ASIAN JOURNAL OF STATISTICS AND APPLICATIONS

VOL NO. 2, ISSUE. 2, PAGE NO. 65-83 YEAR 2025

ISSN: 3048-7455

https://DOI:10.47509/AJSA.2025.V02I02.05

Estimating Short-Term Air Pollution Mortality Effects via Functional Data Models with Temporal Scale Incongruence

GEORGE, Franciosalgeo[1,2], DEY, Sagnik[3,4], GHOSH, Santu[5,6]

[1] Research Scholar, Manipal Academy of Higher Education, Manipal, India
[2] Division of Biostatistics and Epidemiology, St John's Research Institute, St. John's National Academy of Health Sciences, Bengaluru 560034, India
[3] Centre for Atmospheric Sciences, Indian Institute of Technology Delhi, Delhi 110016, India
[4] Adjunct Faculty, Department of Health Policy and Management, Korea University, Seoul, Republic of Korea
[5] Department of Biostatistics, St John's Medical College, Bengaluru 560034, India
[6] Adjunct Faculty, Manipal Academy of Higher Education, Manipal, India

ABSTRACT

Background: Inconsistency in the temporal scale of exposure and outcome measurements is a major bottleneck in air pollution epidemiology, complicating causal inference from data collected by different regulatory authorities. Standard analytical methods often aggregate higher-resolution exposure data to match lower-resolution outcome data, which can obscure acute associations. In this study, we addressed this temporal incongruence by applying Functional Data Analysis models to improve the estimation of epidemiological associations.

Methods: We used data from Delhi, India (2013–2017), including daily counts of non-trauma, all-cause mortality and hourly ozone measurements. First, daily mortality was modelled using scalar-on-function regression (SoFR) on hourly ozone, adjusting for time, temperature, and relative humidity. Then, as a demonstration, daily mortality data were aggregated to monthly counts and modelled on daily average of ozone; estimates were compared against generalized additive models (GAMs) fitted to daily mortality and daily ozone.

Results: Hourly analyses revealed distinct diurnal patterns, with the strongest associations observed during the early afternoon. The largest effect was identified for the 7-day cumulative average at hour 14, corresponding to a 1.073% increase in mortality (95% CI: 0.672, 1.473) per 10 $\mu g/m^3$ increase in ozone. Stratified analyses showed higher estimates for males, particularly in the 45–64 and $\geq$ 65 year age groups. In monthly models, the SoFR estimates fell within the 95% confidence intervals of the GAM estimates obtained from daily mortality.

Conclusions: FDA models effectively overcome information loss caused by standard methods that aggregate data to align exposure and outcome scales. Our analysis identified a critical risk window for ozone-related mortality and showed the FDA's utility where only aggregated health data are available.

KEYWORDS

Functional Data, Scalar-on-Function Regression, Ozone exposure, Mortality.

Corresponding author. Email: santu.g@stjohns.in

1. Introduction

Air pollution is a major environmental risk factor for human health [1–3]. In 2021, air pollution was linked to more than one in every eight deaths worldwide, and 6% of pollution-related deaths were attributed to ozone (O_3) [2]. Short-term exposure to O_3 has been consistently linked to increases in daily mortality [4]. The health impacts of air pollution are particularly severe in low- and middle-income countries, where pollution levels are often higher and populations may be more vulnerable due to limited access to healthcare and other socioeconomic factors [2].

To study this relationship, researchers rely on datasets that provide measurements of air pollution and health outcomes over time. Typically, pollution data are collected at high temporal resolutions, such as hourly measurements. However, health outcomes, particularly mortality or morbidity data, are often reported at broader temporal resolutions, such as daily or monthly aggregates. This mismatch in temporal scales presents a significant challenge for accurately quantifying the health impacts of pollution [5].

The standard approach to address this issue has been to aggregate pollution data to match the temporal scale of the health data. In instances where health data are available only as monthly aggregates, as in Maji et al. [6], daily pollution measurements are often reduced to summary measures such as monthly averages or percentiles to align with monthly outcomes. While this simplifies the analysis, it introduces several limitations. Using summary measures can obscure critical information about short-term variations and peaks, which may have disproportionate health impacts [7]. As a result, the use of aggregated pollution data can lead to biased estimates, underestimating or overestimating the true health risks of pollution exposure.

Even when health outcomes are available at the daily scale, similar challenges arise if pollution data are reduced to daily averages. Many studies examining the short-term effects of air pollution on mortality use the daily average value of pollutant concentrations as the exposure variable. These daily summaries are then linked to daily counts of mortality using time-series models such as Poisson regression, often adjusting for seasonality, weather, and long-term trends. This approach has led to consistent findings across many settings and forms the basis for air quality regulations in countries around the world [4]. However, relying on daily average exposure can mask important within-day variations in pollution levels. For example, pollutant concentrations often peak during certain hours of the day—such as morning traffic hours or midday sunlight for ozone—and these time-specific exposures may have stronger health effects than others. Daily averages may smooth over these peaks, leading to exposure misclassification and potentially underestimating the true health risks [5, 8]. Additionally, the human body's vulnerability to air pollution fluctuates throughout the day due to circadian rhythms, daily activities, and environmental conditions. By using daily averages, traditional models assume a constant effect across all hours, which may not reflect biological reality [9, 10].

An alternative approach to overcome this limitation is the Functional Data Analysis (FDA) model, which allows researchers to retain the temporal granularity of pollution data while modelling its relationship with health outcomes [11, 12]. FDA treats pollutant concentrations as continuous functions over time rather than discrete

values, enabling a more nuanced analysis of their effects. This method has been used successfully in studies examining the relationship between hourly pollution data and daily health outcomes [7, 13]. By leveraging the flexibility of FDA, it becomes possible to analyse mismatched temporal scales without losing important information about short-term variations in pollution levels.

A limited number of studies from high-income countries have explored the use of FDA in environmental epidemiology. For instance, Arisido et al. [13] applied functional regression to investigate how hourly ozone exposure affected daily hospital admissions in Milan, revealing stronger associations at specific hours of the day. Similarly, Masselot et al. [7] used functional linear models to examine the impact of hourly temperature patterns on cardiovascular mortality in Montreal. Their approach helped uncover critical time-of-day effects—particularly during morning and evening hours—that were not evident using standard models.

