

数据分析沙龙（北京站）第10期

时空统计与时空数据可视化

王江浩

(wangjh@reis.ac.cn)

中国科学院地理科学与资源研究所
资源与环境信息系统国家重点实验室

2013年7月29日 @ 人民大学

Topics

1 时空统计学概述

2 R中的时空数据分析方法

- ✧ 时间序列分析
- ✧ 空间统计
- ✧ 时空统计

3 时空数据可视化

- ✧ R中的时空数据可视化
- ✧ Web中的时空数据可视化

4 时空数据可视化应用案例

1 时空统计学概述

时空统计学的发展

- 经典统计学 (classical statistics)
 - 历史久远，以概率论为基础，应用广泛；
 - 假设条件：独立同分布。

- 空间统计学 (spatial statistics)
 - 空间对象，单次观测；
 - 空间自相关性，空间异质性；各向异性

- 时空统计学 (spatiotemporal statistics)
 - 时空序列观测；
 - 时空自相关性；时空异质性；时空交互性。



空间统计学

□ 研究内容

从数据模型划分 (Cressie 1993)

1. 点模式分析——从空间位置的特定视角
2. 格网与区域分析——与空间的区域模型相关
3. 地统计学建模——应用于支撑数据的连续场视角

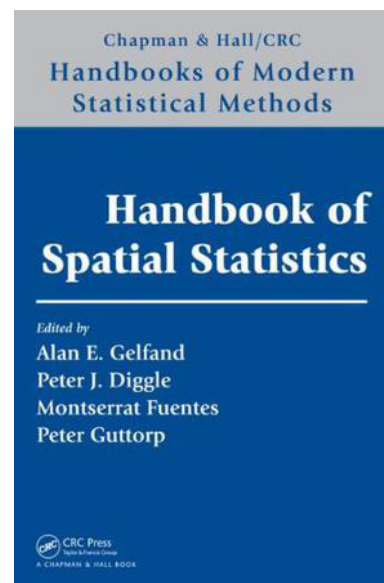
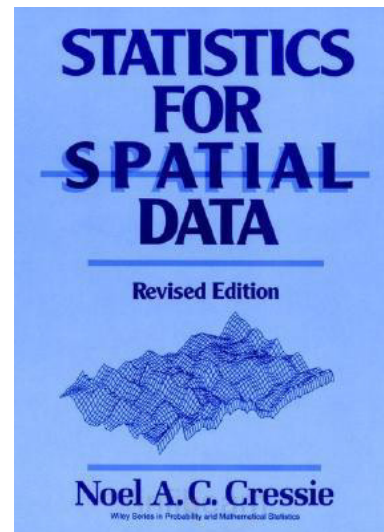
从研究内容划分 (Gelfand 2010)

1. Continuous spatial variation——包括地统计学，层次模型，非平稳非高斯随机场等
2. Discrete spatial variation ——包括空间自相关分析、疾病制图和空间计量经济
3. Spatial point pattern——空间点模式分析
4. Spatio-temporal process——时空统计与数据同化
5. Addition topic——包括多变量随机过程，支撑改变，空间聚合等

□ BOOKS

Cressie, N. A. C. (1993). Statistics for spatial data. New York, Wiley.

Gelfand, A. E. (2010). Handbook of spatial statistics. Boca Raton, CRC Press.



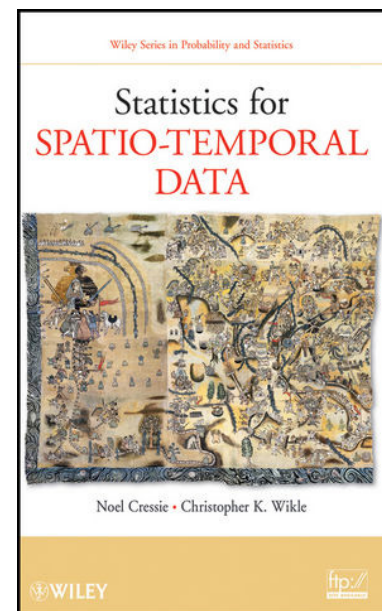
时空统计学

□ 研究内容

- 以空间统计学和时间序列分析为基础
- 探索性时空数据分析
- 基本的时空统计模型
 - 时空协方差函数构建
 - 时空克里格
 - 随机差分方程
 - 空间过程的时间序列分析
 - 时空点过程分析
- 层次动态时空模型
-

□ BOOKS

- Cressie, N. A. C. and C. K. Wikle (2011). Statistics for spatio-temporal data. Hoboken, N.J., Wiley.
- Finkenstädt, B., L. Held, et al. (2007). Statistical methods for spatio-temporal systems. Boca Raton, FL, Chapman & Hall/CRC.



时空统计学应用

□ 地球系统科学

- 气候，气象模拟与预报
- 地质、海洋、大气等

□ 地理信息系统、遥感与全球定位系统（3S）

- GIS：采集、储存、管理、运算、分析、显示时空地理信息。
- RS：卫星遥感；航空遥感；地基遥感等
- GPS：定位，时空轨迹

□ 全球变化与生态环境

- 全球变化；生态系统；环境监测；环境变化等

□ 公共环境健康

- 传染病模型；公共健康与环境

□ 社会经济领域

- 区域经济格局

2 R中的时空数据分析

- ✧ 时间序列分析
- ✧ 空间统计
- ✧ 时空统计

R 中的时空数据分析

□ 按照类型分类

■ TimeSeries

<http://cran.r-project.org/web/views/TimeSeries.html>

■ Spatial

<http://cran.r-project.org/web/views/Spatial.html>

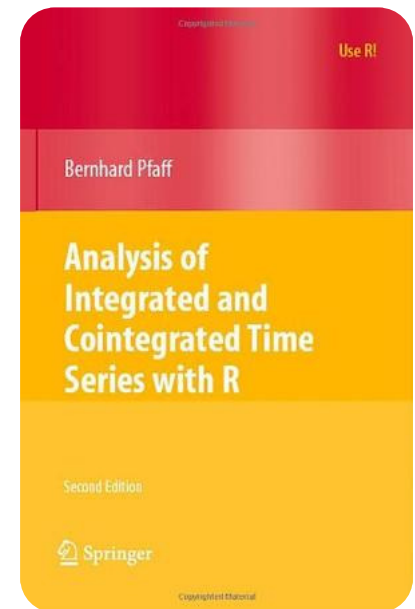
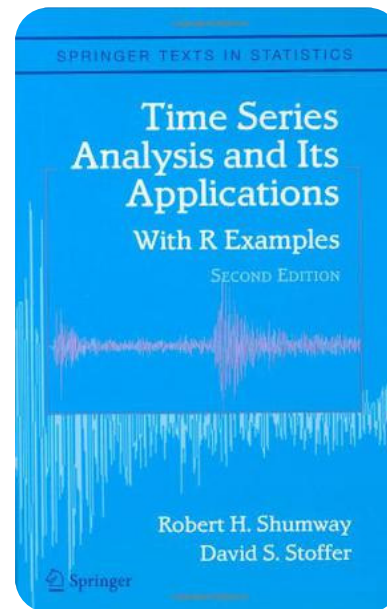
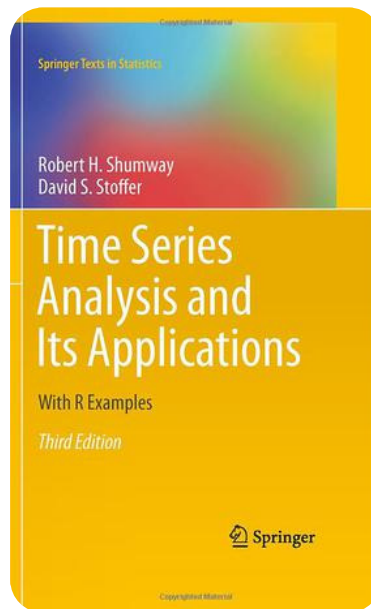
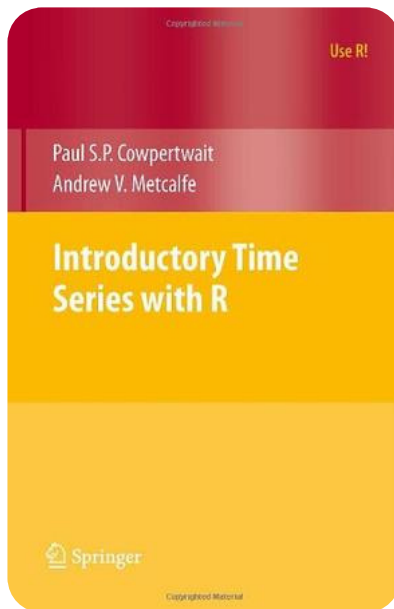
■ Spatiotemporal

<http://cran.r-project.org/web/views/SpatioTemporal.html>

```
install.packages("ctv")  
library("ctv")  
install.views("TimeSeries")  
update.views("TimeSeries")
```

TimeSeries — 时间序列分析

- ❑ **Maintainer:** Rob J. Hyndman
- ❑ **Package stat. :** 131
- ❑ **Core package :** `stat`, `forecast`, `tseries`, `xts`, `zoo`
- ❑ **Application:** Econometrics and Finance



TimeSeries — 时间序列分析

□ 时间和日期数据结构 **stat**

`ts`, `POSIXct`, `POSIXlt`

□ 预测和单变量建模 **forecast** & **tseries**

`ets()`; `ar()`; `arima()`; `arma()`

□ 频域分析 **wavelets**

`spectrum()`

□ 分解与滤波 **stat**

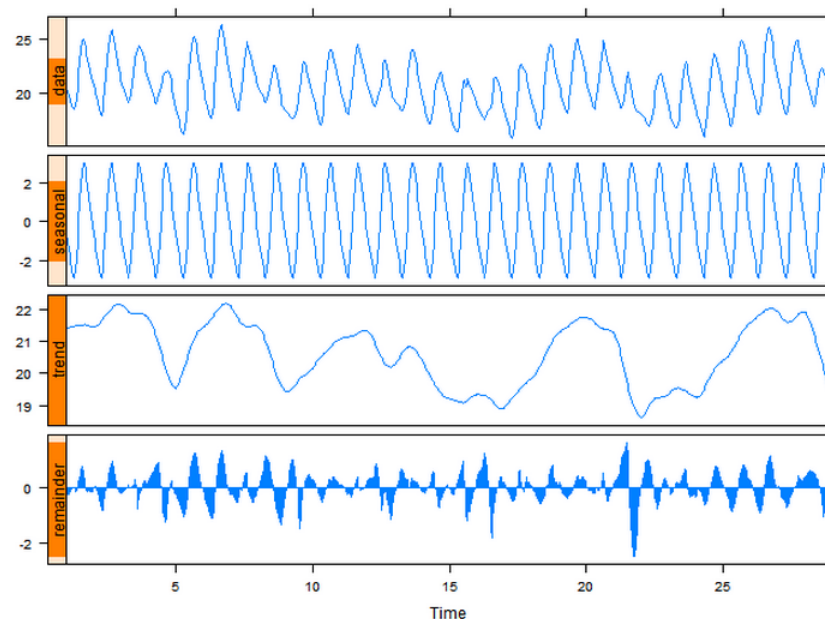
`filter()`; `decompose()`; `stl()`

□ 多元时间序列分析

□ 非线性时间序列分析

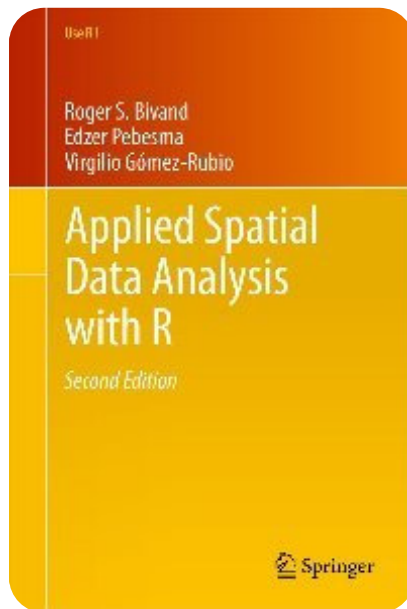
□ 动态回归模型

□



Spatial — 空间数据分析

- ❑ **Maintainer:** Roger Bivand
- ❑ **Package stat. :** 128
- ❑ **Core packages :** `sp`, `geoR`, `gstat`, `maptools`, `raster`, `RandomFields`, `rgdal`, `spacetime`, `spdep`, `splancs`
- ❑ **Mailing list :** R-SIG-Geo ; OSGeo



<http://www.asdar-book.org/>

Overview of R spatial packages

□ 空间数据处理步骤

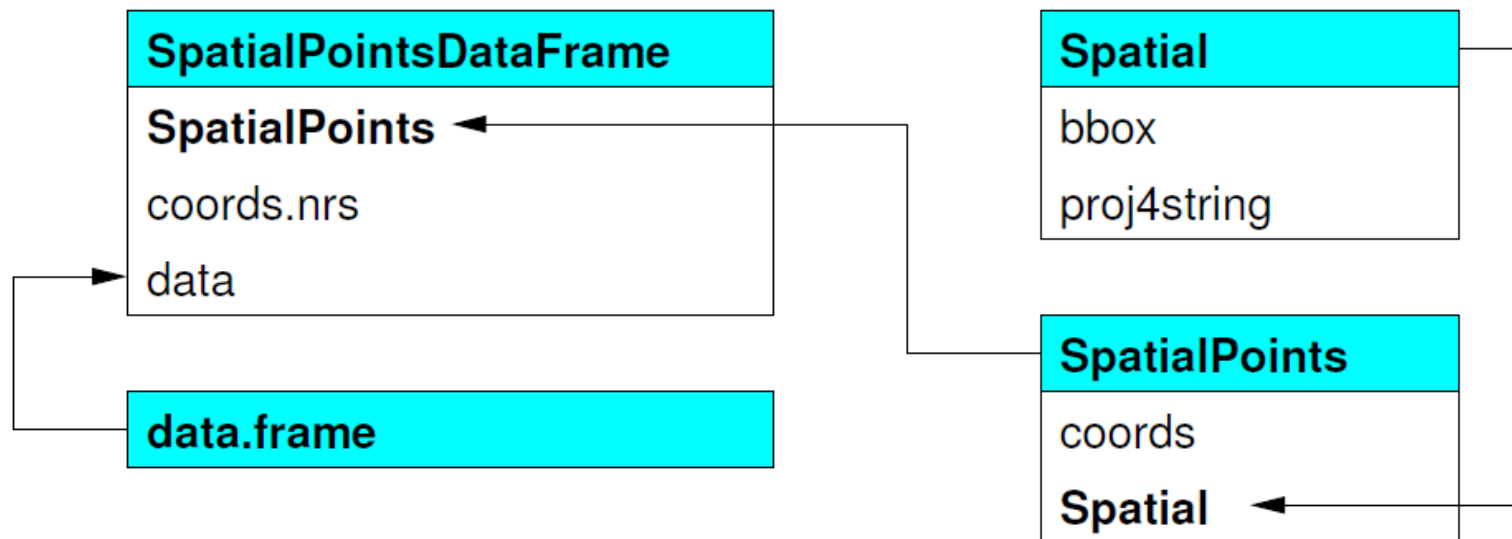
1. R中读入数据
2. 数据清理，探索性分析
3. 进行空间统计分析
4. 输出数据

