OGH talk: Analysing spatial point patterns with spatstat

Ege Rubak

2 September 2021

Plan for the session

- Very short introduction (these slides).
- Live tutorial with examples of analysis.
- Questions and discussion.

The spatstat package

- 25 years development by mainly Adrian Baddeley with co-author Rolf Turner from the beginning and more recently I joined the team.
- 100,000+ lines of R code, 30,000+ lines of C code, 2,500+ exported objects and 1000+ documentation files.
- Recently split into sub-packages `spatstat.xxxx`, which the now almost empty package `spatstat` then `Depends` on.

The spatstat design

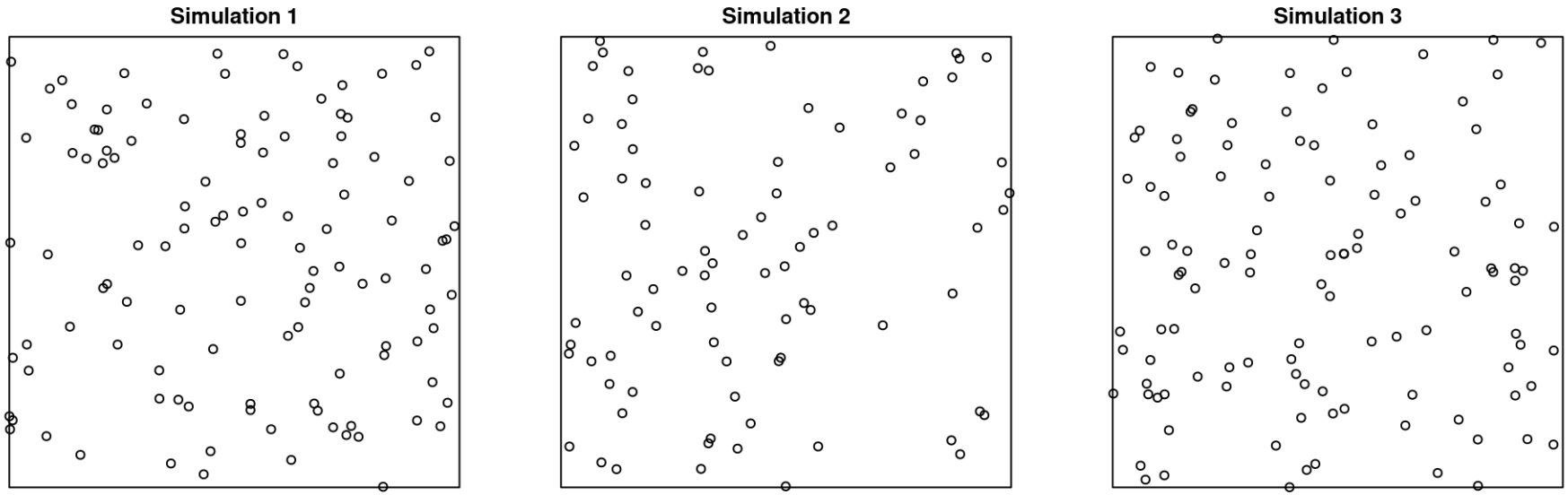
- Strong focus on backwards compatibility.
- Follows base R design and uses base R graphics.
- Extensive usage of the S3 object system.
- Has its own S3 objects for data etc.
- Detailed documentation.
- Use `?spatstat` to get an overview and find hidden gems.

What to use spatstat for?

- You **can use** `spatstat` to describe/summarise any given point set with things like
    - Pairwise distances, nearest neighbour distances, empty space distances, Dirichlet/Voronoi tessellations, …
- However, `spatstat` really focuses on **statistical inference** for phenomena that generate **random locations** (point processes).

