# Investigating the associations among environmental stressors and their links to COVID-19 incidence

## Research in Germany and Brazil from 2020 to 2022

Doctoral thesis

to obtain a doctorate (PhD)

from the Faculty of Medicine

of the University of Bonn

**Leona Hoffmann**

from Gütersloh

2026

Written with authorization of
the Faculty of Medicine of the University of Bonn

First reviewer: Prof. Dr. Matthias Schmid

Second reviewer: PD Dr. Sabine Wüst

Day of oral examination: 10.04.2026

From the Institute of Medical Biometry, Informatics and Epidemiology

This dissertation is dedicated to Professor Jörn Rittweger, whose high expectations and tireless support both challenged and encouraged me. Through his critical comments and thought-provoking questions, he consistently pushed me to my limits and inspired me to grow beyond them. His influence will always resonate in this work and throughout my future path. I am deeply grateful for having had such a thoughtful and inspiring academic supervisor and mentor.

**Table of Contents**

## List of Abbreviations

| | |
|---|---|
| AOK | Allgemeine Ortskrankenkasse |
| BW | Baden-Württemberg |
| COVID-19 | Corona Virus Disease 2019 |
| DE | Germany |
| DLNM | Distributed Lag Nonlinear Model |
| GAM | Generalized Additive Model |
| LISA | Local Indicators of Spatial Association |
| MRF | Markov Random Field |
| $NO_2$ | Nitrogen Dioxide |
| $O_3$ | Ozone |
| $PM_{2.5}$ | Fine Particulate Matter with a diameter of 2.5 µm or smaller |
| $PM_{10}$ | Fine Particulate Matter with a diameter of 10 µm or smaller |
| Prec | Precipitation |
| SARS-CoV-2 | Severe Acute Respiratory Syndrome Coronavirus 2 |
| RHumid | Relative Humidity |
| RR | Relative Risk |
| Temp | Temperature |
| UV | Ultraviolet Radiation |
| VP | Vapor Pressure |
| WHO | World Health Organization |

## 1. Abstract

**Background:** The COVID-19 pandemic has highlighted the importance of environmental stressors in influencing viral transmission. While existing research has primarily focused on the short-term effects of meteorology and air pollution, there is less evidence on how these factors relate to the occurrence of COVID-19 in different climates. This dissertation addresses this gap by examining the associations between environmental stressors and COVID-19 incidence in Germany and Brazil over a three-year period, covering both rural and urban areas.

**Methods:** This work is based on three peer-reviewed papers. The first research paper of this work analyzed the interdependencies among environmental stressors in Baden-Württemberg, Germany, using daily data from 2010 to 2018. The second research paper examined the associations between specific stressors and COVID-19 incidence in the same region during the pandemic (2020-2022) using generalized additive models (GAM). In the third research paper, the analysis was expanded to Brazil, incorporating relative humidity and applying both GAM and distributed lag nonlinear models (DLNM) to monthly data from 2020 to 2022.

**Results:** Temperature and $PM_{2.5}$ were identified as the most consistently influential environmental stressors across different regions, although their impacts varied by pandemic phase and climatic conditions. In temperate Germany, lower temperatures were associated with higher COVID-19 incidence, while in certain climate zones of Brazil higher temperatures and $PM_{2.5}$ linked with increased COVID-19 case numbers.

**Discussion:** The findings highlight the complex and context-dependent associations between environmental stressors and COVID-19 incidence. This research enhances our understanding of the environmental determinants of COVID-19 and highlights the need for studies that are more diverse in terms of geography and climate. Further interdisciplinary research is needed to better understand the causal links between environmental stressors and disease incidence, which is crucial for managing future outbreaks and addressing public health risks in a changing climate.

## 2. Introduction

The Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) and the related Corona Virus Disease 2019 (COVID-19) have caused the most significant pandemic of the 21st century with profound local and global consequences. The World Health Organization (WHO) (2023) reported 729 million confirmed COVID-19 cases and 6.7 million deaths worldwide from 2020 to 2022. More than five years after the WHO classified the pandemic, it is still essential that scientific research improves our understanding of the virus transmission and its interactions with the environment. This knowledge is vital for learning from the current pandemic experience, enhancing preparedness for future outbreaks, and informing evidence-based policy decisions.

### 2.1 Background

Brazil and Germany represent contrasting climatic, geographical, social, economic, and political conditions, offering a valuable basis for a comparative analysis. The COVID-19 pandemic negatively impacted both countries. Brazil reported its first case on February 26, 2020 (Giovanetti et al., 2022), one day before Germany (Schilling et al., 2021). While incidence rates varied, the virus's spread and its mutations followed similar patterns. Particularly the Omicron variant became dominant in Germany and Brazil in January 2022. Numerous factors, including environmental conditions, population density, mobility patterns, healthcare capacity, government policies, socioeconomic factors, and public behavior have influenced the virus. A considerable factor was the association between COVID-19 incidence and environmental stressors, including air pollution and meteorological conditions. Understanding how the COVID-19 disease interacts with environmental stressors across climatic, geographical and socio-political contexts is essential.

To evaluate vaccination progress (Figure 1A), the number of people fully vaccinated is divided by the total population. The data originate from Mathieu et al. (2021) and the World Health Organization (WHO) (2020), with major processing by Mathieu et al. (2020). In 2021, Brazil's vaccination rate was about two months behind Germany's vaccination rate. Germany achieved a vaccination rate of 50 % in August 2021 (56.91 %), Brazil in October 2021 (49.57 %). By March 2022, Brazil surpassed Germany in vaccination rates. In

December 2022, 82.99 % of Brazil's population was fully vaccinated, compared to 75.57 % in Germany. Both countries exceeded the European average of 66.01 % and the global average of 63.48 % (Mathieu et al., 2020). Brazil and Germany implemented various measures to contain the virus. The Stringency Index is a composite measure that uses nine response indicators, such as school closures, workplace closures, and travel bans, converting them to a value between 0 and 100 and originates from Hale et al. (2021) with minor processing by Mathieu et al. (2020). Figure 1B shows the Stringency Index over the three-year period for both countries. Germany and Brazil experienced local peaks in their Stringency Indexes in March 2020, at 76.85 and 75.00, respectively. These values were later surpassed, with Germany reaching its highest level of 83.51 in January 2021 and Brazil reaching 80.54 in May 2020. Over time, Germany's index decreased in distinct steps to 11.11, while Brazil’s declined more steadily to 22.22 by December 2022. Both countries showed Stringency Index values in comparable ranges. The incidence rates of new infections per 100,000 people per year differed between Brazil and Germany, see Figure 1C. For comparability, these rates are shown quarterly. Baden-Württemberg (BW), Germany, consistently reported higher figures than Brazil. Both study areas peaked in Q1 2022, with Germany's incidence being 9.6 times higher. However, differences in testing behaviors and availability across both countries limit direct comparisons, as many cases remain unreported (Martins et al., 2023; Vandenberg et al., 2021). In Brazil, the number of tests conducted was remarkably lower than in Germany (Hasell et al., 2020).