Despite the availability of high-resolution hourly pollution data from multiple monitoring stations in Indian cities, this information is often underutilised in health studies. Most analyses rely on daily average values, which can miss important variations in pollution levels within a day. As a result, the specific hours when pollution may have the strongest impact on health might go unnoticed.

There are two common scenarios where FDA can provide substantial advantages in assessing the acute impact of air pollution on human health. First, when daily aggregated health outcome data are available alongside hourly pollution measurements, instead of averaging hourly values into a daily mean, FDA enables the full diurnal pattern of exposure to be retained. This allows for the identification of specific hours of the day when pollutant concentrations have the greatest impact on health. Second, in situations where only monthly aggregated health data are available, but daily pollution measurements exist, FDA can be used to link daily pollution curves with monthly outcomes, thereby preserving short-term variability that would otherwise be lost in monthly averages. In this study, we demonstrate the application of Functional Data Analysis (FDA) to estimate the short-term effects of air pollution on health. Specifically, we consider two scenarios: (i) associating hourly air pollution data with daily counts of health endpoints, and (ii) linking monthly aggregated counts of health outcomes to daily average air pollution levels.

2. Methodology

2.1. Mortality Data

We used daily all-cause mortality data for Delhi from January 1, 2013, to November 30, 2017. The data were obtained from the Municipal Corporation of Delhi and included details such as the date and place of death, age, sex, and residential address of the deceased. Although information on the cause of death was available, it was not classified using the international medical coding system and hence was not used in the analysis. Deaths due to accidents and suicides were clearly identified and excluded to focus on non-trauma, all-cause mortality.

2.2. Pollutant Data

Hourly O_3 data in Delhi were downloaded from the Central Control Room for Air Quality Management data repository (https://airquality.cpcb.gov.in/ccr/#/caaqm-dashboard-all/caaqm-landing/caaqm-data-repository). Data for the period from January 1, 2013, to November 30, 2017, were extracted for all available continuous ambient air quality monitoring stations located within the Delhi area. These stations are operated by the Central Pollution Control Board (CPCB), the Delhi Pollution Control Committee (DPCC), and other authorised agencies. To derive a city-wide exposure estimate, we computed the hourly average of pollutant concentrations across all stations reporting valid data at each time point.

2.3. Meteorological Data

Daily data on mean temperature, wind speed, solar radiation, and relative humidity (RH) for Delhi were downloaded from the online portal of the India Meteorological Department (http://dsp.imdpune.gov.in).

2.4. Statistical Methods

2.4.1. Functional Data Analysis Framework

We applied the Functional Data Model to examine the associations between mortality and time-varying pollution exposures. In this framework, ozone measurements (either hourly or daily) are treated as discrete observations of smooth underlying functions. These functions were estimated using thin-plate regression spline basis expansions, with smoothness controlled by a penalty term and smoothing parameters estimated by restricted maximum likelihood (REML) through a standard mixed-model framework [14].

2.4.2. Scenario 1: Hourly Ozone and Daily Mortality

Let $X_d(t)$ represent the concentration of a given pollutant at hour $t \in [0, 23]$ on day d, observed on a 24-hour grid. The observed hourly values $X_d(1)$, $X_d(2), .., X_d(24)$ were treated as evaluations of the function $X_d(t)$ on the 24-point grid. These functions entered the scalar-on-function regression (SoFR) model through the functional linear term $\int_t X_d(t)\beta(t)dt$.

Daily mortality was modelled using a log link under a quasi-likelihood framework, to allow for overdispersion. Assuming a Poisson distribution for Y_d, the model was specified as:

$$Y_d \sim \text{Poisson}(\mu_d), \text{ for } d = 1, \ldots, D$$

$$\begin{aligned} \log(\mu_d) = \alpha + \int_t X_d(t)\beta(t)\, dt + f_1(RH_d) + f_2(WS_d) \\ + f_3(SR_d) + f_4(Tmean_d) + f_5(Day_d) \end{aligned} \tag{1}$$

Here, Y_d is the daily non-trauma all-cause mortality on day d, and α is the intercept. $X_d(t)$ is the smooth hourly exposure function (for O3), $f_i(\ .)$ represent penalized smooth functions for relative humidity, wind speed, solar radiation, temperature, and time, estimated using thin-plate regression splines with smoothing parameters estimated by Restricted Maximum Likelihood (REML) estimation algorithm. $\beta(t)$ is a time-varying coefficient function capturing the hour-specific effects of pollution.

The coefficient function $\beta(t)$, which captures the hour-specific effects of ozone, was represented by a thin-plate regression spline basis expansion:

$$\beta(t) = \sum_{l=1}^{L} b_l \theta_l(t) \tag{2}$$

where $\theta_l(t)$ are also penalized-spline basis functions and b_l are coefficients estimated via penalized likelihood. Smoothness of $\beta(t)$ was controlled through a penalty parameter, which was estimated automatically within REML[15, 16].

2.4.3. Scenario 2: Daily Ozone and Monthly Mortality

To demonstrate the application of functional data analysis (FDA), we calculated the total monthly mortality over the study period, which span 59 months. This analysis combined daily ozone data with monthly mortality counts.

Let $X_m(t)$ denote the 24-hour average concentration on day t of month m. To capture intra-month variation, the daily ozone values were treated as evaluations of the functional predictor $X_m(t)$ across each month.

We then fit a SoF regression model to explore the relationship between monthly mortality and ozone functions, along with meteorological factors. The model is specified as follows:

$$Y_m \sim \text{Poisson}(\mu_m), \text{ for } m = 1, ..., M$$

$$\begin{aligned} \log(\mu_m) = \alpha &+ \int_t P_m(t)\eta(t)dt + \int_t RH_m(t)\theta(t)dt + \int_t Temp_m(t)\kappa(t)dt \\ &+ \int_t WS_m(t)\psi(t)dt + \int_t SR_m(t)\omega(t)dt \end{aligned} \tag{3}$$

Here, Y_m represents the mortality count for month m, and α is the intercept. $P_m(t)$, $RH_m(t)$, $Temp_m(t)$, $WS_m(t)$ and $SR_m(t)$ are the functional representations of daily pollutant concentrations, relative humidity, temperature, daily wind speed and solar radiation for month m, respectively. The coefficient functions $\eta(t)$, $\theta(t)$, $\kappa(t)$, $\psi(t)$ and $\omega(t)$ were modelled using a penalised spline, with smoothness parameters estimated by the REML algorithm as described in Scenario-1.

