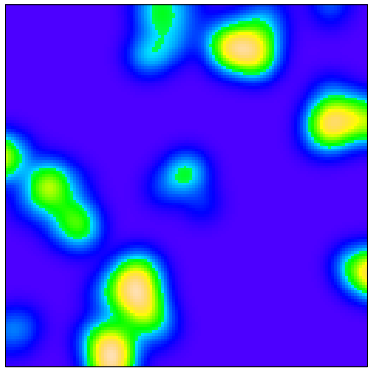
clustering revisited

Finding clusters – why do we need statistics – can't we just see them?

Let's see...

clustered pattern (artificial data)



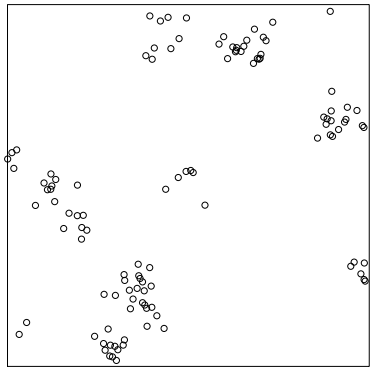


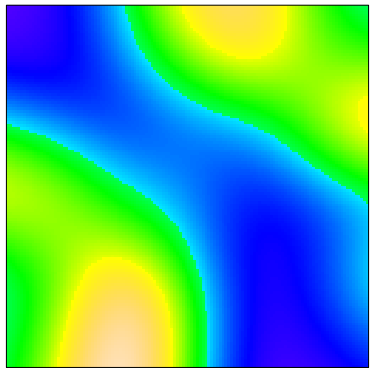
clustering revisited

Finding clusters – why do we need statistics – can't we just see them?

Let's see...

clustered pattern (artificial data)



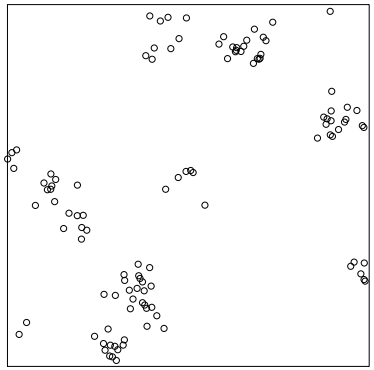


clustering revisited

Finding clusters – why do we need statistics – can't we just see them?

Let's see...

clustered pattern (artificial data)

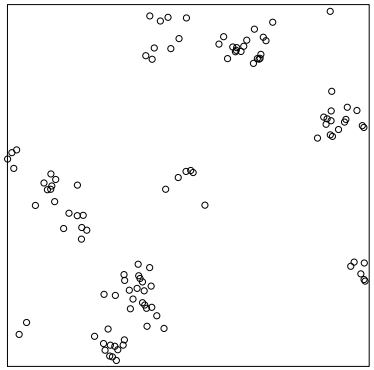


clustering revisited

Finding clusters – why do we need statistics – can't we just see them?

Let's see...

clustered pattern (artificial data)



clustering revisited

- simple smoothing/interpolation techniques can find clusters

clustering revisited

- simple smoothing/interpolation techniques can find clusters
- **however:** degree of smoothing and plotting method determines if we can find them or not...

clustering revisited

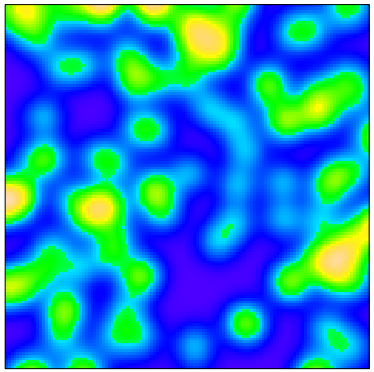
- simple smoothing/interpolation techniques can find clusters
- **however:** degree of smoothing and plotting method determines if we can find them or not...
- not good...

clustering revisited

Finding clusters – are they really there?

