# Spatial Methods for Health and Population Research

## GEOG 8104: Special Topics in Quantitative Geography

**Mondays 2:15 – 4:50**
**Derby Hall Room 1116**
**Prof. Elisabeth D. Root**

### General Introduction

This course is a graduate-level seminar which surveys current topics in health and population research with a focus on theoretical/conceptual motivations and technical approaches for analyzing spatially explicit data for social science research. Throughout the seminar we will review the broad field of spatial data analysis and the range of issues that arise when analyzing georeferenced data. We will read and discuss selections from the current literature and critique how spatial data and statistical methods are being integrated into population and health research. We will also gain hands-on experience with a variety of spatial statistical methods such as spatial regression, geographically weighted regression, and cluster analysis.

### Course Objectives

By the end of this course, students will be able to:

1. Critique spatial analytic methods in selected papers,
2. Present arguments about appropriate and inappropriate spatial research methods for a given research problem,
3. Recognize the complexities inherent in using spatial data and choose appropriate methods for data analysis, and
4. Implement spatial statistical methods using simple population or health datasets.

I have not designed this as a GIS course but throughout the semester you will have plenty of opportunity to learn *ArcGIS* and other software, namely *SatScan* (cluster analysis) and *R* (spdep) for spatial analysis. *R* is an extremely flexible program that facilitates exploratory spatial data analysis, spatial regression modeling, geographically weighted regression, cluster methods and much much more. All of these software packages are free, so we don't need to do any specific work in the lab. You may use any other software you wish (such as Stata or SAS).

### Prerequisites

Some prior experience with GIS software is useful but not a prerequisite. What is a prerequisite for this course is a solid understanding of multivariate statistics and some experience with advanced regression analysis. It is also important to have some experience with the mathematical notation typically used for statistical learning.

### What I Think About Teaching & Research

My philosophy of teaching is that students need to learn to learn. This may sound strange, after all haven't we been in school a really long time? Don't we already know how to learn? The answer to this question is yes, we do know how to learn in some ways but college should require greater depth and breadth of thinking, especially as a graduate student. What I found

out during my PhD is that I needed to teach myself a great deal because there were often no classes on the methods or subjects in which I became interested. So, I think it's very important for graduate students to learn how to find the information they need through research and how to apply it to real world problems.

The other thing I found out after working in both the private sector and at a University is that very little research is done by one person. Most of the best science is conducted by teams of researchers who brainstorm and discuss and learn together. In fact, many of the big federal funding agencies (NIH, NSF, EPA) are increasingly looking to fund multidisciplinary research collaborations. Learning in this class will therefore focus on research, reading, listening, communication and collaboration with other students.

As graduate students, many of you plan to continue on to academic positions. In these positions, you will be expected to teach! So, I have also added a component to this class where you will have the opportunity to teach your peers.

**Course Format**

Our class meetings will be organized around three rotating activities: lecture (by me), student paper presentations and discussion (by you) and lab or hands-on exercises (in groups). Typically, graduate seminars don't have much professor-led lecture. I have experimented in the past and have found that for this class, where a significant amount of technical material is being presented, some lecture is appropriate and useful.

**Assessment**

You will be assessed on both your individual work in the class and on group work. The importance of each component will be as follows:

| | |
|---|---|
| Group lab write ups | 60% |
| Individual paper presentation | 15% |
| Discussion lead | 15% |
| Participation and engagement | 10% |

<u>Group lab write-ups</u>

Students are required to complete 4 lab assignments during the semester. I will not be handing you a "script" for how to do a statistical analysis during the labs. I will provide you with a dataset and some programming advice and ask you to figure out how to use the software packages we learn to run models and answer a set of broad questions. Lab write-ups are expected to look very similar to a journal article's Results and Discussion section. I will pass out a grading rubric before the first lab so you understand how labs will be graded. I expect that you will work in groups of 2 to get the lab done, which means that you should turn in one lab report per group.

In general, lab write-up should include:

1) A clear statement of your research question (just state it, no need for a paragraph of prose),
2) A methods section to explain what data you used and what statistical tests you performed and why,

3) A results section including the relevant tables and graphs that best summarize your output and your interpretation of these (MORE IS NOT BETTER - choose wisely), and
4) A discussion of your results (really no more than a few paragraphs). Keep in mind that the discussion section interprets your results, it doesn't just restate them.