We explored both single-day lag effects (lag 0–6) and cumulative average exposures (2–7 day means) to capture delayed and prolonged impacts of ozone on mortality. For each lag or cumulative average, a separate SoFR model was fitted, and the corresponding time-varying coefficient functions $\beta(t)$ were estimated. The effects of O_3 on mortality

were considered significant with a p-value < 0.05 for a given exposure metric. To assess stability and comparability across models, we identified the most precise significant positive slope as the "best effect" for each exposure metric. These best estimates are reported in the results. The entire analysis was carried out under the R platform [17] with packages 'mgcv' [16] and 'refund' [15].

3. Results

3.1. Descriptives of Data

Figure 1 displays the distribution of hourly ozone concentrations over the study period. A clear diurnal pattern is evident: concentrations remain low during the night and early morning, rise sharply after 09:00, and peak between 13:00 and 15:00 before gradually declining after sunset. Across the study period, the 24-hour average daily ozone concentration ranged from 6.87 to 114.98 $\mu g/m^3$, with a median of 32.10 $\mu g/m^3$. Daily all-cause, non-trauma mortality counts ranged from 29 to 319 deaths per day, with a median of 194 deaths. A total of 354,963 deaths were recorded during the study period.

3.2. Estimates of $\beta(t)$ from SoFR: Hourly O_3 and Daily Mortality

Figure 2 presents the estimated hourly-varying effects $\beta(t)$ of ozone on daily mortality across lags 0–6. The effect curves exhibit notable time-of-day variation, with positive associations emerging during the midday to afternoon hours (around 12:00–15:00) at shorter lags, particularly at lags 1 and 2. At longer lags, the estimates tend to flatten and widen.

Figure 3 shows the results for cumulative averages of ozone exposure (2–7 days). The curves reveal more consistent diurnal patterns compared to single-day lags, with clear peaks during early afternoon and, in some cases, late evening hours. The strongest effects are generally observed between 12:00 and 15:00.

Table 1 summarizes the best hourly effect estimates $\beta(t)$ of ozone on daily mortality across single-day lags (0–6) and cumulative averages (2–7 days). For single-day lags, the strongest association was observed at lag 0 during hour 13, corresponding to a 0.45% (95% CI: 0.23–0.66) increase in mortality per 10 $\mu g/m^3$ increase in ozone. Significant effects persisted at lags 1 and 2, with peaks around hours 13–14, although the magnitude declined at longer lags. In contrast, cumulative averages showed progressively stronger associations, with the maximum effect observed for the 7-day average at hour 14, corresponding to a 1.07% (95% CI: 0.67–1.47) increase in mortality.

3.3. Gender and age-specific estimates of $\beta(t)$ from SoFR

Figure 4 and Figure 5 show the estimated hourly effect curves $\beta(t)$ of ozone on daily mortality stratified by gender. Across lags 0–3 (Figure 4), both males and females exhibited broadly similar temporal patterns, with effects fluctuating around the null during night and early morning hours and rising during midday and afternoon. At lag 1 and lag 2, more pronounced positive estimates were observed around 12:00–15:00 hours, with males generally showing slightly higher point estimates than females. When cumulative averages of exposure were considered (Figure 5), consistent peaks appeared during the midday hours across both sexes for 2–to 7-day averages.

Figure 6 and Figure 7 present the age-stratified estimates. For lags 0–2 (Figure 6), effect sizes were small among children aged 0–4 and individuals aged 5–44, with estimates fluctuating around zero across most hours. In contrast, clearer positive effects were observed in the 45–64 and ≥65 age groups, particularly during afternoon hours. At lag 1, individuals aged 5–44 also showed elevated point estimates around midday, while the ≥65 group exhibited sustained positive estimates into the evening. When cumulative averages were examined (Figure 7), stronger and more consistent associations emerged in the older age groups. Across 2–7-day averages, the 45–64 and ≥65 groups showed distinct peaks around midday (12:00–15:00), with the ≥65 group additionally displaying a secondary rise during late-night hours. The younger age groups (0–4 and 5–44) continued to show wide uncertainty bands, though some modest increases were noted during midday hours.

3.4. Estimates of $\beta(t)$ from SoFR: Daily O_3 and Monthly Mortality

Figure 8 presents the estimated coefficients for the impact of ozone on mortality (mortality being aggregated at the monthly level) using the scalar-on-function regression (SoFR) model, compared with estimates from a generalised additive model (GAM) with a quasi-Poisson family based on daily mortality and daily average O_3. The black curve shows the SoFR estimates across the days of the month, with the light grey band indicating the 95% confidence interval. The horizontal solid line denotes the coefficient of the 24-hour average O_3 on daily mortality counts, with its surrounding grey band representing the 95% confidence interval. In contrast, the dashed line indicates the null value. The maximum coefficient from the SoFR model (0.62; 95% CI: 0.04–1.20) lies within the 95% confidence interval of the GAM estimate (0.83; 95% CI: 0.31–1.35). This comparison demonstrates that the SoFR approach yields results consistent with those obtained from daily mortality data, underscoring the utility of FDA in settings where only monthly mortality information is available.

4. Discussion

This study applied the FDA to investigate the association between ozone exposure and mortality in Delhi between January 2013 and November 2017. Using SoFR models, we examined the time-varying effects of hourly ozone on daily mortality across single-day lags (0–6 days) and cumulative averages (2–7 days), while adjusting for meteorological and temporal confounders. We also evaluated effect modification by sex and age, and extended FDA to analyse daily 24-hour average ozone exposures against monthly mortality counts, validating the approach against standard GAMs using daily mortality counts.

