Geomedicine

Opportunities of using spatial information to move toward more precision in public health

Spatial approaches and clusters

An introduction for clinicians

stephane.joost@giraph.org
idris.guessous@giraph.org

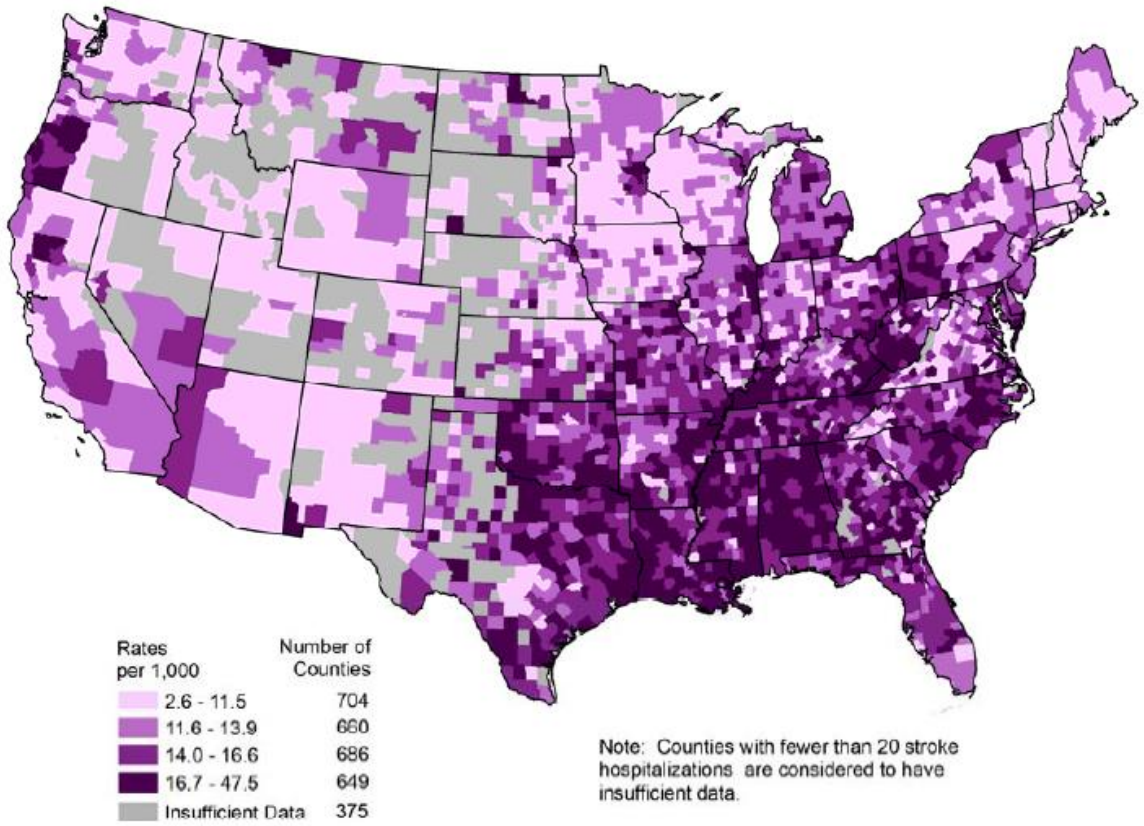
Laboratory of Geographic Information Systems (LASIG, EPFL)
Geographic Information Research and Analysis for Public Health (GIRAPH)
Unit of Population Epidemiology, (UEP, HUG)

SGAIM/SSMIG/SSGIM
Die Zukunft ist jetzt
Le futur c'est maintenant
2. SGAIM FRÜHJAHRSKONGRESS, 3.– 5. Mai 2017
2ÈME CONGRÈS DE PRINTEMPS DE LA SSMIG, 3 – 5 mai 2017
SwissTech Convention Center, Lausanne

2ème congrès de printemps de la Société Suisse de Médecine Interne Générale (SSMIG) **3-5 mai 2017,** SwissTech Convention Center, Lausanne

Introduction

- Geography, maps and statistics
- Most often, geographic maps are produced to represent health data
- Medical information is transmitted through choropleth/thematic maps
- i.e. surface or point administrative units are colored according to the variable of interest

