Assessing the Association of Air Pollution and Social-economic Factors with Asthma in Massachusetts

Project Description

Asthma is an illness that affects the respiratory tract and airways that carry oxygen into and out of the lungs. Air pollution could lead to increased asthma prevalence. The aim of this project is to examine the association of air pollution (fine particulate matter ($PM_{2.5}$) and ozone (O_3)), social-economic factors (distance to highways and median household income) with asthma pediatric prevalence in Massachusetts. The project evaluates the spatial distribution of asthma pediatric prevalence, as well as the spatial distribution of air pollutants values, distance to highways and median household income in Massachusetts. The multiple spatial regression is done in Geoda to measure how and to what extent each variables would affect asthma pediatric prevalence.

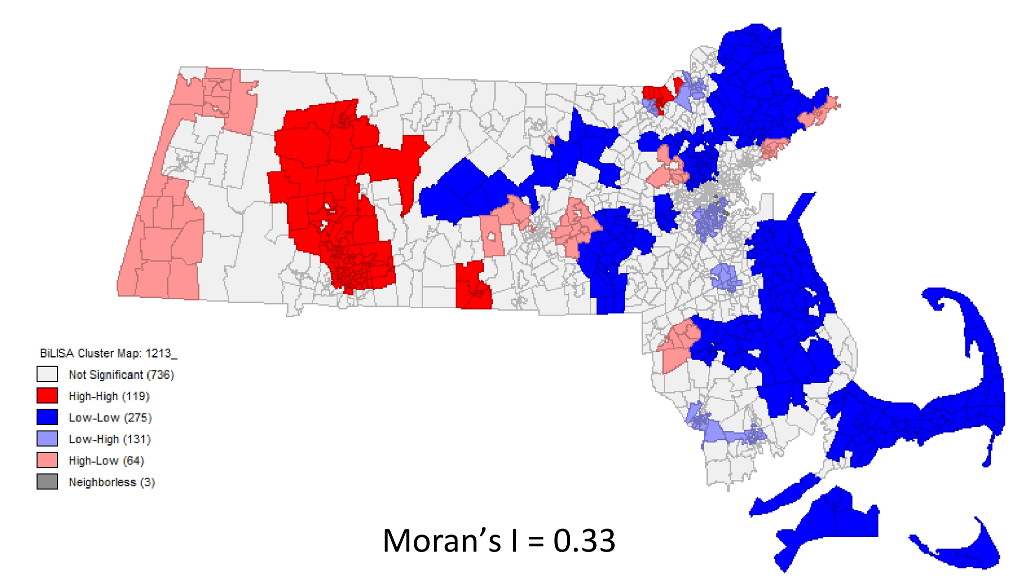
Data and Methodology

The asthma pediatric prevalence data is collected from Bureau of Environmental Health. The data is limited to students in grades K-8 in public and private schools in Massachusetts in 2014. Air quality data was taken from US EPA's Air Quality System (AQS) at the various monitoring stations located in different counties of Massachusetts in 2014. Air pollution concentrations of $PM_{2.5}$ and ozone were collected from 23 and 19 monitoring stations respectively. The characteristics of the raw data collected from the website are daily average (24 hours) concentrations of $PM_{2.5}$ and daily maximum 8 hours average concentrations of ozone. The daily data for each monitoring station were used for determination of annual average concentrations. The median household data is taken from American Community Survey, 2014 estimates and the highway data is from the MassGIS.

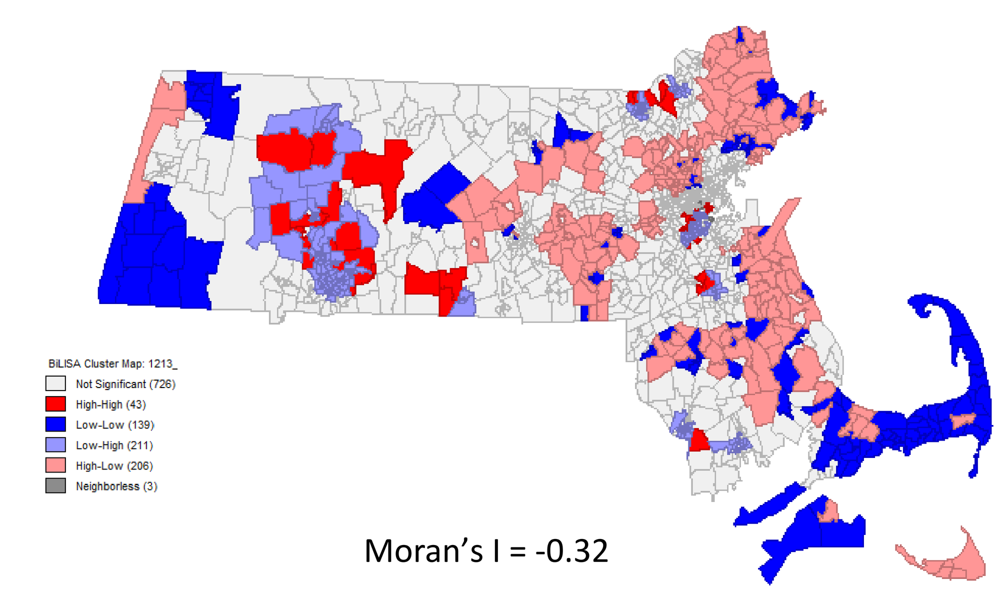
All air pollution point data were entered into ArcGIS by displaying the longitude and latitude of monitoring stations. The Ordinary Kriging method was used to estimate the spatial distribution of asthma prevalence, ozone and $PM_{2.5}$ for Massachusetts in 2014. In Kriging, a smooth surface is estimated from irregularly spaced data points based on the assumptions that the spatial variation in the feature (asthma, ozone and $PM_{2.5}$) is homogeneous over the domain depends only on the distance between sites. The median household data was joined to census tract. The distance to highway was done by using Euclidean distance. The value of above five variables was joined to census tract using spatial join in order to evaluate the spatial regression of four dependent variables and asthma prevalence.