Our results reveal distinct diurnal patterns in the association between ozone and mortality. Effects were strongest in the early afternoon hours (12:00–15:00), particularly at lags 1 and 2, with flatter and less precise estimates at longer lags. Cumulative averages (2–7 days) produced more stable and pronounced associations, again concentrated around midday. Stratified analyses showed broadly similar temporal patterns in both males and females, though males exhibited slightly sharper peaks. Age-stratified models indicated that associations were minimal in younger groups (0–4 and 5–44 years) but stronger in older adults (45–64 and ≥65 years), with clearer peaks around midday.

These findings are consistent with prior work highlighting the importance of intra-day variability in ozone exposure. For example, Arisido et al. [13] reported stronger

associations between hourly ozone and hospital admissions in Milan during afternoon hours, while Masselot et al. [7] observed time-of-day-specific temperature effects on cardiovascular mortality in Montreal that were not captured by daily averages. The midday ozone peaks observed in our study align with periods of maximum photochemical activity, driven by sunlight—a well-established mechanism of ozone formation [18]. Ozone levels in Delhi also show marked seasonal variation, with higher concentrations during the warmer months when stronger solar radiation and higher temperatures enhance photochemical production, and lower concentrations during winter and monsoon periods due to reduced sunlight and washout processes [19]. In addition, Delhi's ozone precursors—primarily nitrogen oxides (NOx) and volatile organic compounds (VOCs)—originate largely from traffic emissions, solvent use, small-scale industrial activity, and fuel combustion. Recent measurement studies indicate substantial ambient VOC concentrations consistent with these sources; for example, Mandal et al. [20] reported chemical signatures at urban Delhi sites that strongly point to vehicular exhaust and combustion processes. Together, these atmospheric and emission characteristics help explain the temporal patterns of ozone exposure observed in Delhi. By leveraging hourly data, the FDA uncovered patterns that would be masked by daily averages, highlighting critical time windows of risk.

A key contribution of this study is demonstrating how the FDA can address mismatched temporal scales. We examined two scenarios: (i) hourly ozone linked to daily mortality and (ii) daily ozone linked to monthly mortality. By modelling exposures as smooth functional curves, the FDA preserves intra-day and intra-month variability that would otherwise be lost to aggregation. In the monthly analysis, SoFR estimates were comparable to those from GAMs using daily mortality and fell within the GAM 95% confidence intervals. This consistency supports the robustness of the FDA when only coarser health data are available, a common limitation in many low- and middle-income settings.

Although we excluded deaths from accidents and suicides, reliance on all-cause mortality without cause-specific coding restricts inferences about underlying health pathways. Exposure misclassification remains possible, even when averaging across multiple stations. This limitation also reflects the use of a city-wide averaged exposure, which may not fully capture neighbourhood-level variability in ozone levels due to the uneven spatial distribution of monitoring stations across Delhi.

Since FDA can identify hours when ozone is most strongly associated with mortality, such information could eventually help guide sensitive groups to limit outdoor activity during elevated-risk periods. However, recommendations cannot rely on ozone concentrations alone, as health risks depend on integrated exposure rather than instantaneous peaks. In addition, any guidance must also account for the combined effects of multiple pollutants. In summary, this study demonstrates the utility of the FDA for air pollution epidemiology in two contexts: identifying diurnal windows of elevated risk from hourly ozone exposure and bridging mismatched temporal scales between daily exposures and monthly mortality. Our findings indicate that ozone-related mortality risks in Delhi are concentrated during early afternoon hours, particularly among older adults, and that the FDA provides reliable estimates even when only aggregated health outcome data are available. These insights underscore the importance of high-resolution exposure data and flexible analytical methods for advancing air pollution health research and guiding targeted public health interventions.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgement

SD acknowledges funding from the Clean Air Fund (Grant Number 001307) and the Fellowship of the Vipula and Mahesh Chaturvedi Chair in Policy Studies.

Conflict interest

All other 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.

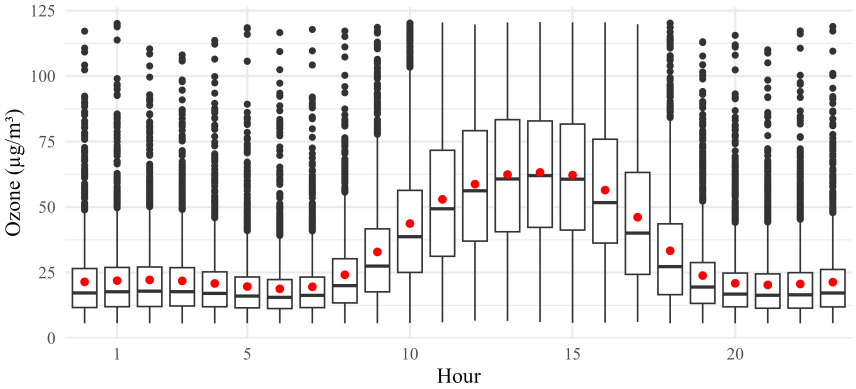

Figure 1. Boxplots of ozone concentrations by hour.

