
– Brownfield development
Spatial Patterns of Pb Hotspots

- Sources of Pb:
  - Traffic (lead additive in gas)
  - Burning of peat & coal for home heating
  - Solid waste
- Spatial cluster (regional hotspots) – sites with high Pb concentration surrounded by high concentration sites.
- Spatial outlier (individual hotspots) – sites with high Pb concentration surrounded by low concentration sites.
- Cool spots
Data

- 166 surface soil samples (0-10 cm depth)
- Stratified random sampling (1 sample per 0.25 km²)

One-point Descriptor
Moran's I (Two-point Descriptor)

\[ I = \left( \frac{n}{\sum_i \sum_j w_{ij}} \right) \left( \frac{\sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2} \right) \]

- \(x_i\) and \(x_j\): values observed at locations \(i\) and \(j\)
- \(\bar{x}\): average of all \(x_i\)
- \(w_{ij}\): weight between locations \(i\) and \(j\)

Examples of \(w_{ij}\)

- \(w_{ij} = 1 / d_{ij}\)
- \(w_{ij} = 1\) if \(i\) touches \(j\), else 0 (local Moran's I)

- +1: clustering (positive spatial autocorrelation)
- 0: random
- -1: dispersion (negative spatial autocorrelation)

**Conceptualization of Spatial Relationships**

 Specifies how spatial relationships between features are conceptualized.

- Inverse Distance—The impact of one feature on another feature decreases with distance.
- Inverse Distance Squared—Same as Inverse Distance, but the impact decreases more sharply over distance.
- Fixed Distance Band—Everything within a specified critical distance is included in the analysis, everything outside the critical distance is excluded.
- Zone of Indifference—A combination of Inverse Distance and Fixed Distance Band. Anything up to a critical distance has an impact on your analysis. Once that critical distance is exceeded, the level of impact quickly drops off.
- Polygon Contiguity (First Order)—The neighbors of each feature are only those with which the feature shares a boundary. All other features have no influence.
- Get Spatial Weights From File—Spatial relationships are defined in a spatial weights file. The path name to the spatial weights file is specified in the Weights Matrix File parameter.

**Value of Distance Band:**  
> sampling interval  
< half of the smaller dimension of the study area
What can you tell about the difference in these patterns?

Test for Statistical Significance

- T-test (observed data have a normal distribution)
  - Z score, p value ($\alpha$)
- (Conditional) permutation
  - Rearrangement of existent elements
  - Check for pseudo significance
    $$PS = \frac{(M+1)}{(R+1)} \times 100\%$$
    - $M$: number of instances that meet certain criteria
    - $R$: total number of permutation instances
A. Random Pattern

Global Moran's I Summary
Moran's Index: 0.000019
Z Score: 0.987484
p-value: 0.323405

B. Positive Spatial Autocorrelation

Global Moran's I Summary
Moran's Index: 0.057173
Z Score: 6.564235
p-value: 0.000000
C. Negative Spatial Autocorrelation

Global Moran's I Summary
Moran's Index: -0.036682
Z Score: -2.593673
p-value: 0.009496

Local Moran’s I: Local Indicators of Spatial Association (LISA)
Transform Data to Normal Distribution

- Box-Cox (Power/Log) Transformation (see eq. 2)
- Normal score transformations (see page 214)

Results – Distance Bands

<table>
<thead>
<tr>
<th>Distance</th>
<th>1000 m</th>
<th>5000 m</th>
<th>2000 m</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Results – Data Transformation

<table>
<thead>
<tr>
<th>Original</th>
<th>Normal Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Box-Cox</td>
<td></td>
</tr>
</tbody>
</table>

Results – Outliers

- Not recommended to remove outliers
- Instead, replace their values with the upper bound of the Box-and-Whiskers plot.
Moran's scatterplot (available in Geoda)

- X axis: variable value at a location (e.g., Pb level)
- Y axis: spatial weighted average variable values of the neighbors of that location.

![Moran Scatterplot: Regional Income in 1989](image)

Slope = Moran's I = 0.7453