Modeling effects of the social and physical environment on health

a spatial perspective

Basile CHAIX
Inserm U707

RECORD
A study of geographic disparities in health

Instituts thématiques Inserm
Institut national
de la santé et de la recherche médicale

UPMC PARIS UNIVERSITAS

prévention santé IPC

SPATIAL PERSPECTIVE: RATIONALE

Contextual analysis: associations between contextual exposures and individual outcomes adjusted for individual confounders, often from multilevel models

Early contextual studies

Geographic distribution described in terms of within-neighborhood correlation (multilevel models)

Explanatory contextual variables measured within administrative neighborhoods

Overall, territory fragmented into disconnected administrative areas

Spatial perspective

- Spatially structured or unstructu-red variability?
- Spatial range of correlation?

- Towards personal exposure areas?
- Optimal spatial scale of measurement?

Introduce spatial continuity in the measurement of exposures and modeling of their effects

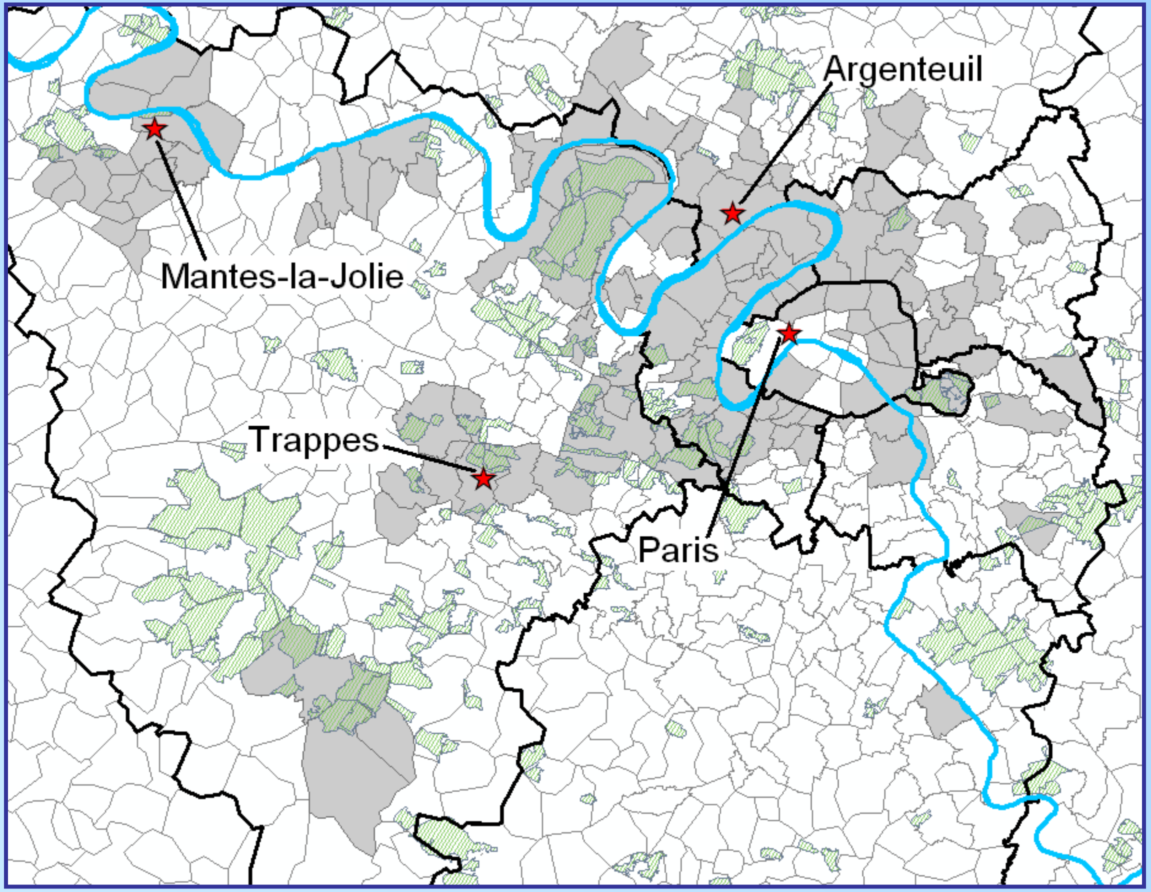
RECORD : STUDY TERRITORY

RECORD
A study of geographic disparities in health

- Participants recruited during general health checkups in 2007-2008
- 4 recruitment sites
- **7292** 30–79 year old participants
- 111 municipalities + 10 Paris subareas