□ 空间统计核心包分类

目的	所需包
输入输出数据	rgdal, maptools, maps RArcInfo, spgrass6, GRASS, ncdf4
空间数据清理和操作	sp, maptools, raster
空间统计分析	Spatstat, gstat, geoR, spdep, splancs

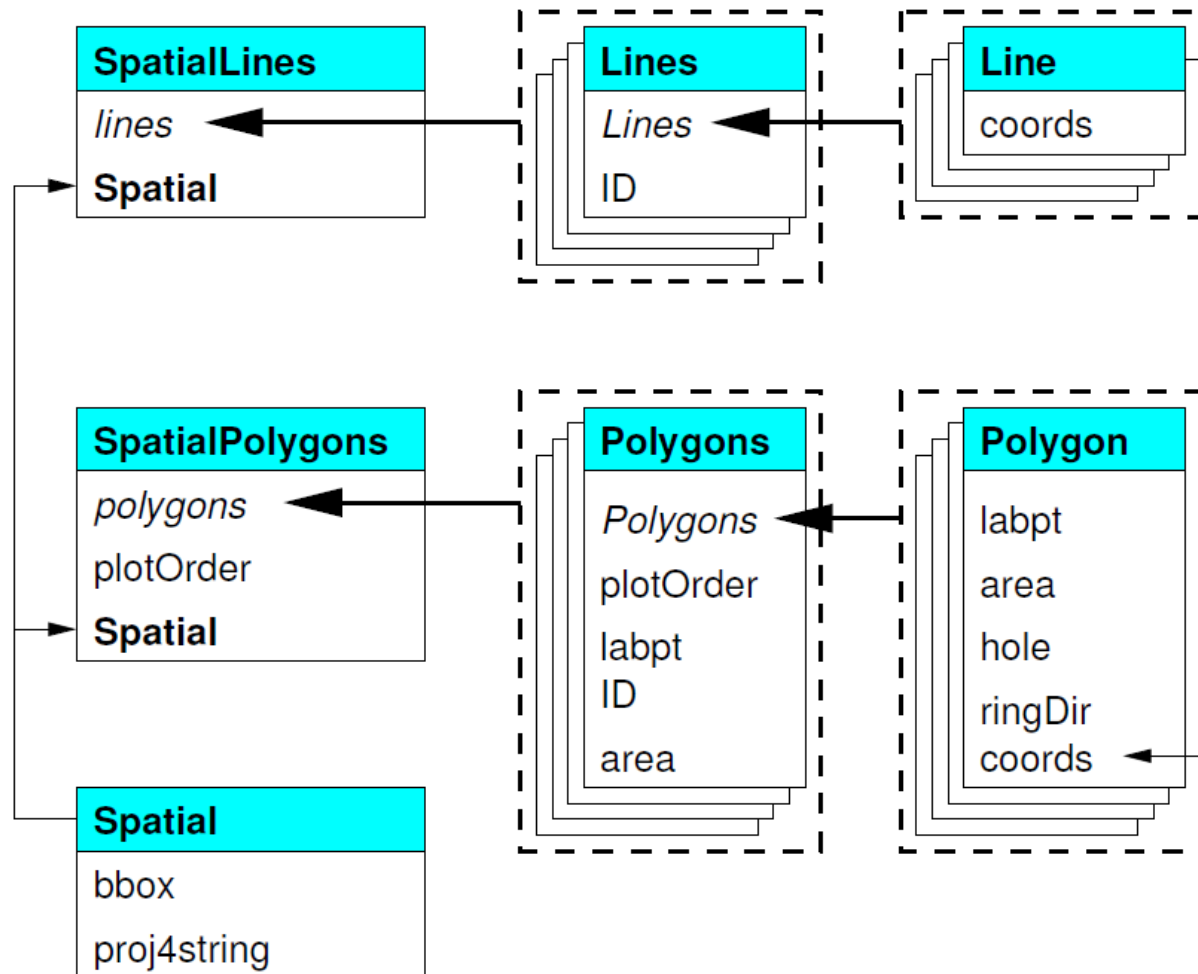
Spatial data structure

- ❑ **sp** 基础包中对空间数据，包括点、线、面、栅格，定义空间数据结构。
- ❑ 超过100包引用**sp**
- ❑ Spatial Points



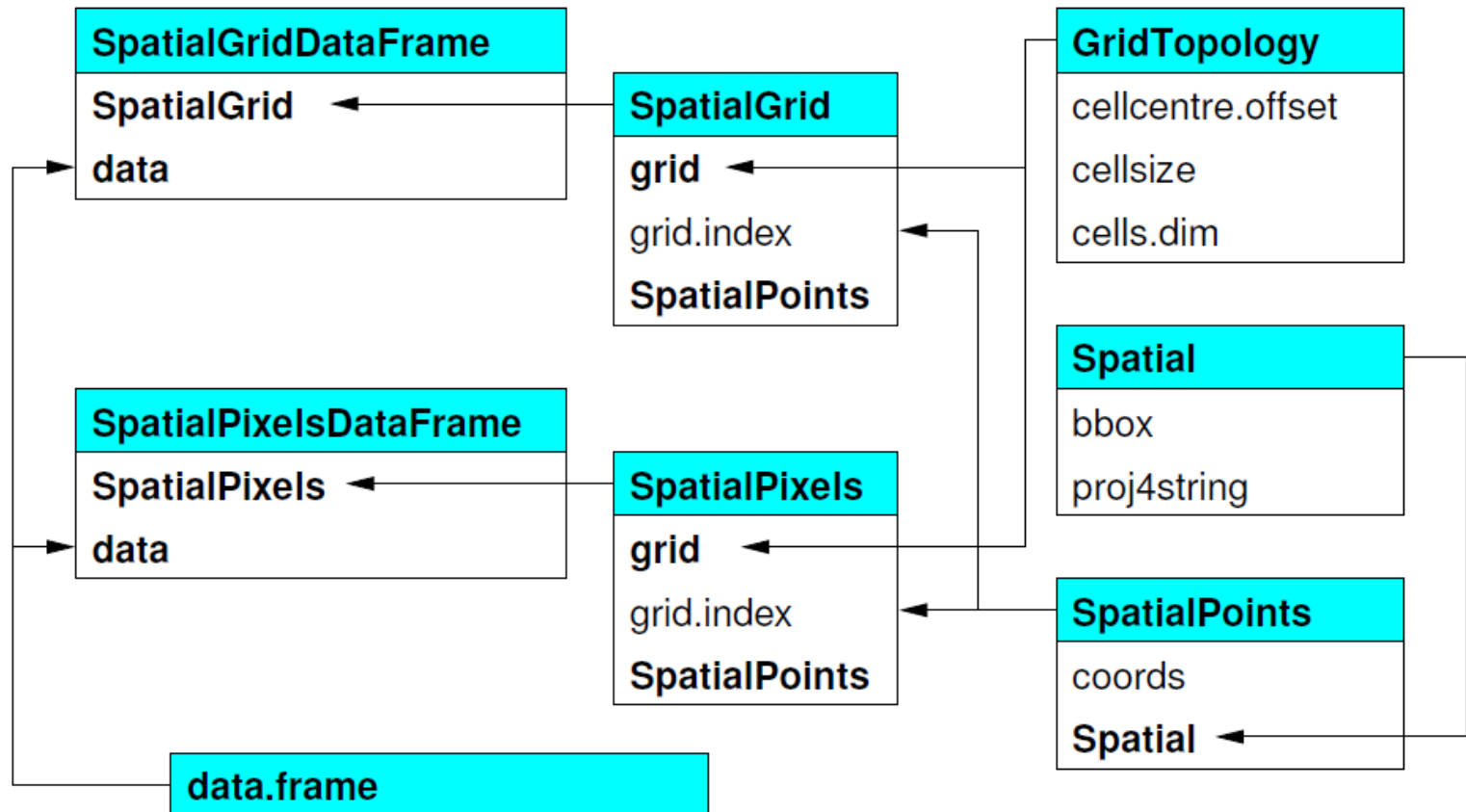
Spatial data structure

□ Spatial Polygons



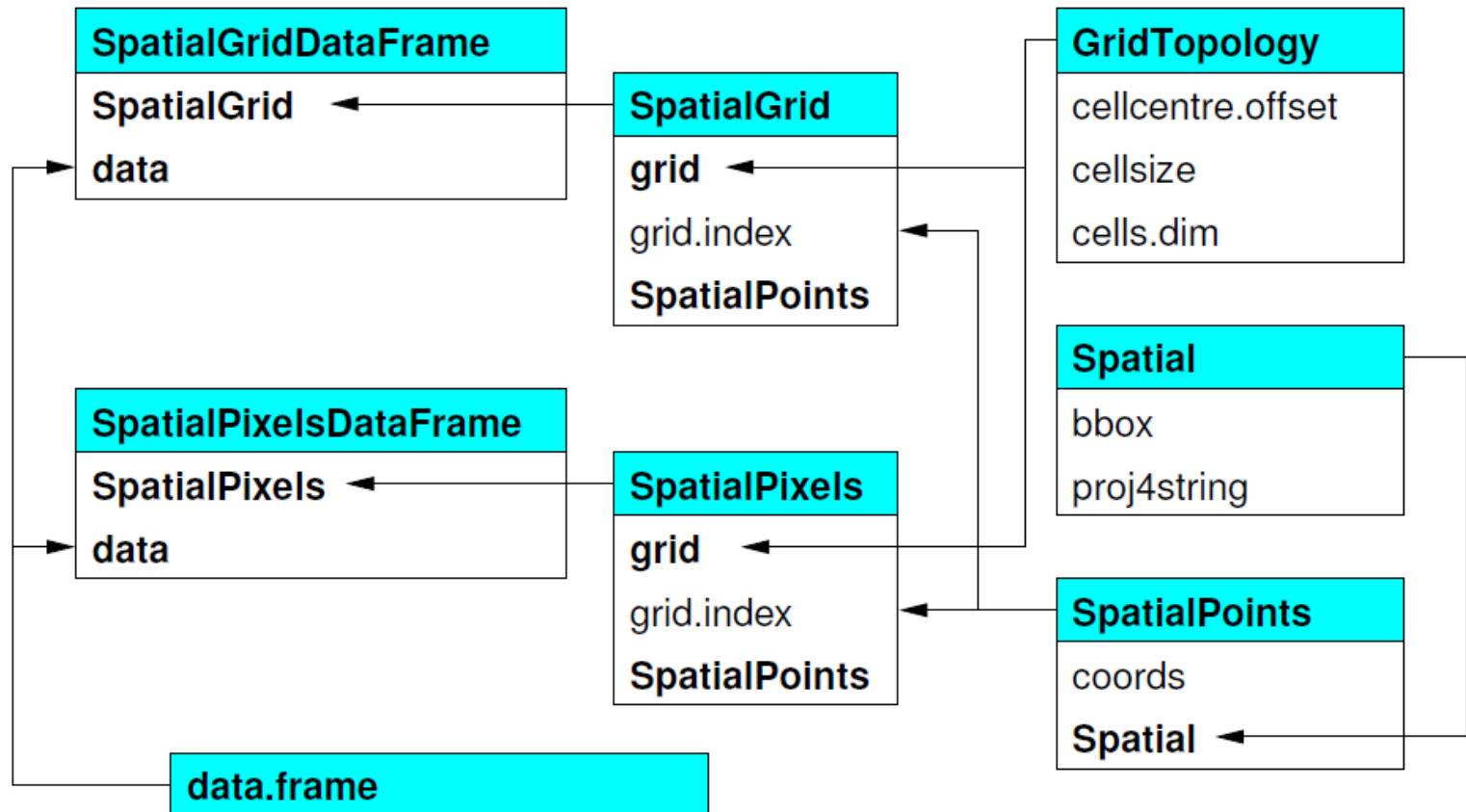
Spatial data structure

□ Spatial grid and pixel classes



Spatial data structure

□ Spatial grid and pixel classes



Read/write spatial data

□ `rgdal`: Bindings for the Geospatial Data Abstraction Library ([GDAL](#))

- `rgdal`
- `readGDAL()` / `writeGDAL()`
- `readOGR()` / `writeOGR()`

□ Other format

- `maptools`, `shapefile`: shapefile format
- `ncdf`: NetCDF
- `raster`: raster data
- `RArcInfo`: *.e00

GIS interface

□ spgrass6

Interface between GRASS 6+ geographical information system and R



GRASS GIS

The world's leading Free GIS software

□ RPyGeo

ArcGIS Geoprocessing in R via Python



Arc
ESRI **GIS**

□ RSAGA

SAGA Geoprocessing and Terrain Analysis in R

SAGA

System for Automated Geoscientific Analyses



Point pattern analysis

□ **Spatial**

Functions for Kriging and Point Pattern Analysis

□ **Spatstat**

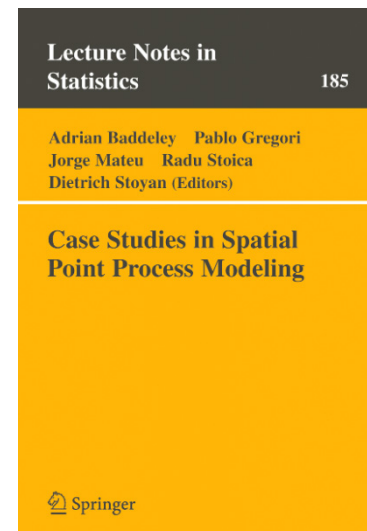
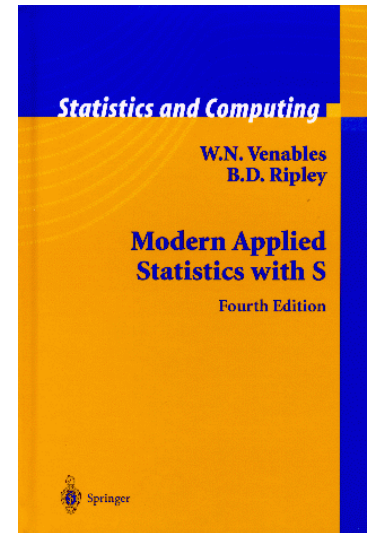
Spatial Point Pattern analysis, model-fitting, simulation, tests