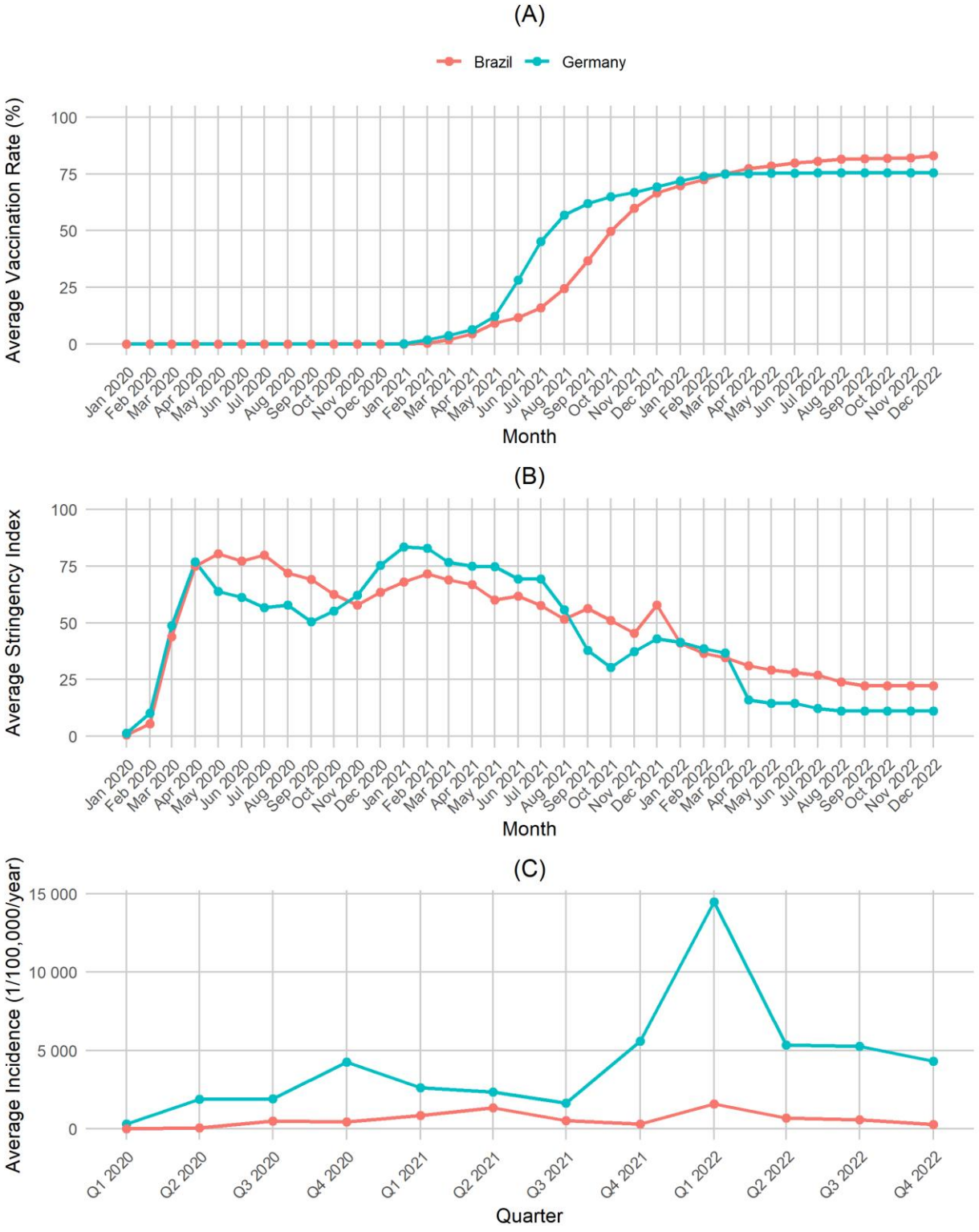


**Figure 1:** Comparison of the average monthly vaccination rate (%) (A), average monthly Stringency Index (B), and average quarterly COVID-19 incidence (1/100,000/year) (C) over the three year period for Brazil and Baden-Württemberg (BW), Germany. The incidence data for BW is sourced from Hoffmann et al. (2025a), while the data for Brazil comes from Hoffmann et al. (2025b). The data of the Stringency Index originates from

Hale et al. (2021) with minor processing by Mathieu et al. (2020) and the vaccination rate originates from Mathieu et al. (2021) and the World Health Organization (WHO) (2020) with major processing by Mathieu et al. (2020). In Germany, vaccination data and stringency index are available on national level, while incidence data is specific to the federal state of BW, as reported in (Hoffmann et al., 2025a).

## 2.2 Objectives and Research Strategy

The main objective of this work was to explore different perspectives on the associations between environmental stressors and COVID-19 incidence. Hoffmann et al. (2024) examined the interdependencies among several environmental stressors, including temperature (Temp), ultraviolet radiation (UV), precipitation (Prec), vapor pressure (VP), ozone ($O_3$), nitrogen dioxide ($NO_2$), and fine particulate matter with a diameter of 2.5 and 10 µm or smaller ($PM_{2.5}$, $PM_{10}$). This analysis identified key factors for further examination. Building on these insights, Hoffmann et al. (2025a) investigated the associations between the environmental stressors Temp, Prec, $PM_{2.5}$, $NO_2$, $O_3$, and COVID-19 incidence in BW, Germany. Hoffmann et al. (2025b) extended the analysis to Brazil, a country with a distinct climatic and socio-political context, to determine whether these associations remained consistent across regions or depended on the environmental conditions. Overall, this research has deepened our understanding of the virus transmission and its environmental determinants.

Figure 2 provides an overview of the similarities and differences of the three research articles. Hoffmann et al. (2024) and Hoffmann et al. (2025a) focused on the same region and spatial aggregation. They differed in temporal resolution: Hoffmann et al. (2024) analyzed daily data from 2010 to 2018, while Hoffmann et al. (2025a) examined quarterly data from 2020 to 2022 related to the COVID-19 pandemic. Hoffmann et al. (2025b) used monthly data aggregation at the municipal level from 2020 to 2022 across Brazil. Health data limited temporal aggregation in Germany, while $PM_{2.5}$ data availability restricted it in Brazil.

During the research process, several statistical approaches were explored to develop an explanatory model that balances interpretability and minimal complexity while ensuring clarity and practical relevance. This aligns with the philosophical and scientific principle of Occam’s Razor, which states that simpler theories are preferred over more complex ones (Sober, 2015). Various statistical methods were explored but were ultimately not included

in the final manuscripts. Principal Component Analysis was dismissed due to its lack of transparency regarding the effects of individual environmental stressors (James et al., 2013). Interaction terms tested in Hoffmann et al. (2025a) were not retained because they did not improve the model and complicated interpretation. Machine learning methods, such as Random Forests and Neural Networks, were discussed with experts on a conference but not pursued, as the priority was to provide clear insights into the links between environmental stressors and COVID-19 incidence (Hastie et al., 2009). For Hoffmann et al. (2025a) and Hoffmann et al. (2025b), a generalized additive model (GAM) was chosen for its flexibility in capturing nonlinear associations while maintaining interpretability (Wood, 2017). The model is well established in the literature for COVID-19 research (Nazia et al., 2022; Zheng et al., 2023). A certain level of complexity was retained to ensure a robust explanatory model. A Markov Random Field (MRF) is used to model spatial dependencies, natural and penalized splines are applied to capture nonlinear associations, and an offset term is included for population adjustment (Wood, 2017). The models evolved throughout the research process. For Hoffmann et al. (2025b), natural splines rather than penalized splines were identified as the preferred choice for consistent results in Brazil. Precipitation from Hoffmann et al. (2024) and Hoffmann et al. (2025a) was replaced by relative humidity (RHumid) in Hoffmann et al. (2025b), reflecting its significant impact on well-being and the associations of the parameters on tropical and subtropical climate (Auler et al., 2020; Prata et al., 2021). A Distributed Lag Nonlinear Model (DLNM) was added in Hoffmann et al. (2025b) to assess time-lagged effects, and improve the understanding of the associations between environmental stressors and health outcomes. Spatial and temporal analyses were emphasized across all three papers to ensure comprehensive modeling of both dimensions.