= **1915** neighborhoods

- Data:
  - 2-h long checkup
  - Questionnaires
  - Geocoding & contextual data

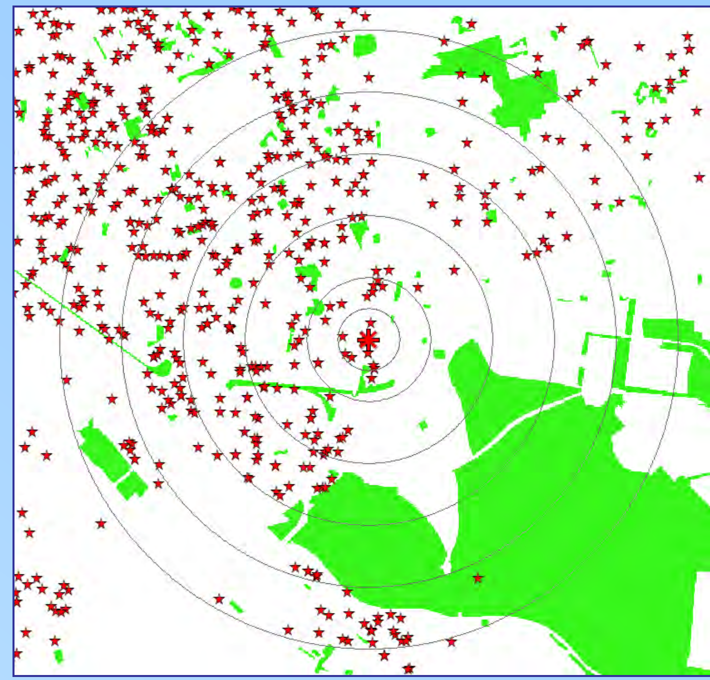


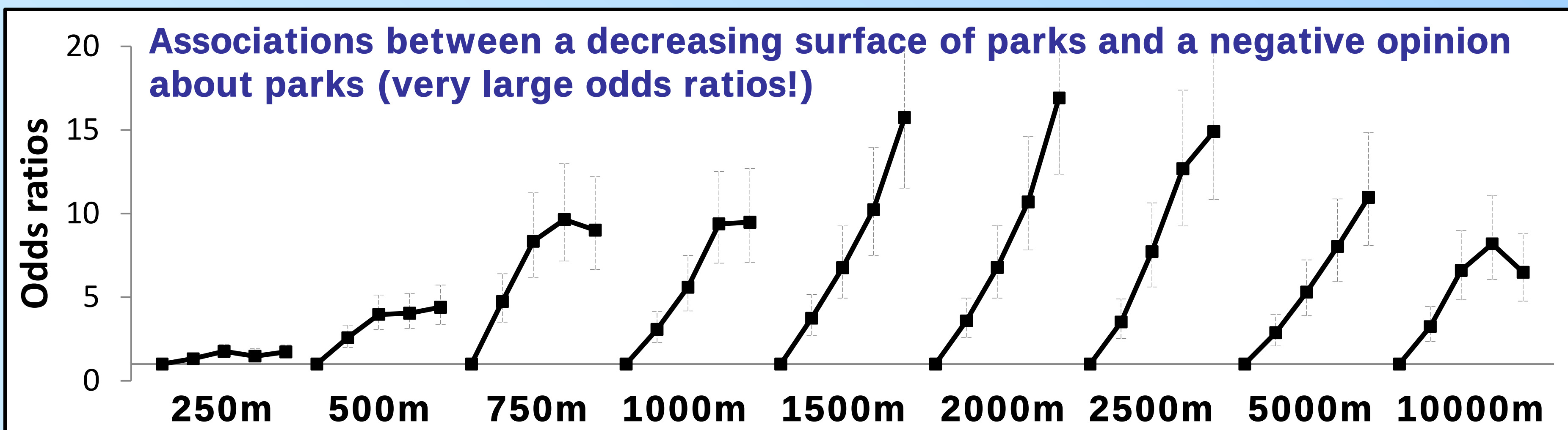
EGO-CENTERED NEIGHBORHOODS

Personal exposure areas assessed as ego-centered neighborhoods

Objective measurement: Surface of parks and green spaces within radiuses of 100-10000 m

Perception of neighborhood problems: “lack of green spaces nearby?”