This whole write-up should be no longer than 3 pages (single spaced, not including tables/graphs); you will be graded down for excessively wordy write-ups and for including unnecessary tables/graphs. Your annotated R code should be included as an Appendix. You will turn in these lab reports via the Dropbox on Carmen. Labs are due the Monday following the lab exercise.

Individual presentation

Students will individually present a paper of their choosing once during the semester. You will sign up for a presentation day, and then conduct a literature search for a peer-reviewed journal article that reflects the topics discussed the prior week. You will submit this paper to me for approval at least 6 days prior to your in-class presentation. Once you have received approval for your selected paper, you will prepare a 10-15 minute presentation, with PowerPoint slides, that "teaches" the article to the other students in the class. This presentation must consist of:

1) A presentation of the research question
2) An explicit explanation of the methods used by the author, presented in a way that you are "teaching" the methodology to the students in the class
3) A short presentation and discussion of the results
4) Some discussion questions for the class which:
    a. explore the assumptions in the author's argument(s), research question(s), and methodology,
    b. evaluate the validity of the methodology used to answer the research question(s) and discuss if there were other "better" ways to conduct the research, and
    c. examine the implications of the argument for other substantive topics.

Discussion Lead

One week out of the semester, you will be asked to prepare for and lead discussion of the assigned readings. I will help, but you must come prepared to help lead discussion and goad your peers into helping you discuss the assigned papers.

Participation and engagement

Students are expected to attend all classes. Attendance, as well as my assessment of how engaged you are in the class, are 10% of your grade.

| Week | Class Dates | Lecture Topic/Readings | Readings |
|---|---|---|---|
| 1 | 8/28 | Introduction, Spatial Thinking, and Spatial Data (Analysis)<br><br>Theoretical Arguments for Spatial Modeling | • Goodchild, MF, L Anselin, RP Appelbaum and B.H. Harthorn. 2000. Toward Spatially Integrated Social Science. *Int Regional Science Review* 23(2): 139-159.<br>• Voss, P. 2007. Demography as a spatial social science. *Population Research and Policy Review* 26: 457-476.<br>• Loftin, C. & S.K. Ward. 1983. A Spatial Autocorrelation Model of the Effects of Population Density on Fertility. *American Sociological Review*, 48(1):121-128.<br>• Galle, O.R., W.R. Gove, & J.M. McPherson. 1972. Population Density and Pathology: What Are the Relations for Man? *Science,* 176:23-30.<br>• Rushton, G. 2003. Public Health, GIS and Spatial Analytic Tools. *Annual Reviews of Public Health* 24: 43-56.<br>• Worrall, L. and Bond, D., 1997. Geographical information systems, spatial analysis and public policy: the British experience. *International Statistical Review* 65(3): 365-379.<br>• Hillier, A. 2007. Why social work needs mapping. *Journal of Social Work Education* 43(2): 205-222.<br>• Downey, L. 2006. Using geographic information systems to reconceptualize spatial relationships and ecological context. *AJS; American journal of sociology*, *112*(2): 567. |
| 2 | 9/4 | NO CLASS – LABOR DAY | |
| 3 | 9/11 | ESDA & Spatial Autocorrelation | TEXTBOOK: Ward, M.D. and K.S. Gleditsch (2008). *Spatial Regression Models. Chapter 1*<br>• Anselin, L., 1989. What is Special About Spatial Data? *Alternative Perspectives on Spatial Data Analysis* (89-4).<br>• Tobler, W.R. 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography* 46(June):234-240.<br>• Getis, A. 2008. A History of the Concept of Spatial Autocorrelation: A Geographer's Perspective. *Geographical Analysis* 40:297-309.<br>• Messner, S.F., L. Anselin, R.D. Baller, D.F. Hawkins, G. Deane, and S.E. Tolnay. 1999. The Spatial Patterning of County Homicide Rates: An Application of Exploratory Spatial Data Analysis. *Journal of Quantitative Criminology* 15(4):423-450.<br>• Tolnay, S.E., Deane, G. & E.M. Beck. 1996. "Vicarious Violence: Spatial Effects on Southern Lynchings, 1890-1919." *American Journal of Sociology* 102(3):788-815.<br>• Poulsen, M, R Johnston, J Forrest. 2010. The intensity of ethnic residential clustering: exploring scale effects using local indicators of spatial association. *Environment and Planning A* 42:874-894. |
| 4 | 9/18 | ESDA & Spatial Autocorrelation | Student Papers |