<http://www.spatstat.org/>

□ **splancs**

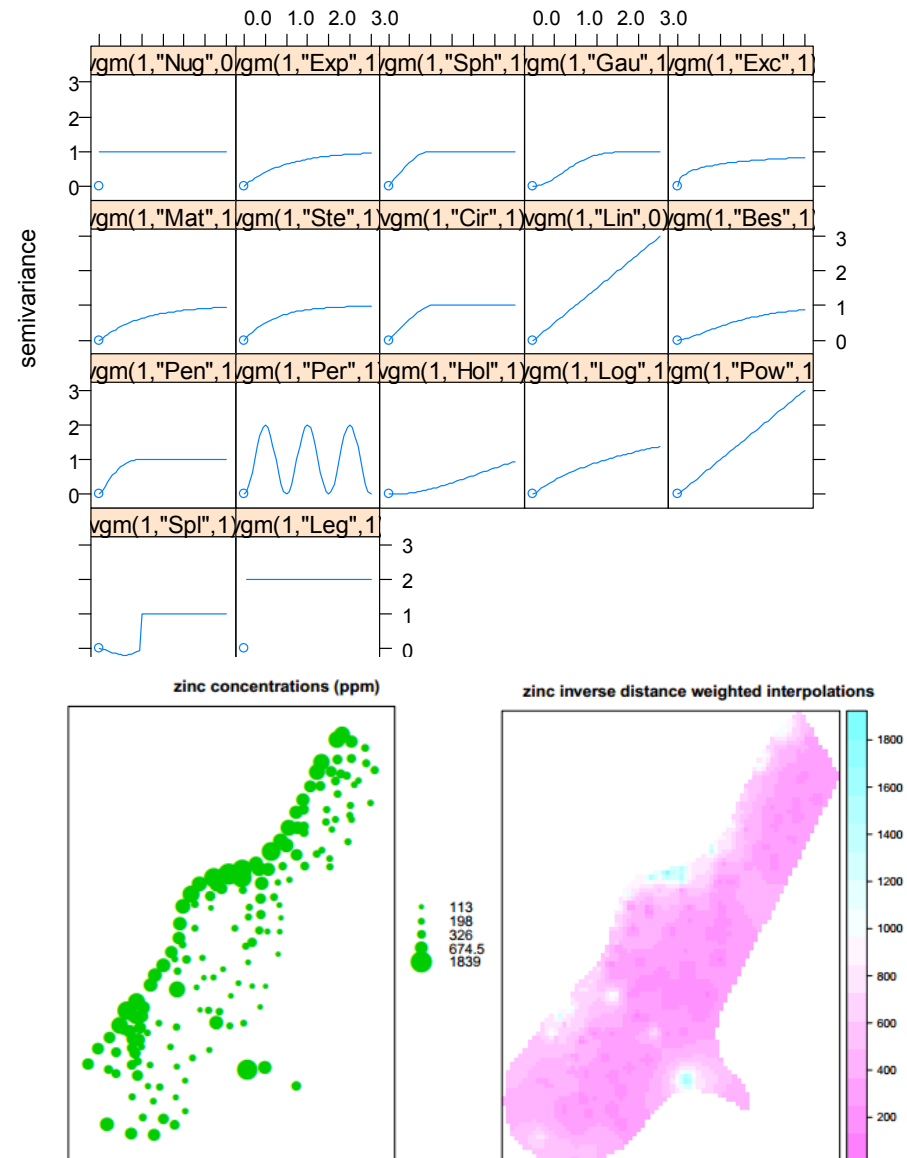
Spatial and Space-Time Point Pattern Analysis

<http://www.maths.lancs.ac.uk/~rowlings/Splancs/>



Geostatistics

- ❑ **gstat**
spatial and spatio-temporal
geostatistical modelling,
prediction and simulation
- ❑ **geoR, geoRglm**
Model-based Geostatistics
- ❑ **RandomField**
Simulation and analysis of
random field



Spatial regression

□ Spdep

Spatial dependence: weighting schemes, statistics and models

1. Classical Regression (CLR)
2. Simultaneous Autoregressive Models (SAR)
3. Conditional Autoregressive Models (CAR)
4. Spatial Lag Model (SLM)
5. Spatial error Model (SEM)

□ nlme

Linear and Nonlinear Mixed Effects Models

□ spgwr

Geographically weighted regression

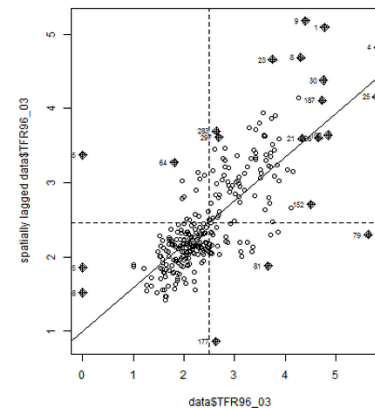
Queen - style



Rook - style



Moran's I scatterplot



Local Indicator of Moran's I



Spatiotemporal – data process and analysis

- ❑ **Maintainer:** Edzer Pebesma
- ❑ **Package stat.** : 46
- ❑ **Core package** : `sp`, `xts`, `spacetime`, `stpp`,
`surveillance`, `gstat`, `RandomFields`, `raster`
- ❑ **Mailing list** : R-SIG-Geo ; OSGeo



Journal of Statistical Software

November 2012, Volume 51, Issue 7.

<http://www.jstatsoft.org/>

spacetime: Spatio-Temporal Data in R

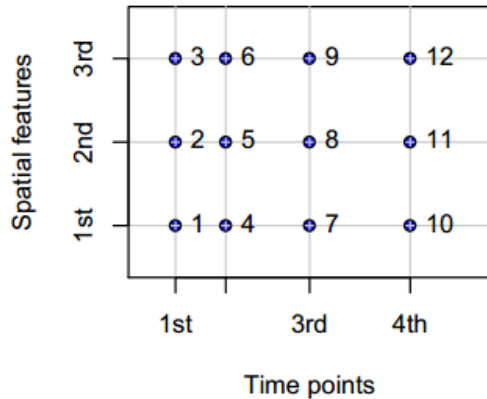
Edzer Pebesma
University of Münster

Spatio-temporal data structure

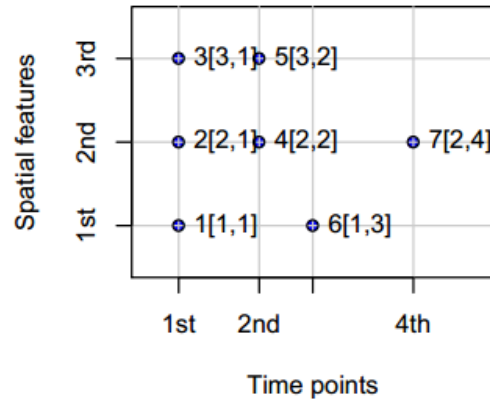
- ❑ `data.frame`[longitude, latitude, time]
- ❑ time-wide table
- ❑ space-wide table
- ❑ Generic classes
 - `spacetime : sp + xts`
- ❑ Dedicated classes
 - Geostatistical data: `SpatioTemporal {STdata}`
 - Gridded/raster data: `raster {rasters}`
 - Lattice data: `surveillance {sts}`
 - Point patterns: `stpp {stpp}`
 - Trajectory data: `adehabitatLT {ltraj}`

spacetime data structure

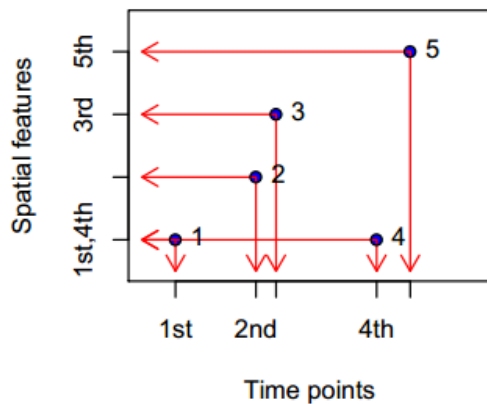
STF: full grid layout



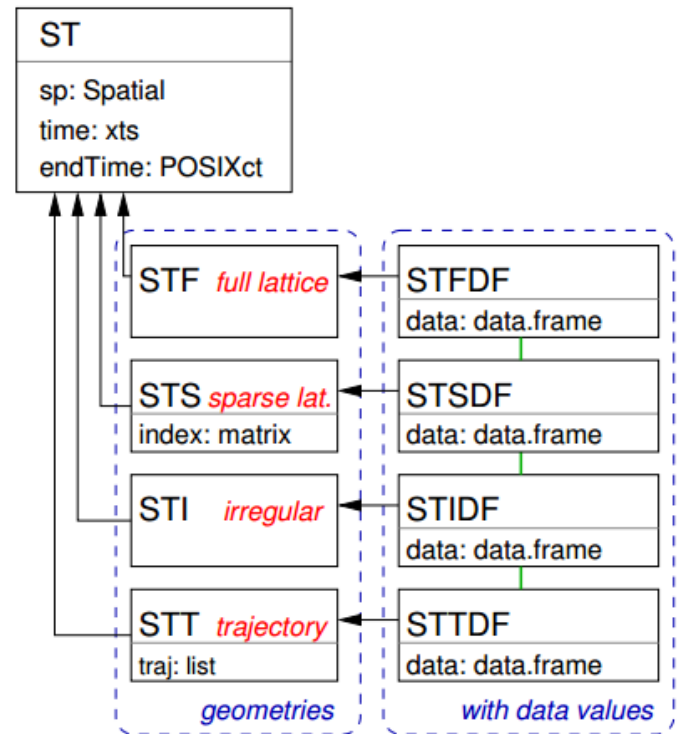
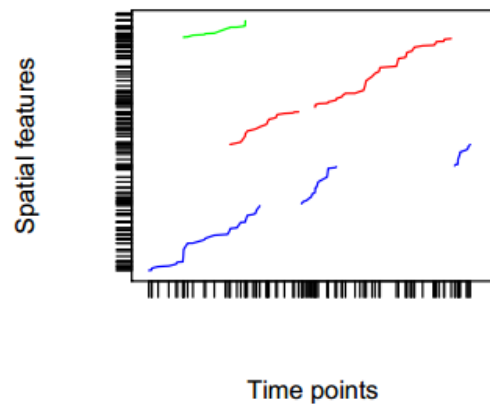
STS: sparse grid layout



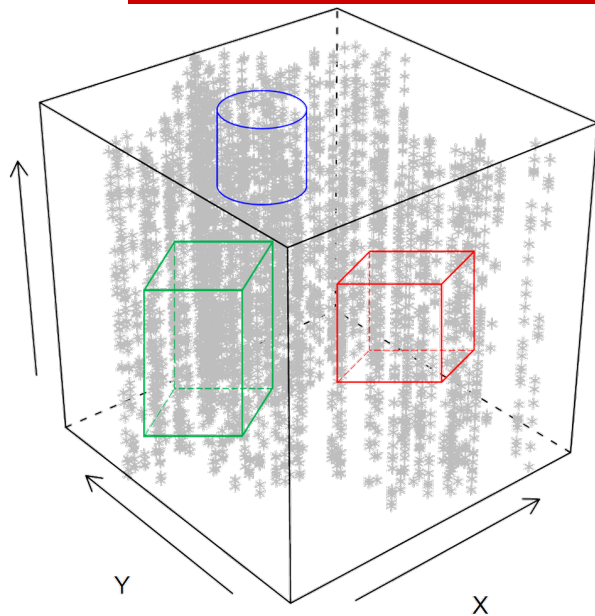
STI: irregular layout



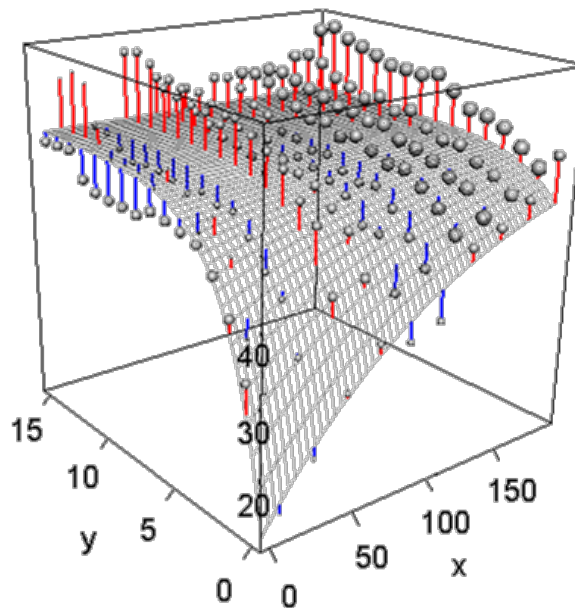
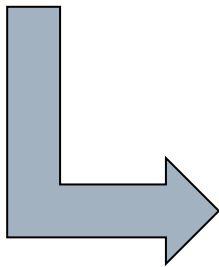
STT: trajectory



