International Summer Academy on Spatial Ecotoxicology and Ecotoxicological Risk Assessment Using an Open Community Approach

#### Dr. Avit Kumar Bhowmik

avit.bhowmik@su.se

September 17, 2016

Stockholm Resilience Centre
Sustainability Science for Biosphere Stewardship



### What do we know?

## What is GIS?



## **GIS**

# **GIS** ? Information ?

# **GIS**? Information ?

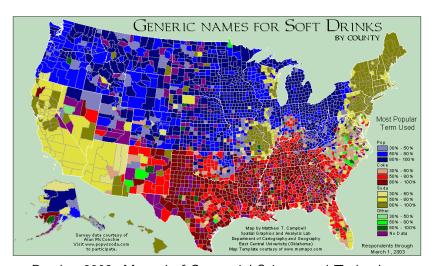
- Geographic
   Parent and Church, 1987. Conf. GIS
- Spatial (Geospatial)
   Anselin, 1989. What is special about spatial data?
- Spatiotemporal
   Burrough and Frank, 1995. Int J GIS

# GIS ? Information ?

- Geographic
   Parent and Church, 1987. Conf. GIS
- Spatial (Geospatial)
   Anselin, 1989. What is special about spatial data?
- Spatiotemporal
   Burrough and Frank, 1995. Int J GIS

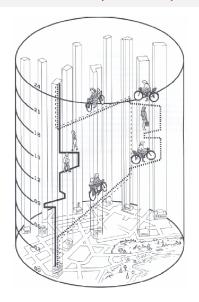
- System
- Science Goodchild, 1992. Int J GIS

## 80% of data are Spatiotemporal



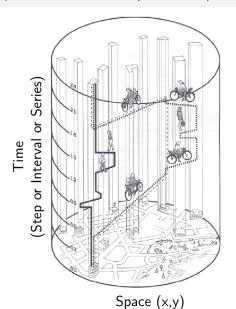
Bossler, 2002. Manual of Geospatial Science and Technology

## Representation of Spatiotemporal Data



Chrisman,1997.Exploring GIS

## Representation of Spatiotemporal Data



Chrisman,1997.Exploring GIS

## Wear the GI Glasses



# What is Spatial Statistics?



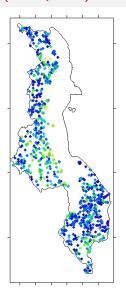
## Spatial(-temporal) Statistics

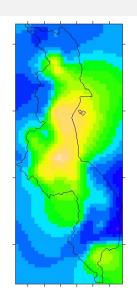
### Experts' Thoughts

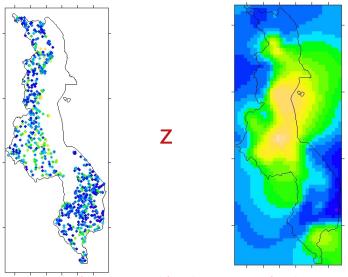
Spatial statistics offers a way of describing the spatial continuity that
is an essential feature of many natural phenomena and provides
adaptations of classical regression techniques to take the advantage of
this continuity

Isaaks and Srivastava, 1989. An Introduction to Applied Geostatistics

 Spatial statistics provides a set of statistical tools for incorporating the spatial coordinates of observations in data processing Goovaerts, 2007. Geostatistics for Natural Resources Evaluation







z is a random process with unique mean and variance  $z(\text{sampled locations}) \approx z(\text{unsampled location})$ 

#### Input

- Set of Points sampled, sparsely distributed in space and time
- Each point represents a measurement of a variable (spatiotemporal attribute) that occurs in that space and time location

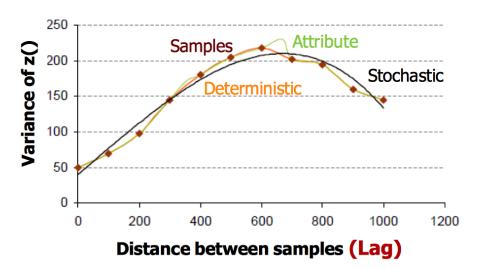
#### Input

- Set of Points sampled, sparsely distributed in space and time
- Each point represents a measurement of a variable (spatiotemporal attribute) that occurs in that space and time location