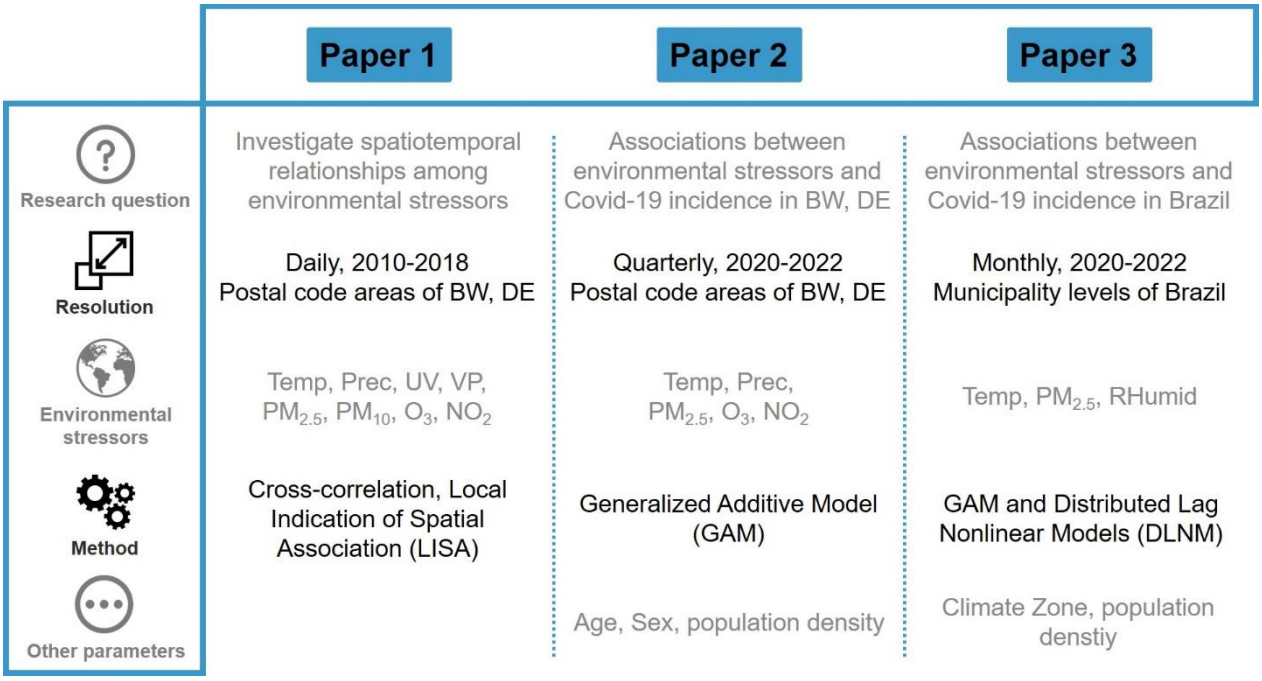


**Figure 2:** Overview of the three scientific papers of the dissertation. Key features are outlined, including the research question, resolution, examined environmental stressors, employed methods, and other relevant parameters.

## 2.3 Added value of this work

At the beginning of the pandemic, the limited available knowledge created a pressing need for insights to save lives and contain the virus. Tan et al. (2023) demonstrated that this urgency often compromised the quality of scientific articles on the impact of weather conditions on the spread of COVID-19, leading to problems with the reliability of their findings. In 2020, peer-review processes were much faster than in 2021, with the most cited papers accepted in an average of just 10 days, while studies had an average observation period of 123 days (Tan et al., 2023). Research during this time mainly focused on urban areas (Houweling et al., 2024; Marquès et al., 2022) and often relied on single-pollutant models, such as $PM_{2.5}$, neglecting other air pollutants (Sheppard et al., 2023). Additionally, the emphasis was frequently on particular vulnerable groups, such as the elderly and children (Monoson et al., 2023). To overcome these limitations, this work focused on a three-year observation period, including the total population of our study areas. A comprehensive analysis of both rural and urban areas was performed to provide a clearer understanding of the associations between environmental stressors and COVID-19 incidence. Including multiple stressors in a single model helps prevent bias from their

intercorrelations (Hoffmann et al., 2024). By integrating health data with satellite-based environmental data, which provide consistent, area-wide, and homogeneous spatial coverage, a robust nationwide data foundation is established (Holloway et al., 2021), overcoming the limitations of sparse, unevenly distributed ground-based monitoring stations and enhancing the reliability and impact of our findings. This research uniquely explores the effects during the Omicron and pre-Omicron periods, contrasting the different influences of environmental determinations on COVID-19 incidence. Furthermore, to our knowledge, this is the first systematic investigation of the associations between COVID-19 incidence and environmental stressors in BW and Brazil from 2020 to 2022, accounting for relevant confounding variables. Our comprehensive approach provides new insights into COVID-19 transmission dynamics and guides evidence-based public health measures.

## 2.4 References

Auler, A, Cássaro, F, Da Silva, V, & Pires, L. Evidence that high temperatures and intermediate relative humidity might favor the spread of COVID-19 in tropical climate: A case study for the most affected Brazilian cities. In: Science of the Total Environment 2020; 729: 139090

Giovanetti, M, Slavov, SN, Fonseca, V, Wilkinson, E, Tegally, H, Patané, JSL, Viala, VL, San, EJ, Rodrigues, ES, & Santos, EV. Genomic epidemiology of the SARS-CoV-2 epidemic in Brazil. In: Nature Microbiology 2022; 7(9): 1490-1500

Hale, T, Angrist, N, Goldszmidt, R, Kira, B, Petherick, A, Phillips, T, Webster, S, Cameron-Blake, E, Hallas, L, & Majumdar, S. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). In: Nature human behaviour 2021; 5(4): 529-538

Hasell, J, Mathieu, E, Beltekian, D, Macdonald, B, Giattino, C, Ortiz-Ospina, E, Roser, M, & Ritchie, H. A cross-country database of COVID-19 testing. In: Scientific data 2020; 7(1): 345

Hastie, T, Tibshirani, R, Friedman, JH, & Friedman, JH. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. New York: Springer, 2009

Hoffmann, L, Gilardi, L, Antoni, T, Baltruweit, M, Bittner, M, Breitner, S, Dally, S, Erbertseder, T, Hawighorst-Knapstein, S, Schmitz, MT, Schneider, R, Wüst, S, & Rittweger, J. Modulation of COVID-19 incidence by environmental stressors is variant between pre-Omicron and Omicron periods. In: Scientific Reports 2025a; 15(1): 27636

Hoffmann, L, Gilardi, L, Schmitz, M-T, Erbertseder, T, Bittner, M, Wüst, S, Schmid, M, & Rittweger, J. Investigating the spatiotemporal associations between meteorological conditions and air pollution in the federal state Baden-Württemberg (Germany). In: Scientific Reports 2024; 14(1): 5997

Hoffmann, L, Jacobson, L, Erbertseder, T, Berger, M, Rittweger, J, & Schneider, R. Associations between COVID-19 Incidence and Environmental Stressors in Brazil: A Nationwide Study from 2020 to 2022. In: Frontiers in Environmental Health 2025b; 4: 1635503

Holloway, T, Miller, D, Anenberg, S, Diao, M, Duncan, B, Fiore, AM, Henze, DK, Hess, J, Kinney, PL, & Liu, Y. Satellite monitoring for air quality and health. In: Annual review of biomedical data science 2021; 4(1): 417-447

Houweling, L, Maitland-Van der Zee, A-H, Holtjer, JC, Bazdar, S, Vermeulen, RC, Downward, GS, & Bloemsma, LD. The effect of the urban exposome on COVID-19 health outcomes: a systematic review and meta-analysis. In: Environmental research 2024; 240: 117351

James, G, Witten, D, Hastie, T, & Tibshirani, R. An introduction to statistical learning. Springer, 2013

Marquès, M, & Domingo, JL. Positive association between outdoor air pollution and the incidence and severity of COVID-19. A review of the recent scientific evidences. In: Environmental research 2022; 203: 111930

Martins, JP, Siqueira, BA, Sansone, NMS, & Marson, FAL. COVID-19 in Brazil: a three-year update. In: Diagnostic Microbiology and Infectious Disease 2023: 116074

Mathieu, E, Ritchie, H, Ortiz-Ospina, E, Roser, M, Hasell, J, Appel, C, Giattino, C, & Rodés-Guirao, L. A global database of COVID-19 vaccinations. In: Nature human behaviour 2021; 5(7): 947-953

Mathieu, E, Ritchie, H, Rodés-Guirao, L, Appel, C, Giattino, C, Hasell, J, Macdonald, B, Dattani, S, Beltekian, D, Ortiz-Ospina, E, & Roser, M, 2020: COVID-19 Pandemic. https://ourworldindata.org/coronavirus (Access date: 2025/04/14)

Monoson, A, Schott, E, Ard, K, Kilburg-Basnyat, B, Tighe, RM, Pannu, S, & Gowdy, KM. Air pollution and respiratory infections: the past, present, and future. In: Toxicological Sciences 2023; 192(1): 3-14

Nazia, N, Butt, ZA, Bedard, ML, Tang, W-C, Sehar, H, & Law, J. Methods used in the spatial and spatiotemporal analysis of COVID-19 epidemiology: a systematic review. In: International journal of environmental research and public health 2022; 19(14): 8267

Prata, D, Rodrigues, W, Bermejo, PHDS, Moreira, M, Camargo, W, Lisboa, M, Reis, GR, & de Araujo, HX. The relationship between (sub) tropical climates and the incidence of COVID-19. In: PeerJ 2021; 9: e10655

Schilling, J, Tolksdorf, K, Marquis, A, Faber, M, Pfoch, T, Buda, S, Haas, W, Schuler, E, Altmann, D, & Grote, U. Die verschiedenen Phasen der COVID-19-Pandemie in