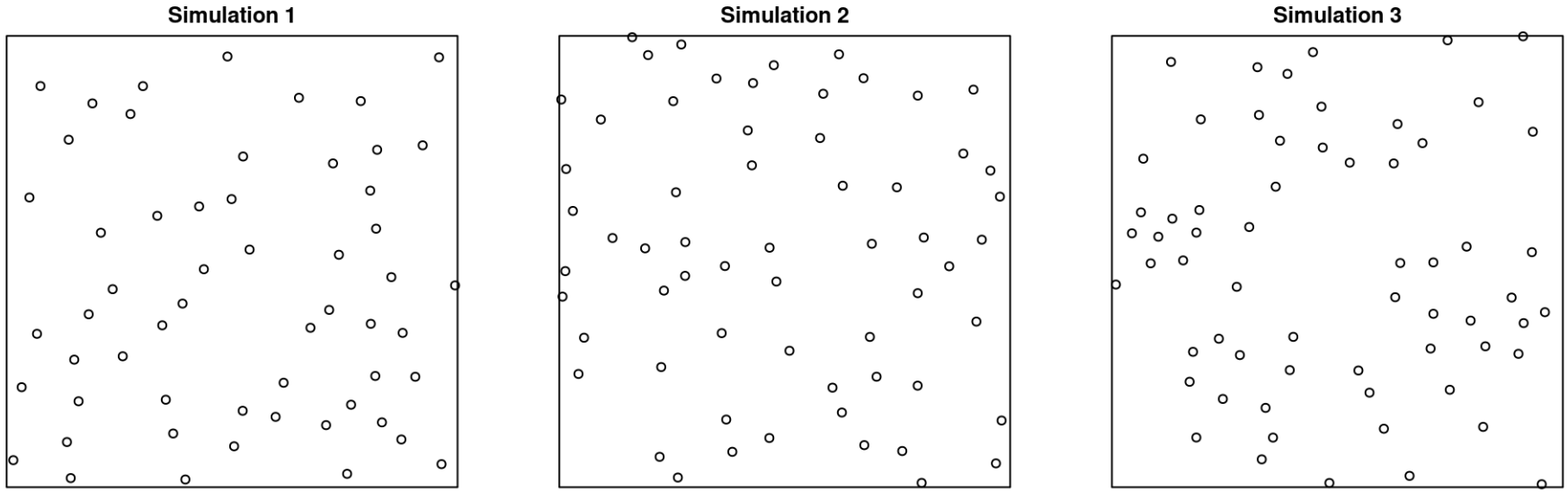
Complete spatial randomness (Poisson process)

```
library(spatstat)
set.seed(42) # Reproducibility
Xpois <- rpoispp(100, nsim = 3)
plot(Xpois, main = "")
```



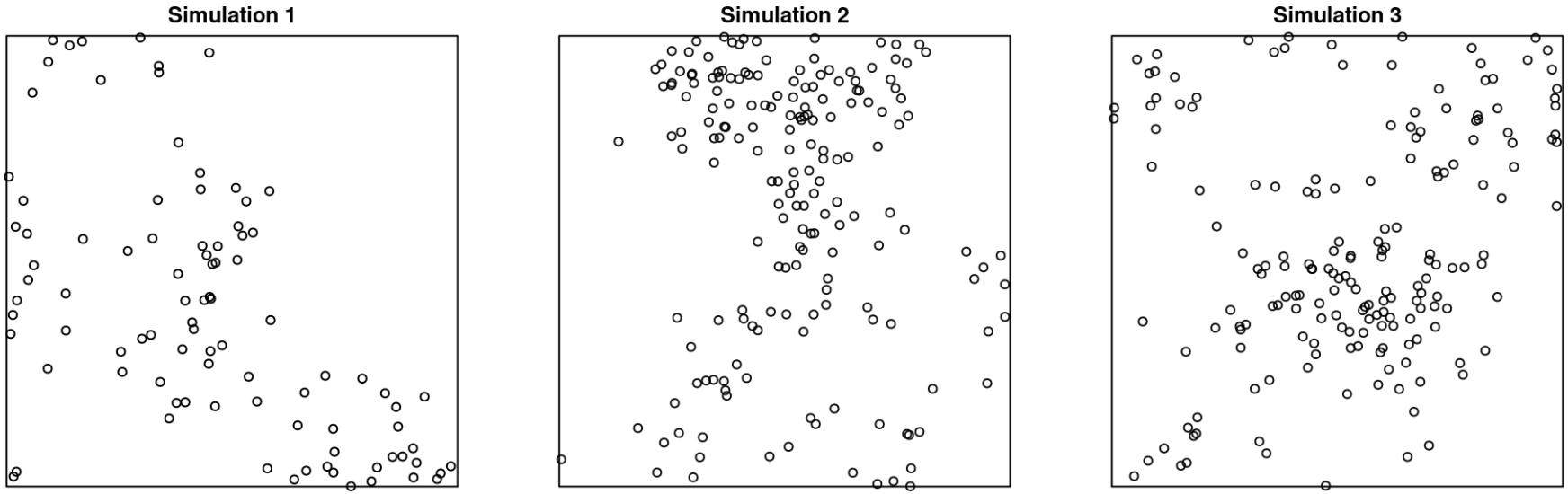
Hard core Gibbs process

```
Xhc <- rHardcore(beta = 100, R = .05, nsim = 3)
plot(Xhc, main = "")
```