| Week | Class Dates | Lecture Topic/Readings | Readings |
|---|---|---|---|
| 5 | 9/25 | Lab: ESDA & Spatial Autocorrelation | None<br>Lab Write-up Due 10/6 by 5pm |
| 6 | 10/2 | Spatial Regression Models | TEXTBOOK: Ward, M.D. & K.S. Gleditsch. 2008. *Spatial Regression Models.* Chapters 2, 3 & 4. Waller, L.A. and Gotway, C.A., 2004. *Applied Spatial Statistics for Public Health Data.* Chapter 9, pp. 325-370.<br>• Voss, P.R., D.D. Long, R.B. Hammer, & S. Friedman. 2006. County Child Poverty Rates in the U.S.: A Spatial Regression Approach. *Population Research and Policy Review* 25: 369-391.<br>• Emch, M., M. Ali, J.K. Park, M. Yunus, D.A. Sack & J.D. Clemens. 2006. Relationship between neighborhood-level killed oral cholera vaccine coverage and protective efficacy: evidence for herd immunity. *International Journal of Epidemiology* 35: 1044-50.<br>• Sparks, P.J., & C.S. Sparks. 2010. An Application of Spatially Autoregressive Models to the Study of US County Mortality Rates. *Population, Space and Place* 16:465-481.<br>• Burnett, J.W. & D.J. Lacombe. 2012. Accounting for Spatial Autocorrelation in the 2004 Presidential Popular Vote: A Reassessment of the Evidence. *The Review of Regional Studies* 42:75-89. |
| 7 | 10/10 | Spatial Regression Models | Student Papers |
| 8 | 10/16 | (More) Spatial Regression Models | TEXTBOOK: Waller, L.A. and Gotway, C.A., 2004. *Applied Spatial Statistics for Public Health Data.* Chapter 9, pp. 370-409.<br>• Wall, M.M. 2004. "A close look at the spatial structure implied by the CAR and SAR models." *Journal of Statistical Planning and Inference* 121(2): 311-324.<br>• Trogdon, J.G. & T. Ahn. 2015. "Geospatial Patterns in Human Papillomavirus Vaccination Uptake: Evidence from Uninsured and Publicly Insured Children in North Carolina." *Cancer Epidemiol Biomarkers Prev*; 24(3); 595–602.<br>• Webster, T., Vieira, V., Weinberg, J. and Aschengrau, A., 2006. "Method for mapping population-based case-control studies: an application using generalized additive models." *International Journal of Health Geographics* 5(1): 1.<br>• Vieira, Verónica, et al., 2005. "Spatial analysis of lung, colorectal, and breast cancer on Cape Cod: an application of generalized additive models to case-control data." *Environmental Health* 4(1): 1. |
| 9 | 10/23 | Lab: Spatial Regression | None<br>Lab Write-up Due 11/3 by 5pm |

| Week | Class Dates | Lecture Topic/Readings | Readings |
|---|---|---|---|
| 10 | 10/30 | Geographically Weighted Regression | TEXTBOOK: Fotheringham, S.A., C. Brunsdon & M. Charlton. 2002. *Geographically Weighted Regression*. Chapters 1 & 2.<br>• Wheeler, D. & M. Tiefelsdorf. 2005. Multicollinearity and Correlation among Local Regression Coefficients in Geographically Weighted Regression. *Journal of Geographical Systems* 7:161-187.<br>• O'Loughlin, J. & F. Witmer. 2011. The Localized Geographies of Violence in the North Caucasus of Russia, 1999-2007. *Annals of the Association of American Geographers* 101(1): 178-201.<br>• Grillet, M-E., R. Barrera, J-E Martinez, J. Berti & M-J Fortin. 2010. Disentangling the Effects of Local and Global Spatial Variation on a Mosquito-Borne Infection in a Neotropical Heterogeneous Environment. *American Journal of Tropical Medicine and Hygiene* 82(2): 194–201.<br>• Fotheringham, A.S., Charlton, M.E. and Brunsdon, C. 2001. Spatial variations in school performance: a local analysis using geographically weighted regression. *Geographical and Environmental Modelling* 5(1): 43-66. |
| 11 | 11/6 | Geographically Weighted Regression | Student Papers |
| 12 | 11/13 | Lab: Geographically Weighted Regression | None<br>Lab Write-up Due 11/27 |
| 13 | 11/20 | Spatial Cluster Methods | TEXTBOOK: Haining, R. 2003. *Spatial Data Analysis: Theory and Practice.* Chapter 7<br>• Kulldorff M. 1997. A spatial scan statistic. *Communications in Statistics: Theory and Methods* 26: 1481-1496<br>• Kulldorff, et al. 1998. Evaluating Cluster Alarms: A Space-Time Scan Statistic and Brain Cancer in Los Alamos, New Mexico. *Am J Public Health* 88: 1377-1380.<br>• Wheeler, D.C. 2007. A comparison of spatial clustering and cluster detection techniques for childhood leukemia incidence in Ohio, 1996–2003. *Int J of Health Geographics* 6(1): 1.<br>• Sheehan, T.J., DeChello, L.M., Kulldorff, M., Gregorio, D.I., Gershman, S. and Mroszczyk, M., 2004. The geographic distribution of breast cancer incidence in Massachusetts 1988 to 1997, adjusted for covariates. *Int J of Health Geographics*, *3*(1), p.1.<br>• Stevenson, J.R., Emrich, C.T., Mitchell, J.T. and Cutter, S.L. 2010. Using building permits to monitor disaster recovery: A spatio-temporal case study of coastal Mississippi following Hurricane Katrina. *Cartography and Geographic Information Science* 37(1): 57-68.<br>• Pickle, L.W. and Su, Y. 2002. Within-state geographic patterns of health insurance coverage and health risk factors in the United States. *American Journal of Preventive Medicine* 22(2): 75-83. |