Spatio-temporal data analysis

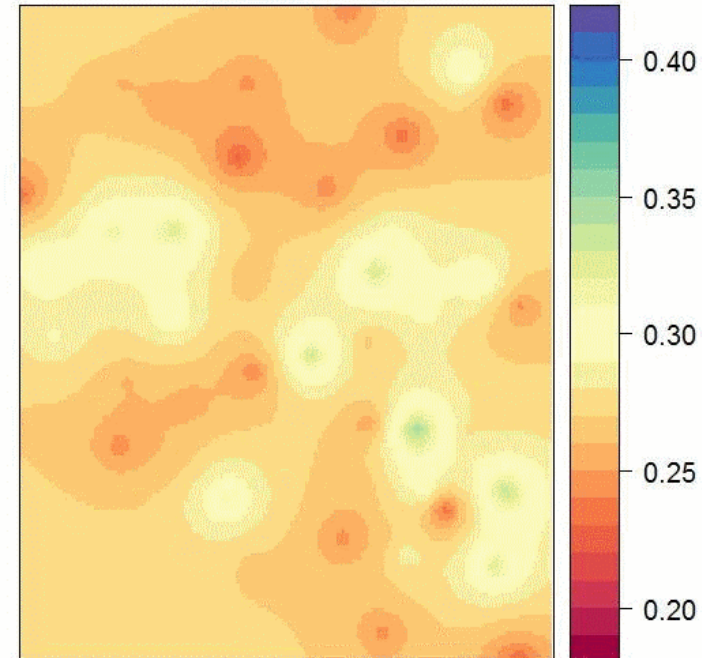


Observations

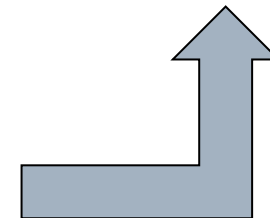


Spatiotemporal variogram

2012-07-09 00:00:00



Spatiotemporal kriging interpolation



Spatio-temporal data analysis

□ 地统计

`spacetime`, `SpatioTemporal`, `RandomFields`,
`spBayes`, `Stem`, `spcopula`, *et al.*

□ 点模式分析

`splanCS`, `stpp`, `stppResid`, `stam`, `ptproc`

□ 格数据分析

`surveillance`, `plm`, `splm`, `sphet`, `nlme`

□ 轨迹数据分析

`adehabitatLT`, `trip`, `tripEstimation`, `diveMove`,
`move`, *et al.*

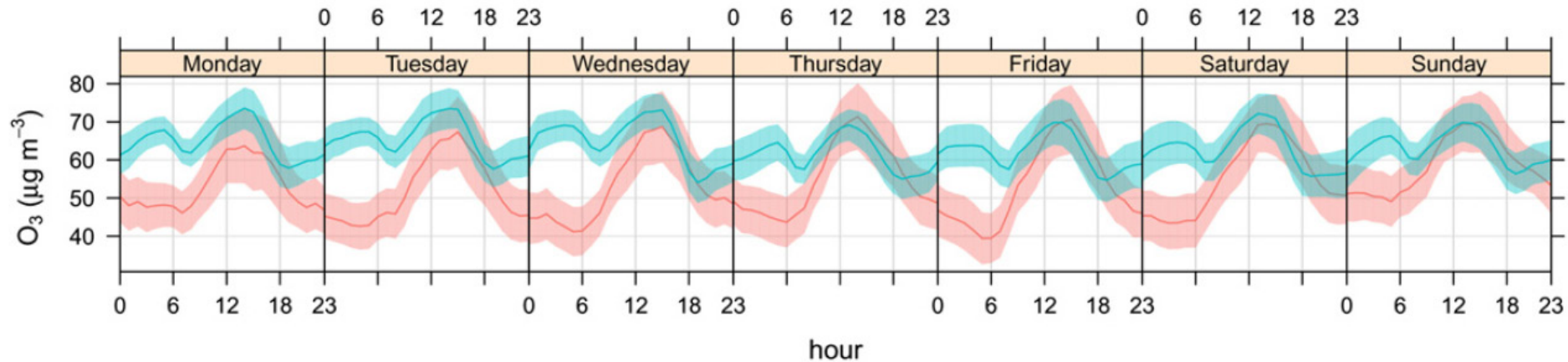
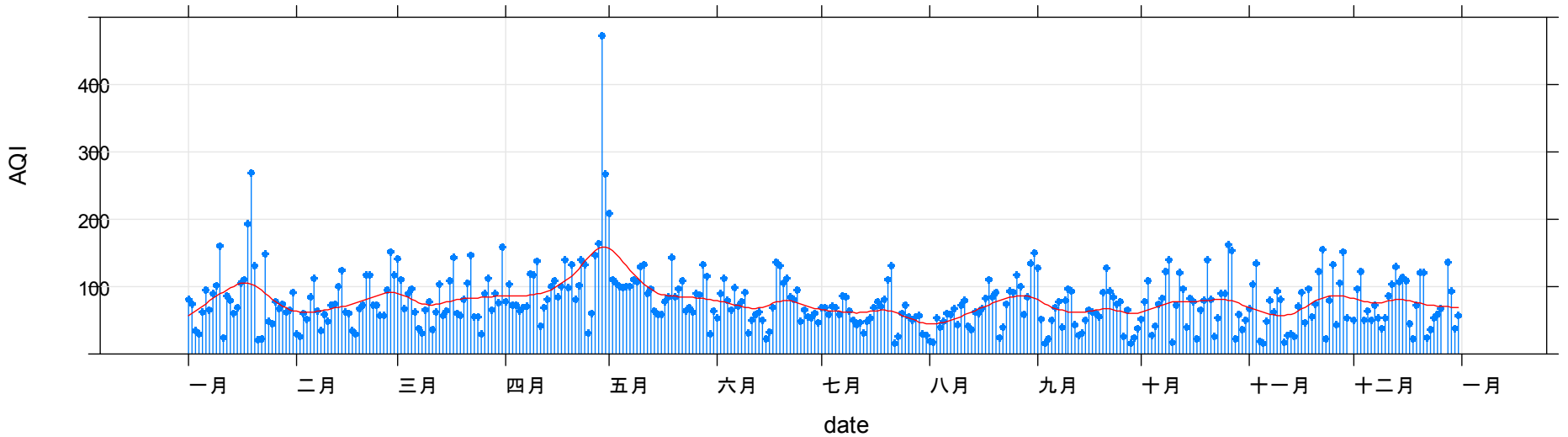
3 时空数据可视化

- ✧ R中的时空数据可视化
- ✧ Web中的时空数据可视化

Time series data visualization

```
lattice {xyplot() }
```

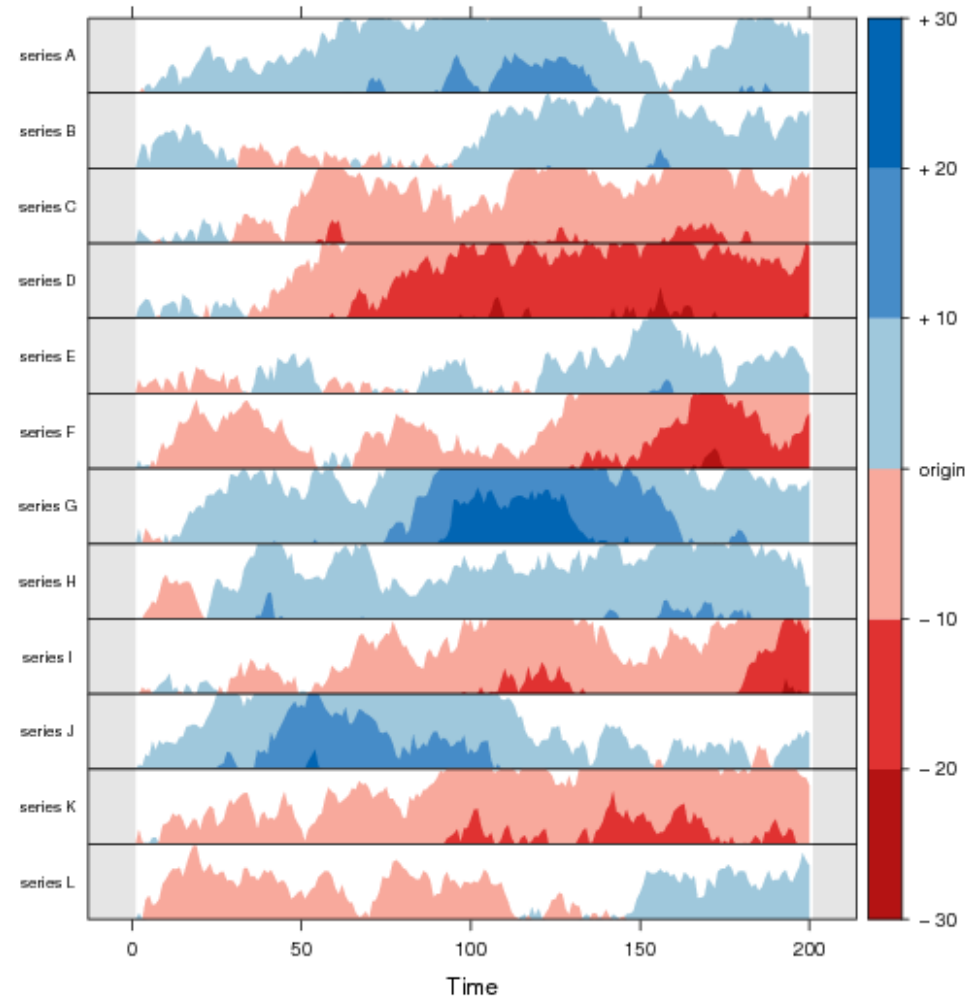
2012 Daily AQI of Beijing



TS visualization

latticeExtra
{horizonplot() }

```
horizonplot(data, horizonscale =  
10, colorkey = TRUE, layout = c(1,  
12), strip.left = FALSE, ylab =  
list(rev(colnames(dat)), rot = 0,  
cex = 0.7))  
+ layer_(panel.fill(col =  
"gray90"),  
  panel.xblocks(..., col =  
"white"))
```



Calendar Map

openair

`{calendarPlot() }`



Adobe Acrobat
Document

```
breaks <- c(0, 50, 100, 150,
200, 300, 500)
```

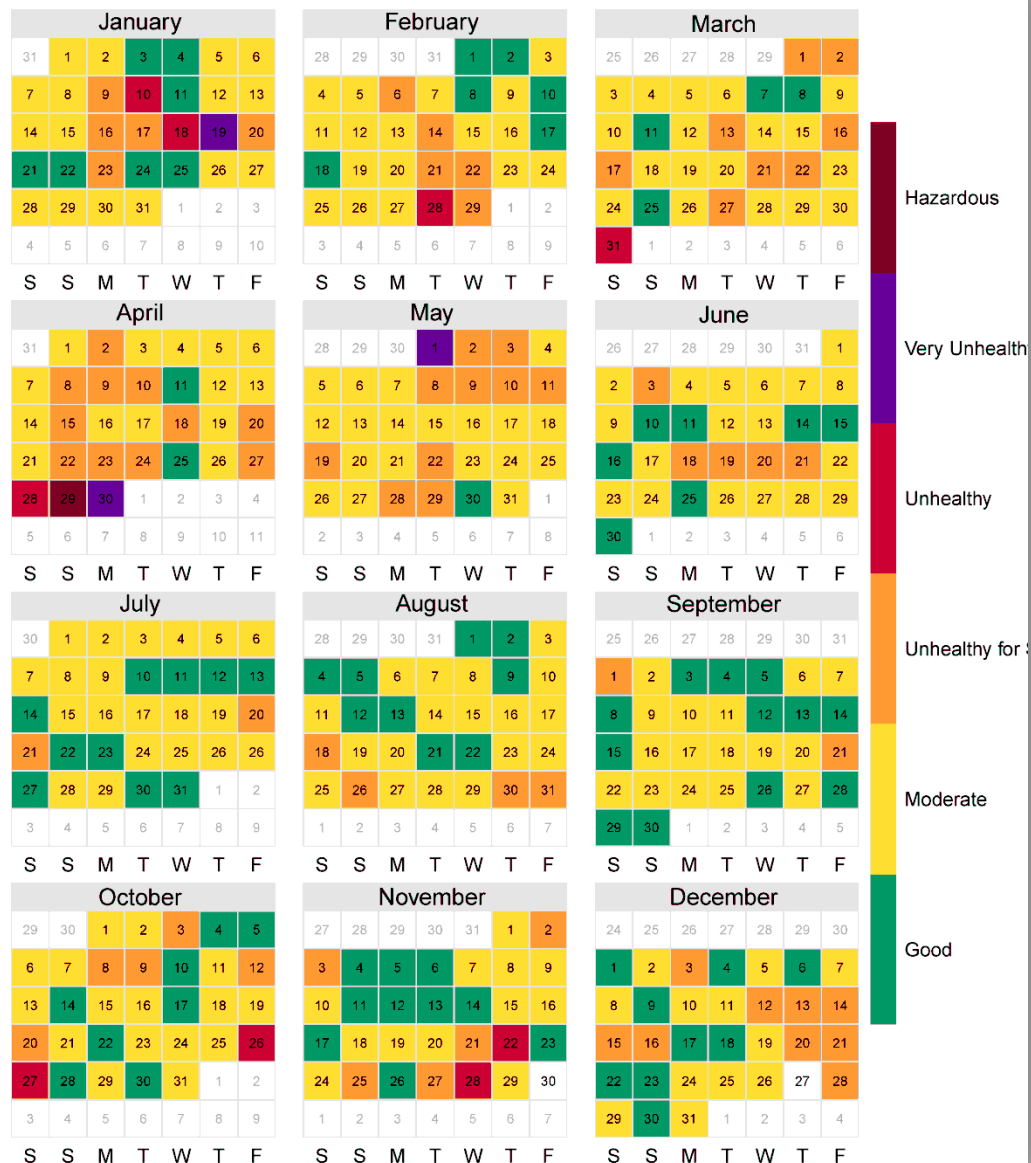
```
labels <- c("Good", "Moderate",
"Unhealthy for Sensitive
Groups", "Unhealthy", "Very
Unhealthy", "Hazardous")
```

```
colors <- c("#009966",
"#ffde33", "#ff9933",
"#cc0033", "#660099",
"#7e0023")
```