Deutschland: Eine deskriptive analyse von Januar 2020 bis Februar 2021. In: Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz 2021; 64(9): 1093-1106

Sheppard, N, Carroll, M, Gao, C, & Lane, T. Particulate matter air pollution and COVID-19 infection, severity, and mortality: A systematic review and meta-analysis. In: Science of the Total Environment 2023; 880: 163272

Sober, E. Ockham's razors: a user's manual. Cambridge University Press, 2015

Tan, L, & Schultz, DM. Weather effects on the spread of COVID-19: characteristics and critical analysis of the first and second years of scientific research. In: Bulletin of the American Meteorological Society 2023; 104(8): E1345-E1371

Vandenberg, O, Martiny, D, Rochas, O, van Belkum, A, & Kozlakidis, Z. Considerations for diagnostic COVID-19 tests. In: Nature Reviews Microbiology 2021; 19(3): 171-183

Wood, S. Generalized additive models: an introduction with R. New York: Chapman and Hall/CRC, 2017

World Health Organization (WHO), 2020: WHO COVID-19 Dashboard. https://covid19.who.int/ (Access date: 2025/04/15)

World Health Organization (WHO), 2023: WHO Coronavirus (COVID-19) dashboard > Data. https://data.who.int/dashboards/covid19/data (Access date: 2025/02/04)

Zheng, Y, Guo, Z, Wu, Z, Wen, J, & Hou, H. Comparisons of different statistical models for analyzing the effects of meteorological factors on COVID-19. In: Frigid Zone Medicine 2023; 3(3): 161-166

# 3. Publications

## 3.1 Publication 1: Investigating the spatiotemporal associations between meteorological conditions and air pollution in the federal state Baden-Württemberg (Germany)

DOI: https://doi.org/10.1038/s41598-024-56513-4

www.nature.com/scientificreports

scientific reports

Check for updates

OPEN

# Investigating the spatiotemporal associations between meteorological conditions and air pollution in the federal state Baden-Württemberg (Germany)

Leona Hoffmann[1]✉, Lorenza Gilardi[2], Marie-Therese Schmitz[3], Thilo Erbertseder[2], Michael Bittner[2], Sabine Wüst[2], Matthias Schmid[3] & Jörn Rittweger[1,4]

**When analyzing health data in relation to environmental stressors, it is crucial to identify which variables to include in the statistical model to exclude dependencies among the variables. Four meteorological parameters: temperature, ultraviolet radiation, precipitation, and vapor pressure and four outdoor air pollution parameters: ozone ($O_3$), nitrogen dioxide ($NO_2$), particulate matter ($PM_{2.5}$, $PM_{10}$) were studied on a daily basis for Baden-Württemberg (Germany). This federal state covers urban and rural compartments including mountainous and river areas. A temporal and spatial analysis of the internal relationships was performed among the variables using (a) cross-correlations, both on the grand ensemble of data as well as within subsets, and (b) the Local Indications of Spatial Association (LISA) method. Meteorological and air pollution variables were strongly correlated within and among themselves in time and space. We found a strong interaction between nitrogen dioxide and ozone, with correlation coefficients varying over time. The coefficients ranged from negative correlations in January (−0.84), April (−0.47), and October (−0.54) to a positive correlation in July (0.45). The cross-correlation plot showed a noticeable change in the correlation direction for $O_3$ and $NO_2$. Spatially, $NO_2$, $PM_{2.5}$, and $PM_{10}$ concentrations were significantly higher in urban than rural regions. For $O_3$, this effect was reversed. A LISA analysis confirmed distinct hot and cold spots of environmental stressors. This work examined and quantified the spatio-temporal relationship between air pollution and meteorological conditions and recommended which variables to prioritize for future health impact analyses. The results found are in line with the underlying physico-chemical atmospheric processes. It also identified postal code areas with dominant environmental stressors for further studies.**

**Keywords** Air pollution, Environmental stressors, Meteorological data, Cross-correlation, LISA

**Abbreviations**

| | |
|---|---|
| BVOC | Biogenic volatile organic substances |
| BW | Baden-Württemberg |
| CAMS | Copernicus Atmosphere Monitoring Service |
| CCF | Cross-correlation function |
| COVID-19 | Coronavirus SARS-CoV-2 |
| $CO_2$ | Carbon dioxide |
| C3S | Copernicus Climate Change Service |
| DLR | German Aerospace Center |
| ECMWF | European Centre for Medium-Range Weather Forecasts |
| EEA | European Environment Agency |
| LISA | Local indication of spatial association |

[1]Institute of Aerospace Medicine, German Aerospace Center (DLR), Cologne, Germany. [2]German Remote Sensing Data Center, German Aerospace Center (DLR), Weßling, Germany. [3]Institute of Medical Biometry, Informatics and Epidemiology, University Hospital Bonn, Bonn, Germany. [4]Department of Pediatrics and Adolescent Medicine, University Hospital Cologne, Cologne, Germany. ✉email: leona.hoffmann@dlr.de

| | |
|---|---|
| $NO_2$ | Nitrogen dioxide |
| $NO_x$ | Nitrogen oxide |
| $O_3$ | Ozone |
| PCA | Principal component analysis |
| $PM_{2.5}$ | Particulate matter with a diameter of 2.5 micrometers or smaller |
| $PM_{10}$ | Particulate matter with a diameter of 10 micrometers or smaller |
| Prec | Precipitation |
| sd | Standard deviation |
| r | Correlation coefficient |
| Temp | Temperature |
| UV | ultraviolet radiation |
| VOC | Volatile organic compound |
| VP | Vapor pressure |
| WHO | World Health Organization |

The human-made climate change threatens human health not only through extreme meteorological conditions, but also through polluted air that often accompanies it[1,2]. There are concrete plans for the reduction of air pollutant concentrations with guidelines from the World Health Organization (WHO)[3,4] and air quality directive of the European Parliament and the Council[5]. However, there is still a need to better understand outdoor air pollution and their internal relations to prevent potential misinterpretation of the outcomes. In the past, numerous studies were conducted on the effect of air pollution on human health[6–10]. All of them confirm that outdoor air pollution harms human health. Air pollution causes acute and chronic health effects and affects various systems, and organs[8].

Air pollutants are often released in conjunction, such as nitrogen dioxide ($NO_2$), carbon dioxide ($CO_2$) and particulate matter from combustion processes. The dispersion and deposition of air pollutants through meteorological factors is subject to variation given by emission sources, chemical transformations and average atmospheric lifetime. As a consequence there are spatiotemporal covariations between the various pollutants and meteorological variables.

One of the pollutants is particulate matter with a diameter up to 2.5 µm ($PM_{2.5}$) which is a variable that has been frequently studied in the literature and has often shown to have a strong negative effect on health. Short-term as well as long-term exposure to $PM_{2.5}$ has an impact on mortality and morbidity[6]. The evidence supports the possibility that both $PM_{2.5}$ and $PM_{10}$ (particulate matter with a diameter up to 10 µm) are associated with increased mortality from all causes, cardiovascular disease, respiratory disease, and lung cancer[11,12]. In Europe in 2020, exposure to $PM_{2.5}$ concentrations above the WHO guideline level of 5 $\mu g/m^3$ resulted in 275,000 premature deaths[13]. Most individuals residing in Germany inhabit polluted regions[14]. Air pollution negatively impacts virus-transmitting infections such as influenza[15,16] and COVID-19[17–19]. Irrespective of which specific health outcome is considered, understanding the interdependencies between environmental variables and the selection of environmental stressors is a key aspect in epidemiological analyses. Previous studies that examined the relationship between meteorological and air pollution variables were conducted in Cairo (Egypt)[20], China[21], Rome (Italy)[22] and Stuttgart (Germany)[23], for example. Studies on the spatiotemporal variability of tropospheric ozone and nitrogen dioxide are available for Athens (Greece)[24], major cities in India[25] as well as in form of worldwide reviews[26].

In a study on the link between influenza and air pollution[15], a strong correlation was found between some of the environmental stressors considered, which included air pollutants and meteorological variables. Other studies[27,28] made use of the principal component analysis (PCA) method for dimension reduction[29]. However, PCA combines pollutants to create principal components. Understanding the individual coefficients can be difficult because they lack interpretability[30]. PCA is unsuitable for our analysis as we want to focus on the specific relationships between and among meteorological conditions and air pollution. Therefore, we opted for temporal and spatial analysis techniques that allow for a more conclusive interpretation of the environmental stressors and investigate the internal dependencies.