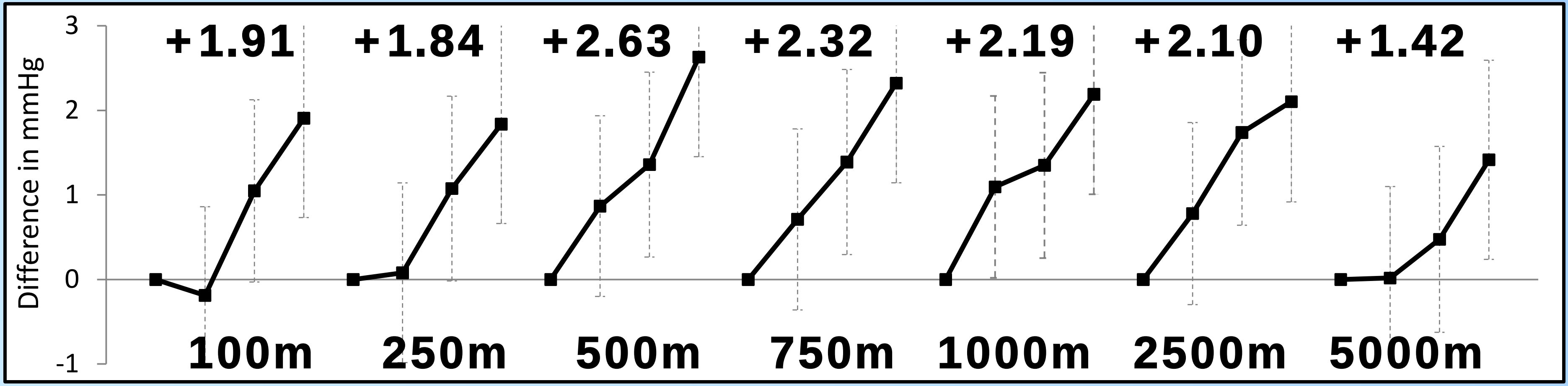


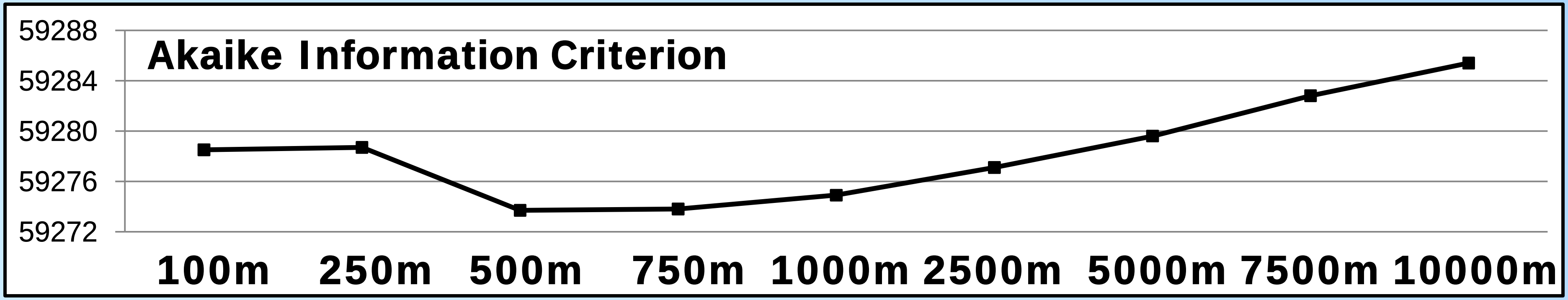
Chaix, Merlo, Evans, Leal, Havard. Soc Sci Med 2009;69:1306-1310.