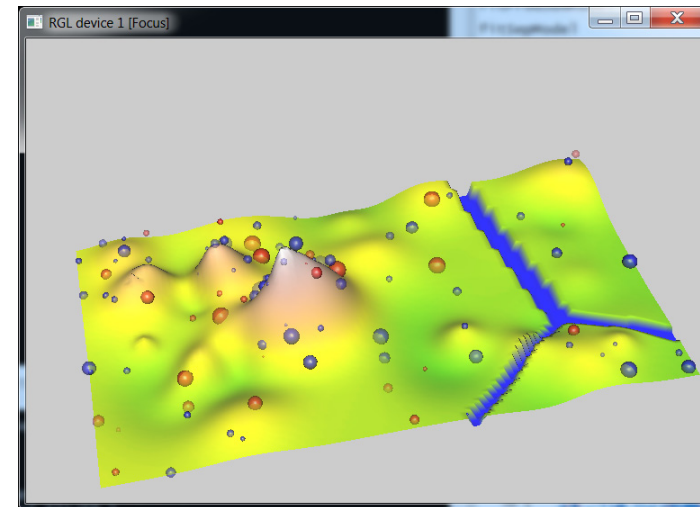
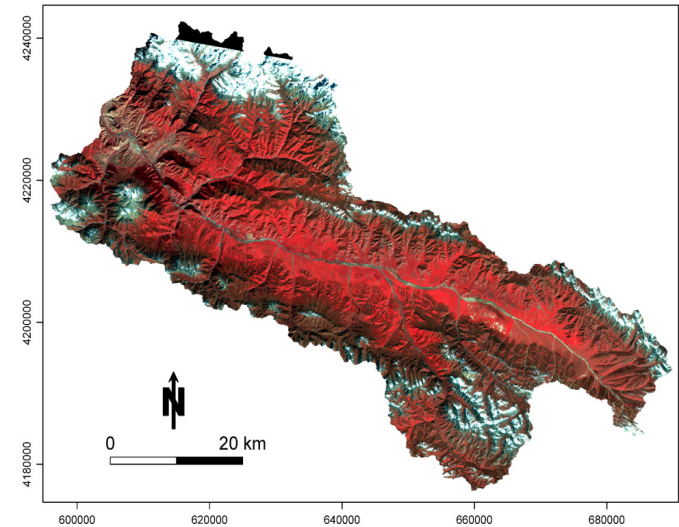
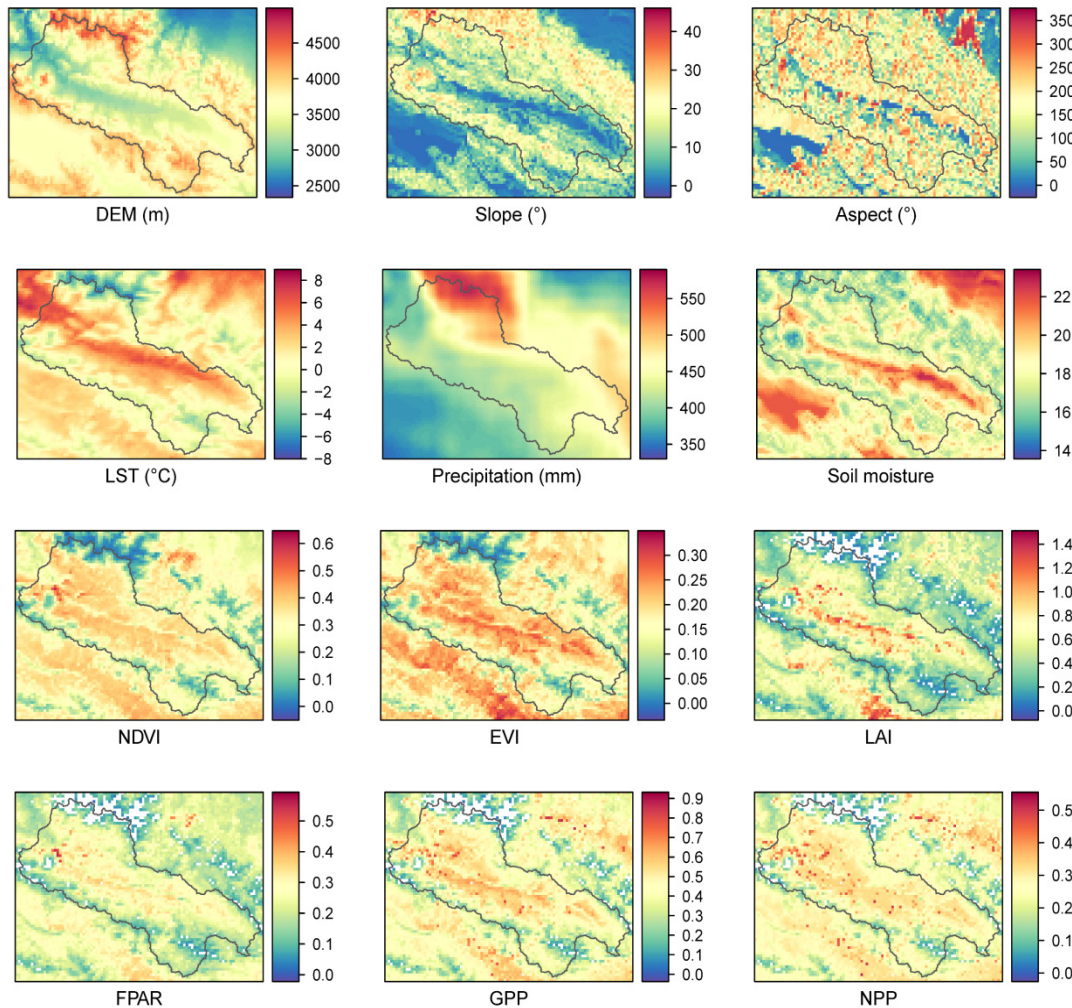
X11()

```
calendarPlot(data.sub,
pollutant = "AQI", year=2012,
breaks=breaks, labels=labels,
cols=colors, main="Beijing
2012 Daily Air Quality Index
(AQI) ")
```

Beijing 2012 Daily Air Quality Index (AQI)



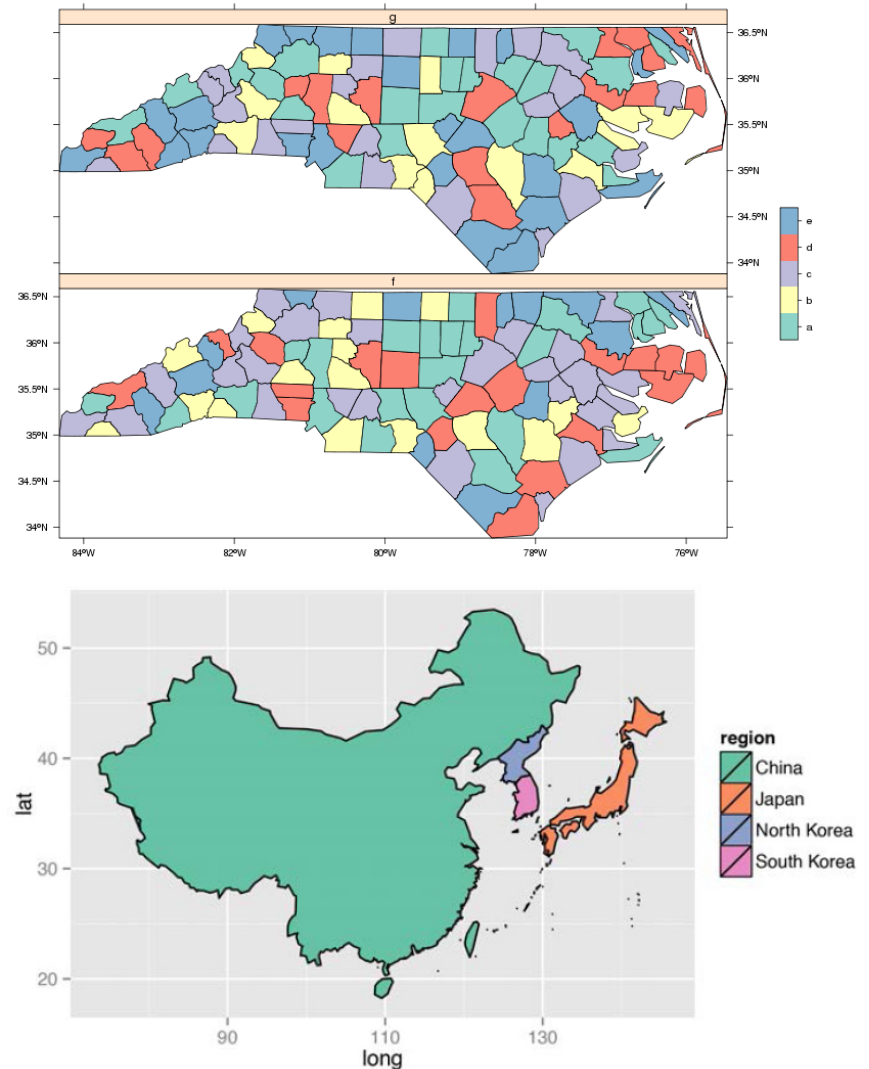
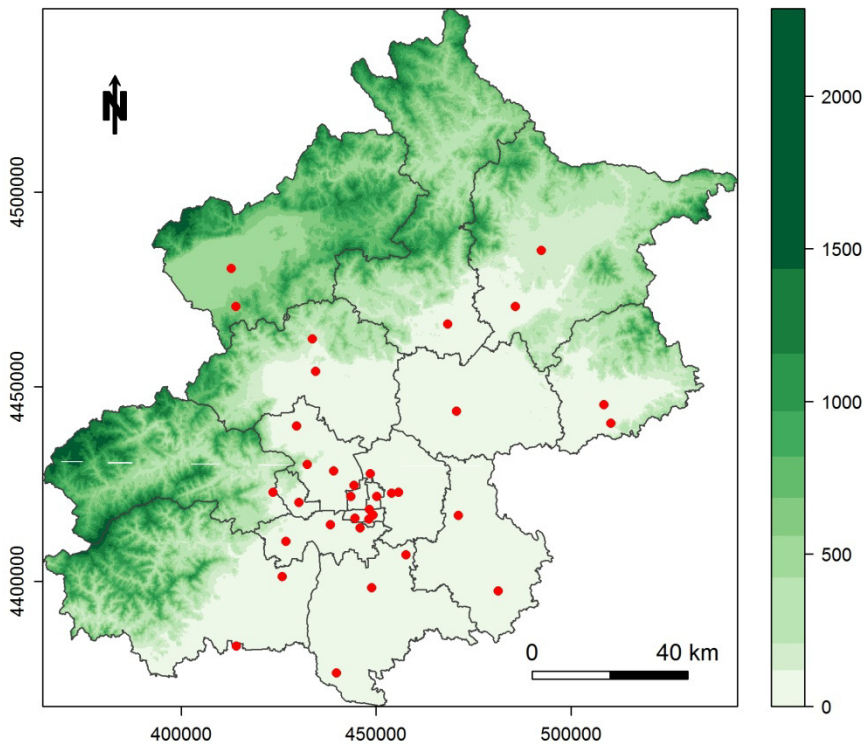
Raster data visualization



- `sp {spplot()}; raster {image() }`
- `lattice({levelplot(), contourplot(), wireframe() }`
- `rgl {surface3d}`

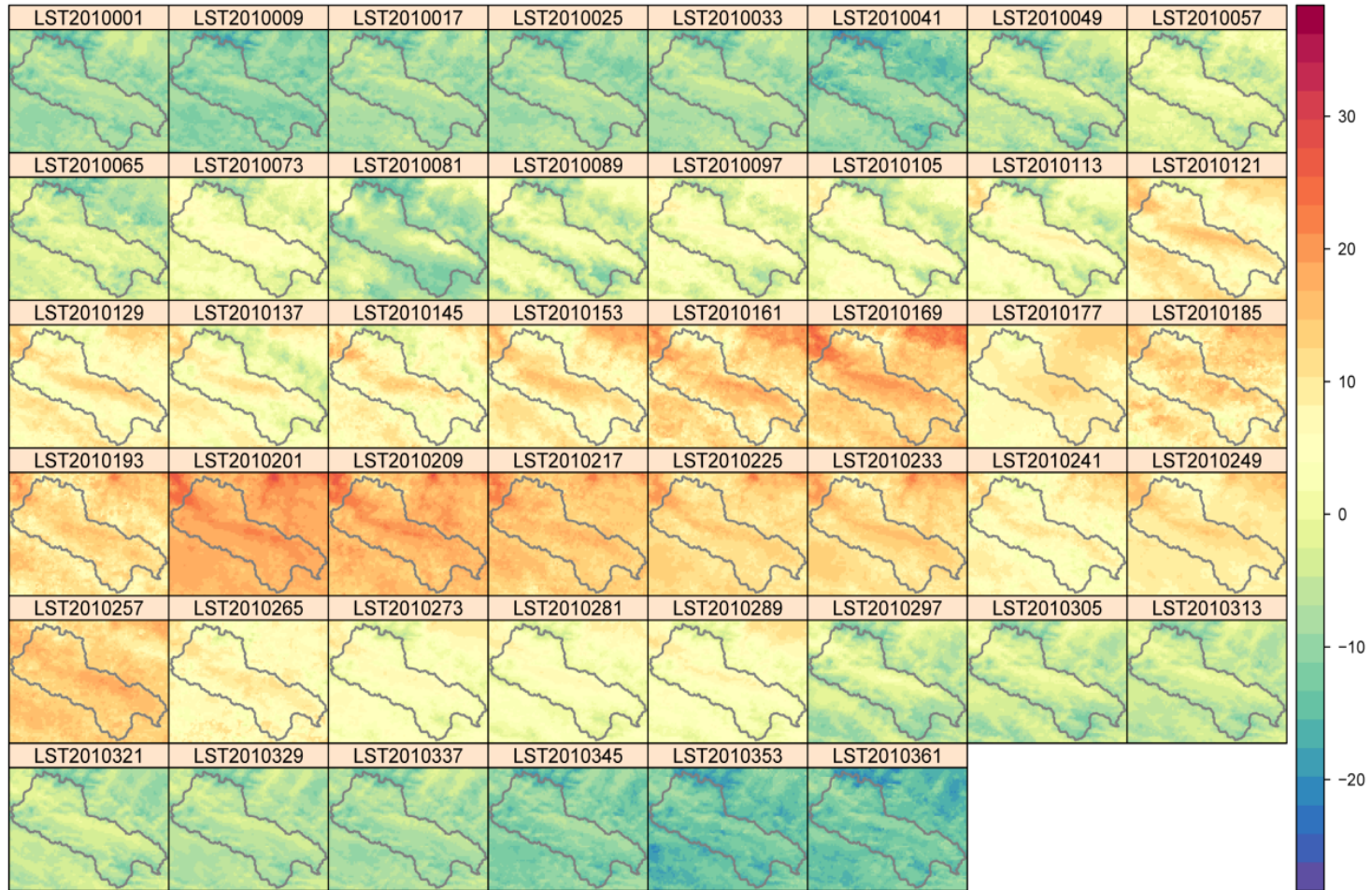
Vector data visualization

```
sp {spplot()}\nlattice {xyplot()}\nggplot2 {geom_polygon}
```