Similarly...

another pattern... (artificial data)

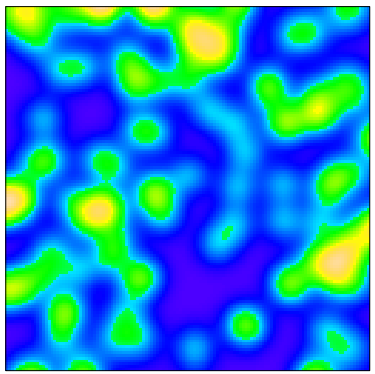


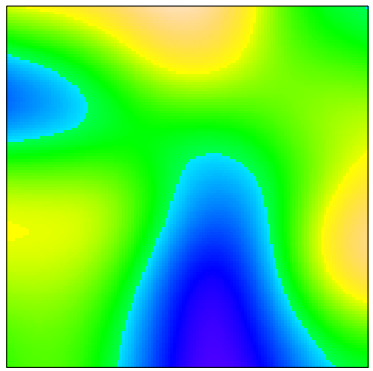
clustering revisited

Finding clusters – are they really there?

Similarly...

another pattern... (artificial data)



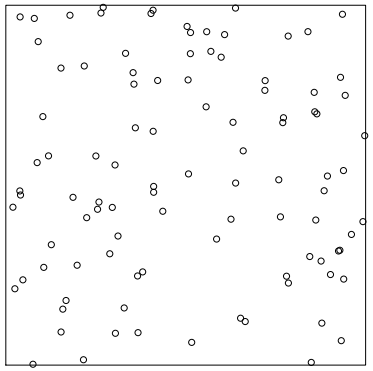


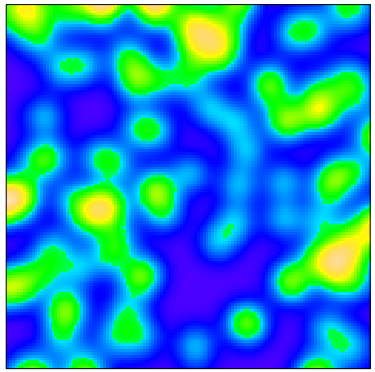
clustering revisited

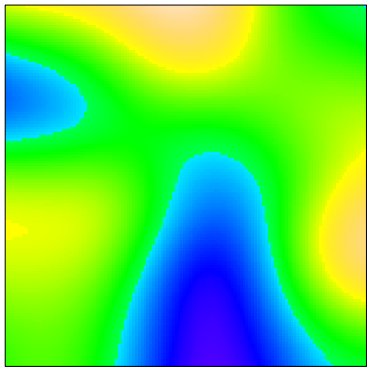
Finding clusters – are they really there?

Similarly...

another pattern... (artificial data)







in summary...

- we need to be clear about what we mean by a cluster
- spurious clusters – we need to be able to decide when a hotspot is a real hotspot

in summary...

- we need to be clear about what we mean by a cluster
- spurious clusters – we need to be able to decide when a hotspot is a real hotspot

$\Rightarrow$ But **why** are the clusters there?

in summary...

- we need to be clear about what we mean by a cluster
- spurious clusters – we need to be able to decide when a hotspot is a real hotspot

$\Rightarrow$ But **why** are the clusters there?
$\Rightarrow$ Are there reasons for a hotspot occurring in a specific location?

in summary...

- we need to be clear about what we mean by a cluster
- spurious clusters – we need to be able to decide when a hotspot is a real hotspot

$\Rightarrow$ But **why** are the clusters there?
$\Rightarrow$ Are there reasons for a hotspot occurring in a specific location?

in summary...

- we need to be clear about what we mean by a cluster
- spurious clusters – we need to be able to decide when a hotspot is a real hotspot

$\Rightarrow$ But **why** are the clusters there?
$\Rightarrow$ Are there reasons for a hotspot occurring in a specific location?

$\Rightarrow$ this is where statistical models come in...

the case for a (statistical) model

a statistical model allows us to use **inference**
we can

- decide if observed patterns are real patterns
- relate observed data to potential explanatory factors

the case for a (statistical) model

a statistical model allows us to use **inference**
we can

- decide if observed patterns are real patterns
- relate observed data to potential explanatory factors
- decide how relevant ("significant") each of these factors is

spatial statistical models

- decide if observed clusters are real clusters
- relate observed spatial patterns to potential explanatory factors – in space
- provide mechanisms to tell us how relevant these are

spatial statistical models

- decide if observed clusters are real clusters
- relate observed spatial patterns to potential explanatory factors – in space
- provide mechanisms to tell us how relevant these are

this implies...