NEIGHBORHOOD EDUCATION AND SBP

Association between neighborhood education in quartiles (circular areas of various radiuses) and systolic blood pressure

(model adjusted for age, gender, study center, antihypertensive med., education, unemployment, dwelling ownership, country of birth)

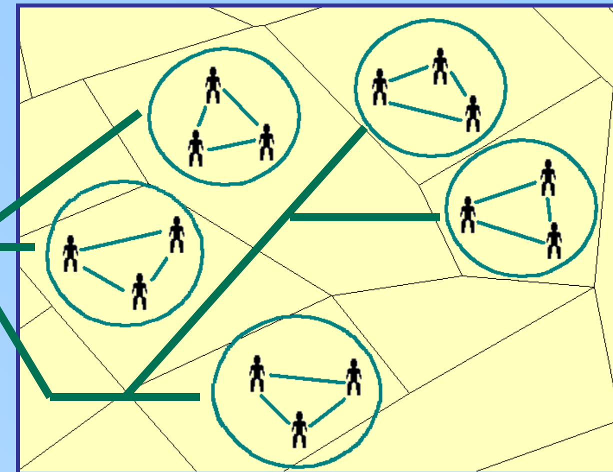




SPATIAL RANDOM EFFECT MODELS FOR INDIV. DATA

Multilevel model

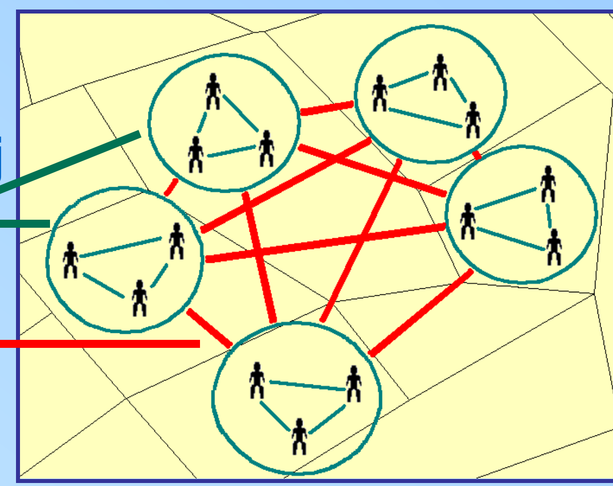
$$y_{ij} = \alpha + \beta X_{ij} + \gamma Z_j + \boxed{u_j} + e_{ij}$$



→Does not account for correlation between adjacent/nearby neighborhoods

Spatial random effect model

$$y_{ij} = \alpha + \beta X_{ij} + \gamma Z_j + \boxed{s_j} + \boxed{u_j} + e_{ij}$$



(often intractable at the individual level)

Motivations

- Improved control of residual autocorrelation
- Information on the spatial distribution of health phenomena
- Information supporting the interpretation of fixed effects

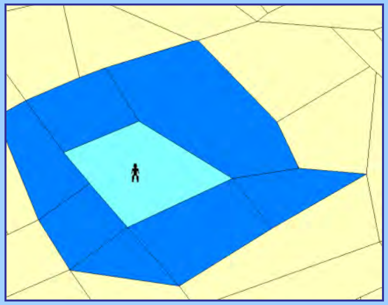
SPATIAL MODELING OF SBP

$SBP_{ij} = \alpha + \beta X_{ij} + u_j + s_j + e_{ij}$

-Unstructured effect : $u_j \sim N(0, \sigma_u)$

-Structured CAR effect: $s_j \sim N(\Sigma s_j / n_j, \sigma_s / n_j)$

Spatial contiguity



Explaining structured/unstructured variations in blood pressure

	Unstructured variance	Structured variance	% of structured variance	Intra-neighborhood correlation
Model with age & gender	4.4 (0.4, 9.6)	5.5 (3.3, 8.2)	56% (32, 94)	3.8% (2.0, 5.8)
+ individual SES variables	3.6 (0.1, 8.7)	2.8 (1.1, 4.8)	44% (17, 96)	2.5% (0.9, 4.5)
+ neighborhood SES	3.5 (0.2, 8.4)	1.9 (0.9, 3.4)	36% (14, 89)	2.1% (0.7, 4.1)
+ risk factors	1.4 (0.1, 5.4)	1.6 (0.7, 2.9)	54% (18, 95)	1.4% (0.5, 3.2)

Warning #1: We use empirical marginal variances

Warning #2: Separability of the structured/unstructured effects??