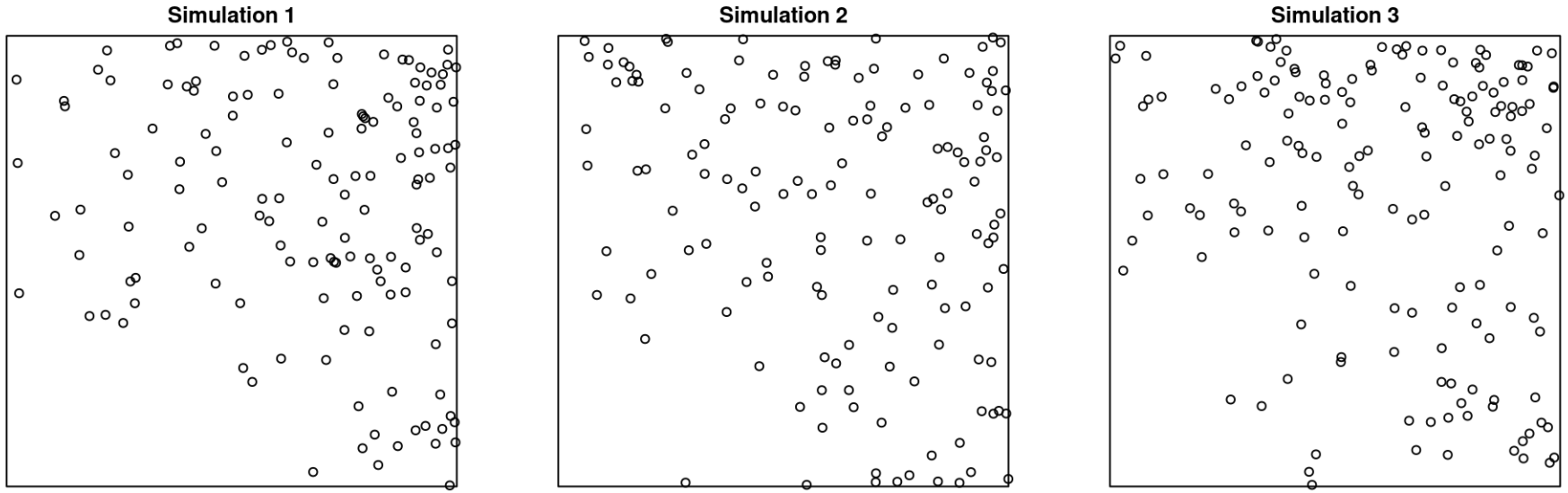
Thomars cluster process

```
Xthomas <- rThomas(kappa = 5, mu = 20, scale = .1, nsim = 3)
plot(Xthomas, main = "")
```



Inhomogeneous Poisson process

```
lambda <- function(x,y){200*(x^2+y^2)}
Xinhom <- rpoispp(lambda, nsim = 3)
plot(Xinhom, main = "")
```



Separating intensity and interaction

- Intensity is a first moment property.
- Interaction is a higher moment property (inter-point correlation).
- They are confounded and without further assumptions it is impossible to separate them in general.
- Often a approach like in time series is used:
  - First, model the mean (trend, seasonality).
  - Second, model the interaction after accounting for the mean model.

How to learn more and report bugs

- Explore `?spatstat` which includes lists of commonly (and less commonly) used functions.
- Get the book. Unfortunately we don't have a license to share an online version as many authors have nowadays. Maybe this will change with a second edition. There are three free sample chapters at https://book.spatstat.org/
- Ask questions on stackoverflow under the spatstat tag
- Report bugs or make feature requests on GitHub. (If possible find the right sub package repo to put the issue under.)