The aim of this analysis is to better characterize and identify the spatiotemporal relationships among environmental parameters. Our analysis takes into account multiple stressors and their spatial and temporal connections across the entire state of Baden-Württemberg. The scope is to reduce the number of variables needed in the epidemiological analysis and therefore simplify them and avoid biased results caused by correlated factors.

In the past, research has focused either on analyzing specific cities or studying one environmental factor affecting larger regions. This study is valuable because it considers multiple environmental stressors and covers at the same time the cross-sectional region of Baden-Württemberg (BW). Due to the combination of urban and rural areas including mountainous and river regions, Baden-Württemberg is particularly well suited for analyzing the spatial-temporal relationships between air quality and meteorological conditions. In addition, this paper provides decision support for the selection of environmental variables for future analyses. Understanding how different stressors are interconnected can offer valuable insights to aid in future health impact analyses and assist other researchers in related fields.

Through our spatial and temporal analysis, we have identified distinct differences and similarities in terms of spatiotemporal patterns in environmental stressors. Based on these findings, we suggest prioritizing certain variables for further investigation. Additionally, we categorized postal code areas into specific groups based on their environmental stressor pattern, providing a spatial delineation.

## Methods

### Study area

Baden-Württemberg is Germany's third most populated state, with an area of approximately 36 thousand km$^2$ and a population of approximately 11 million. Geographically, the state is located in southwestern Germany and includes urban and rural areas. The state capital Stuttgart is the largest city in BW with about 630 thousand inhabitants and is located in the center of the state. Other large cities are Mannheim (310 thousand inhabitants), Karlsruhe (308 thousand inhabitants), and Freiburg (231 thousand inhabitants). Especially in the south, southwest, and southeast of the state, there are many rural and mountain regions, including the Black Forest, the Lake Constance, the edge of the Alps, and the low mountain range Swabian Alb. Of all 1101 communities of BW, 586 have fewer than 5 thousand inhabitants[31].

We obtained a shapefile of the postal code areas in BW from the Esri Germany database[32] in order to handle the postal code areas.

The population density provides information about the number of inhabitants per km$^2$ and was calculated from the available data as follows:

$$\text{population density} = \frac{\text{number of inhabitants}}{\text{postal code area (km}^2\text{)}} \tag{1}$$

Based on a map of population density from the German Federal Institute for Population Research (Bundesinstitut für Bevölkerungsforschung)[33], we decided to introduce four population density categories. Figure 1 provides a graphical overview of the distribution of the population density categories for BW at postal code level.

Unsurprisingly, postal code areas near cities such as Stuttgart and Mannheim are assigned to category 4, that means densely populated areas. The areas of Black Forest, Swabian Alb, and northeastern of BW predominantly have a population density below 151 inhabitants per km$^2$.

### Environmental variables

Eight environmental variables were considered in more detail. These were split into four meteorological parameters [temperature (*Temp*), precipitation (*Prec*), vapor pressure (*VP*) and UV radiation (*UV*)] and four outdoor air pollution parameters [ozone ($O_3$), nitrogen dioxide ($NO_2$), particulate matter ($PM_{2.5}$, $PM_{10}$)]. Air pollution surface concentrations were retrieved from the European air quality reanalysis dataset provided by the Copernicus Atmosphere Monitoring Service (CAMS)[34]. The data source of the meteorological parameters was the ERA5 reanalysis dataset as provided by the Copernicus Climate Change Service (C3S) of the European Centre for Medium-Range Weather Forecasts (ECMWF)[35]. The datasets had a native spatial resolution of 0.1° × 0.1° and a temporal resolution of one hour. The variables were geographically aggregated to postal code areas as described in[36]. A detailed evaluation of uncertainties of the applied data can be found in[37–39]

Table 1 gives an overview of the environmental stressors. Daily mean values between 2010 and 2018 were available for each variable in the table, aggregated at the postal code level in BW and resulting in a total of 3,931,252 data. Next is a brief explanation of the environmental variables[40,41]. The variable *Temp* was measured two meters above the ground. *Prec*, expressed in mm/day, represented the depth of water if it were evenly

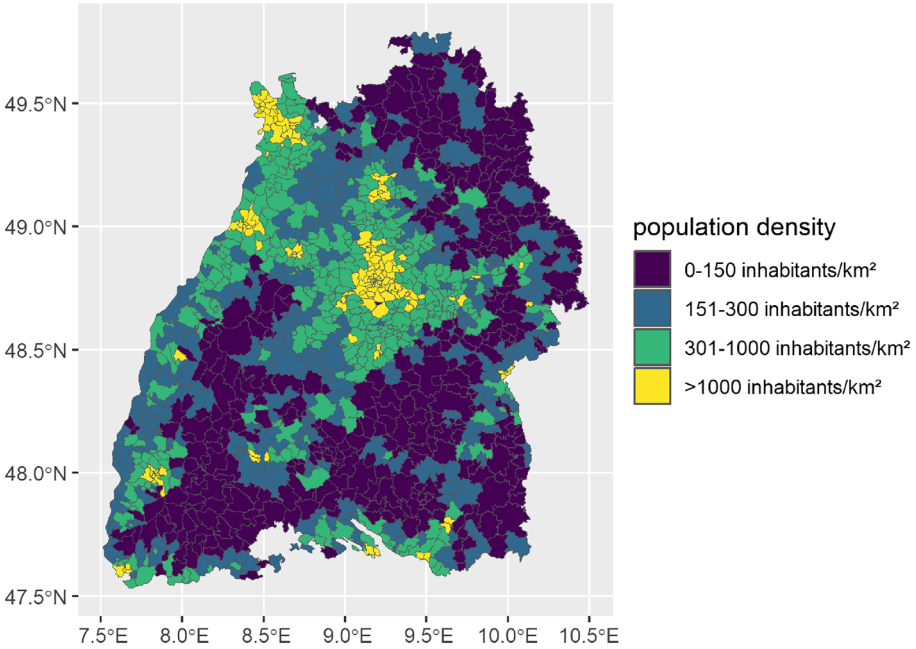


**Figure 1.** Postal code areas in BW categorized by population density.

| Variable | Parameter | Value |
|---|---|---|
| Temp (°C) | Mean (sd) | 9.7 (7.7) |
| | Median [min, max] | 9.8 [−18.9, 30.9] |
| Prec (mm/day) | Mean (sd) | 3.4 (6.0) |
| | Median [min, max] | 0.9 [0, 99.3] |
| VP (hPa) | Mean (sd) | 10.0 (4.2) |
| | Median [min, max] | 9.3 [1.0, 24.6] |
| UV (W) | Mean (sd) | 15.4 (9.7) |
| | Median [min, max] | 14.1 [0.4, 37.0] |
| $O_3$ ($\mu g/m^3$) | Mean (sd) | 51.5 (23.0) |
| | Median [min, max] | 52.4 [0.3, 149.0] |
| $NO_2$ ($\mu g/m^3$) | Mean (sd) | 12.2 (7.3) |
| | Median [min, max] | 10.4 [1.07, 66.4] |
| $PM_{2.5}$ ($\mu g/m^3$) | Mean (sd) | 11.0 (6.7) |
| | Median [min, max] | 9.5 [0.7, 72.3] |
| $PM_{10}$ ($\mu g/m^3$) | Mean (sd) | 14.9 (8.4) |
| | Median [min, max] | 13.2 [0.9, 83.5] |

**Table 1.** Overview of environmental stressors. Overview of meteorological data and air pollutants according to the statistical parameters mean, standard deviation (sd), median, minimum (min) and maximum (max) value. The variables cover all postal code areas in BW and are based on daily measurements from 2010 to 2018.

distributed over the area under consideration. The variable *VP* was a variable constructed from $2m$ dewpoint temperature, as expressed by the following empirical formula[42]:

$$e = 6.112 * \exp(\frac{17.67 * T_d}{T_d - 243.5}) \tag{2}$$

where e was the vapor pressure in hectopascal (hPa) and $T_d$ was the dew point temperature in °C. For *UV*, the unit was converted from $J/m^2$ to $W/m^2$ by dividing the integration time in seconds, resulting in a mean value of 15.4 $W/m^2$ in the processed data. $O_3$ is a colorless and toxic gas in the atmosphere close to the ground (troposphere). $O_3$ is one of the main components of photochemical smog and is produced by complex photochemical processes during intense sunlight. $NO_2$ is a reactive nitrogen compound that is commonly released from the combustion of fuels in the transportation and industrial sectors. $PM_{2.5}$ and $PM_{10}$ are not single pollutants but a mixture of many components such as sulfates, nitrates, ammonia, sodium chloride, black carbon, mineral dust, and water. Depending on the size of the particles, a distinction is made between $PM_{10}$ and $PM_{2.5}$. Particulate matter is generated, in particular, by combustion processes in motor vehicles, power plants, small combustion plants, domestic heating as well as in metal and steel production.