DISTANCE-BASED SPATIAL STRUCTURE

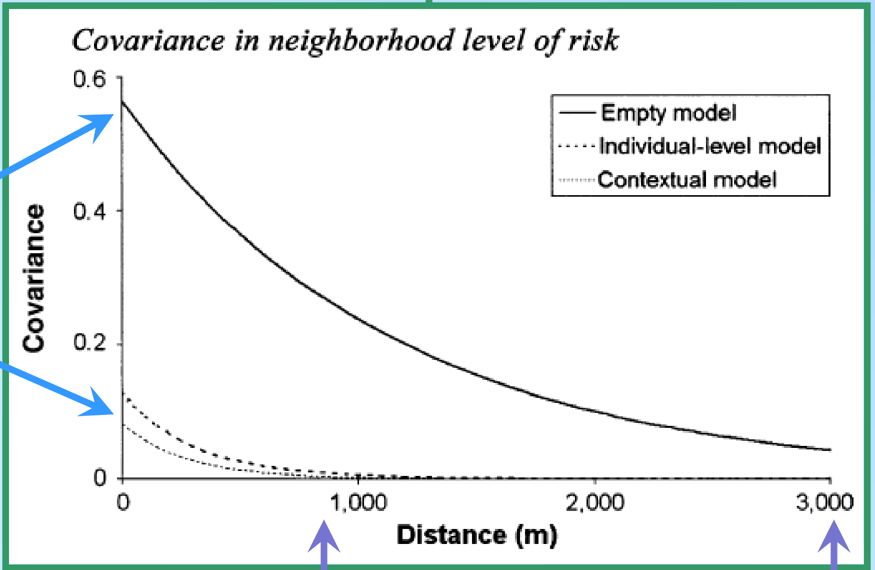
Mental disorders related to psycho-active substance, Malmö, 2001

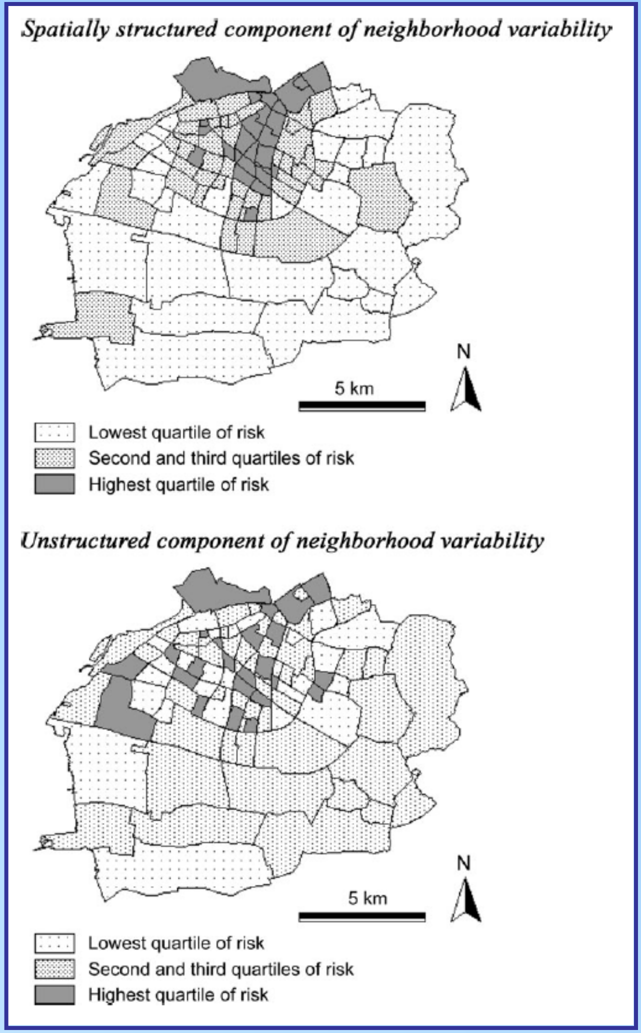
Hierarchical geostatistical model

$$\text{logit}(p_j) = \beta_0 + u_j + s_j$$

$$\text{Var}(u_j) = \sigma_u^2$$

$$\text{Var}(s_j) = \sigma_s^2 \quad ; \quad \text{Corr}(s_k, s_l) = \exp(-\phi\, d_{kl})$$



