Spatial Analysis for Population Health Research

Guangqing Chi

Associate Professor of Rural Sociology and Demography
Faculty Director, Computational and Spatial Analysis Core, PRI/SSRI

The CTSI BERD seminar on Recent Topics in Research Methods, Penn State
April 19, 2016
Population Health research in the Era of Big Data

1) Individual Electronic Medical Records (EMRs)
2) Contextual data
3) Behavioral data collected by mobile devices
4) Social media data

Spatial analysis and interdisciplinary research
Outline

1. GIS and spatial analysis
2. Example 1: Chi’s population research and GIS
3. Example 2: P3 projects
4. Example 3: Penn State Cancer Institute’s healthcare catchment area
5. Example 4: County Health Ranking
6. The Computational and Spatial Analysis Core
GIS and spatial analysis
“GIS is a computer based technology which provides the tools for collecting, editing, storing, retrieving, analyzing and displaying spatial data, and can create new information from existing data.”

Software Components of a GIS

1) Data input and verification
   - Getting information into the GIS in a format that the software will recognize. Data quality is an issue.

2) Data storage and database management
   - Considering location, topology, and attributes of the data.

3) Data transformation
   - Data analysis and generation of new information.

4) Data output and presentation
   - Type of output media
   - How the final product be used?

5) Interaction with the end user
   - Need to present the data in a format which is easily understood. This includes both spatial and nonspatial presentations.
GIS

Spatial Analysis

Spatial Data Analysis

“Spatial Statistics”

Spatial point data analysis

Lattice data analysis

Geostatistics

Spatial interaction data analysis
Example 1:
Chi’s population research and GIS
Building upon existing studies of population change

Demography

Sociology
(Rural sociology & community sociology)

Geography
(Population geography & transportation geography)

Regional science

Planning
(Urban & regional planning)
The framework of my research in demography

- Natural environment
- Transportation infrastructure
- Land use & development
- Socioeconomic conditions
- Government policies
The framework of my research in demography

Natural environment

Transportation infrastructure

Land use & development

Socioeconomic conditions

Government policies

Forests
Water
Wetlands
Public lands
Riverbank/Lakeshore/Coastline
Viewsheds
The framework of my research in demography

- Natural environment
- Transportation infrastructure
- Land use & development
- Socioeconomic conditions
- Government policies

Population change
The framework of my research in demography

- Natural environment
- Population change
- Transportation infrastructure
- Land use & development
- Socioeconomic conditions
- Government policies

Spatial & Temporal
The framework of my research in demography
An example of the power of spatial analysis and interdisciplinary approach for population research

Spatial Variation in Amenity-Driven Rural Development: Implications of Economic Climate, Transportation Infrastructure, and Land Use
Prior research (cont.)

1. Regional developmental theories
2. The equilibrium theory
3. The life-cycle literature
4. Roback’s general equilibrium formulation of household and firm locations

“The past 75 years reveal that [natural amenities in terms of] climate and landscape were the most important factors [of migration].” – Mark Partridge, 2014.
But why are there still so many places that have rich natural amenities but cannot attract people to live there?
Conceptual framework

- Natural amenities
- Rural population change
- Transportation infrastructure
- Land use & development
- Socioeconomic conditions
- Government policies

Spatial & Temporal
Natural Amenities in Wisconsin
1. Natural amenities have impacts on rural development:
   a) population growth

<table>
<thead>
<tr>
<th>Natural Amenities</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Golf courses</td>
<td>4.20E-7**</td>
</tr>
<tr>
<td>Lakeshore/riverbank/coastline</td>
<td>-0.001***</td>
</tr>
</tbody>
</table>


1. **Natural amenities have impacts on rural development:**

b) **in-migration**

The effects of national amenities on in-migration, 1995–2000, Wisconsin

<table>
<thead>
<tr>
<th>Natural Amenities</th>
<th>Spatial Lag Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forest</td>
<td>$-0.044^{***}$</td>
</tr>
<tr>
<td>Water</td>
<td>0.023</td>
</tr>
<tr>
<td>Wetland</td>
<td>$-0.020$</td>
</tr>
<tr>
<td>Public land</td>
<td>0.027$^{†}$</td>
</tr>
<tr>
<td>Riverbanks/lakeshores/coastlines</td>
<td>$8.171E-5$</td>
</tr>
<tr>
<td>Golf courses</td>
<td>$-1.716E-8$</td>
</tr>
<tr>
<td>Viewsheds (12.5%–20%)</td>
<td>0.116$^{***}$</td>
</tr>
</tbody>
</table>

