Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Modifiable areal unit problem and environmental factors of COVID-19 outbreak

Yaqi Wang a, Qian Di b,⁎

a Research Center for Public Health, Tsinghua University, Beijing 100084, China
b Vanke School of Public Health, Tsinghua University, 100084, Beijing, China

HIGHLIGHTS

• Modifiable areal unit problem (MAUP) arises when aggregating data into districts, leading to statistical bias.
• Real data of atmospheric NO2 and COVID-19 outbreak in China is analyzed to illustrate the presence of MAUP.
• Atmospheric NO2 shows different directions of association with COVID-19 mortality under different aggregation strategies.

ARTICLE INFO

Article history:
Received 19 April 2020
Received in revised form 1 June 2020
Accepted 3 June 2020
Available online 6 June 2020

Keywords:
COVID-19
NO2
Modifiable areal unit problem
Statistical bias
Spatial analysis

ABSTRACT

Several recent studies have explored the association between environmental factors, such as temperature, humidity, and air pollution, and the severity of the COVID-19 outbreak by analyzing the statistical association at the district level. However, we argue that the modifiable areal unit problem (MAUP) arises when aggregating disease and environmental data into districts, leading to bias in such studies. Therefore, in this study, we analyzed the association between environmental factors and the number of COVID-19 death cases under different aggregation strategies to illustrate the presence of MAUP. We used real-world COVID-19 outbreak data from the Hubei and Henan Provinces and studied their association with atmospheric NO2 levels. By fitting linear regression models with penalized splines on NO2, we found that the association between COVID-19 mortality and NO2 varies when data were aggregated (1) at the city level, (2) under two different aggregation strategies, and (3) at the provincial level, indicating the presence of MAUP. Therefore, this study reminds researchers of the presence of MAUP and the necessity to minimize this problem while exploring the environmental determinants of the COVID-19 outbreak.

© 2020 Elsevier B.V. All rights reserved.
framework: researchers analyzed the statistical association between numbers of COVID-19 cases at the city, regional, or national level with respect to a specific environmental factor. Apart from environmental factors, a recent study analyzing the association between COVID-19 cases and the prevalence of the Bacillus Calmette–Guérin (BCG) vaccination employed the same approach (Berg et al., 2020; Sala and Miyakawa, 2020). By finding the association between disease mortality and vaccine policy in different countries, the researchers suggested that BCG vaccination reduced the risk of COVID-19. However, we argue that these conclusions are dependent on the aggregation of environmental factors.
and disease data. Therefore, we are writing this correspondence to remind researchers of a common flaw, namely the modifiable areal unit problem (MAUP) that is associated with studies on the environmental factors of the COVID-19 outbreak.

The modifiable areal unit problem is a statistical bias that arises when aggregating point measurements into districts. Summary measures, such as mean and standard deviation, are influenced by the boundaries of the aggregation districts. Consequently, the regression outcome is also influenced by the aggregation strategy. For example, disease is negatively associated with temperature at the individual level (Fig. 1); however, after aggregating to the district level, the relationship between disease prevalence and temperature can be positive, negative, or null, depending on district boundaries.

We used real-world NO₂ concentration and COVID-19 outbreak data in the Hubei and Henan Provinces, China, to illustrate the presence of MAUP. By fitting regression models, we studied the association between COVID-19 death cases and daily NO₂ change (1) at the city level, (2) under two different aggregation strategies, and (3) at the provincial level (Figs. 2a and 3a). The dose-response curves for the four scenarios showed different shapes with different regression coefficients (Figs. 2b and 3b). The NO₂–COVID-19-death relationship was negative at the provincial level in Hubei Province, indicating a protective effect of NO₂ on COVID-19 mortality; however, a positive relationship was observed under aggregation strategy two (Fig. 2b). Similar differences were discovered in the Henan Province, where the positive city-level association became negative when the aggregation strategy was employed (Fig. 3b). The purpose of this real-world example is not to explore the relationship between COVID-19 and NO₂ but to demonstrate that such a relationship can be influenced by aggregation level and aggregation strategy.

The modifiable areal unit problem causes unreliable analytical results and encourages false conclusions regarding the dependence of transmission on certain factors, unnecessary control measures, or the unrealistic hope that warm weather or the BCG vaccine will impede COVID-19 transmission. Solutions for minimizing MAUP and achieving reliable analytical results include: (1) Conducting epidemiological studies at the individual level with decent exposure assessment and case-control or cohort study designs; although the finest data in this article was at the city level, MAUP still exists when aggregating individual data to the city level. Therefore, the Center for Disease Control and other institutions with individual tracking data may wish to try the preferred individual analysis method. (2) Combining epidemiological evidence with biological evidence; lab results on virus stability under different temperature and humidity conditions (van Doremalen et al., 2020) would be a good supplement for epidemiological findings. (3) Attempting to conduct the study at the finest spatial scale possible and listing this as a possible limitation if (1) is not possible.