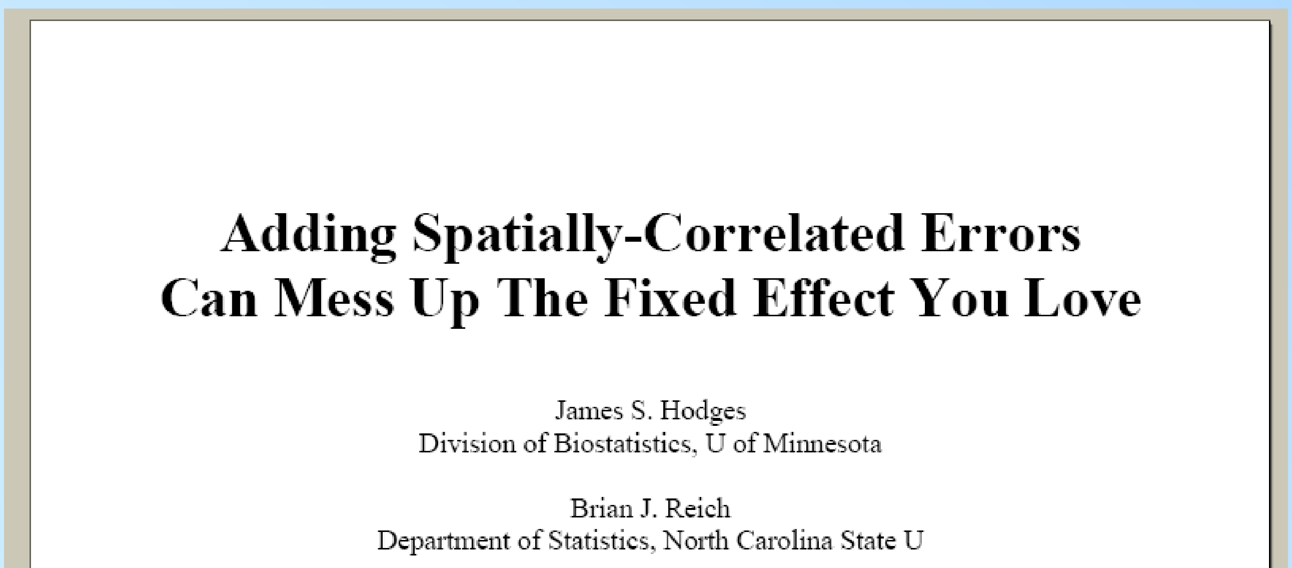


Chaix, Merlo, Subramanian, Lynch, Chauvin. Am J Epidemiol, 2005;162:171-182

SPATIAL RANDOM EFFET AS A SOURCE OF BIAS?

Collinearity between the fixed effects and the spatial random effect may cause a significant bias when:

- there are strong geographic variations in the outcome
- the fixed effect variables are themselves spatially autocorrelated
- the fixed effect variable and the spatial random effect capture variations on a comparable spatial scale


Adding Spatially-Correlated Errors Can Mess Up The Fixed Effect You Love

James S. Hodges
Division of Biostatistics, U of Minnesota

Brian J. Reich
Department of Statistics, North Carolina State U

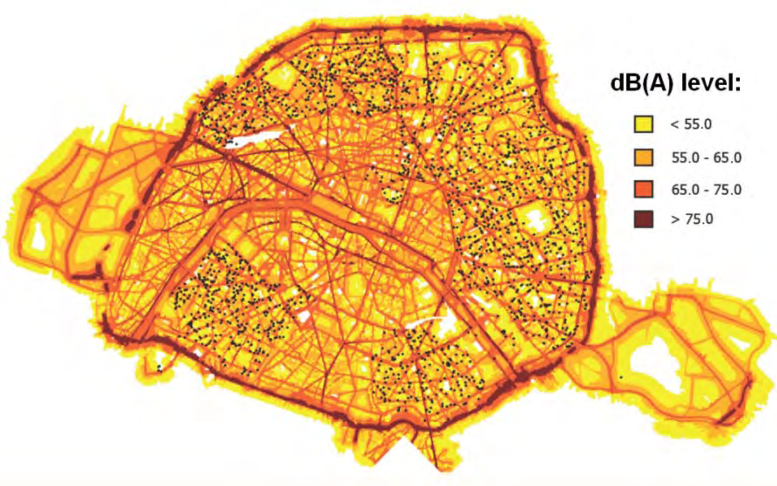
Reich BJ, Hodges JS, Zadnik V.