1. Natural amenities have impacts on rural development:
   c) educational attainment

Temporal variation of the effects that natural amenities and economic forces have on talent share (% bachelor’s degree) over time

But...
2. Natural amenities play a stronger role in a good economy, but economic conditions play a stronger role in a weak economy

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Forest</td>
<td>0.002</td>
<td>-0.006</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>Water</td>
<td>0.077</td>
<td>0.051</td>
<td>0.065</td>
<td>0.060</td>
</tr>
<tr>
<td>Unemployment rate in 1980</td>
<td>-0.130*</td>
<td>-0.133*</td>
<td>-0.128*</td>
<td>-0.128*</td>
</tr>
<tr>
<td>Income in 1980</td>
<td>-1.79E-7</td>
<td>-9.08E-7</td>
<td>-6.81E-7</td>
<td>-7.82E-7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Forest</td>
<td>0.052**</td>
<td>0.042*</td>
<td>0.061**</td>
<td>0.061**</td>
</tr>
<tr>
<td>Water</td>
<td>0.070</td>
<td>0.011</td>
<td>0.005</td>
<td>0.004</td>
</tr>
<tr>
<td>Unemployment rate in 1990</td>
<td>0.155</td>
<td>0.097</td>
<td>0.093</td>
<td>0.094</td>
</tr>
<tr>
<td>Income in 1990</td>
<td>9.38E-7</td>
<td>3.18E-7</td>
<td>1.59E-7</td>
<td>1.48E-7</td>
</tr>
</tbody>
</table>

3. The impacts of natural amenities on rural development are conditional on transportation infrastructure

The effects of natural amenities and transportation on in-migration in Wisconsin, 1995–2000

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Spatial lag model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forest</td>
<td>–0.044***</td>
</tr>
<tr>
<td>Water</td>
<td>0.023</td>
</tr>
<tr>
<td>Wetlands</td>
<td>–0.020</td>
</tr>
<tr>
<td>Public lands</td>
<td>0.027†</td>
</tr>
<tr>
<td>Riverbanks/lakeshores/coastlines</td>
<td>8.171E–5</td>
</tr>
<tr>
<td>Golf courses</td>
<td>–1.716E–8</td>
</tr>
<tr>
<td>Viewsheds (12.5%–20%)</td>
<td>0.116***</td>
</tr>
<tr>
<td>Accessibility index 1</td>
<td>–0.006*</td>
</tr>
<tr>
<td>Accessibility index 2</td>
<td>–1.432e–4</td>
</tr>
</tbody>
</table>

Accessibility indices were generated by PFA. Index 1 mainly represents distances to highways, airports, and metropolitan areas, and highway density. Index 2 mainly represents journey to work.

4. Natural amenities have a trade-off relationship with land use and development in affecting rural development.
Land Developability in the Continental USA, 2011, Census Tract Level
Land Developability in Interactive WebGIS Map

http://www.landdevelopability.org
Land Developability in 2006, Centre County = 60%
5. The impacts vary spatially across urban, suburban, exurban, and rural remote areas.

Urban-rural classifications:
1. The U.S. Census Bureau’s 2000 Census Urbanized Areas.
3. The Beale Code of USDA ERS.
5. The impacts vary spatially across urban, suburban, exurban, rural remote areas (cont.)

The spatial variation of natural amenity effects on immigration, 1995–2000, Wisconsin

<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>All</th>
<th>Urban</th>
<th>Suburban</th>
<th>Rural-Adjacent</th>
<th>Rural-Exurban</th>
<th>Rural-Remote</th>
</tr>
</thead>
<tbody>
<tr>
<td>The proportion of forest area</td>
<td>−</td>
<td>−</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The proportion of water area</td>
<td></td>
<td></td>
<td>−</td>
<td>+</td>
<td></td>
<td>−</td>
</tr>
<tr>
<td>The proportion of wetland area</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>The proportion of public land area</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Riverbanks/lakeshores/coastlines</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Golf courses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Viewsheds (12.5%–20%)</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

6. Counterurbanization in remote rural areas: amenities become valuable only when accessible such as through public lands

What is counterurbanization?