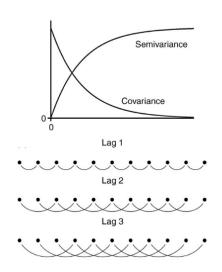
## Output

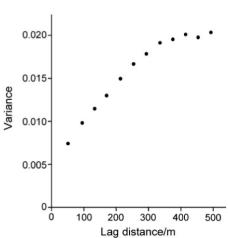
- Spatial Data Model
- Computer/mathematical representation that allows one to perform estimations and/or simulations for attribute values at spatial/temporal locations not sampled

## Spatial(-temporal) Variability

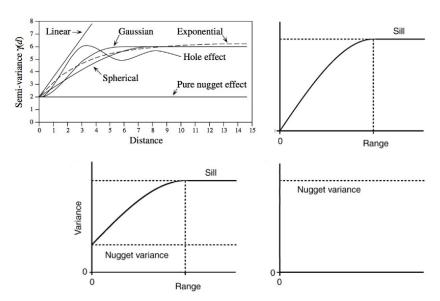


## Spatial(-temporal) Variability



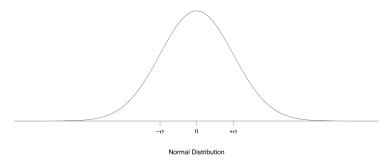


## Spatial(-temporal) Variogram



## Stochastic or Geostatistical Interpolation

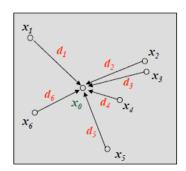
- A probability distribution function is associated to its probable values
- Uncertainties can be associated to its estimation
- e.g. Kriging
- Minimization of estimation variance (error)



## **Deterministic Interpolation**

- An unique value is associated to its spatial location
- No uncertainty is associated to its estimation
- e.g. Inverse Distance Weighting (IDW)

$$Z^*(x_0) = \frac{\sum_{i=1}^n \frac{1}{d_i^2} z(x_i)}{\sum_{i=1}^n \frac{1}{d_i^2}}$$



## Learn more about Spatial(-temporal) Statistics







George Christakos

#### Spatio-temporal geostatistics using gstat



#### Introduction

Since gstat package version 1.0-0, a dependency of gstat on the R package spacetime was introduced, allowing the code in gstat to exploit spatio-temporal data structures from that package. This vignette describes the possibilities and limitations of the package for spatio-temporal geostatistics.

To understand some of the possibilities and limitations, some knowledge of the history of the software is needed. The original gattat software (Pebesma and Wesseling, 1998) was a standalone computer program written in around 25,000 lines of C code, and would do gesetatistical modelling, prediction and simulation. The gatat R package (Pebesma, 2004) consisted mostly of an R interface to this C code, together with convenience functions to use R's modelling interface (formulas', see Tela) and graphic capabilities (trellis graphics in package lattice to show cross variogram as matrix plots; interaction with variogram clouds using base plots).

Starting 2003, a group of programmers developed a set of classes and methods for dealing with spatial data in R (points, lines, polygons, grids), which was supported by the publications of the well-known ASDAR book (Bivand et al.

## Learn more about Spatial(-temporal) Statistics



Enablased in 1995, the Journal of Statistical Software publishes artifacts, book reviews, code supports, and software reviews on the subject of statistical software and algorithms. The contents are fonly a called to soles. For both unities and rode supports the surpers code in published along, with the paper. Statistical software in the key list between statistical enhabet and their applications as practice. Software that makes that list in the processor of the journal, and easy to radiated as, for instance, took for large service companing, database technology, desilops companing, distributed systems, the World Wale Web, superduchte research, artifacts gain desirants and netholoded presents.

We attempt to present research that demonstrates the joint evolution of computational and statistical methods and techniques. Implementations can use languages such as C. C++, S. Fortson, Java, PHP, Python and Ruby or environments such as Mathematica, MATLAB, R. S-PLUS, SAS, Stata, and XLISP-STAT.



### The slides, scripts, materials and data are available from: https://github.com/AvitBhowmik/SA16



Learning by doing!



## Thank You!