Mapping Cluster and Outlier

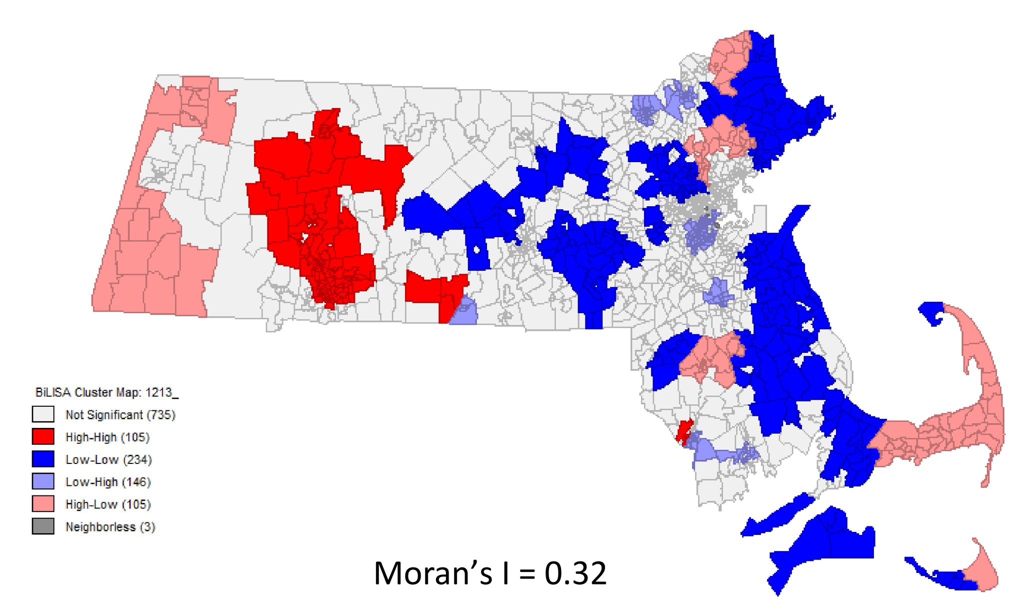
PM2.5 Bivariate Lisa Cluster Map



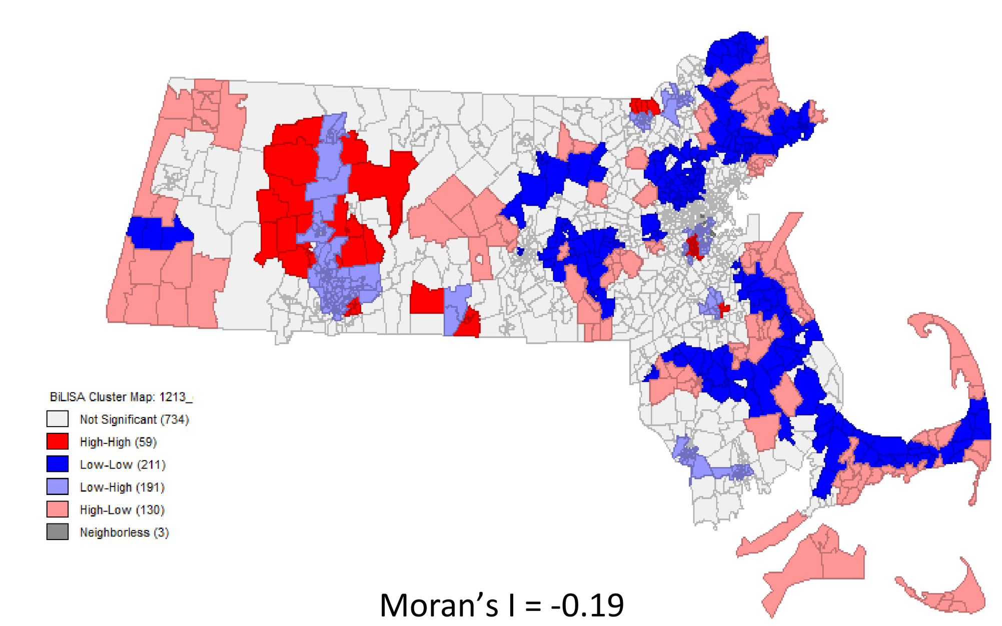
Household Income Bivariate Lisa Cluster Map



