Modeling impacts of traffic, air pollution, and weather conditions on CPD mortality[1]

Cong Cao,[a,b] Jan Morten Dyrstad [b], Colin P. Green [b]

[a] Linde center for Science, Society, and Public Policy, Division of the Humanities and Social Sciences, California Institute of Technology, 1200 E California Blvd, Pasadena, CA 91125, USA

[b] Department of Economics, Norwegian University of Science and Technology, Høgskoleringen 1, 7491 Trondheim, Norway

Abstract:

Aims: Cardiopulmonary disease (CPD) is a leading cause of death worldwide. Increasing evidence shows that air pollution and exposure to various weather conditions have important contributory roles. Understanding the interaction of these factors is difficult due to the complexity of the relationship between CPD, air pollution, and environmental factors; Methods: This paper uses regression models and machine learning approaches to explore these relationships, and investigate whether meteorological factors and air pollution have a synergistic effect on CPD. We use daily data from 2009-2018 from four cities representing the heterogenous climate conditions in Norway: the far north, the west coast, mid-Norway, and the south-east; Results: We demonstrate the importance of the interaction between weather and air pollution associated with higher CPD mortality, as is exposure to air pollution in the form of NOx and particulate matter. This impact is seasonal. Traffic is also positively related to CPD mortality, which may be caused indirectly through increased pollution. We demonstrate that machine learning outperforms regression models in terms of the accuracy of predicting CPD mortality; Conclusions: The inclusion of rich lagged structures and interactions between environmental factors are both important but can lead to overfitting of traditional models; Since these cities are not large cities by international standards, it is surprising that environmental factors have such obvious impacts on CPD mortality. CPD mortality shows a clear negative trend, implying an improvement in the public health situation.

Keywords: Machine Learning, Air Pollution, Mortality, Health Policy

Word count for the whole manuscript, excluding the abstract, references, figures, and tables：2449.

[1] **Acknowledgment:** Cong Cao is grateful for the research funding from Statens vegvesen (the Norwegian public roads administration) and acknowledges data received from the European Union's Horizon 2020 Project EXHAUSTION (Grant JO: 820655). She is also grateful for the comments that were made by participants at the Norwegian Public Health Institution Method Development and Analytics Department's meetings, the NAERE Conference (Uppsala, Sweden), the Urban Transitions Conference (Sitges, Spain) in 2022, and the Statistics Norway Research Department (Oslo, Norway) in 2023. E-mail: congc@caltech.edu