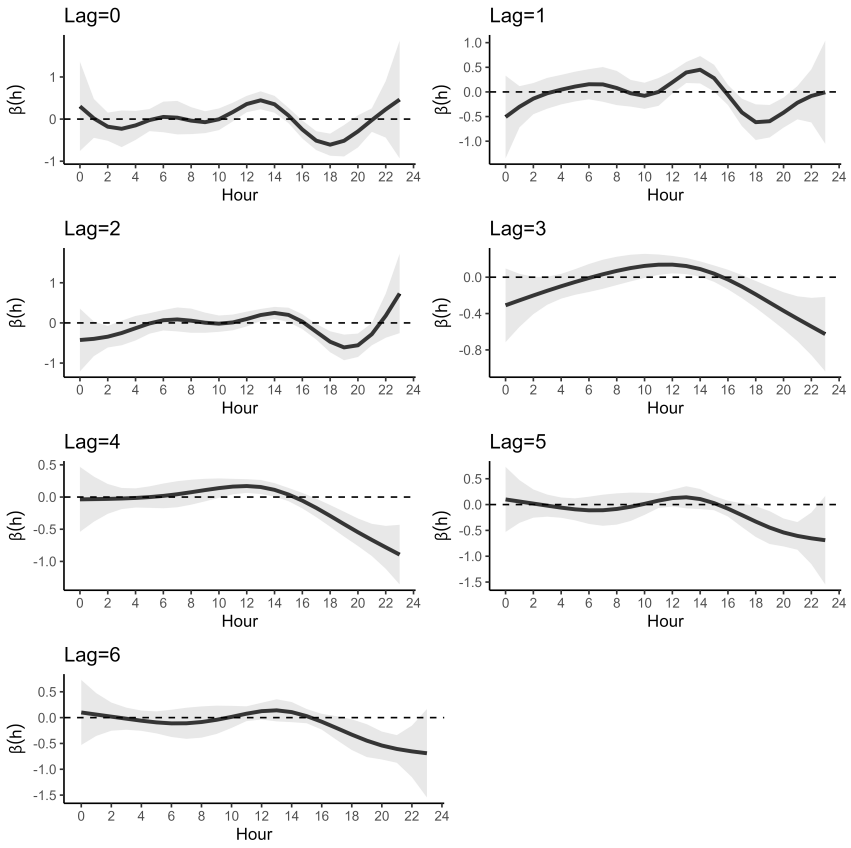

Figure 2. Hourly-varying effect estimates $\hat{\beta}(t)$ of O_3 on daily mortality across lags (Lag 0–6). Values are in % increase in mortality for a 10-unit increase of pollutant.

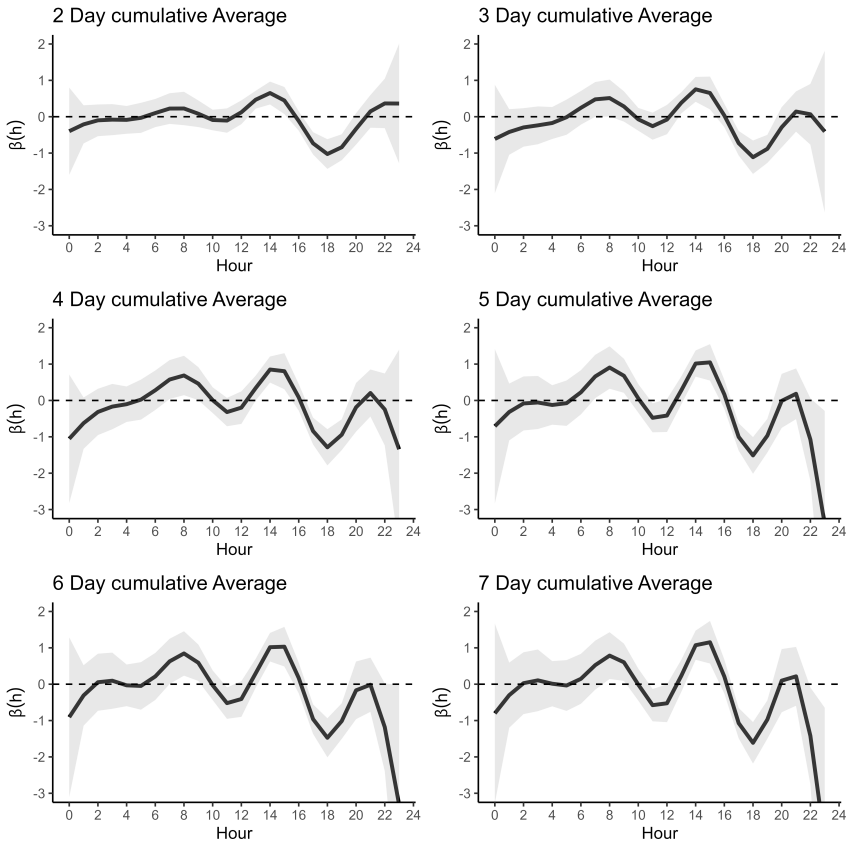

Figure 3. Hourly-varying effect estimates $\hat{\beta}(t)$ of O_3 on daily mortality across different cumulative averages (2-day to 7-day cumulative average). Values are in % increase in mortality for a 10-unit increase of pollutant.

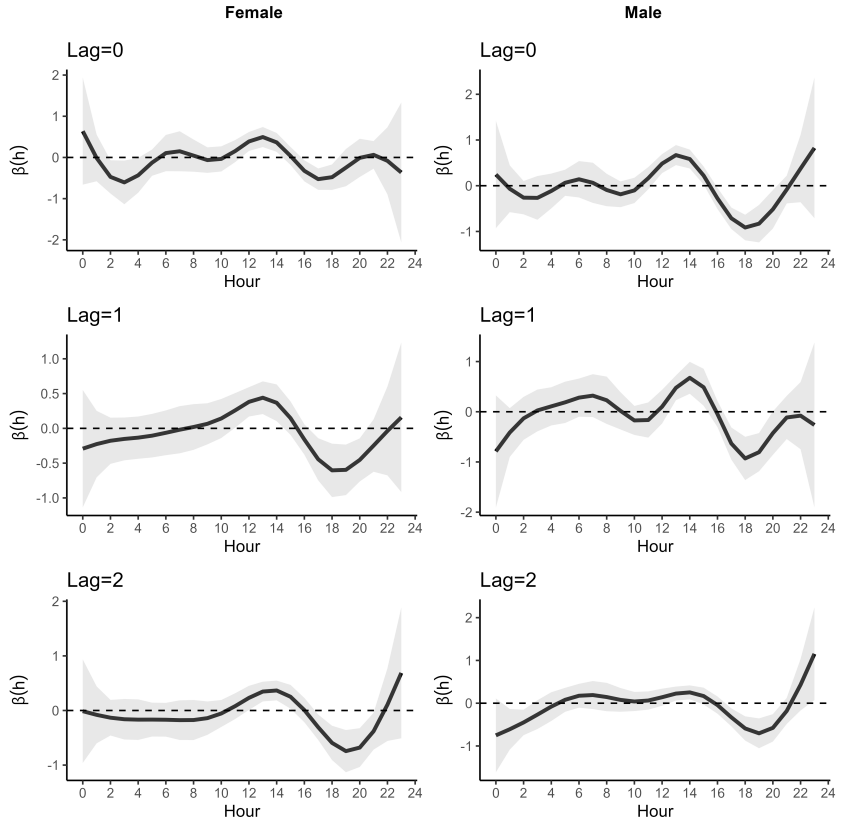

Figure 4. Hourly-varying effect estimates $\hat{\beta}(t)$ of O_3 on daily mortality across lags (Lag 0–2) stratified by gender. Values are in % increase in mortality for a 10-unit increase of pollutant.