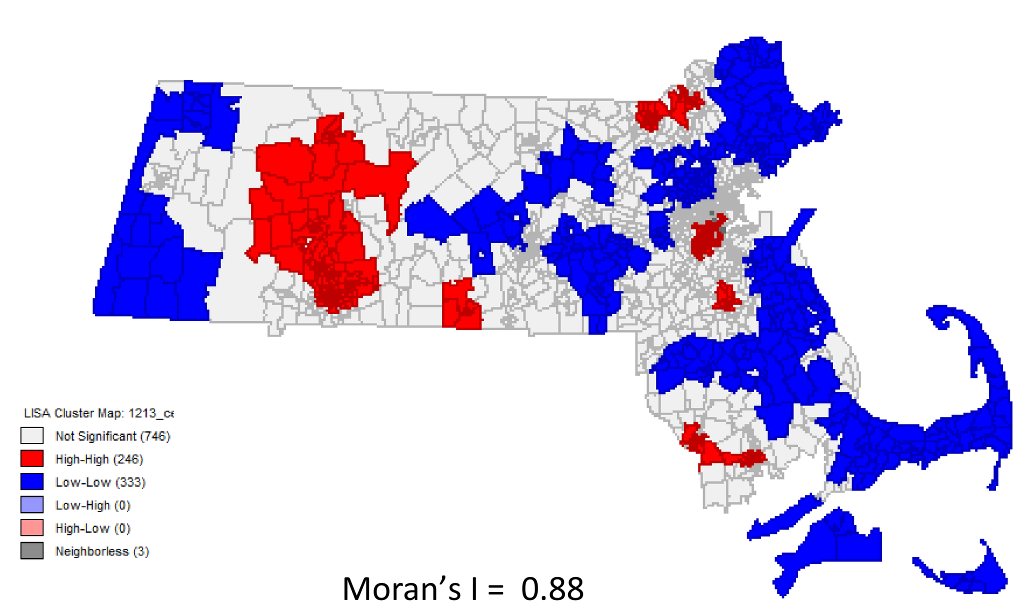
Ozone Bivariate Lisa Cluster Map



Distance to Highway Bivariate Lisa Cluster Map



Asthma Univariate Lisa Cluster Map



Spatial Regression Analysis

Spatial error model was used according to the diagnostic test for variables in classic model. The independent variable is asthma prevalence, and dependent variables are $PM_{2.5}$, ozone, distance to highway, and median household income.

Dependent Variables	Coefficient	Probability
Constant	-4.45	0.46
$PM_{2.5}$ Concentration	2.88	0.00
Ozone Concentration	-2.88	0.00
Distance to Highway	-0.27	0.00
Household Income	-0.04	0.00
Lambda	0.22	0.00

Conclusion

The univariate cluster map of asthma prevalence indicates asthma is mainly clustered in Hampden, Hampshire, and Franklin County. The Moran's I is 0.88, indicating the cluster of asthma is intensive. According to the bivariate cluster map for the four dependent variables with asthma prevalence, the High-High cluster of ozone and $PM_{2.5}$ are very similar, which means the area of high asthma prevalence surrounded by high $PM_{2.5}$ concentration is also surrounded by high ozone concentration.

The regression result shows that four dependent variables are all significant in this analysis (probability less than 0.05). Ozone concentration, distance to highway and household income are negatively correlated with asthma prevalence whereas $PM_{2.5}$ is positively correlated with asthma prevalence. The coefficient indicates that each additional 1μg/m^3 of $PM_{2.5}$ concentration is associated with an increase of 2.88 asthma prevalence, while 1ppb of ozone concentration, one mile distance to highway, or one thousand dollars of median household income is associated with a decrease of 2.88, 0.27 or 0.04 asthma prevalence respectively. The probability of Lambda is less than 0.05, indicating the relative importance of the spatial context to the model. The coefficient of Lambda is 0.22, indicating the dependent variables are not relatively clustered in the model.

Limitation

1. Avalability of uniformly distributed air pollution and health data. Air monitoring stations in Massachusetts was not uniformly distributed and might lead some errors in air pollution prediction level.
2. Asthma data in this study only records asthma cases that occur inside schools, but some student might choose to visit private doctor and those cases may not be listed in the data.
3. Asthma data does not reflect severity of asthma problems.
4. The variables in the regression analysis is not enough, and more significant dependent variables should be included in the future.

Cartography: Yi Zhong

Data: 15 Dec. 2017

Projected Coordinate System: GCS_WGS_1984

Source: MassGIS 2010; ACS 2014 estimates; EPA Air Quality 2014; Bureau of Environmental Health

Tufts UNIVERSITY

Mapping Variables