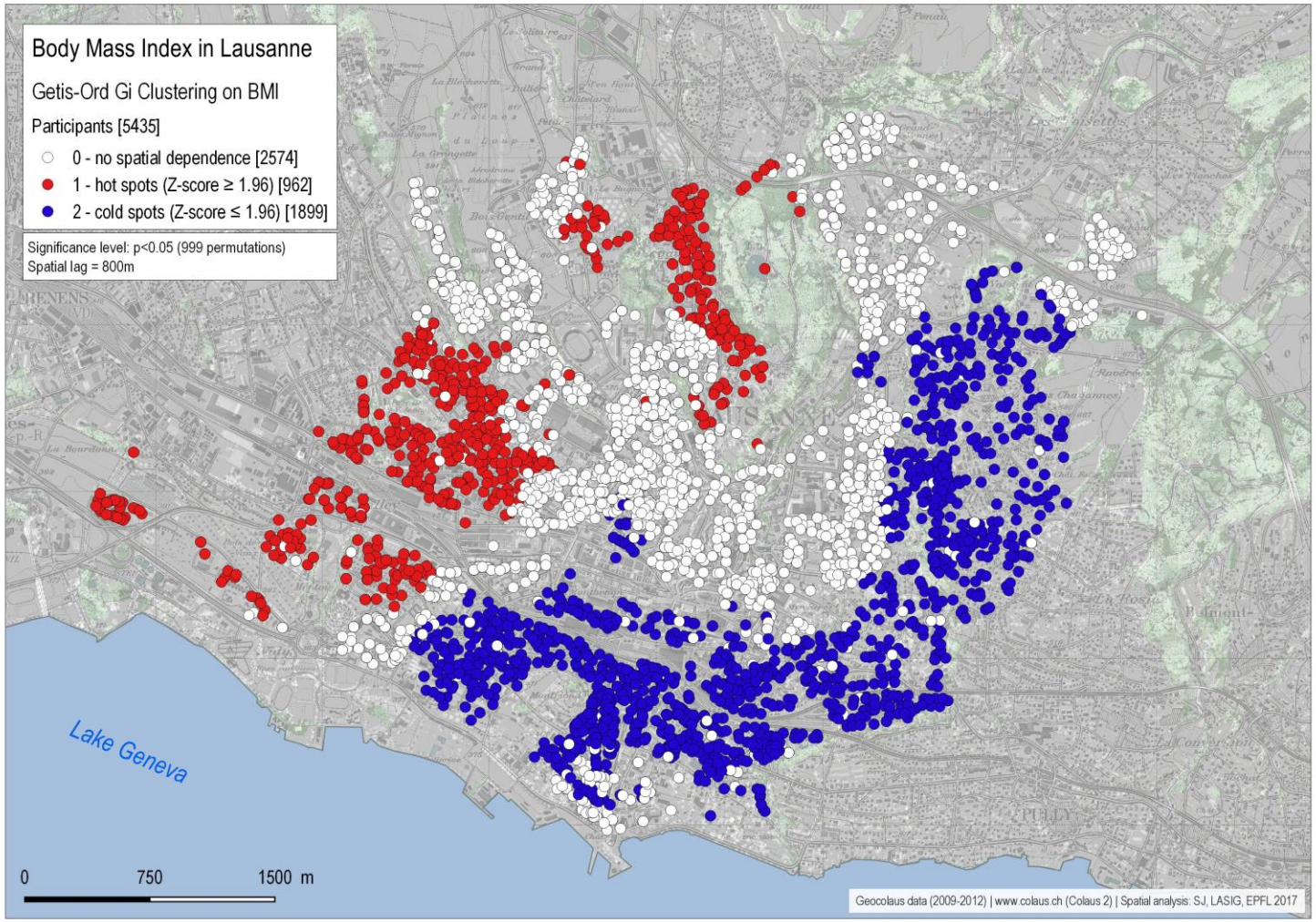


Age-adjusted stroke hospitalization rates among Medicare beneficiaries aged 65 or older in 2005-2006 at the county level

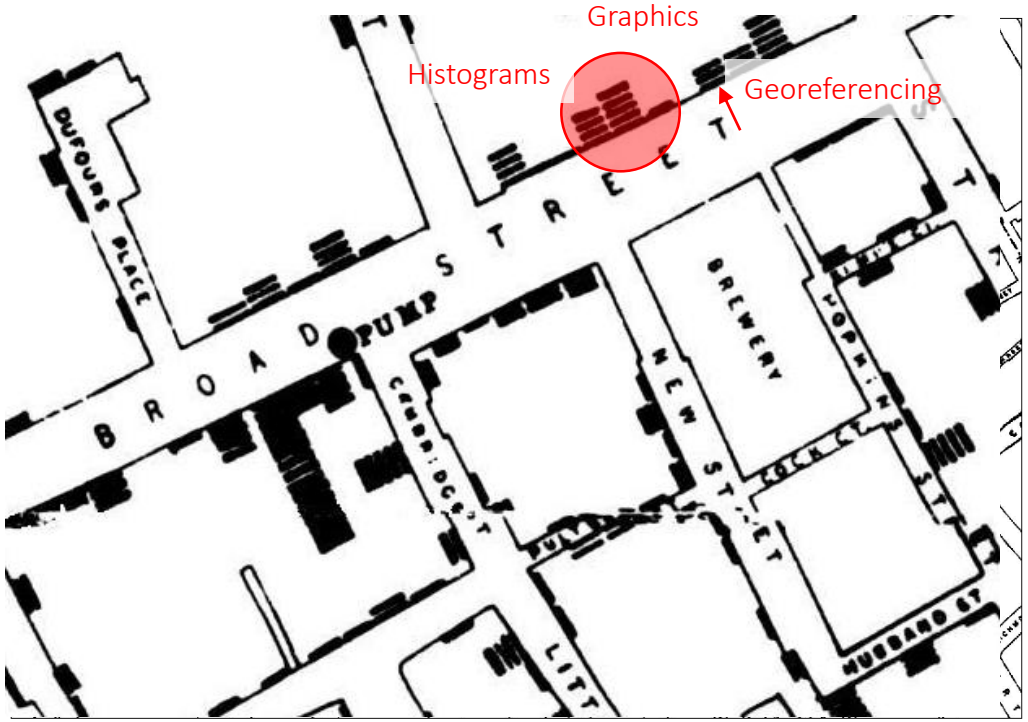
Schieb et al. (2012) doi: 10.1161/STROKEAHA.112.669705

Introduction

- In this presentation I will stress:
  - the importance of **analysing** health data by explicitly including geographic characteristics (distances, co-location)
  - the potential and power of spatial statistics to detect specific patterns in the geographic distribution of disease occurrences (make visible the invisible)