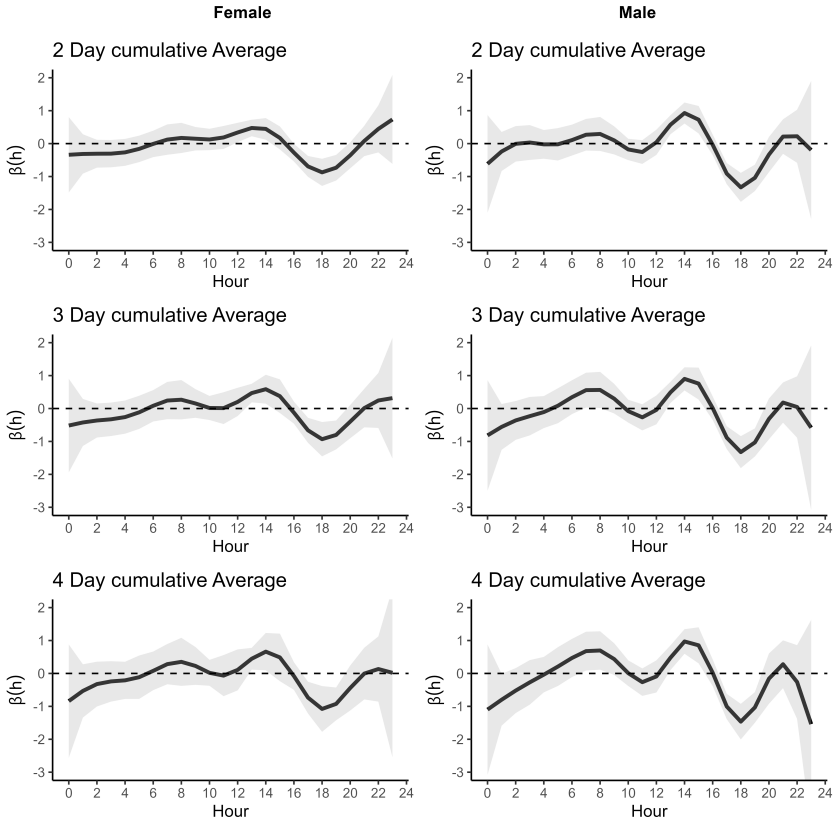

Figure 5. Hourly-varying effect estimates $\hat{\beta}(t)$ of O_3 on daily mortality across different cumulative averages (2-day–4-day cumulative average) stratified by gender. Values are in % increase in mortality for a 10-unit increase of pollutant.

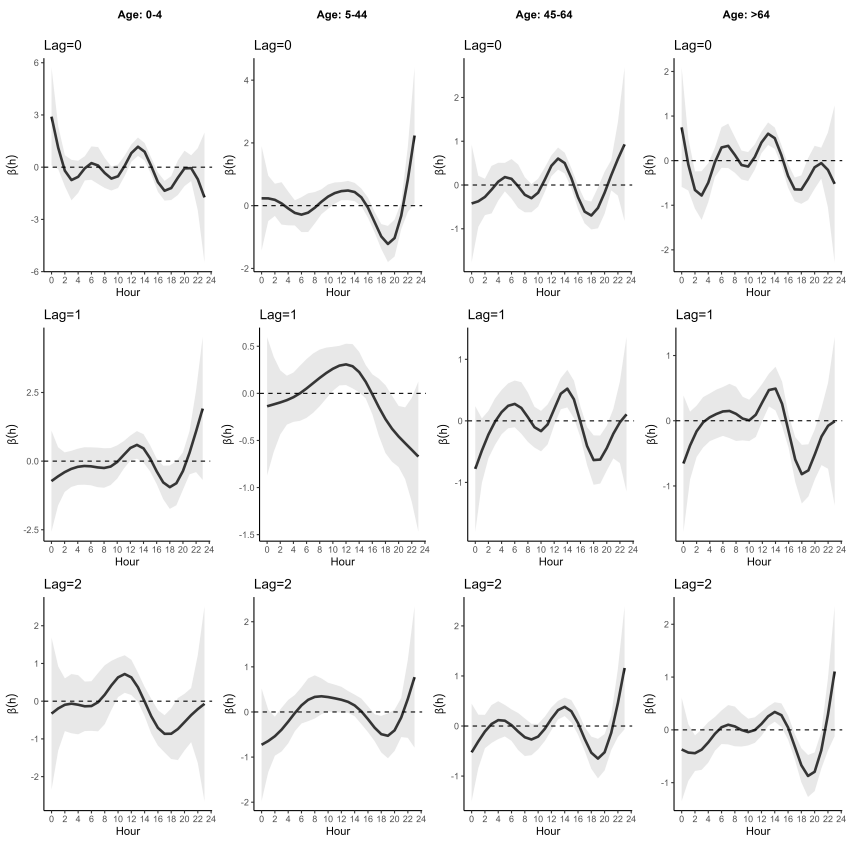

Figure 6. Hourly-varying effect estimates $\hat{\beta}(t)$ of O_3 on daily mortality across lags (Lag 0–2) stratified by age. Values are in % increase in mortality for a 10-unit increase of pollutant.

