report

Title of the report

Introduction

World Health Organization estimates 7 million death annually related to air pollution(WHO 2020). Air quality plays an vital role in public health. In order to improve air quality, this study aims to pinpoint the factors of air quality quantitatively. The factors under study are rainfall(inch), population density(per square mile), income per capita(dollar), added value of companies(dollar) and adjacency to coast(binary). To shed light on the relation between air quality and its factors, we adopts better subset selection algorithm. A better subset selection is favourable due to its better fit in terms of smaller residual sum of squares. The algorithm yields a better fit than a subset without optimization as the result of its monotonicity(**???**). Another strong motivation for a subset selection is the avoidance of unnecessary measurement or sampling. Smaller sets of variables may reduce measurement cost.
This study attempts to answer: Are societal impact stronger than geographic impact to air quality in 1970's California? (Ruoying: Here I group our variables into two groups. One is societal, the other is geographic. Open to discussion!!) We implemented better subset selection based on the algorithm of maximization in majorization. As the optimal subset is unknown, we ran the selection with all possible variable number. In this report, the implementation details in **??**Method) follows the introductio and preprocessing of data in **??**Data),

Data

Previous study has proven or suggested the relation between the five chosen variables and air quality. Both natural and anthropogenic events attribute to air quality in the atmosphere. A difference between generating air pollutant and distributing air pollutant is noted. The generation of air pollution relates to production and consumption from society (Baklanov, Molina, and Gauss 2016). Accordingly, this study captures the relation by population density, income per capita and added values from companies. Furthermore, the distribution of air pollution mainly depends on the wind field (Leelőssy et al. 2014), which is quantified by the variable of adjacency to coast and rainfall in this study as they both reflect the wind field's condition.
From the econometrics dataset, this study examines five independent variables and one dependent variable (**???**).

Method

Result

Reference

Baklanov, Alexander, Luisa T Molina, and Michael Gauss. 2016. "Megacities, Air Quality and Climate." *Atmospheric Environment* 126: 235–49.

Leelőssy, Ádám, Ferenc Molnár, Ferenc Izsák, Ágnes Havasi, István Lagzi, and Róbert Mészáros. 2014. "Dispersion Modeling of Air Pollutants in the Atmosphere: A Review." *Central European Journal of Geosciences* 6 (3): 257–78.

WHO. 2020. "Air Pollution." World Health Organization. https://www.who.int/health-topics/air-pollution#tab=tab_1.