S 470/670 Small project 1

Due Oct 6th, 2017

Background

A 2015 paper published in PNAS was titled "Methane emissions from natural gas infrastructure and use in the urban region of Boston, Massachusetts." The abstract:

> Methane emissions from natural gas delivery and end use must be quantified to evaluate the environmental impacts of natural gas and to develop and assess the efficacy of emission reduction strategies. We report natural gas emission rates for 1 y in the urban region of Boston, using a comprehensive atmospheric measurement and modeling framework. Continuous methane observations from four stations are combined with a high-resolution transport model to quantify the regional average emission flux, 18.5 ± 3.7 (95% confidence interval) g $CH_4 m^{-2} y^{-1}$. Simultaneous observations of atmospheric ethane, compared with the ethane-to-methane ratio in the pipeline gas delivered to the region, demonstrate that natural gas accounted for ∼60-100% of methane emissions, depending on season. Using government statistics and geospatial data on natural gas use, we find the average fractional loss rate to the atmosphere from all downstream components of the natural gas system, including transmission, distribution, and end use, was 2.7 ± 0.6% in the Boston urban region, with little seasonal variability. This fraction is notably higher than the 1.1% implied by the most closely comparable emission inventory.

The paper is here:
`http://www.pnas.org/content/112/7/1941.full.pdf`

Data

The data can be downloaded here: `tdx.doi.org/10.7910/DVN/28530`, or you can download from Canvas → Files → Small Projects → Small Project 1. Find the description of the datasets in the `README` file. To get data into R, save it as a .csv file, then use `read.csv()`.

Questions

- **Q1:** Consider the dataset `COP_CH4_Obs_Mod_Bg_Sep2012-Aug2013.csv`.

  (i) How does CH_4 observations in the afternoon (16-21 UTC) change with four seasons at the location Copley Square (COP)? Consider spring as March, April and May; summer: JJA; fall: SON; winter: DJF.

  (ii) Make a time series plot for *daily* CH_4 observations at COP. Fit a trend component and determine if there are other strong time components in daily CH_4 observations at COP. One option to find the daily average of CH_4 observations is to use `ddply()` from R library `plyr`. You can also use other functions you prefer.

- **Q2:** Consider the dataset `BU_C2H6_CH4_5min_May-Jun2014.csv` for spring observations and `BU_C2H6_CH4_5min_Oct2012-Jan2013.csv` for fall-winter observations. Fit a linear model using C_2H_6 as response variable and CH_4 as explanatory variable for spring and fall-winter, respectively. Summarize all your findings about model fitting and compare the results in spring and fall-winter.

Directions

- Work in groups of three or four (or two). One person per group should submit a set of answers (making sure that everybody's name is on the submission.) The submission should include a PDF and your code.

- The maximum length of your write-up, *including graphs*, is six pages. Your write-up should aim to convince someone with reasonable knowledge of statistics of your main points.

Grading

You'll be graded on the following criteria:

- **Answer to the questions.**

- **Presentation.** This include clarity and correctness of the graphs and clarity and correctness of the writing. Spelling and grammar will be a small but non-zero proportion of the grade.

- **Reproducibility.** Miguel will attempt to reproduce all your graphs and numerical results from the materials you provide. If he can, you get all the points for this; if he can't, you get no points.