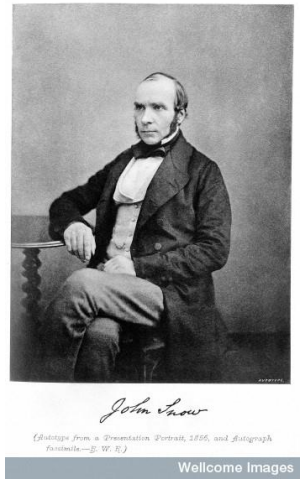


John Snow, cholera outbreak and clusters (1854)



Cholera death locations with proportional circles showing the number of deaths. The pumps are displayed in blue.
By Robin Wilson, http://blog.rtwilson.com/

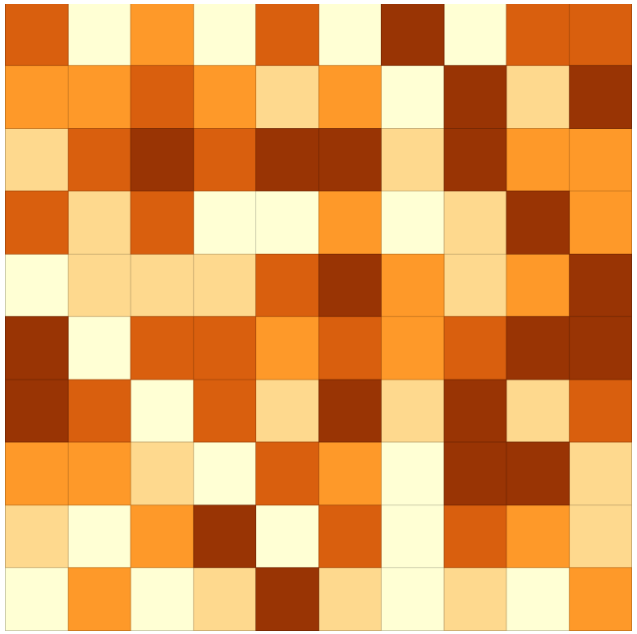
Modern map of that area, using the Open StreetView data
Data available here:
http://blog.rtwilson.com/john-snows-cholera-data-in-more-formats/

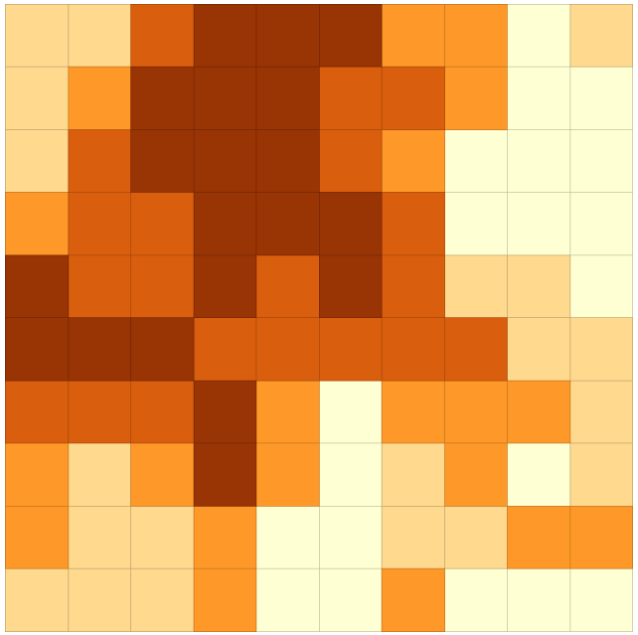


- Clusters reveal spatial dependence – effect on the territory
- Depends on a co-located infected water pump – cause
- Removing the handle of the pump stopped the outbreak
- How to detect spatial dependence and measure it?

Spatial dependence

- Main objective: identify patterns in the geographic space
- Is the variable of interest randomly distributed?