| Week | Class Dates | Lecture Topic/Readings | Readings |
|---|---|---|---|
| 15 | 11/27 | Lab: Spatial Cluster Analysis | None<br>Lab Write-up Due 12/8 |
| 14 | 12/4 | Spatial Cluster Methods | Student Papers |
| | | | **FINAL PROJECT DUE 12/4** |

**Reading Materials**

Below, I have some suggestions for books and workbooks that may be useful for you, especially if you want more complex "textbook style" readings or step-by-step instructions for how to conduct specific analyses. Some of these focus on R, some focus on ArcGIS. If you need to learn some skills, these are my suggestions.

Workbooks

- Kurland KS, Gorr WL. 2007. *GIS Tutorial for Health*. ESRI Press: Redlands, CA. ($$)

- Parker RN, Asencio EK. 2008. *GIS and Spatial Analysis for the Social Sciences: Coding, Mapping, and Modeling*. Routledge/Taylor & Francis: New York, NY. ($$)

- Anselin L. 2005. *Spatial Regression Analysis in R: A Workbook.* University of Illinois, Urbana-Champaign: Spatial Analysis Laboratory. Available online: http://geodacenter.asu.edu/system/files/rex1.pdf

- Harris R. 2009. Introduction to Geographically Weighted Regression. University of Bristol: School of Geographical Sciences & CMPO. Available online: http://www.bris.ac.uk/cmpo/events/2009/segregation/gwr.pdf. And here's the data: http://www.bris.ac.uk/cmpo/events/2009/segregation/southeastdata.csv.

  *The expectation is that if you need to you work thorough these workbooks at your own pace and use these as a way to build up your confidence and abilities in handling geospatial data. If you want, I can identify the "must do" tutorials. There is some duplication of general areas across these workbooks and there is no need to do all exercises (except for honing your own skills). However, one should note that while there is overlap there are also differences.*

Book and other materials

- Bivand, R, EJ Pebesma and V Gomez-Rubio. *Applied Spatial Data Analysis with R*. New York: Springer.
  Supplementary material: http://www.asdar-book.org/

- Cromley E and S McLafferty. 2011. *GIS and Public Health* (2nd ed). New York: The Guilford Press.

- Fotheringham AS, Brunsdon C, Charlton M. 2002. *Geographically Weighted Regression*. West Sussex, England: John Wiley and Sons.
  Supplementary material: http://ncg.nuim.ie/ncg/GWR/software.htm

- Singer JD and JB Willett. 2003. *Applied Longitudinal Data Analysis.* Oxford, New York: Oxford University Press.
  Supplementary material: http://gseacademic.harvard.edu/alda/

- de Smith MJ, Goodchild MF, Longley PA. 2006-2008. *Geospatial Analysis: A Comprehensive Guide to Principles, Techniques and Software Tools*. Available online at: http://www.spatialanalysisonline.com/. This is perhaps the most comprehensive single on-line source for material on both concepts and methods available.

- Waller LA and CA Gotway. 2004. *Applied Spatial Statistics for Public Health Data*. New York: John Wiley and Sons.
  Supplementary material: http://www.sph.emory.edu/~lwaller/WGindex.htm

- Ward MD and KS Gleditsch. 2008. *Spatial Regression Models*. Thousand Oaks, CA: Sage Publications.