Mapping Air Quality Trends
Robin Osekowsky, Josh Ramos, Michael Ostertag, Tajana Rosing
rosekows@uscd.edu, jdr023@ucsd.edu, mosterta@eng.ucsd.edu, tajana@eng.ucsd.edu

Introduction
- Respiratory problems have been linked to poor air quality
- Health issues can be prevented with increased awareness of air quality in daily life
- Goal is to use predictive modelling to map trends using the limited data obtained from 12 Environmental Protection Agency (EPA) sites in San Diego County and participatory sensing
- Part of MetaSense project to make air quality data more available on a day-to-day basis as a preventative measure against respiratory health problems

Materials
- Electrochemical sensor: Samples CO, NO2, & ozone every 5 sec
- GPS enabled
- Connected via Bluetooth to smartphone

Methods

Collect Data → Calibration and Processing → Prediction and Mapping

Collect Data
- Chose locations with somewhat regular spacing
- Collected sensor readings for 2 minutes
- Averaged values across time period

Calibration and Processing
- Calibrated the sensors when they were co-located with EPA sites
- Create linear regression model with data collected after calibration period
- Use model to convert our data to ppb/ppm

Prediction and Mapping
- Goal is to solve following equation for each point in grid across area of interest
  \[ \hat{Z}(s_0) = \sum_{i=1}^{N} \lambda_i Z(s_i) \]
  where \( Z(s_i) = \) concentration measured at location \( s_i \) and \( s_0 = \) location of interest
- Weights (\( \lambda_i \)) determined by kriging\(^1\), a method of statistical mapping optimized for geospatial data, accomplished through use of open-source software, PyKrige\(^2\)
- Weights mimic trends across area

Results

Comparison of EPA and Calibrated Sensor Measurements over Time

![Estimated CO Distribution](image1)

![Estimated O3 Distribution](image2)

Average CO Error : 37.49%
Average O3 Error : 33.13%
- Despite high error, models capture general trends which are similar for different pollutants

Conclusions & Future Work
- We tried several calibration models and found one to fit best. We will continue to investigate better combinations of features
- We validated a method for creating an interpolated map of pollutant concentrations
- In the future, we will reduce the error by incorporating land-use regression and transient pollution models
- We also would like to expand to other forms of measurement which would allow us to produce a more accurate model through sensor fusion

Citations

Acknowledgments
This material is based upon work supported by the National Science Foundation under Grant No. CNS-1339335 and Grant No. CNS-1446912 (MetaSense Project).
Thank you to Aditi Mavalankar and Christine Alvarado for their involvement.