Effects of residual smoothing on the posterior of the fixed effects in disease-mapping models.

Biometrics 2006;62:1197-206.

Associations with outdoor noise exposure in dB(A) in Paris, France

	Multilevel model	CAR model
Proportion of high educated residents (vs. low)		
Mid-low	**0.51** (–0.08, 1.12)	**–0.05** (–0.52, 0.42)
Mid-high	**1.46** (0.78, 2.15)	**0.31** (–0.25, 0.89)
High	**2.02** (1.78, 2.86)	**0.55** (–0.17, 1.27)
Unemployment rate (vs. low)		
Mid-low	**0.56** (–0.10, 1.22)	**0.10** (–0.37, 0.58)
Mid-high	**1.28** (0.49, 2.07)	**0.50** (–0.10, 1.11)
High	**1.62** (0.67, 2.56)	**0.84** (0.11, 1.56)
Mean dwelling value (vs. low)		
Mid-low	**0.23** (–0.39, 0.85)	**0.14** (–0.30, 0.57)
Mid-high	**1.10** (0.41, 1.78)	**0.53** (0.02, 1.03)
High	**1.49** (0.75, 2.24)	**0.23** (–0.36, 0.81)
Proportion of foreign-born residents(vs. low)		
Mid-low	**0.55** (–0.06, 1.16)	**0.09** (–0.47, 0.66)
Mid-high	**1.49** (0.84, 2.15)	**0.44** (–0.18, 1.06)
High	**1.53** (0.79, 2.27)	**0.14** (–0.59, 0.88)



Sabrina HAVARD, Basile CHAIX, unpublished work in progress...

Neighborhood factors associated with participation in the RECORD Study

	Multilevel model	CAR normal model
Distance to the center (vs. high)		
Mid-high	1.19 (1.09, 1.30)	1.14 (0.97, 1.33)
Mid-low	1.45 (1.32, 1.58)	1.20 (0.98, 1.44)
Low	1.75 (1.60, 1.91)	1.31 (1.04, 1.61)
Proportion of the area covered by buildings (vs. high)		
Mid-high	1.13 (1.03, 1.23)	1.07 (0.98, 1.17)
Mid-low	1.26 (1.14, 1.39)	1.07 (0.97, 1.20)
Low	1.37 (1.23, 1.51)	1.08 (0.96, 1.21)
Mean building height (vs. high)		
Mid-high	1.11 (1.03, 1.21)	1.03 (0.96, 1.12)
Mid-low	1.27 (1.16, 1.39)	1.11 (1.00, 1.22)
Low	1.27 (1.15, 1.40)	1.11 (0.99, 1.24)
Mean dwelling value (vs. low)		
Mid-low	1.10 (1.00, 1.21)	1.02 (0.93, 1.13)
Mid-high	1.11 (1.00, 1.24)	1.02 (0.91, 1.14)
High	1.23 (1.09, 1.39)	1.10 (0.97, 1.25)
Median income (vs. low)		
Mid-low	1.20 (1.09, 1.32)	1.14 (1.03, 1.26)
Mid-high	1.29 (1.14, 1.45)	1.18 (1.05, 1.34)
High	1.39 (1.20, 1.60)	1.29 (1.11, 1.50)
Proportion of the active population looking for work (vs. low)		
Mid-low	1.01 (0.93, 1.10)	1.11 (1.02, 1.21)
Mid-high	1.18 (1.06, 1.31)	1.43 (1.28, 1.60)
High	1.31 (1.15, 1.47)	1.68 (1.47, 1.93)

Correlation between the variable and the spatial random effect

R = 0.35

R = 0.30

R = 0.17

R = 0.03

R = 0.05

R = –0.21

BRIEF CONCLUSION

Spatial modeling in social epidemiology

- aim: taking space into account as a continuum
- relevant to both the measurement of exposures and the modeling of their effects

Strengths of the approach

- provides information on the spatial scale of:
  - health variations: relevant to public health
  - contextual effects: etiological relevance

Drawbacks of the approach

- time-consuming to implement
- biases related to spatial random effects