<http://rspatial.r-forge.r-project.org/gallery/>

Spatio-temporal data visualization



```
lattice {levelplot() }
```

Spatio-temporal visualization

```
spacetime {stplot() }
```

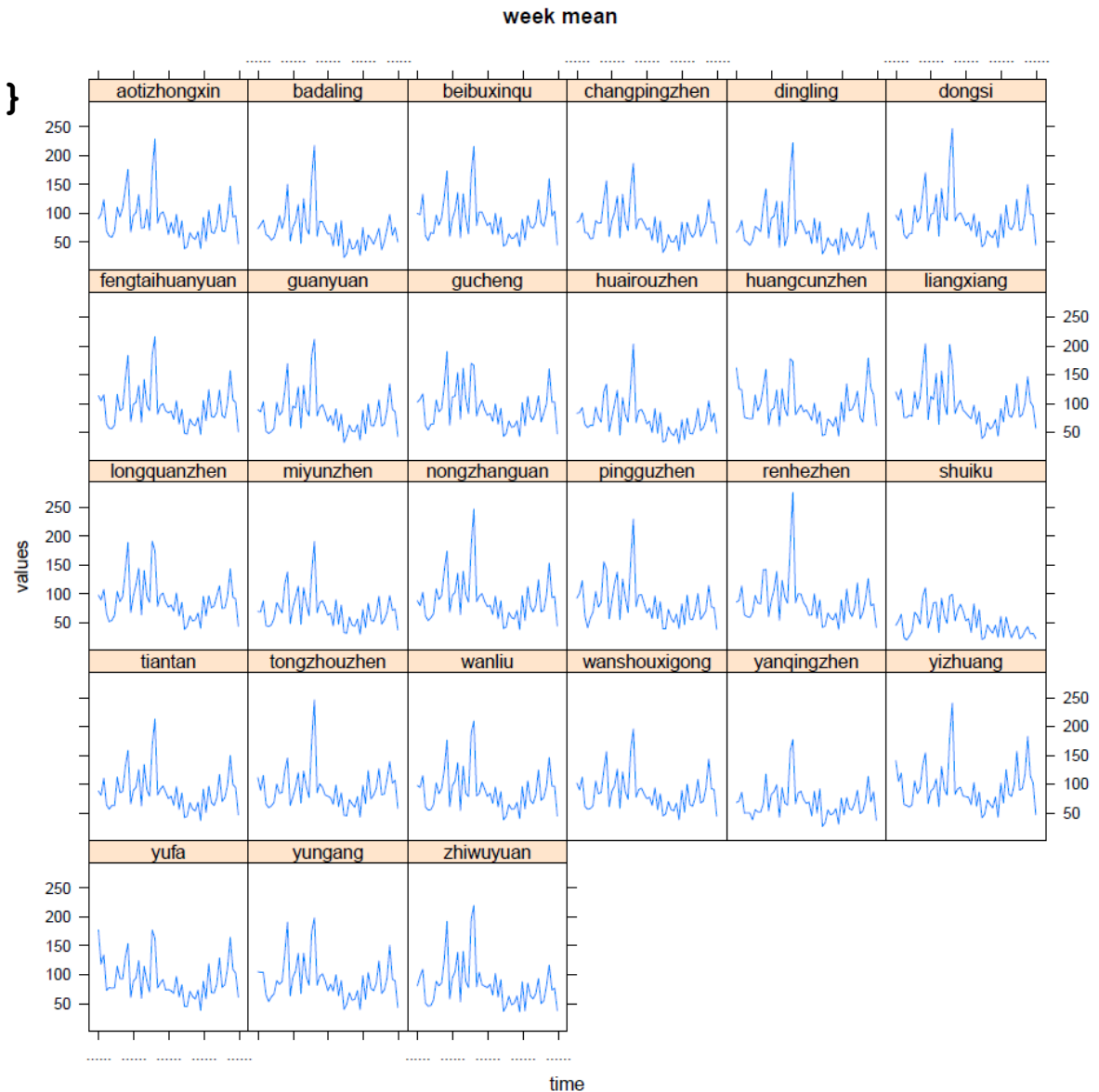
MODE:

= "tp"

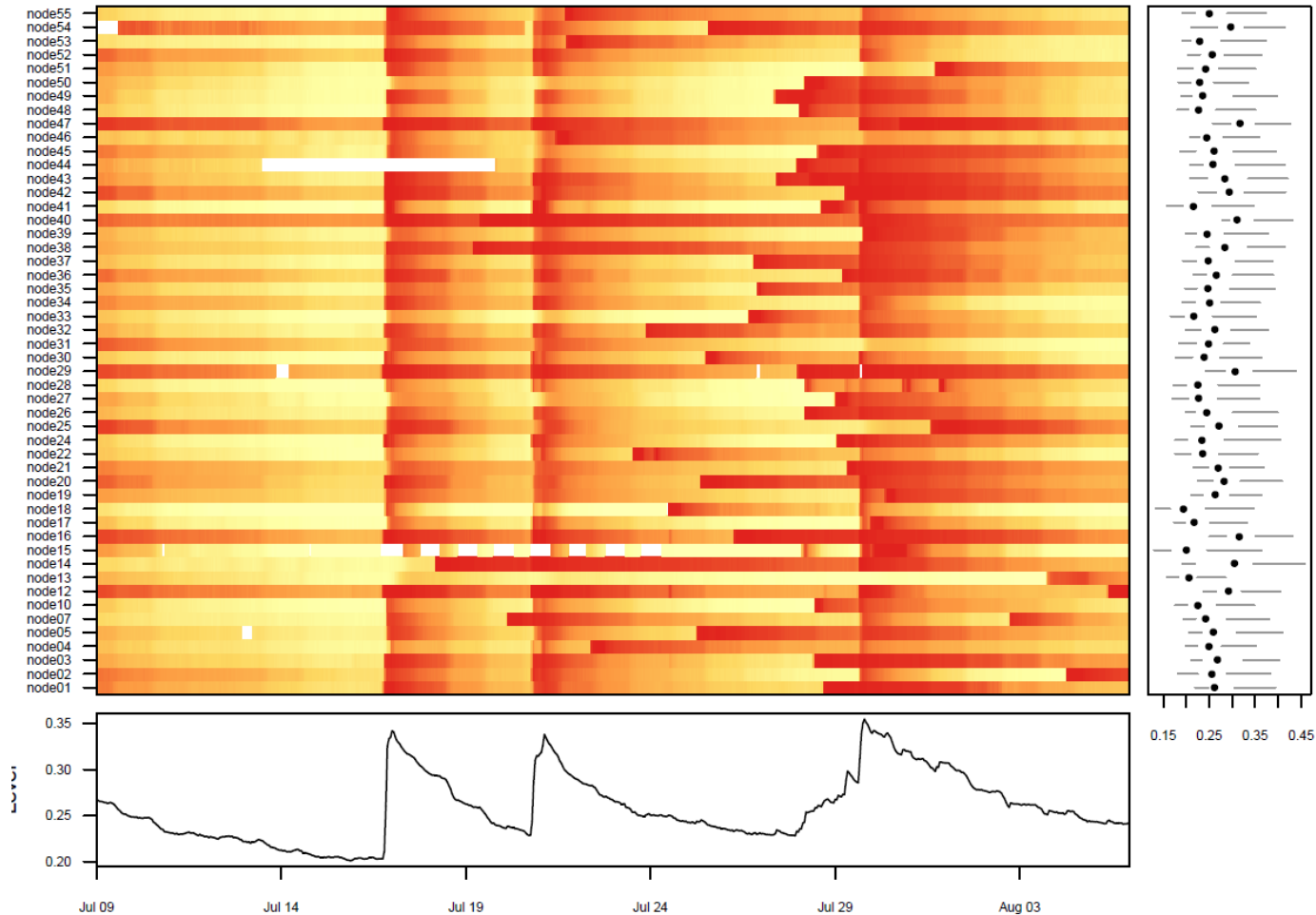
= "ts"

= "xt"

= "xy"



Spatio-temporal data visualization

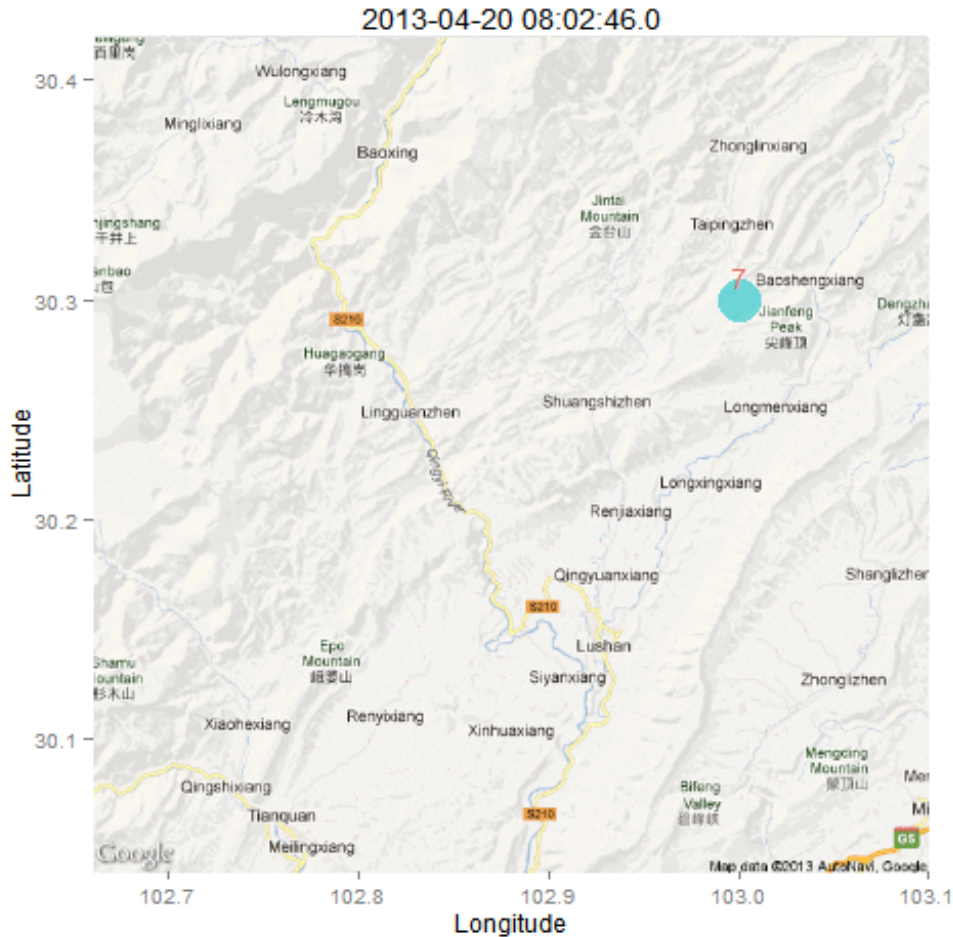


`mvtsplot {mvtsplot() }`

Peng, R. (2008). "[A method for visualizing multivariate time series data.](#)" Journal of Statistical Software 25(1): 1-17.

Spatio-temporal data visualization

animation



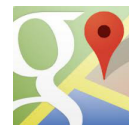
```
gridSVG  
{gridToSVG() }
```

<http://oscarperpinan.github.io/spacetime-vis/images/vLine.svg>

3.2 Web中的时空数据可视化

利用网络地图可视化空间数据

□ Google Map or Earth : <https://maps.google.com/>



□ Baidu Map : <http://developer.baidu.com/map/>



□ Leaflet : <http://leafletjs.com/>



□ OpenLayers : <http://openlayers.org/>



□ OpenStreetMap : <http://www.openstreetmap.org/>



□ ArcGIS online : <http://www.arcgis.com/home/>



□ Cloudmate : <http://cloudmade.com/>



□ Ploymaps : <http://polymaps.org/>



Google Map

- [Google Map APIs](#)
- [Google Map Gadget](#)
 - 通过插件的形式将自己的数据嵌入到Google Map
- KML
 - 上传KML文件，在google map中加载kml文件，返回的地址，如
https://maps.google.com/maps?q=http://globalsoilmap.net/data/E33_S14_ORCDRC_1_gif.kml
- Fusion Table
 - CSV文件直接生成地图
- [igoogle 插件](#)，直接嵌入网页