## Statistical methods

Pearson's correlation coefficient was used to determine the pairwise correlations between environmental stressors. The strength of correlation is classified as follows: $r > 0.9$ almost perfect, $0.7 < r \leq 0.9$ very large, $0.5 < r \leq 0.7$ large, $0.3 < r \leq 0.5$ moderate, $0.1 < r \leq 0.3$ small and $r < 0.1$ trivial[43]. From the pearson correlation coefficient emerges the concept of cross- and autocorrelations. In short, a cross-correlation examines the relationship between two or more parameters over time or in space, whereas an autocorrelation examines the relationship to itself (also possible in time and space). The concept of temporal cross-correlation was used to make more precise statements about the temporal internal dependencies of environmental stressors[22]. In a cross-correlation function (ccf), two time series, $x(t)$ and $y(t)$, are examined for correlations with a time offset[44,45]. The formula can be represented as follows:

$$\text{ccf} = \frac{\sum_{t=1}^{N-1}[(x(t) - \bar{x}) * (y(t - \text{lag}) - \bar{y})]}{\sqrt{\sum_{t=1}^{N-1}(x(t) - \bar{x})^2}\sqrt{\sum_{t=1}^{N-1}(y(t - \text{lag}) - \bar{y})^2}} \tag{3}$$

where $\bar{x}$ and $\bar{y}$ denote the mean over time of the corresponding series, respectively. The time series $x(t)$ is fixed, and $y(t \pm lag)$ has a time lag, which is possible in both directions, i.e., $x$ leads $y$ or $x$ lags $y$. In a resulting correlation plot, horizontal lines represent the individual correlations of the two time series with the respective time lag.

A local indication of spatial association (LISA) model was used for the spatial analysis. The LISA statistic measures the degree of autocorrelation between a geographical location and its neighbors, identifying so-called hot and cold spots. For instance, hot spots refer to areas with significantly high values that are surrounded by postal code regions with high values. This method was developed by Luc Anselin[46] and, among others, applied in the context of air pollution[16,47]. Several statistics represent the measure of spatial autocorrelation[48,49]. Here, we used the local Moran statistic, whereby the statistic was applied to each environmental stressor individually. The local Moran's $I$ statistic was given as follows:

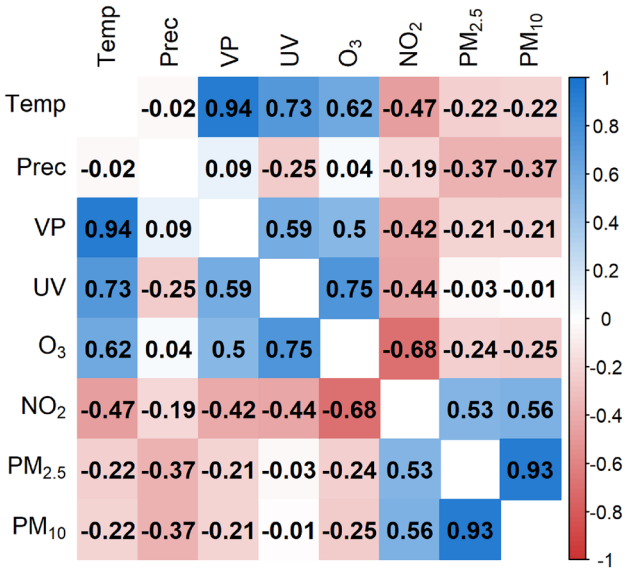


**Figure 2.** Pearson correlation matrix based on daily measurements from 2010 to 2018 across BW. The more intense the color, the stronger the correlation between the two variables, whereas the color blue indicated positive correlations and red negative correlations.

$$I_i = \frac{(x_i - \bar{X})}{S_i^2} \sum_{j=1, j\neq i}^{n} w_{i,j}(x_j - \bar{X}) \tag{4}$$

where $x_i$ was the stressor concentration at location $i$, $x_j$ was the concentration of spatial lag $j$ (neighbors), and $\bar{X}$ was the global mean of the environmental stressor. The spatial weight between $i$ and $j$ was described by the matrix $w_{i,j}$, and the total number of observations was $n$. $S_i^2$ was a constant for all locations with:

$$S_i^2 = \frac{\sum_{j=1, j\neq i}^{n} (x_j - \bar{X})^2}{n-1} \tag{5}$$

Moran's test uses a null hypothesis of randomly dispersed data. All statistical analyses were conducted using R[50], and maps and figures 1 to 9 were generated in R version 4.3.0.

## Results

A matrix of pairwise Pearson's correlation coefficients between the different environmental parameters was shown in Fig. 2.

The correlations represented associations across the entire observation period, that means over all days from 2010 to 2018 and total BW. As expected the temperature and vapor pressure (correlation coefficient (r) 0.94) and $PM_{2.5}$ and $PM_{10}$ ($r = 0.93$) had a strong positive correlation. In addition, there were other relatively strong correlations. The correlation coefficient between *UV* radiation and $O_3$ was 0.75. This means that the stronger the radiation was, the higher the $O_3$ concentration. $O_3$ and $NO_2$ are negatively correlated ($r = -0.68$). If one air pollutant was low, the other was high.

### Temporal analysis

Since this part focuses on temporal relationships, the spatial separation into postal code areas was neglected for this section and the values were averaged over time. It was investigated whether the pairwise Pearson correlations differ across months. We carefully examined correlation plots for all months and decided to include the first correlation plot of each quarter (January, April, July, and October) in the manuscript. Within the manuscript, we do not display all the months explicitly, as some of them represent transitions between the extreme figures shown. The correlation matrices for every month are available in the Supplementary information. The results for the four months January, April, July and October were presented in Fig. 3.

In January, there were many negative correlations. In April, the correlation matrix was primarily characterized by lower correlations. In July, stronger correlations dominate and in October, again, many weaker correlations were visible. The matrices of October and April were overall very similar. They seemed to be intermediate states between the more extreme correlations in January and July.

For $O_3$, the sign of the correlation coefficient with *Prec* was not constant. The strength of the correlation differed in January ($r = 0.59$), April ($r = -0.13$), July ($r = -0.43$) and October ($r = 0.21$). The medium-large negative correlation between *UV* and *Prec* was almost constant over the months. *UV* and $O_3$ were slightly correlated