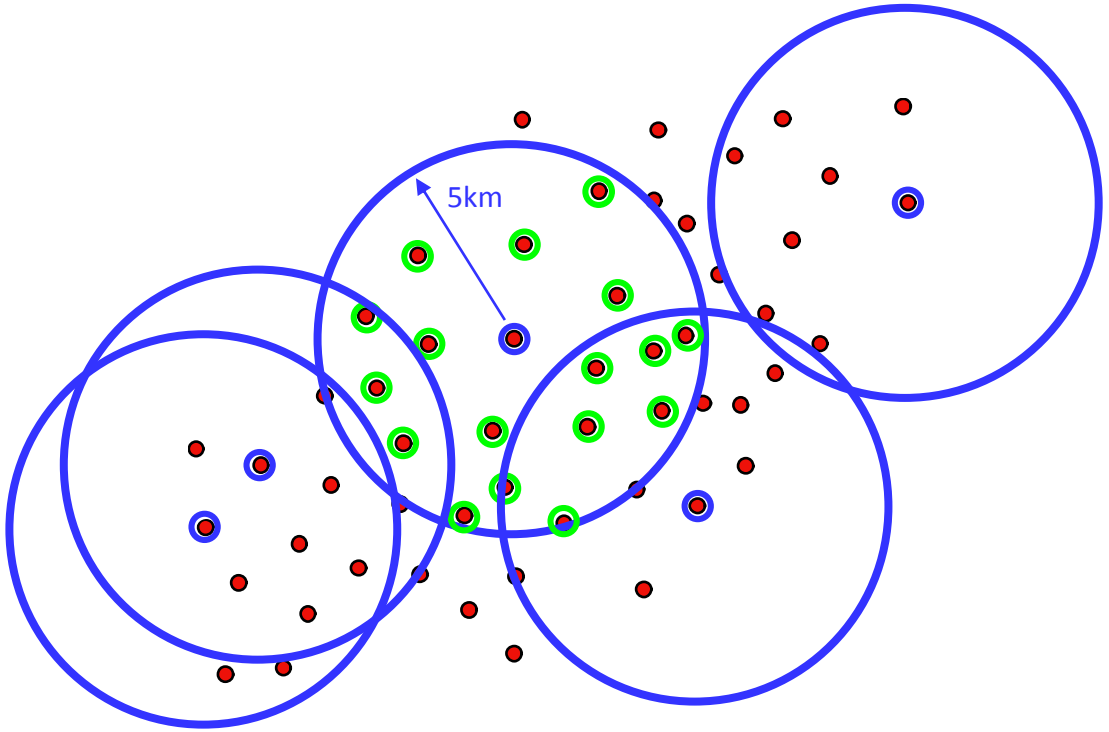


- Or spatially dependent?

- Are the spatial patterns observed robust to random permutations?
- Explore data to find out what is the range of influence of the spatial dependence ?

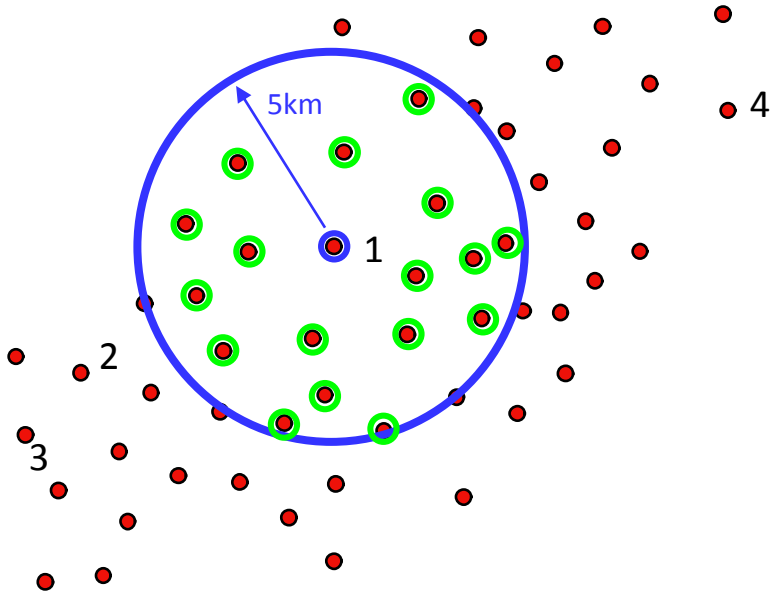
Neighborhood

- One among several measures of spatial autocorrelation: Moran's I
- This measure of spatial dependence is characterized by a correlation between the values of neighboring geographic units for a given variable



Moran's I as a coefficient of regression

- At point 1, the value of the variable of interest (e.g. BMI) is 20 while the mean of its neighborhood within 5km is 22



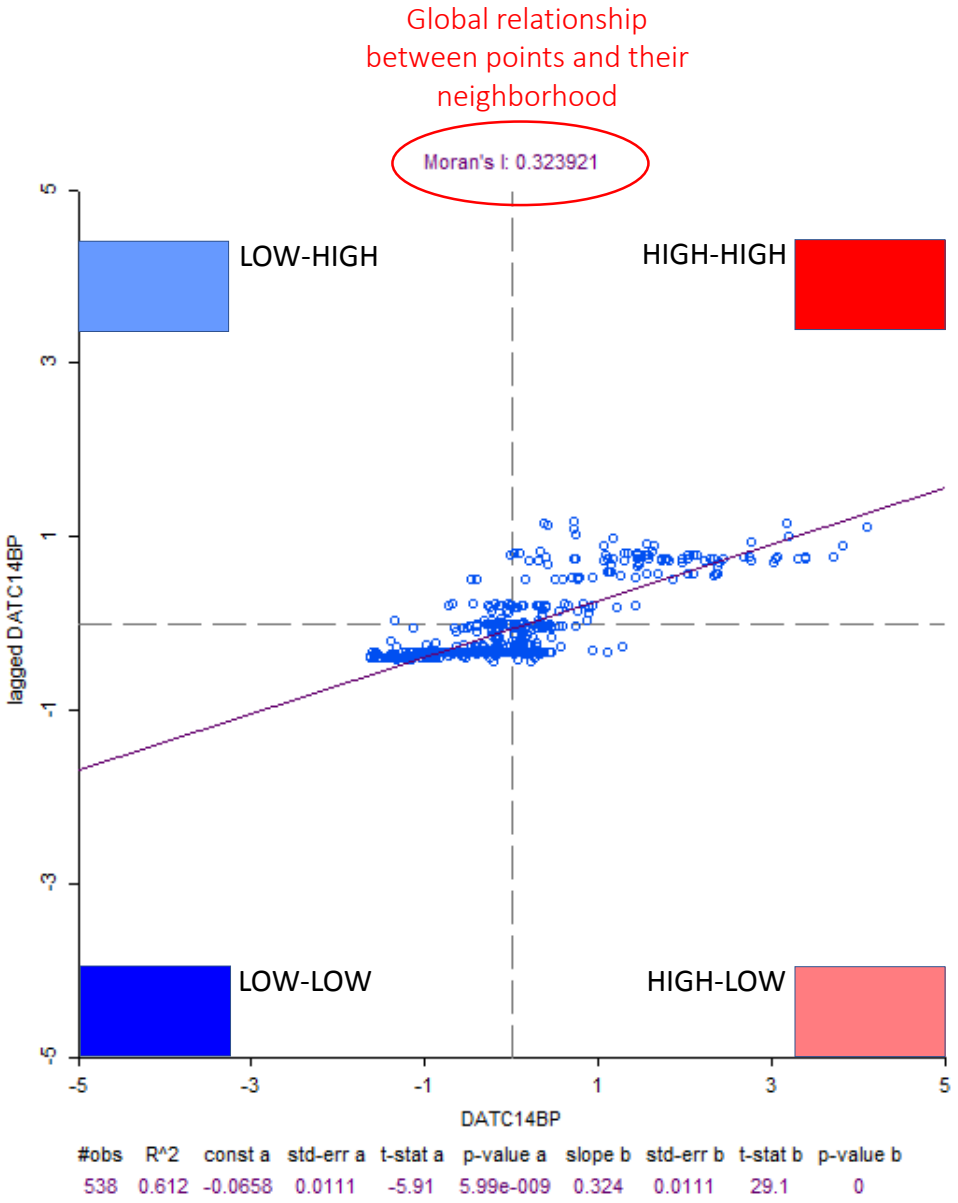
Point	Observed value	Weighted value
1	20	Mean of = 22