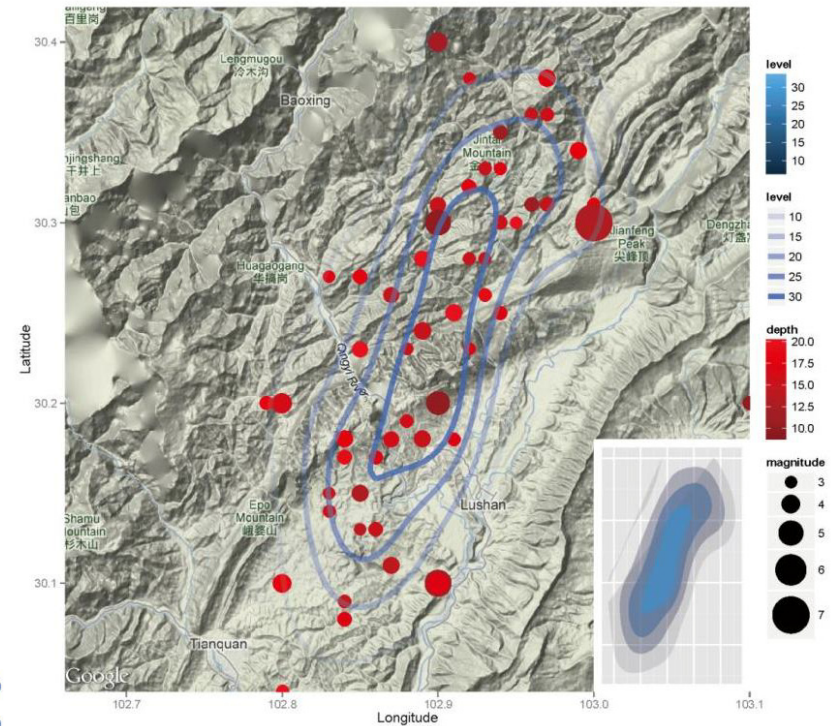
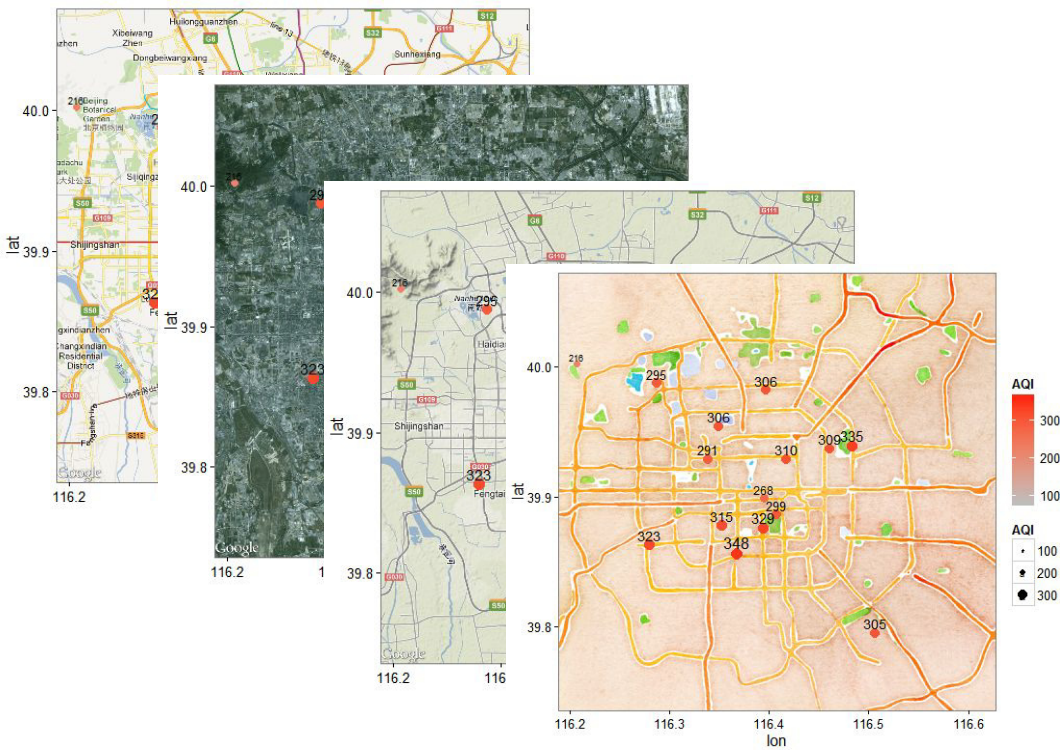
Google Map & R

□ ggmap

Kahle, D. and H. Wickham "[ggmap: Spatial Visualization with ggplot2.](#)" [The R Journal.](#)

□ Data Sources

1. Google Maps
2. OpenStreetMap
3. Stamen Maps
4. CloudMade Maps



<http://jianghao.github.io/earthquake/spacetime/ppa.htm>

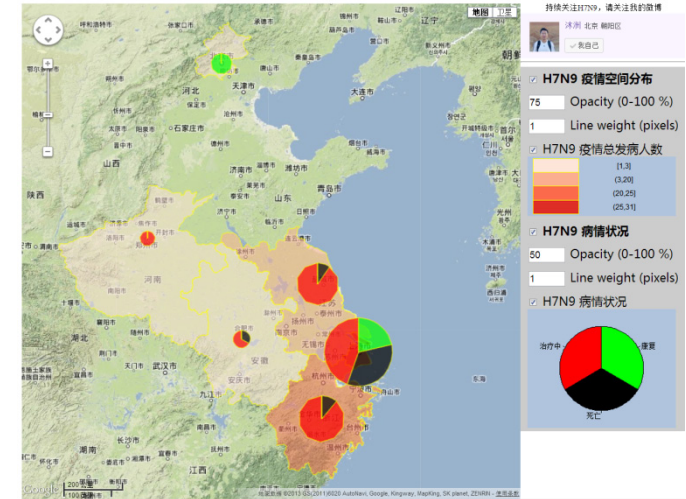
Google Map & R

□ plotGoogleMaps

- [Demo1](#) 面
- [Demo2](#) 线
- [Demo3](#) 点
- [Demo4](#) 点

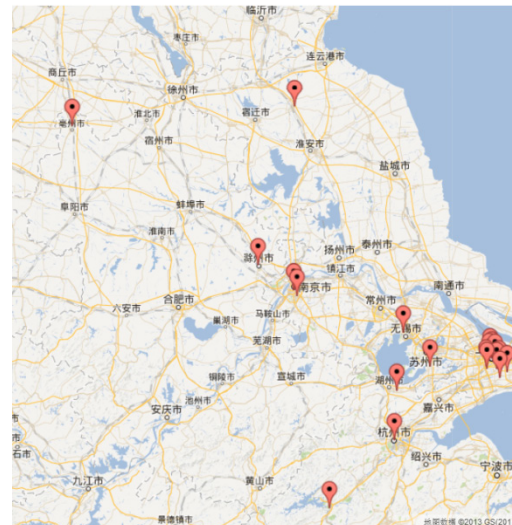
H7N9 禽流感疫情信息

更新时间: 2013年4月19日 12:00 AM

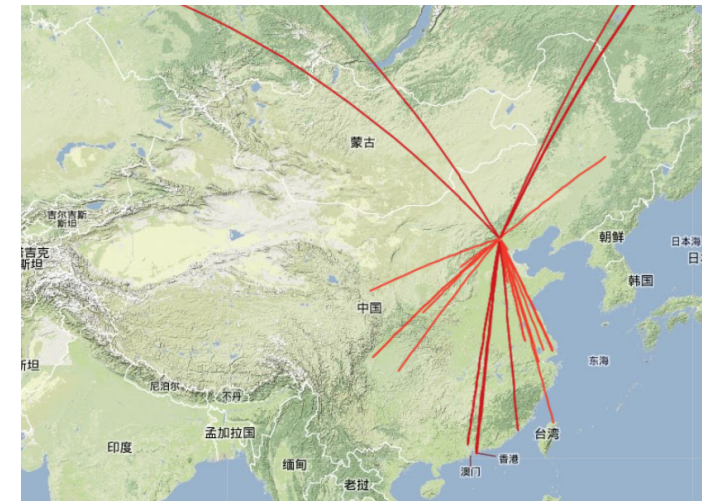
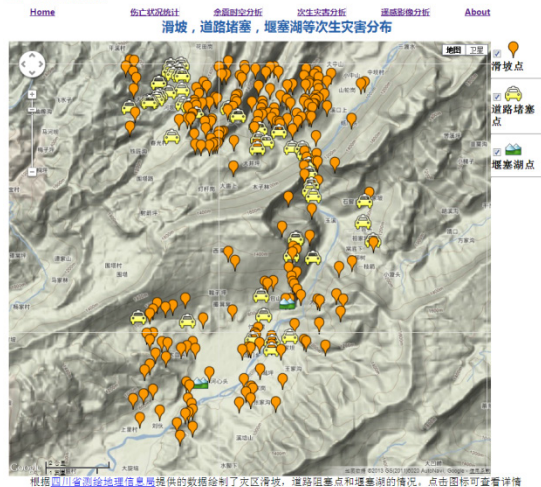


H7N9 禽流感疫情信息

更新时间: 2013年4月8日 星期一 16:00 PM



四川雅安地震信息

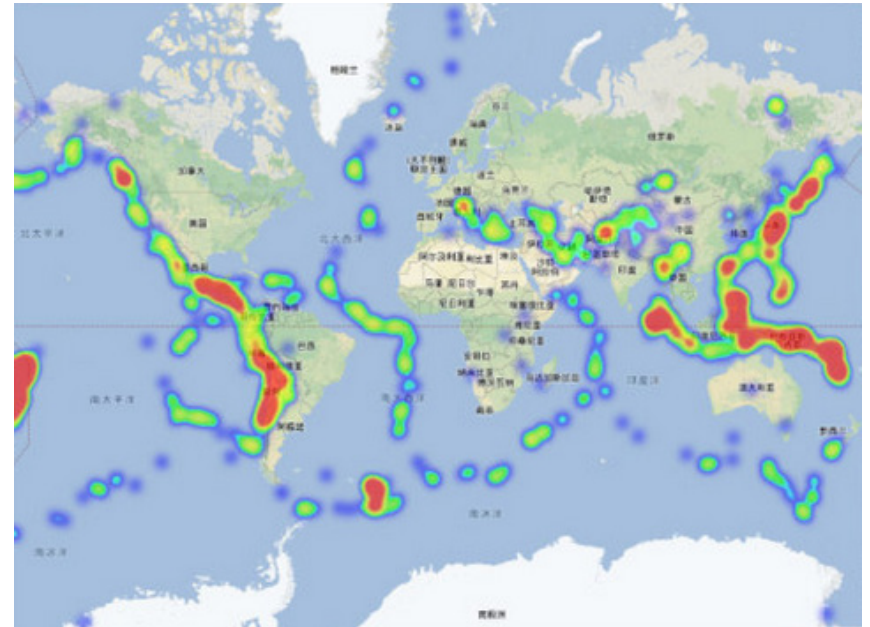
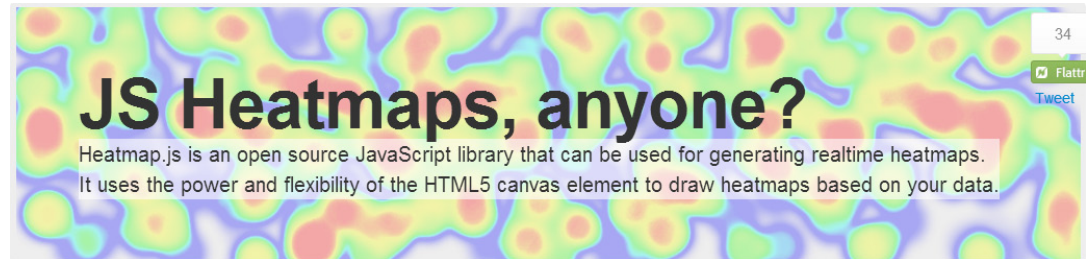


Google Map & R

□ heatmap.js

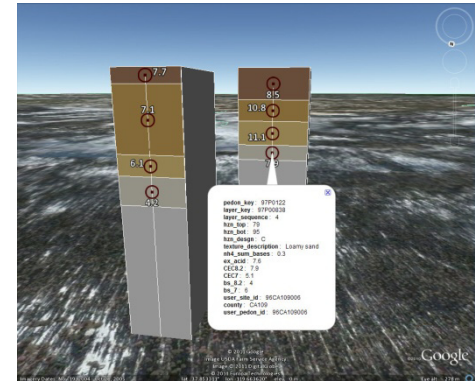
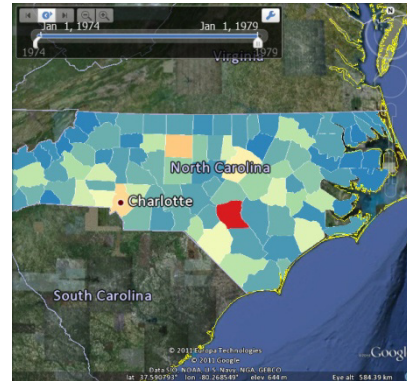
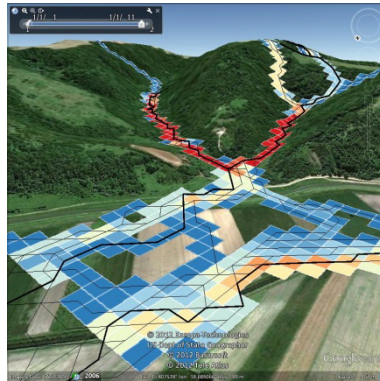
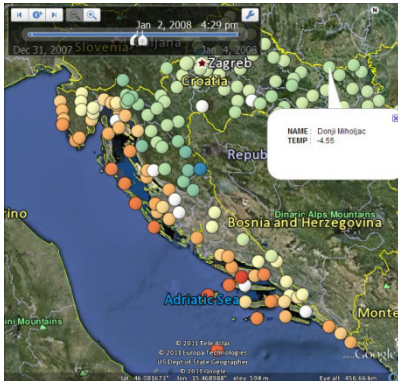
Demo1 (H7N9)

Demo2 (雅安地震)



Google Earth & R

plotKML



thematicmapping.org



Thematic Mapping API

Create KML based thematic maps with a few lines of JavaScript

Web Map & R

- ❑ `RgoogleMaps`: like `ggmap`
- ❑ [`osmar`](#): OpenStreetMap and R
- ❑ `OpenStreetMap`: Access to open street map raster images

Html5 + CSS3 + JS & R

□ googleVis

- `vignette('googleVis')`
- `demo('AnimatedGeoMap')`
- `demo('EventListener')`
- `demo('WorldBank')`
<http://bost.ocks.org/mike/nations/>
- `demo('googleVis')`



□ rchart

interface for R users to create interactive visualizations using polychart.js

demo : <http://ramnathv.github.io/rCharts/>



- gg2v : Render ggplot2 graphics using vega
- rHighcharts; rNVD3; clickme; rVega

Shiny

- ❑ Shiny: <http://www.rstudio.com/shiny/>
- ❑ Shiny [demos](#) on Github
 - [Reconstruct Gene Networks](#)
 - [WebGL Demo](#)
 - [2010 US Census Shiny App](#)

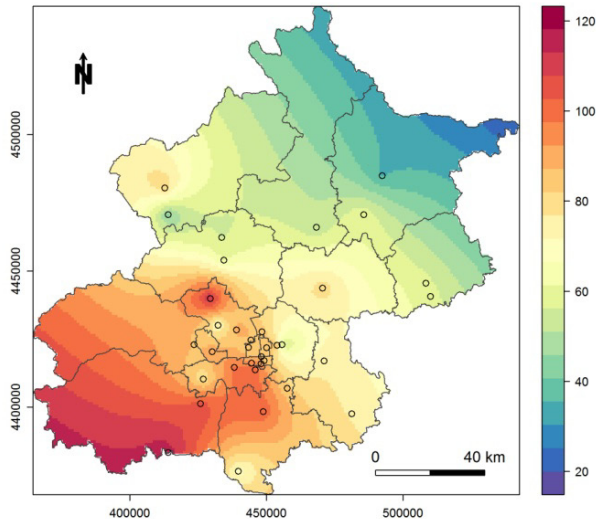
Web Visual

- ❑ **d3js**: <http://d3js.org>
- ❑ **Processing** (Java): <http://processing.org/>
- ❑ **Many Eyes**: <http://www-958.ibm.com/software/data/cognos/manyeyes/>
- ❑ **Open Flash Chart** (Flash): <http://teethgrinder.co.uk/open-flash-chart/>
- ❑ **simile** (AJAX): <http://simile.mit.edu/>
- ❑ **FLARE** (ActionScript): <http://flare.prefuse.org/>