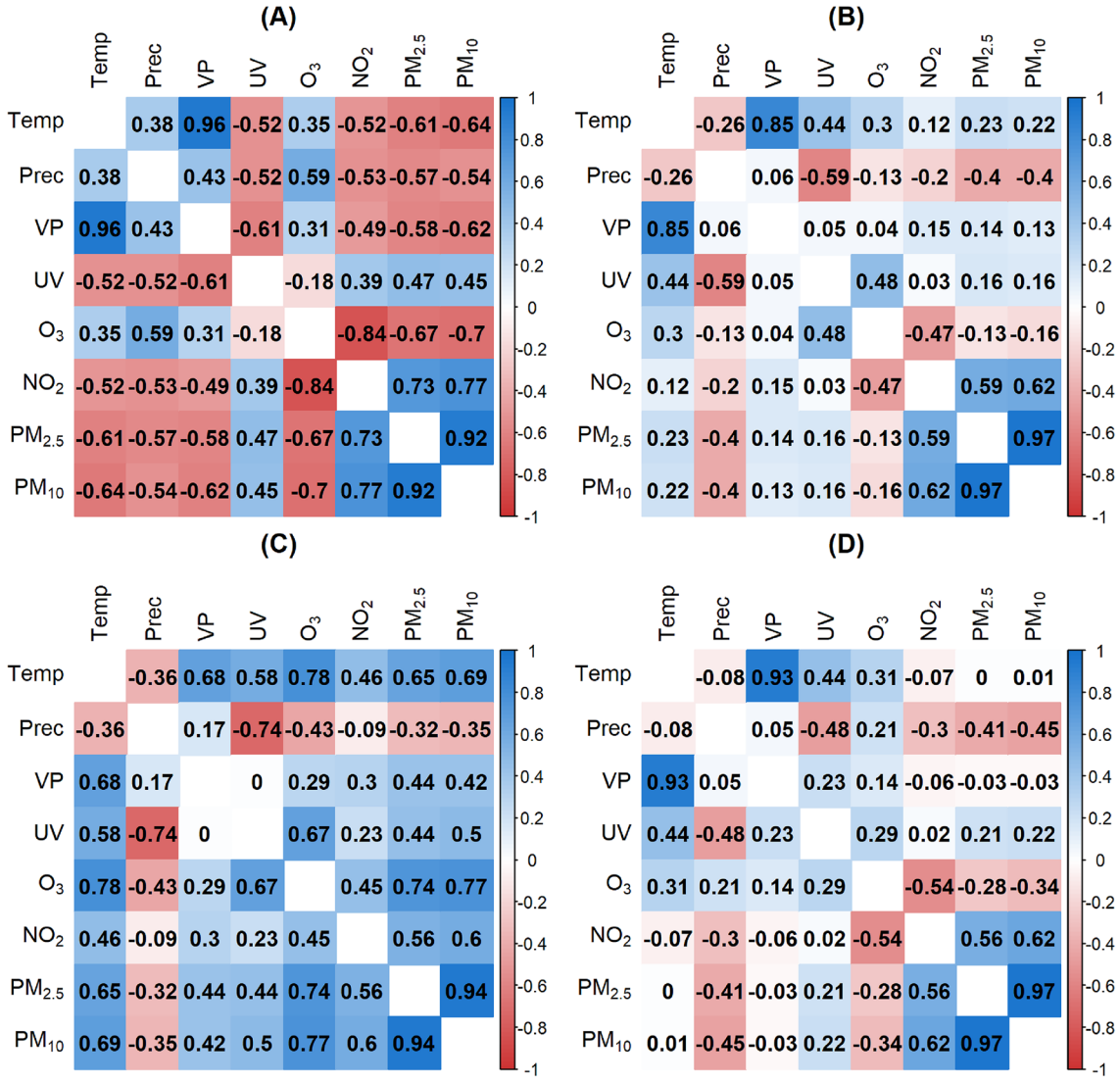


**Figure 3.** Pearson correlation matrix over months for January (**A**), April (**B**), July (**C**) and October (**D**).

in January ($r = -0.18$), moderately positively correlated in April ($r = 0.48$), strongly positively correlated in July ($r = 0.67$), and slightly correlated in October ($r = 0.29$). The variables $O_3$ and $NO_2$ were very strongly negatively correlated in January ($r = -0.84$), moderately negatively correlated in April ($r = -0.47$), moderately positively correlated in July ($r = 0.45$), and again negatively correlated in October ($r = -0.54$). The correlation between $NO_2$ and $PM_{2.5}$, as well as $PM_{10}$, was positively correlated over all months. In January, the correlation was stronger, with values around 0.75, than in the other months, with values around 0.6. To summarize, Fig. 3 showed numerous changes of sign over months. Specifically, 14 of the total 28 correlation coefficients showed a change in sign between January and July.

The correlation between $NO_2$ and $O_3$ was already prominent in Fig. 3. Therefore, the auto- and cross-correlation relationship between the $O_3$ and $NO_2$ variables was depicted in Fig. 4. Overall, the plots by year look very similar, so exemplary, the correlations of the environmental stressors in Fig. 4 were presented for 2018. All graphs had a low point at lag nine and a high point at lag 14. There were noticeable differences in the plots. Specifically, lag 32 in (A) showed a dip that was not present in (B) and (C). In Fig. 4D there was an annual periodicity with the highest positive Pearson correlation coefficient reaching around 0.4 when $NO_2$ comes before $O_3$ and approximately 0.3 when $NO_2$ follows $O_3$. In addition, there was a pattern that repeats about every 13 days in (C).

The output of the cross-correlation function between $NO_2$ and $O_3$ in Fig. 5 was essentially the same as the day-level correlation in Fig. 4. The figures for $NO_2$ and *Temp* as well as $NO_2$ and *UV* looked very similar to $NO_2$ and $O_3$. There was also a change in signs and another high point after about half a year in both directions. A small structure in Fig. 5 was seen between $NO_2$ and *Prec* in the correlation values shifted by months. That means a wave-like structure similar to $NO_2$ and $O_3$ can be observed. In contrast to $NO_2$ and $O_3$, the correlation coefficient starts in the positive range at lag 0, changes to negative with increasing lags, and returns to positive. However, these were relatively weak correlations. A shift made sense content-wise because rain washes the air clean of pollutants. $NO_2$ and $PM_{2.5}$ started at lag 0 with a relatively strong correlation value above 0.5. For positive lag values,

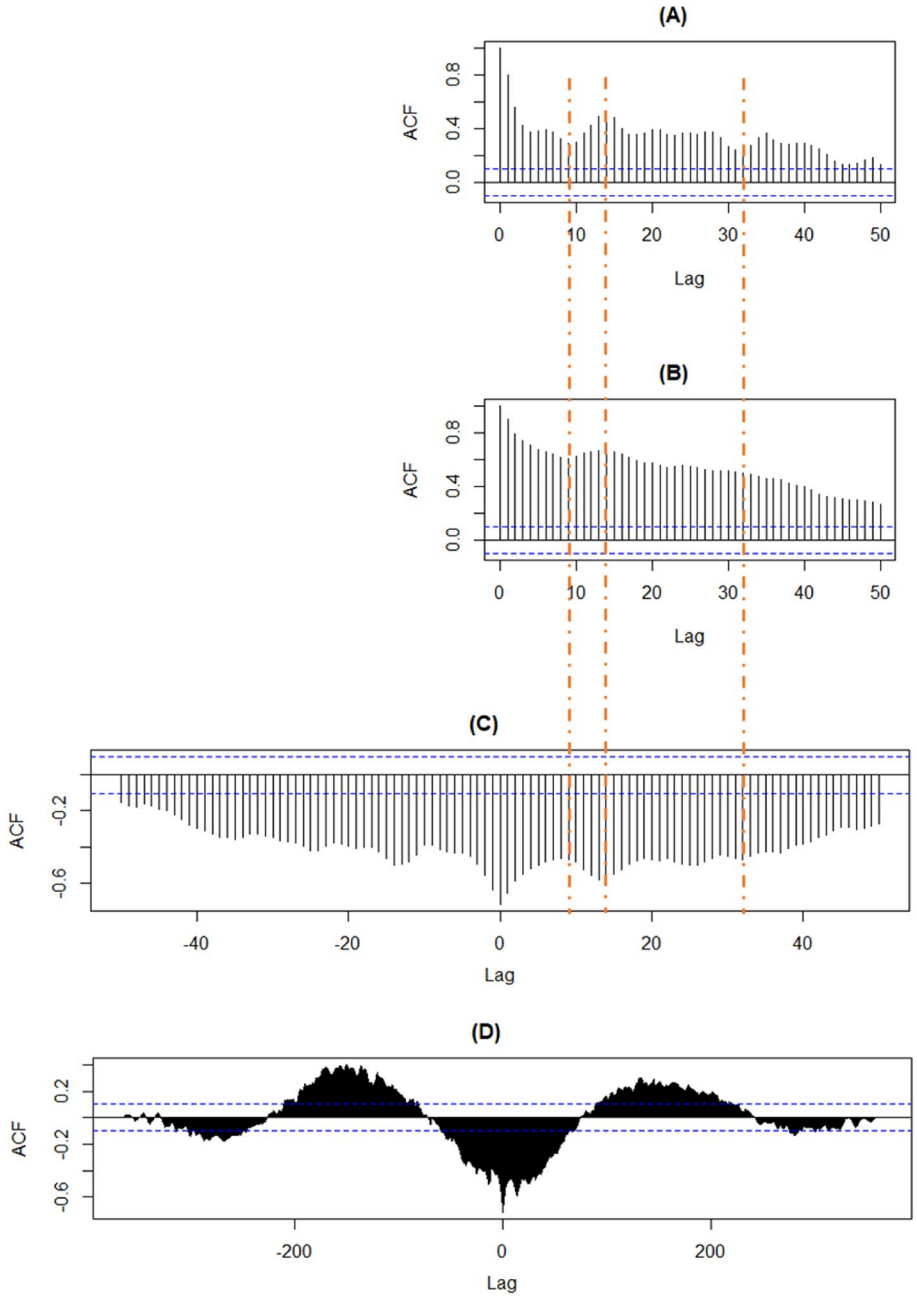


**Figure 4.** Autocorrelation function for $NO_2$ (**A**) and $O_3$ (**B**) of the year 2018. Cross-correlation function (ccf) for $NO_2$ and $O_3$ of the year 2018 with 50 days lag (**C**) and 400 days lag (**D**). The time series $O_3$ was fixed, and the time series $NO_2$ shifted by lags for the ccf function. Plots (**C**) and (**D**) showed the correlation between $NO_2$ at time $t \pm lag$ and $O_3$ at time $t$.

if $NO_2$ lags behind $PM_{2.5}$, there was a similar structure to the cross-correlation of $NO_2$ and *Prec* at low level. The ccf plot for $PM_{10}$ and $PM_{2.5}$ gave a strong correlation at lag 0, followed by no clear structure at the monthly shift.