We obtain two distributions of observed versus weighted values

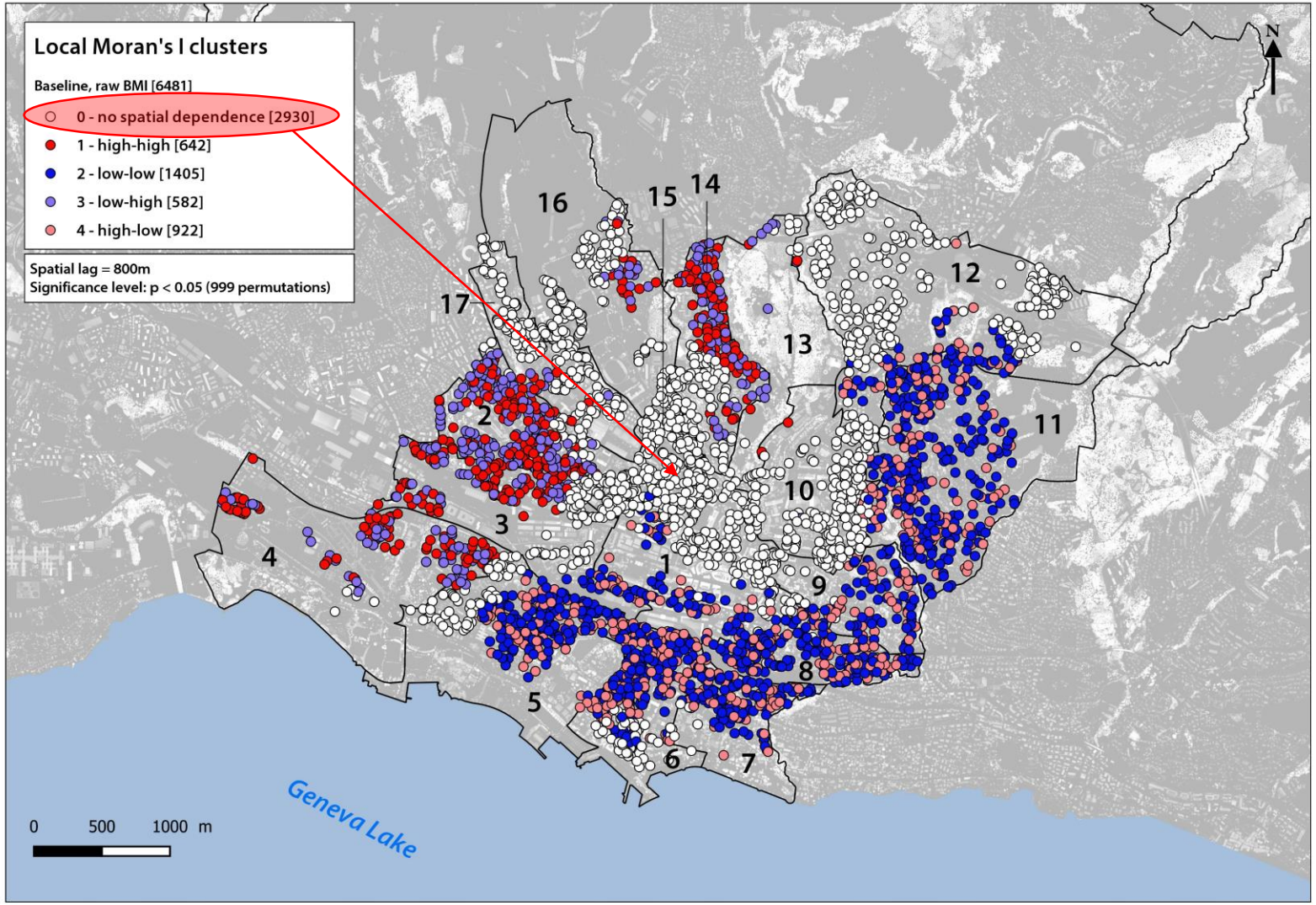
Linear regression
Dependent variable = weighted
Independent variable = Observed