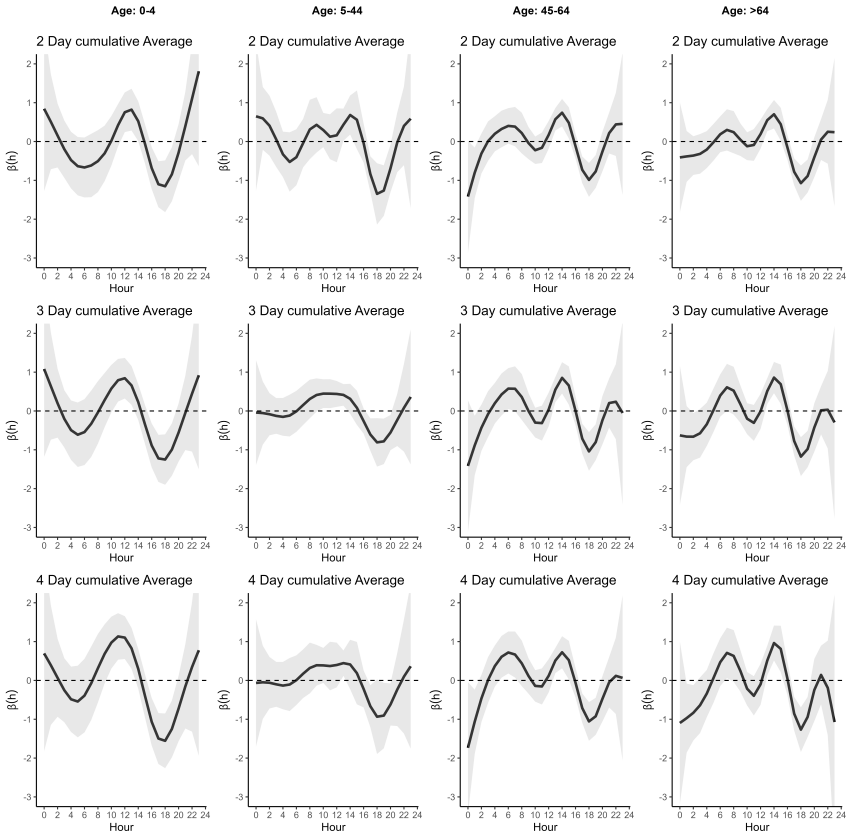

Figure 7. Hourly-varying effect estimates $\hat{\beta}(t)$ of O_3 on daily mortality across different cumulative average (2 day–4-day cumulative average) stratified by age. Values are in % increase in mortality for a 10-unit increase of pollutant.

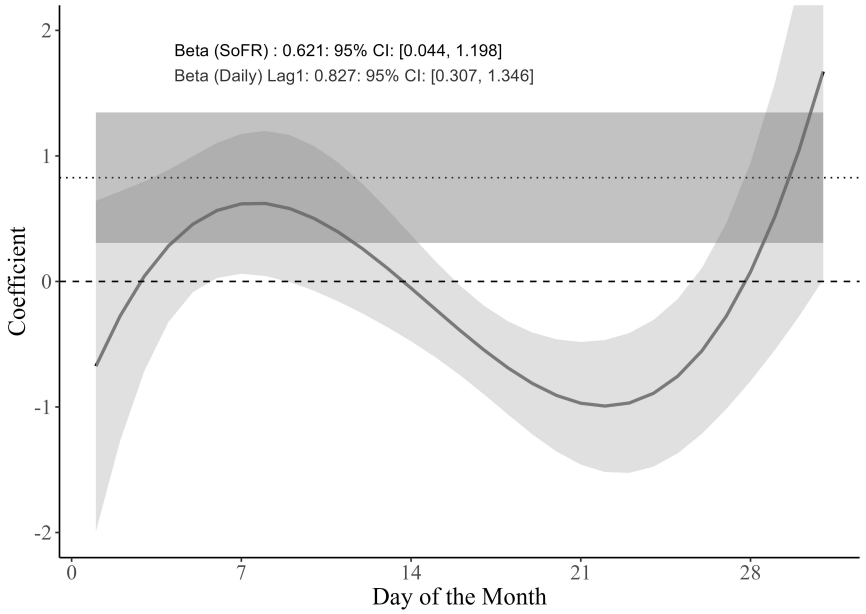

Figure 8. Estimates of $\beta_t(h)$ of Ozone using functional generalized linear model (Values are in % increase in mortality for 10-unit increase of pollutant).

The black curve represents the estimated coefficients ($\beta_t(h)$) from the scalar-on-function regression (SoFR) model for Ozone across the days of the month, with the surrounding light grey shaded area indicating the 95% confidence interval. The horizontal solid line indicates the coefficient estimate from a generalized additive model (GAM) using daily mortality and pollution data, and the medium grey band surrounding it shows its 95% confidence interval. The horizontal dashed line represents the null value ($\beta = 0$) for reference. All estimates are expressed as the percent increase in mortality per 10-unit increase in Ozone concentration.

Lags			Cumulative averages		
Lag	Hr	Best $\hat{\beta}(t)$ (95% CI)	Avg	Hr	Best $\hat{\beta}(t)$ (95% CI)
Lag 0	13	0.446 (0.233, 0.659)			
Lag 1	13	0.395 (0.182, 0.608)	2 days	14	0.651 (0.337, 0.965)
Lag 2	14	0.249 (0.1, 0.398)	3 days	14	0.753 (0.412, 1.095)
Lag 3	12	0.138 (0.042, 0.234)	4 days	14	0.851 (0.493, 1.208)
Lag 4	12	0.171 (0.065, 0.277)	5 days	14	1.015 (0.656, 1.374)
Lag 5	12	0.125 (-0.037, 0.287)	6 days	14	1.018 (0.625, 1.41)
Lag 6	12	0.125 (-0.037, 0.287)	7 days	14	1.073 (0.672, 1.473)

Table 1. Best Hourly Effect Estimates $\beta(t)$ (Hour of Minimum Coefficient of Variation) for Ozone on Daily Mortality Across Lag Days cumulative averages (with 95% Confidence Intervals) (Values are in % increase in mortality for 10-unit increase of pollutant).

References

[1] Patrick E. Brown, Yurie Izawa, Kalpana Balakrishnan, Sze Hang Fu, Joy Chakma, Geetha Menon, Rajesh Dikshit, R. S. Dhaliwal, Peter S. Rodriguez, Guowen Huang, Rehana Begum, Howard Hu, George D'souza, Randeep Guleria, and Prabhat Jha. Mortality Associated with Ambient PM2:5 Exposure in India: Results from the Million Death Study. *Environmental Health Perspectives*, 130(9):097004–1, 2022. ISSN 15529924. .