The diffusion of more affluent “urban refugees” to remote high-quality environments, catering to consumers in the development of both primary and recreational housing (as second, third, and fourth homes) (Halfacree, 2012; Mitchell, 2004).
Remote Rural Areas and Natural Amenities

Remote Rural Wisconsin | Wisconsin as a Whole
--- | ---
Population density | 28 persons/km² | 142 persons/km²
Distance to the nearest metro city | 111 km | 52 km

Summary

1. Natural amenities have impacts on rural development overall
   BUT

2. The impacts are conditional on several factors:
   • Economic climate
   • Transportation infrastructure
   • The availability of land for development and conversion

3. The impacts vary across the urban-rural continuum
   • Natural amenities are appreciated the most in rural areas adjacent to metropolitan areas, but the least in urban areas
   • In remote rural areas, amenities become valuable only when accessible through managed recreational areas
Example 2:
P3 projects: Collaborative projects with Chuang and Hwang
1. PSU Pathway to Partnership (P3) Stage 1: Population health research in the era of Big Data
2. P3 Stage 2: Place and health in Pennsylvania
   • Obesity and Asthma Outcomes
   • Geographic Patterns in Prescription Narcotics
Contextual data (about 50 variables)

1. Demographics
2. Social environment
3. Built environment (including transportation)
4. Natural environment
5. Access to healthcare
6. Food environment
7. Policy environment
Example 3:
Penn State Cancer Institute’s healthcare catchment area
http://redlands.pop.psu.edu/psci/
Example 4:
County Health Ranking
http://www.countyhealthrankings.org
Computational and Spatial Analysis Core (CSA)
The Computational and Spatial Analysis Core

1) To provide multifaceted data support and advice to social scientists.

2) To provide programming support, statistical expertise, and software packages for population research.
The Computational and Spatial Analysis Core

1) To provide multifaceted data support and advice to population and social scientists.

2) To provide programming support, statistical expertise, and software packages for population and social research.

3) To provide expertise, services, and opportunities for novel research collaborations using spatial statistics and analysis.
   • ESDA, geovisualization, webGIS
   • [http://csa.ssri.psu.edu/webgis](http://csa.ssri.psu.edu/webgis)
   • Spatial statistics and analysis
The Computational and Spatial Analysis Core

1) To provide multifaceted data support and advice to social scientists.

2) To provide programming support, statistical expertise, and software packages for population research.

3) To provide expertise, services, and opportunities for novel research collaborations using spatial statistics and analysis.
   - ESDA, geovisualization, webGIS
   - Spatial statistics and analysis

4) To support and strengthen capacity to access and analyze microdata provided by Penn State’s new Federal Statistical RDC.
Emerging Areas: RDC

- Opened in the spring 2014, the Federal Statistical Research Data Center (RDC) is a valuable asset for population researchers.
- The CSA Core staff have access to and experience with public versions of many datasets available in the RDC. Staff can provide support for preliminary analyses for RDC proposals and can reduce startup costs by developing as much of the analysis as possible for later application in the RDC.
- Staff either have already obtained or are in the process of obtaining “Special Sworn Status” for accessing the restricted microdata available in the RDC.
- The CSA Core aims to become the gateway for accessing and analyzing microdata as housed in the Penn State’s RDC.
The Computational and Spatial Analysis Core

1) To provide multifaceted data support and advice to social scientists.
2) To provide programming support, statistical expertise, and software packages for population research.
3) To provide expertise, services, and opportunities for novel research collaborations using spatial statistics and analysis.
   • ESDA, geovisualization, webGIS
   • Spatial statistics and analysis
4) To support and strengthen capacity to access and analyze microdata provided by Penn State’s new Federal Statistical RDC.
5) To promote innovative population and social research using Big Data.
Emerging Areas: Big Data

- the integration and analysis of large spatial, historical, individual, and contextual data;
- social networks and complex systems analysis;
- geo-tagged social media data collection and analysis; and
- population-engineering nexus modeling.
Big Data projects that the CSA Core is involved in

- Kifer, Matthews, Yang, NSF: privacy-preserving technology
- Li, Kifer, Graif, NSF: social flow using the Census’ Local Employment Dynamics data and New York City’s Taxi Trip data
- Van Hook, Chi: food environment using HomeScan data and transportation network data
- Chi, Li, NSF: population-infrastructure nexus
- Henebry, Chi, NASA: climate change and high-elevation communities using population and remote sensing data
The Deliverables

- Maps for publications, grant proposals, and presentations
- Contextual databases drawn from publicly available data sets or other geospatial data sources
- Software programs, interactive mapping websites, spatial and network analysis
Guangqing Chi
gchi@psu.edu
814-865-5553
http://www.landdevelopability.org/chi