- increased understanding of mechanisms that facilitate crime
- prediction... aiding policing decisions

complexity...

spatial statistical models are complex – why?

- assumption of (spatial) independence violated

complexity...

spatial statistical models are complex – why?

- assumption of (spatial) independence violated
- ignoring spatial autocorrelation leads to spurious "reasons"

complexity...

spatial statistical models are complex – why?

- assumption of (spatial) independence violated
- ignoring spatial autocorrelation leads to spurious "reasons"

this implies...

- we need to account for spatial autocorrelation
- this makes things rather complicated (dependence structures)
- ongoing statistical research

complexity...

spatial statistical models are complex – why?

- assumption of (spatial) independence violated
- ignoring spatial autocorrelation leads to spurious "reasons"

this implies...

- we need to account for spatial autocorrelation
- this makes things rather complicated (dependence structures)
- ongoing statistical research

spatio-temporal models – even more complex...

statistical research...?

Why is this of interest to **us**?

- complex spatial and spatio-temporal models are still being developed

statistical research...?

Why is this of interest to **us**?

- complex spatial and spatio-temporal models are still being developed
- computational complexity needs to be dealt with (recent developments very promising...)
- we aim to develop methods that are relevant in practice

statistical research...?

Why is this of interest to **us**?

- complex spatial and spatio-temporal models are still being developed
- computational complexity needs to be dealt with (recent developments very promising...)
- we aim to develop methods that are relevant in practice

$\Rightarrow$ interdisciplinary work

interdisciplinarity ...

symbiosis

interaction among two or more "species" with **mutual benefit**

interdisciplinarity ...

symbiosis

interaction among two or more "species" with **mutual benefit**

$\Rightarrow$ multi-directional interaction!

interdisciplinarity ...

symbiosis

interaction among two or more "species" with **mutual benefit**

$\Rightarrow$ multi-directional interaction!

statisticians' role:

- not "simply" analysing a data set
- not only using a data set as an illustration of a new method

interdisciplinarity ...

symbiosis

interaction among two or more "species" with **mutual benefit**

$\Rightarrow$ multi-directional interaction!

statisticians' role:

- not "simply" analysing a data set
- not only using a data set as an illustration of a new method

BUT:

- obtaining inspiration from applied problems
- contributing to the scientific discourse

interdisciplinarity ...

symbiosis

interaction among two or more "species" with **mutual benefit**

$\Rightarrow$ multi-directional interaction!

statisticians' role:

- not "simply" analysing a data set
- not only using a data set as an illustration of a new method

BUT:

- obtaining inspiration from applied problems
- contributing to the scientific discourse
- communication...

truly multi-disciplinary...

mutual understanding

interaction requires functioning communication...

truly multi-disciplinary...

mutual understanding

interaction requires functioning communication...

interdisciplinary statistician's role:

- aiding communication by:
  - participating in scientific dialogue **on both sides**
  - engaging with contents and relevant scientific questions

truly multi-disciplinary...

mutual understanding

interaction requires functioning communication...

interdisciplinary statistician's role:

- aiding communication by:
  - participating in scientific dialogue **on both sides**
  - engaging with contents and relevant scientific questions
- fostering symbiosis:
  - contributing to both scientific fields
  - stimulating research in both areas

truly multi-disciplinary...

mutual understanding

interaction requires functioning communication...

interdisciplinary statistician's role:

- aiding communication by:
  - participating in scientific dialogue **on both sides**
  - engaging with contents and relevant scientific questions
- fostering symbiosis:
  - contributing to both scientific fields
  - stimulating research in both areas
- developing methods that are
  - statistically interesting and
  - practically relevant

this workshop...

initiate **communication**:

- to help us learn from each other

this workshop...

initiate **communication**:

- to help us learn from each other
- what can our models do?
- what are interesting relevant questions?

this workshop...

initiate **communication**:

- to help us learn from each other
- what can our models do?
- what are interesting relevant questions?

$\Rightarrow$ develop a common language and **common interests**