[2] Health Effects Institute. State of Global Air 2024. Special Report. Technical report, Boston, MA:Health Effects Institute, 2024. URL https://www.stateofglobalair.org/resources/report/state-global-air-report-2024.

[3] World Health Organization. Ambient Air Pollution- A Global Assessment of Exposure and Burden of Disease ISBN: 9789241511353 (2016). 2016.

[4] Rongqi Abbie Liu, Yaguang Wei, Xinye Qiu, Anna Kosheleva, and Joel D. Schwartz. Short term exposure to air pollution and mortality in the US: a double negative control analysis. *Environmental Health*, 21(1):81, 9 2022. ISSN 1476-069X. . URL https://doi.org/10.1186/s12940-022-00886-4https://ehjournal.biomedcentral.com/articles/10.1186/s12940-022-00886-4.

[5] Roger D. Peng and Michelle L. Bell. Spatial misalignment in time series studies of air pollution and health data. *Biostatistics*, 11(4):720–740, 2010. ISSN 14654644. .

[6] Sanjoy Maji, Santu Ghosh, and Sirajuddin Ahmed. Association of air quality with respiratory and cardiovascular morbidity rate in Delhi, India. *International Journal of Environmental Health Research*, 28(5):471–490, 2018. ISSN 13691619. . URL https://doi.org/10.1080/09603123.2018.1487045.

[7] Pierre Masselot, Fateh Chebana, Taha B.M.J. Ouarda, Diane Bélanger, André St-Hilaire, and Pierre Gosselin. A new look at weather-related health impacts through functional regression. *Scientific Reports*, 8(1):1–9, 2018. ISSN 20452322. .

[8] Anqi Jiao, Qianqian Xiang, Zan Ding, Jiguo Cao, Hung Chak Ho, Dieyi Chen, Jian Cheng, Zhiming Yang, Faxue Zhang, Yong Yu, and Yunquan Zhang. Short-term impacts of ambient fine particulate matter on emergency department visits: Comparative analysis of three exposure metrics. *Chemosphere*, 241:125012, 2 2020. ISSN 00456535. . URL https://doi.org/10.1016/j.chemosphere.2019.125012https://linkinghub.elsevier.com/retrieve/pii/S0045653519322519.

[9] Mascia Benedusi, Elena Frigato, Cristiano Bertolucci, and Giuseppe Valacchi. Circadian Deregulation as Possible New Player in Pollution-Induced Tissue Damage. *Atmosphere*, 12 (1):116, 1 2021. ISSN 2073-4433. . URL https://www.mdpi.com/2073-4433/12/1/116.

[10] Jacqueline M. Leung and Micaela E. Martinez. Circadian Rhythms in Environmental

Health Sciences. *Current Environmental Health Reports*, 7(3):272–281, 9 2020. ISSN 2196-5412. . URL https://link.springer.com/10.1007/s40572-020-00285-2.

[11] J.O. Ramsay and B.W. Silvermann. *Functional Data Analysis. Springer Series in Statistics*, volume 40. 2005. ISBN 9780387400808. .

[12] J.O Ramsay, Giles Hooker, and Spencer Graves. Functional data analysis with R and MATLAB- James Ramsay, Giles Hooker, Spencer Graves (auth.)-Springer-Verlag New York (2009), 2009.

[13] Maeregu Woldeyes Arisido. Functional measure of ozone exposure to model short-term health effects. *Environmetrics*, 27(5):306–317, 2016. ISSN 1099095X. .

[14] David Ruppert, M. P. Wand, and R. J. Carroll. *Semiparametric Regression*. Cambridge University Press, 7 2003. ISBN 9780521785167. . URL https://www.cambridge.org/core/product/identifier/9780511755453/type/book.

[15] Jeff Goldsmith, Fabian Scheipl, Lei Huang, Julia Wrobel, Chongzhi Di, Jonathan Gellar, Jaroslaw Harezlak, Mathew W. McLean, Bruce Swihart, Luo Xiao, Ciprian Crainiceanu, Philip T. Reiss, and Erjia Cui. refund: Regression with Functional Data, 8 2010. URL https://cran.r-project.org/package=refund.

[16] Simon N. Wood. *Generalized Additive Models*, volume 1. Chapman and Hall/CRC, 5 2017. ISBN 9781315370279. . URL https://projecteuclid.org/journals/statistical-science/volume-1/issue-3/Generalized-Additive-Models-Rejoinder/10.1214/ss/1177013609.fullhttps://www.taylorfrancis.com/books/9781498728348.

[17] R Core Team. R: A Language and Environment for Statistical Computing, 2023. URL https://www.r-project.org/.

[18] Cheng He, Xiao Lu, Haolin Wang, Haichao Wang, Yan Li, Guowen He, Yuanping He, Yurun Wang, Youlang Zhang, Yiming Liu, Qi Fan, and Shaojia Fan. The unexpected high frequency of nocturnal surface ozone enhancement events over China: characteristics and mechanisms. *Atmospheric Chemistry and Physics*, 22(23):15243–15261, 11 2022. ISSN 1680-7324. . URL https://acp.copernicus.org/articles/22/15243/2022/.

[19] Sachin D. Ghude, S. L. Jain, B. C. Arya, G. Beig, Y. N. Ahammed, Arun Kumar, and B. Tyagi. Ozone in ambient air at a tropical megacity, Delhi: characteristics, trends and cumulative ozone exposure indices. *Journal of Atmospheric Chemistry*, 60(3):237–252, 7 2008. ISSN 0167-7764. . URL http://link.springer.com/10.1007/s10874-009-9119-4.

[20] T.K. Mandal, Pooja Yadav, Mukesh Kumar, Shyam Lal, Kirti Soni, Lokesh Yadav, Ummed Singh Saharan, and S.K. Sharma. Characteristics of volatile organic compounds (VOCs) at an urban site of Delhi, India: Diurnal and seasonal variation, sources apportionment. *Urban Climate*, 49:101545, 5 2023. ISSN 22120955. . URL https://linkinghub.elsevier.com/retrieve/pii/S2212095523001396.