Types of clusters

- After standardization, we obtain a Moran's scattergram
- The distribution of points among the quadrats of the scattergram defines 4 classes
- They correspond to the types of relationship between observed values and weighted values at all locations
- E.g. High-high (red) = high observed values and high weighted values
- The slope defines the level of global spatial autocorrelation (0.32)



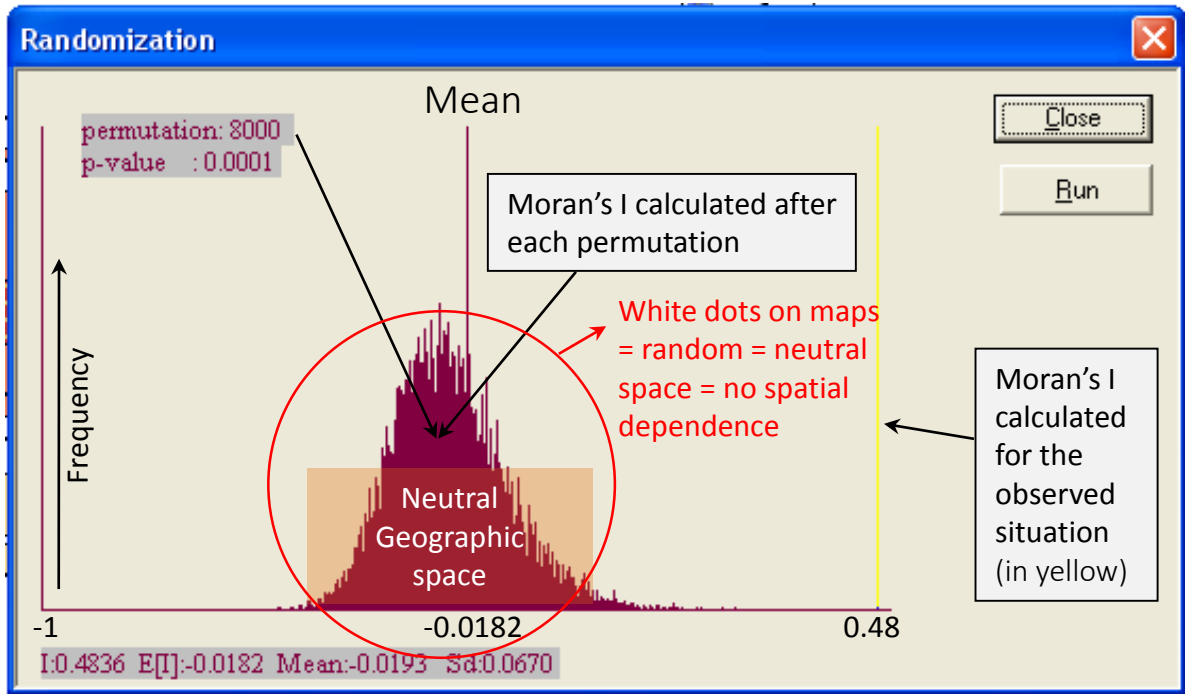
The fifth class



How to identify them?

Significance

- We need to check if the Moran's I obtained is statistically significant
- Does the spatial structure observed and quantified by the Moran's I (here 0.48) persist when BMI values are randomly distributed among all locations (Monte-Carlo method for permutations) ?
- Moran's I is calculated again after each run of random permutations (here 8'000)
- The Moran's I for each run feeds the histogram
- A pseudo p-value is calculated on the basis of the number of random configurations that produce a Moran's I higher or equal to the observed one (yellow bar)

