Abstract
The association between increasing air pollution and adverse health effects has been demonstrated by numerous studies in the fields of epidemiology and occupational health. Carbon monoxide (CO) is an air pollutant commonly released through vehicle emissions. It can serve as an indicator of air pollution levels in specific locations. This project is interested in predicting CO levels at homes of women in the Seattle area in order to determine whether they are predictive of carboxyhemoglobin (COHb) levels in pregnant women. One approach for a CO prediction model is Land Use Regression (LUR). A LUR predictor model will be fit to the data. A spatio-temporal model will be generated in order to see trends in the data. However, this project has focused mostly on developing an improved model which includes both space and time coefficients. This model is known as a spatio-temporal model. The data for model fitting was obtained from the Puget Sound Clean Air Data Bank. Average monthly CO levels were calculated across monitor locations in the greater Seattle area for the period 1996-2006. The geographical predictors of distance to major roadway, population density, and street density within a 500 meter buffer were calculated using 2000 census data. The smooth curves added to the data visually appear to fit better than the LUR predictions. A spatio-temporal model will be fit and used to better capture variation in time and space at each site as well as make future predictions. This model is expected to predict the CO levels at women’s homes more accurately than simply using LUR. Future work will relate CO concentration to human absorbance using the biomarker COHb. COHb levels are predictive of adverse pregnancy outcomes in women. Furthermore, the process of generating the model for CO levels can be applied to other air pollutants.

CO in the Atmosphere and Its Impact on Health
• CO binds to hemoglobin forming COHb, a biomarker of adverse effects (see figure).
• Long-term exposure to air pollution is associated with higher COHb levels, increased incidence of heart failure, and damage to the nervous system.
• Exposure to increased CO levels in utero is connected with preterm delivery, intrauterine growth restriction, and reduced birth weight in diverse geographic settings.1
• Short-term increases in CO exposure can lead to fatigue of skeletal muscles and respiratory infections.

Methods
• Average monthly CO levels were calculated using data at 17 western Washington EPA sites.
• The data were plotted and a smooth curve was fit through the points to show the overall trend at each site.
• A Land Use Regression model was fit in order to make long-term predictions from the model.
• For comparison, a spatio-temporal model will be developed to provide predictions. This model allows predictions to vary over space and time.
• CO predictions will be made at homes of women and we will assess their relationship with maternal prenatal COHb levels, an important biomarker of prenatal exposure to CO.

Land Use Regression Steps
• Calculate GIS covariates: distance to major roadways (classified as A1 = freeways, A2 = highways, and A3 = secondary roads), population density, and street density within a 500 meter buffer at the site. Categorize distance to roadways and street density. Develop the LUR regression model to obtain predicted CO values from the following equation:

\[ \text{CO}_{\text{predicted}} = \beta_0 + \beta_1 \text{distance to A1 or A2 road} + \beta_2 \text{distance to A3 road} + \beta_3 \text{month indicator} + \beta_4 \text{year indicator} \]

where
• \(\beta_0\) is the intercept, \(\beta_1\) and \(\beta_2\) are regression coefficients for the trend in CO levels, \(\beta_3\) and \(\beta_4\) are regression coefficients for the trend in time.

Site Variation
Each site has unique GIS covariates. This is exemplified by the differences in Sites 60, 61, 62.
• Site 61 is a monitor on Beacon Hill. The monitor rests near a golf course and a high traffic area.
• Site 60 resides on 45th Street in the University District. It has the highest population density of any of the sites.
• Site 62 is a monitor on Beacon Hill. This site is near a golf course and park.

These evident differences in site location make it necessary to take each site into account and not generalize across them unless their covariates are similar.

Development of a Spatio-Temporal Model for Predicting CO levels

Discussion and Conclusions, Future Steps
• The LUR predictions miss important features of the data. This is especially evident in site 62 and site 67. The LUR predictions lie far from the actual data points and would therefore be unreliable if used for predictions of CO levels.
• The spatio-temporal model does a better job of capturing how temporal variation varies across the sites. This is because the curve is adjusted differently based on the unique characteristics of each site.
• More work is needed to understand how to incorporate differences across sites simply in the spatio-temporal model. It is proving difficult to fit a complex model when there are a limited number of sites with long time series.

References and Acknowledgements
3.Kind assistance and code from Paul Sampson, Adam Saprio