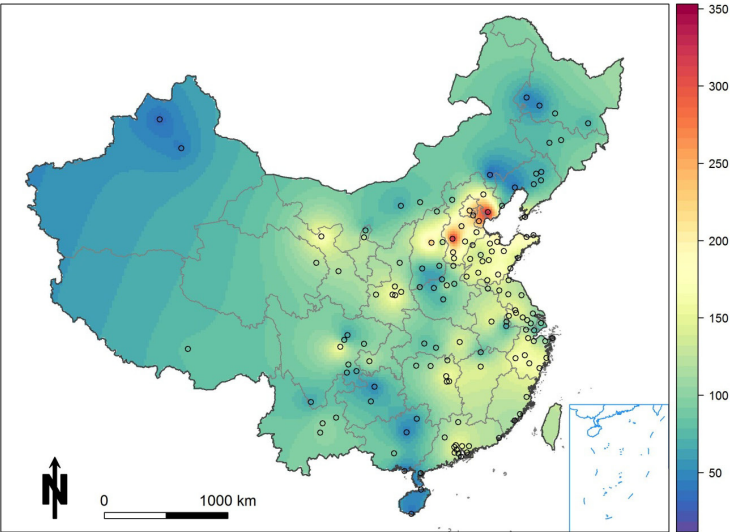
4 时空可视化应用

空气质量指数AQI分析与可视化

Mean AQI = 70



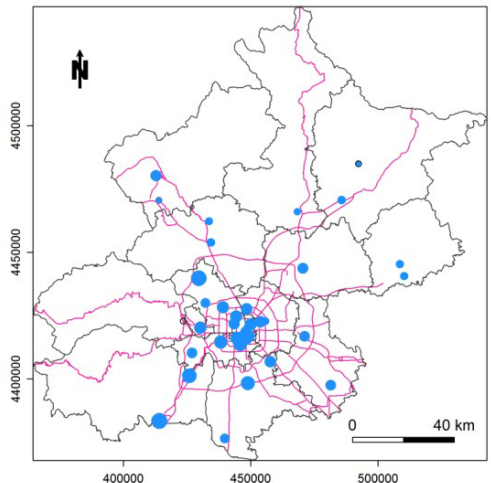
2013-04-03 Wednesday 13:00 Chinese mean AQI = 122



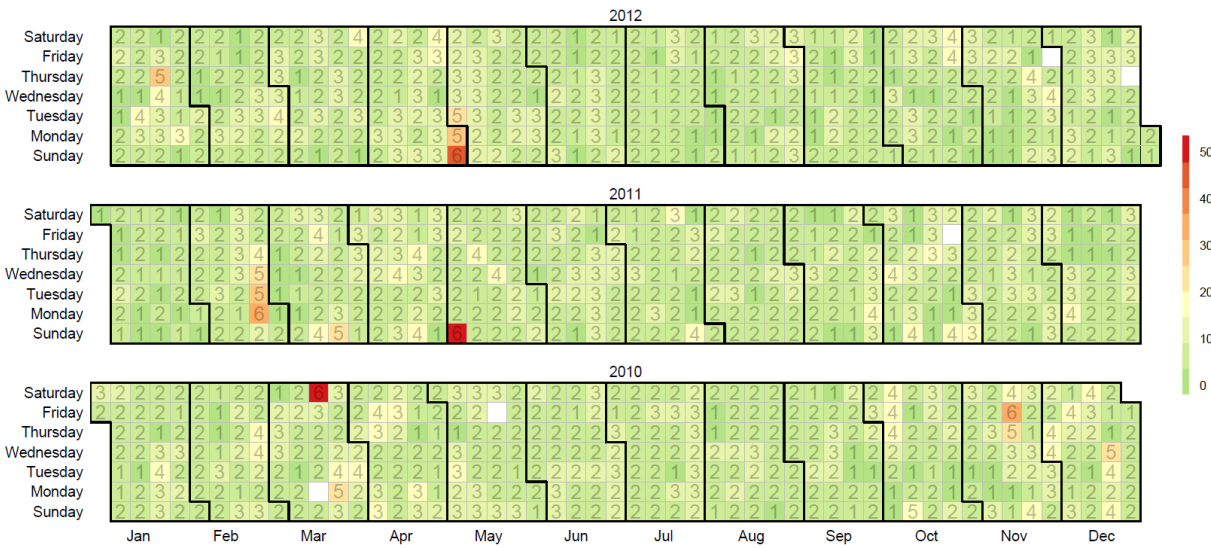
Data source: CNEMC

新浪微博发布
@沐洲

Beijing 2010 - 2012 Daily AQI Calendar Map



Time:03/05/2013 10:00:00



雅安地震信息发布

<http://jianghao.github.io/earthquake/>

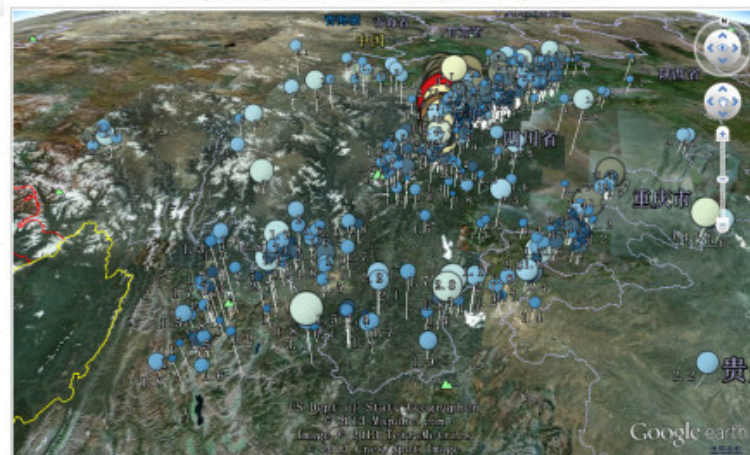
四川雅安地震信息



预览 地震时间序列 地震动画展示 Google Earth三维(所有点) Google Earth三维(>30级) 点模式分析

四川区域2013年4月所有地震三维显示

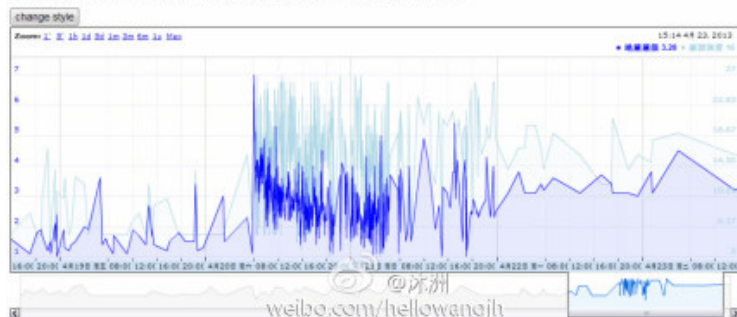
需要Google Earth Plugin, 请下载安装! 或者直接 下载KML文件, 然后在Google Earth中打开。



预览 地震时间序列 地震动画展示 Google Earth三维(所有点) Google Earth三维(>30级) 点模式分析

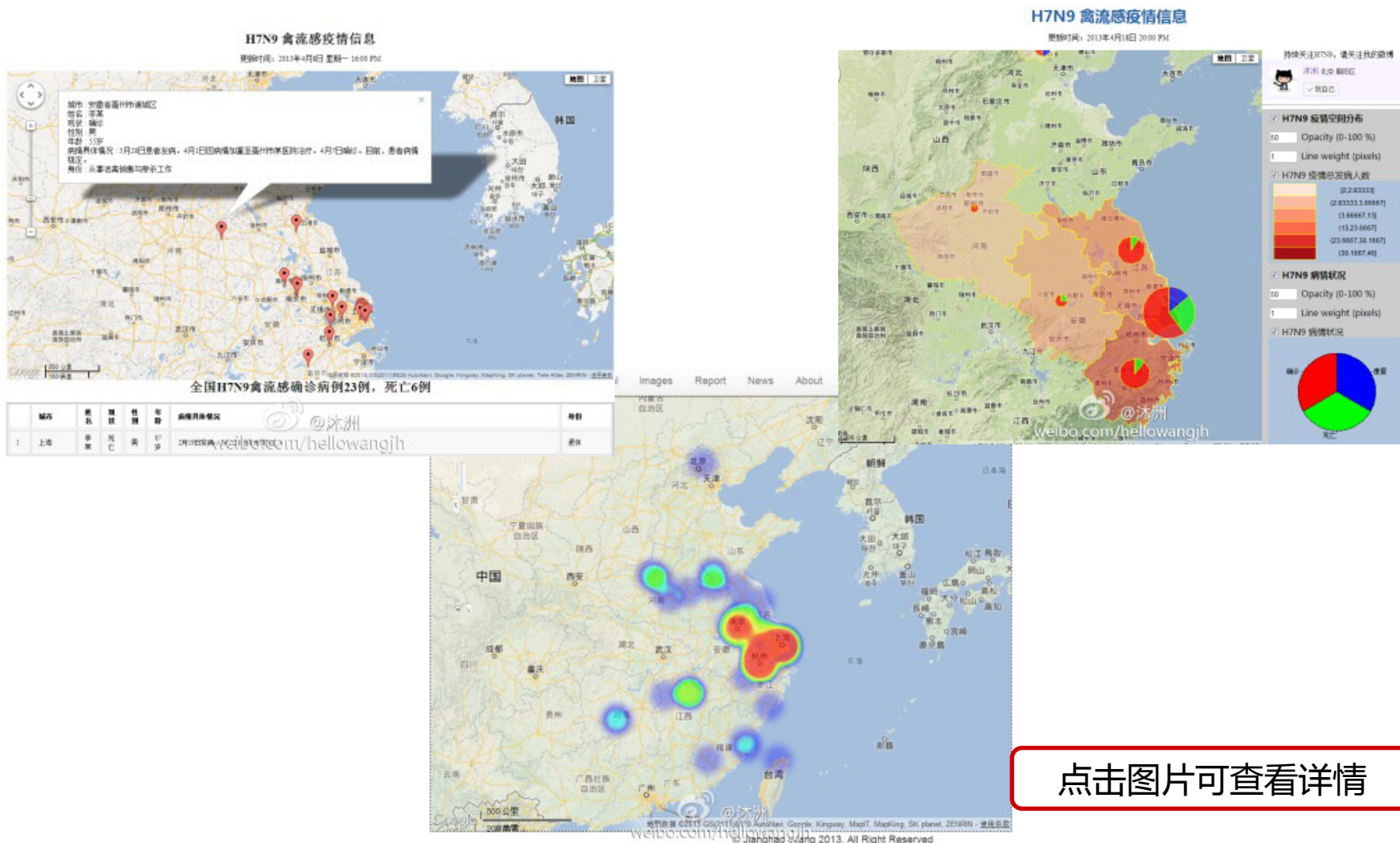
四川地区2013年4月以来地震时间序列分析

查询地震数据共享中心四川地区(经度介于97-106度, 纬度介于26-34度), 自2013年4月1日以来发生的所有地震, 共计1055条记录(数据下载 或在线查看表格), 其中绝大多数为震级强度小于3.0的小地震, 将地震震级按时间序列绘制在下图中, 其中深蓝色表示地震的震级, 浅蓝色表示地震的震源深度, 可以移动图表查询任一次地震的详细信息, 非常明显的, 下图中地震发生地震的时间正是四川雅安地震爆发时间。



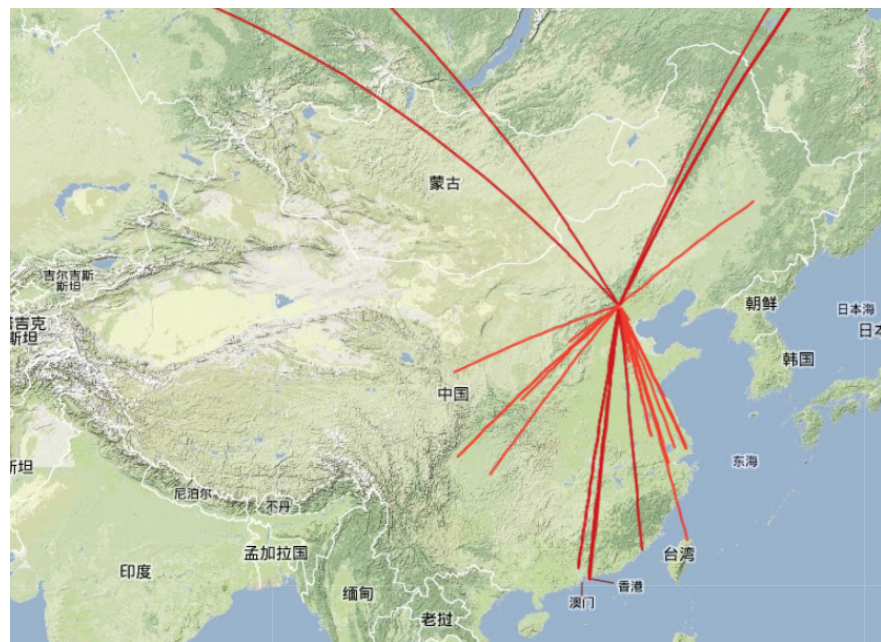
H7N9 信息发布

<http://jianghao.github.io/H7N9/index.html>



微博社交网络分析

<https://github.com/Jianghao/weibo> (待完善)



点击图片可查看详情

科学研究中

□ 主要研究方向

- 时空数据分析
- 遥感影像的不确定性
- 时空地统计学
- 监测网布局优化
- 时空插值与过程模拟
- 。 。 。

□ 论文 90%的内容均由R语言完成

Thanks!

Q & A