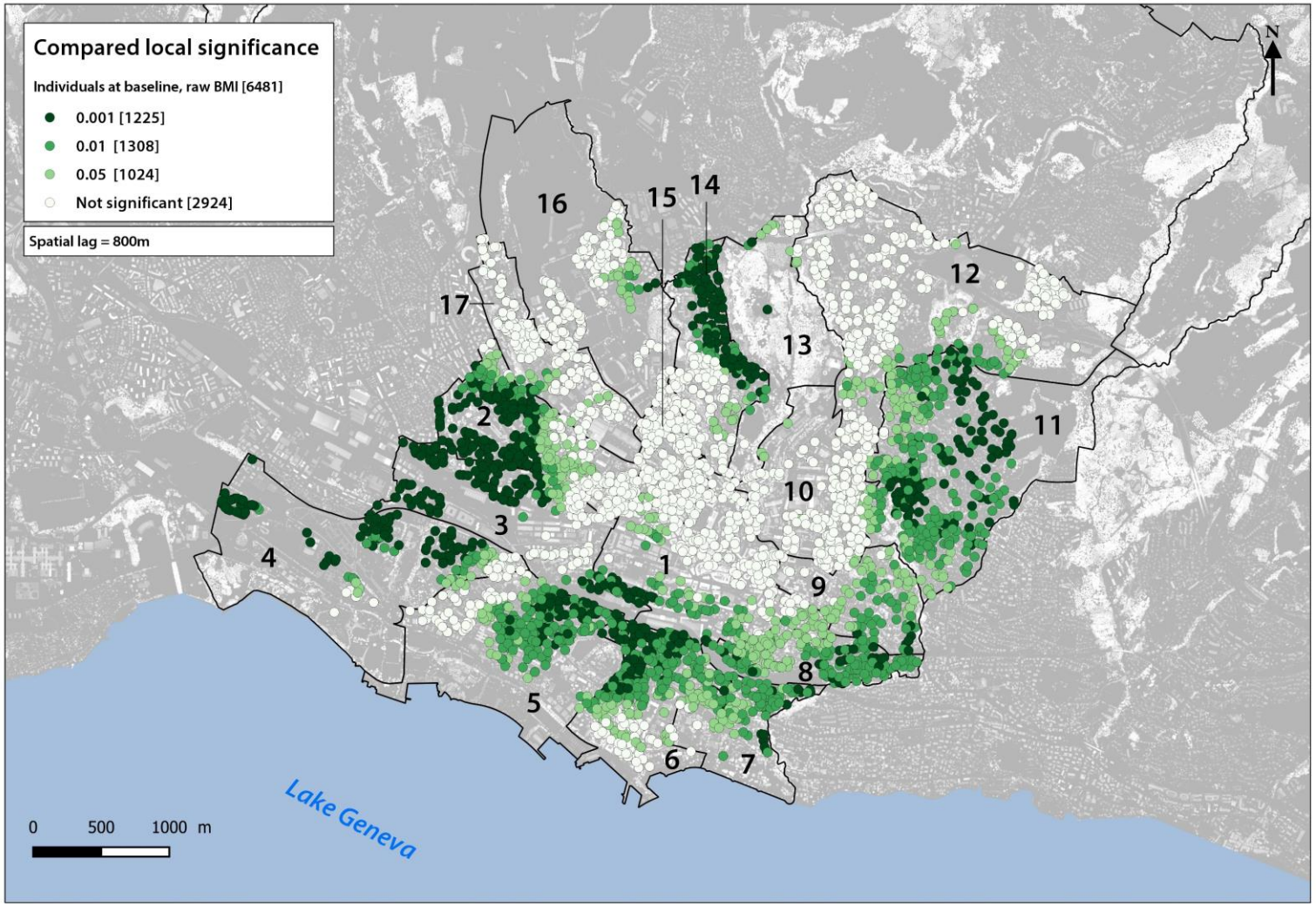


$$\text{p-value} = \frac{Nb\ I_{al} \geq I_{obs} + 1}{Nb\ permutations + 1}$$

$$or \quad \frac{Nb\ I_{al} \leq I_{obs} + 1}{Nb\ permutations + 1}$$

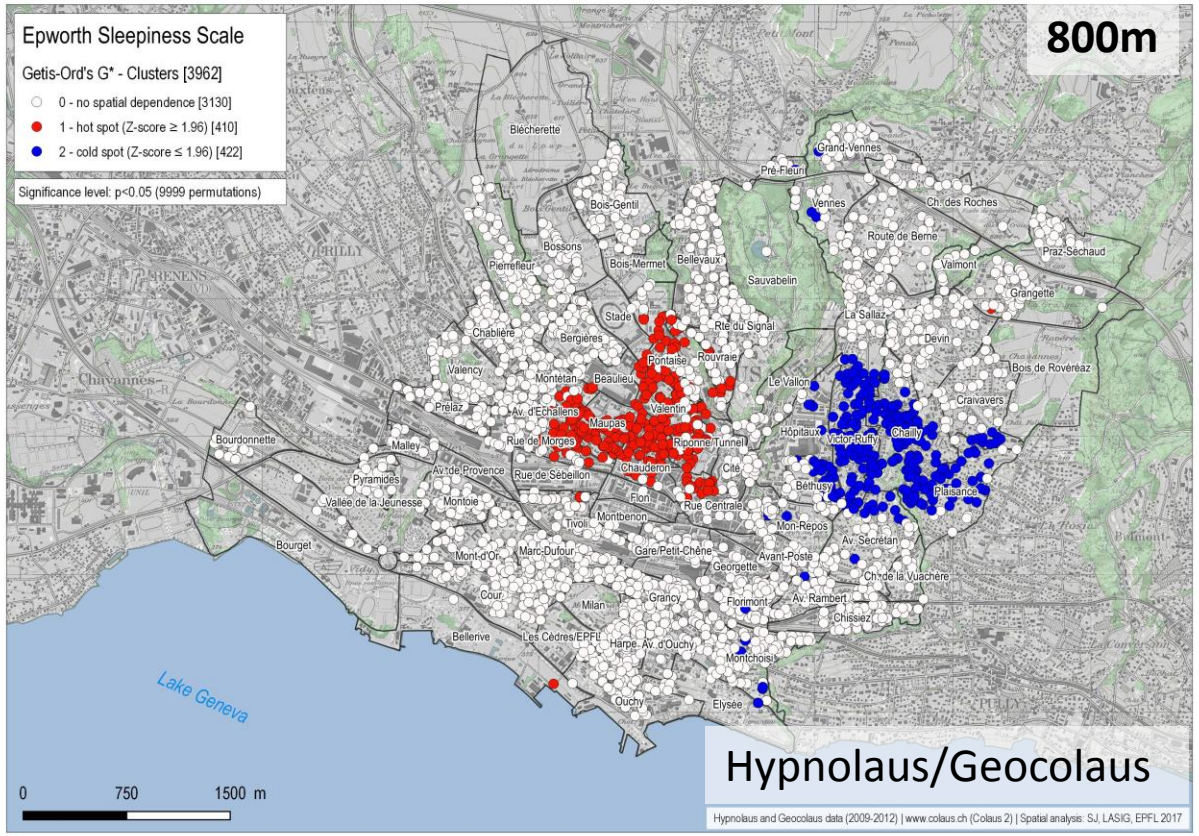
Mapping significance

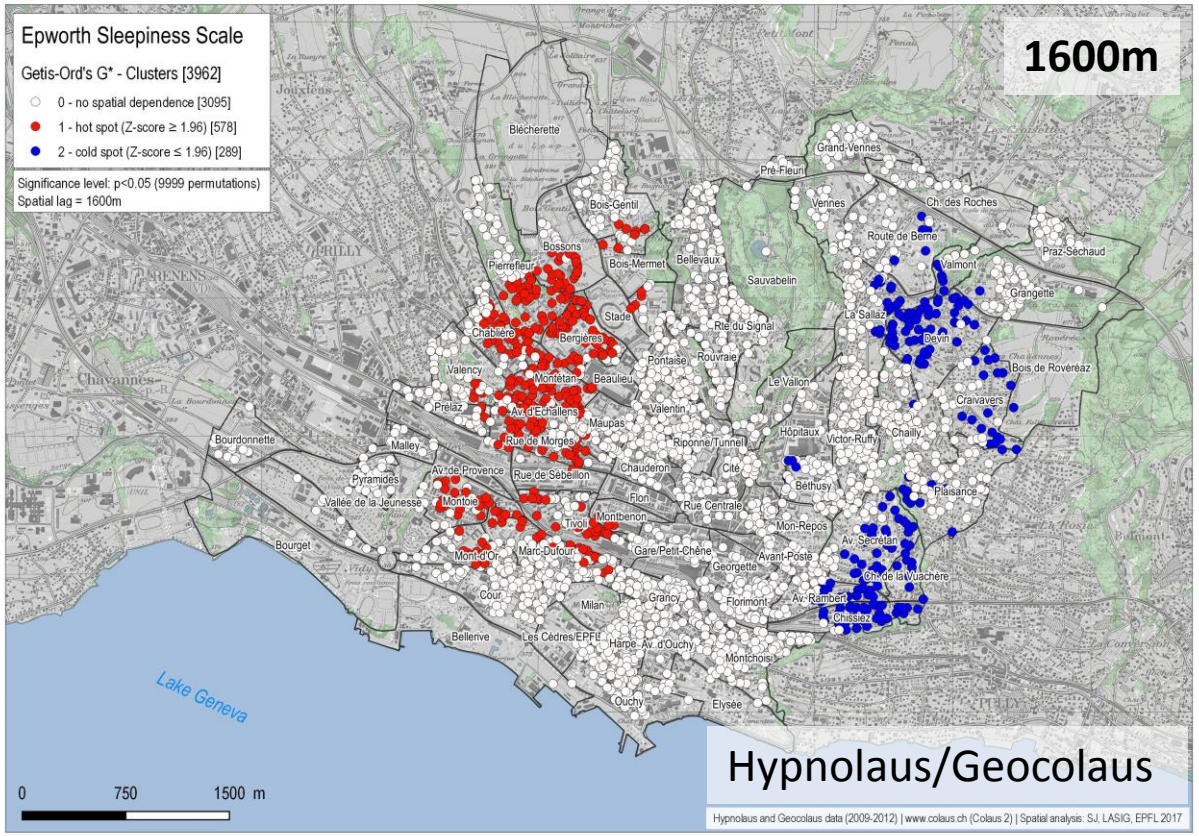
- Local Indicators of Spatial Association (LISA) – a local version of Moran's I
- The pseudo p-value obtained can be mapped to show the level of significance of the local spatial autocorrelation
- This allows to introduce subtleties in the interpretation of the clusters obtained



Scale issue: the size of spatial lags defining neighborhood

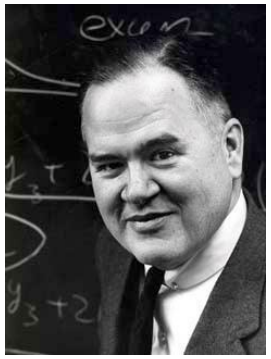
- Important: spatial statistics like Moran's I or Getis-Ord Gi constitute **exploratory** approaches
- Necessary to test several spatial lags to possibly identify different explanatory factors

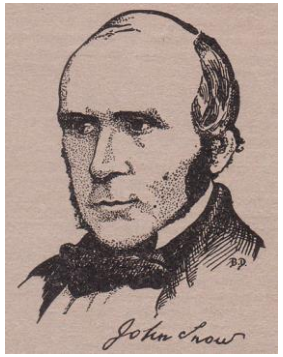




Conclusion

- The measure of spatial dependence is key to detect and visualize spatial patterns in health data
- Spatial statistics can reveal signals that remain hidden using thematic mapping
- On the basis of the clusters highlighted by these exploratory methods, it is then possible to formulate hypotheses about potential environmental or socio-economic causes and to test them with the help of confirmatory statistics
- «Ideas come from previous explorations» John Tukey, The American Statistician, 1980, in «We Need Both Exploratory and Confirmatory»





- This reasoning was applied by John Snow to detect deaths "hot spots" in London, which then allowed him to hypothesize that a particular water pump was infected, and finally to take public health steps to check the cholera epidemic

Thank you for your attention!

Acknowledgments

Duruz S, Mansolino C, Bacchilega B, Salmi A

Marques-Vidal P, Gaspoz JM, Theler JM, Chételat J, Chenal J, Kutalik Z, Cornuz J, Bochud M, Stringhini S, Heinzer R, Haba Rubio J, Waeber G, Vollenweider P

http://www.giraph.org

stephane.joost@giraph.org

idris.guessous@giraph.org