We hope this short correspondence can help other researchers working on this topic to find substantial evidence on environmental contributors and other influential factors of the COVID-19 pandemic.

**Funding**

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**CRediT authorship contribution statement**

**Yaqi Wang:** Methodology, Software, Validation, Formal analysis, Data curation, Writing - original draft, Visualization. **Qian Di:** Conceptualization, Writing - review & editing, Supervision.

**Declaration of competing interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Acknowledgement**

The publication was supported by Beijing Key Laboratory of Indoor Air Quality Evaluation and Control.

**References**

Berg, M.K., Yu, Q., Salvador, C.E., Melani, I., Kitayama, S., 2020. Mandated Bacillus Calmette-Guérin (BCG) vaccination predicts flattened curves for the spread of COVID-19. medRxiv https://doi.org/10.1101/2020.04.05.20054163.

Luo, Wei, Majumder, Maxmina S., Liu, Dianduo, Poirier, Canelle, Mandl, Kenneth D., Lipstich, Marc, et al., 2020. The Role of Absolute Humidity on Transmission Rates of the COVID-19 Outbreak. https://doi.org/10.1101/2020.02.12.20022467.

Ma, Yueling, Zhao, Yadong, Liu, Jiangtao, He, Xiaotao, Wang, Bo, Fu, Shihua, et al., 2020. Effects of temperature variation and humidity on the death of COVID-19 in Wuhan, China. Sci. Total Environ. https://doi.org/10.1016/j.scitotenv.2020.138226.

Notari, A., 2020. Temperature Dependence of COVID-19 Transmission. Oliveiraos, Barbara, Caramelo, Liliana, Ferreira, Nuno C., Caramelo, F., 2020. Role of Temperature and Humidity in the Modulation of the Doubling Time of COVID-19 Cases. https://doi.org/10.1101/2020.03.05.20031872.

Qi, Hongchao, Xiao, Shuang, Shi, Runye, Ward, Michael P., Chen, Yue, Tu, Wei Luo, et al., 2020. COVID-19 Transmission in Mainland China Is Associated With Temperature and Humidity: A Time-series Analysis. https://doi.org/10.1101/2020.03.30.20044099.

Rahman, Md Ariful, Hossain, Md Golzar, Singh, Atul Chandra, Islam, Sayedul, Islam, A., 2020. A Retrospective Analysis of Influence of Environmental/Air Temperature and Relative Humidity on SARS-COV-2 Outbreak. Preprints. https://doi.org/10.20944/PREPRINTS202003.0325.V1.

Sajadi, M.M., Habibzadeh, P., Vintzileos, A., Shokouhi, S., Miralles-Wilhelm, F., Amoroso, A., 2020. Temperature, Humidity and Latitude Analysis to Predict Potential Spread and Seasonality for COVID-19. https://doi.org/10.1123/SSRN.3550308.

Sala, G., Miyakawa, T., 2020. Association of BCG vaccination policy with prevalence and mortality of COVID-19. medRxiv https://doi.org/10.1101/2020.03.30.20044165.

Shi, Peng, Dong, Yingqia, Yan, Huanchang, Li, Xiaoyang, Zhao, Chenkai, Liu, Wei Luo, et al., 2020. The Impact of Temperature and Absolute Humidity on the Coronavirus Disease 2019 (COVID-19) Outbreak - Evidence From China. https://doi.org/10.1101/2020.02.22.20038919.

Tripllett, M., 2020. Evidence that higher temperatures are associated with lower incidence of COVID-19 in pandemic state, cumulative cases reported up to March 27, 2020. medRxiv https://doi.org/10.1101/2020.04.02.20051524.

van Doremalen, Neeltje, Bushmaker, Trenton, Morris, Dylan H., Holbrook, Myndi G., Gamble, Amanda, Williamson, Brandi N., et al., 2020. Aerosol and surface stability of SARS-COV-2 as compared with SARS-COV-1. N. Engl. J. Med. https://doi.org/10.1056/NEJM204973.

Wang, Jingyuan, Tang, Ke, Feng, Kai, Lv, W., 2020a. High Temperature and High Humidity Reduce the Transmission of COVID-19. SSRN. https://doi.org/10.1101/2020.04.02.20051524.

Wang, Mao, Jiang, Alii, Gong, Liqian, Luo, Lina, Guo, Wenbin, Li, Chuyi, et al., 2020b. Temperature Significant Change COVID-19 Transmission in 429 Cities. https://doi.org/10.1101/2020.02.22.2005791.

Wu, Xiao, Nethery, Rachel C., S. B. M., Braun, Danielle, Dominici, F., 2020. Exposure to air pollution and COVID-19 mortality in the United States. medRxiv https://doi.org/10.1101/2020.04.05.20054502.

Zhu, Yongqian, Xie, J., 2020. Association between ambient temperature and COVID-19 infection in 122 cities from China. Sci. Total Environ. 724. https://doi.org/10.1016/j.scitotenv.2020.1382901.