



**Undergraduate Modeling Workshop**  
**May 20-25, 2018**  
**PROJECT ABSTRACTS**

**Group Leader:**  
Yawen Guan

***Project II: Air Quality***

*“Data analysis on air pollutant exposures. (We will use R for data analysis)”*

**Abstract:** Fine particulate matter (PM<sub>2.5</sub>) is a mixture of air pollutants that, at a high concentration level, has adverse effects on human health. An interesting statistics problem is to estimate these pollutant exposures for the entire US, such estimates can be used to inform policy and decision making. During the workshop, we will work on two major source of air quality data that are used by the EPA to estimate pollutant exposures, including monitoring data and the Community Multiscale Air Quality (CMAQ) model. The monitoring stations provide fairly accurate measurements of the pollutants; however, they are sparse in space and take measurements at a coarse time resolution, typically 1-in-3 or 1-in-6 days. On the other hand, the CMAQ model provides daily concentration levels of each component with complete spatial coverage on a grid; these model outputs, however, need to be evaluated and calibrated to the monitoring data. We will explore these air quality data for the summer of 2011 and brainstorm on statistical models to estimate air pollutant exposures.

**Group members: Meixi Chen, Vincent Gonzales, Alan Ji, Chandni Malhotra, Hongyu Mao, Sharon Sung**