## Spatial analysis

The aim of the the LISA analysis[46,47] was to identify the locations of clusters of LISA hot and cold spots of the environmental stressors throughout BW.

According to Fig. 6, regions with very high or low population density are most affected by air pollution variables such as $PM_{10}$, $PM_{2.5}$, $NO_2$ and $O_3$. The meteorological variables had various patterns. A more detailed representation of the spatial associations was given in Fig. 7 with a significance level set to 0.001. Using local cluster maps, the spatial associations between postal code areas were summarized into LISA hot and cold spots.

As already showed in the local significance map, most air pollution variables showed a similar structure. The following applies to $NO_2$, $PM_{10}$, and $PM_{2.5}$: In urban areas (Stuttgart + Mannheim region), many postal code

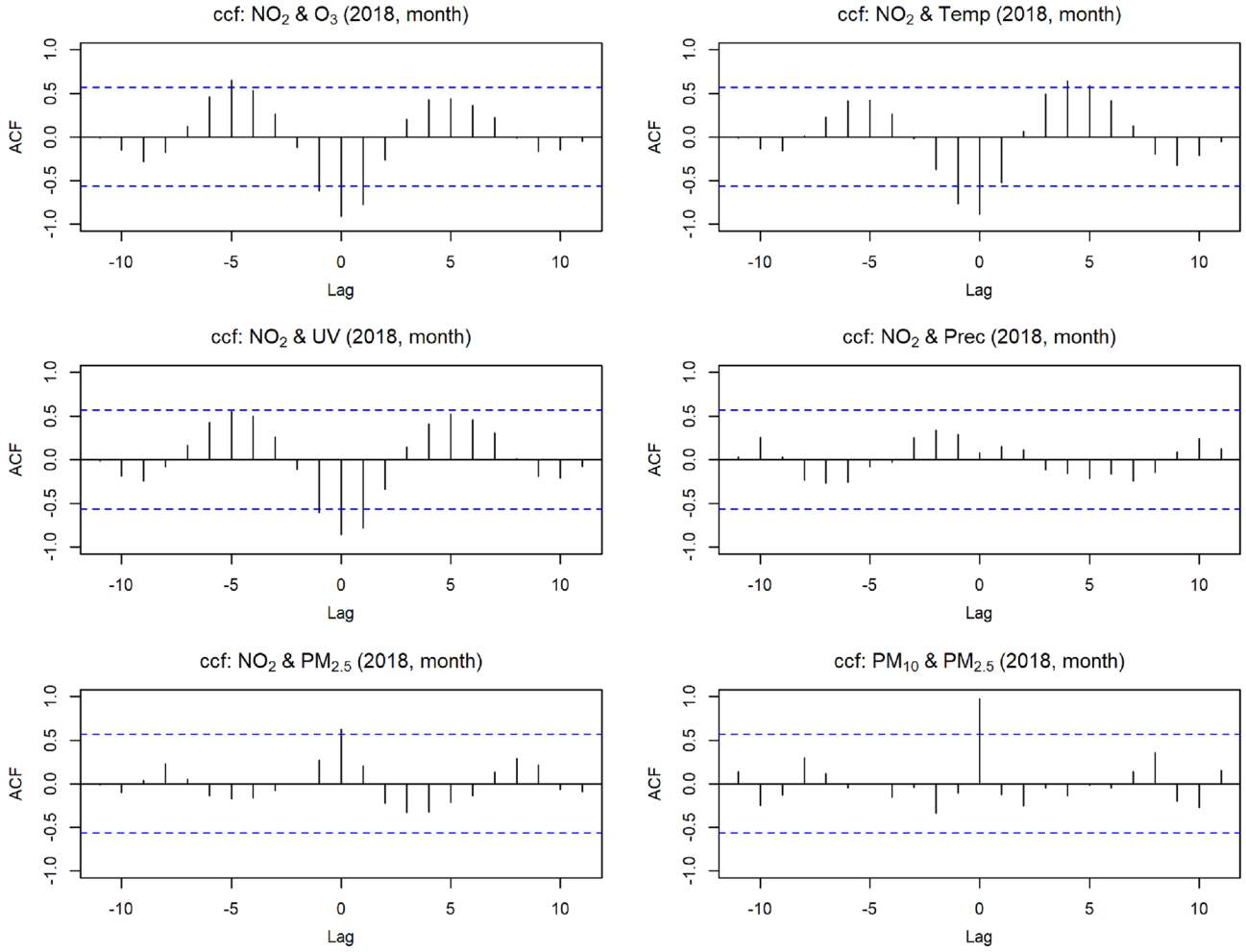


**Figure 5.** Different cross-correlation functions for the year 2018 (lag = month).

areas had neighbors with significantly similar high values. The cold spots geographically contain the Black Forest and parts of the Swabian Alb. For $O_3$, the distribution of LISA hot and cold spots was similar but reversed. Based on the meteorological variables, different postal code regions stand out, with *VP* and *Temp* showing similar trends. Accordingly, LISA hot spots were located at the western edge of BW. LISA cold spots were found along the Schwäbische Alb. A possible geographical connection could be the differences in altitude of the individual areas. For *UV*, LISA cold spots were located exclusively in the north, and LISA hot spots were exclusively in the south of the state. The Local Moran Map for *Prec* looked similar, i.e., LISA cold spots in the north and LISA hot spots in the south. However, no such clear and extreme assignment as for *UV* was recognizable. This could possibly be explained by the altitude, latitude and local air circulations. Furthermore, areas of the Black Forest showed LISA hot spot postal code areas. One could assume a connection between *Temp* and *UV*. However, this map showed no spatial relationship between LISA hot and cold spots for *UV* and *Temp*.

Pairwise correlations between stressors and cross-correlation plots were examined by the population density. Table 2 showed a summary of environmental stressors split by population density categories.

$NO_2$ showed the most remarkable change in concentration across the population density compared to the other environmental stressors with a mean value ranging from 9.6 $\mu g/m^3$ in rural to 16.5 $\mu g/m^3$ in urban areas. The more urban the area was, the higher the $NO_2$ concentration. For $PM_{2.5}$ and $PM_{10}$, a similar but less intense increase was observed. $O_3$ had a reverse effect ranging from 54.4 $\mu g/m^3$ to 47.2 $\mu g/m^3$. The environmental stressors *UV*, *VP*, *Temp*, and *Prec* were mainly constant over space.

Figure 8 showed the Pearson correlation matrix for all data separated by population density category. If we specifically compare the most rural and urban areas (categories 1 and 4), we noticed that $NO_2$ had different correlation values depending on the category. $NO_2$ correlated more negatively with *Temp* ($r = -0.67$), *VP* ($r = -0.6$), and *UV* ($r = -0.5$) in category 1 of population density than in category 4 with correlation coefficients of −0.55 (*Temp*), −0.48 (*VP*) and −0.45 (*UV*). The correlation between $NO_2$ and $O_3$, as well as *Prec*, remained stable mainly across the shifting categories of population density. Between $NO_2$ and $PM_{2.5}$ as well as $PM_{10}$, the positive correlation increased with rising population density from 0.51 and 0.53 (category 1) to 0.61 and 0.63 (category 4). All other correlations showed only minor variations across population density categories.

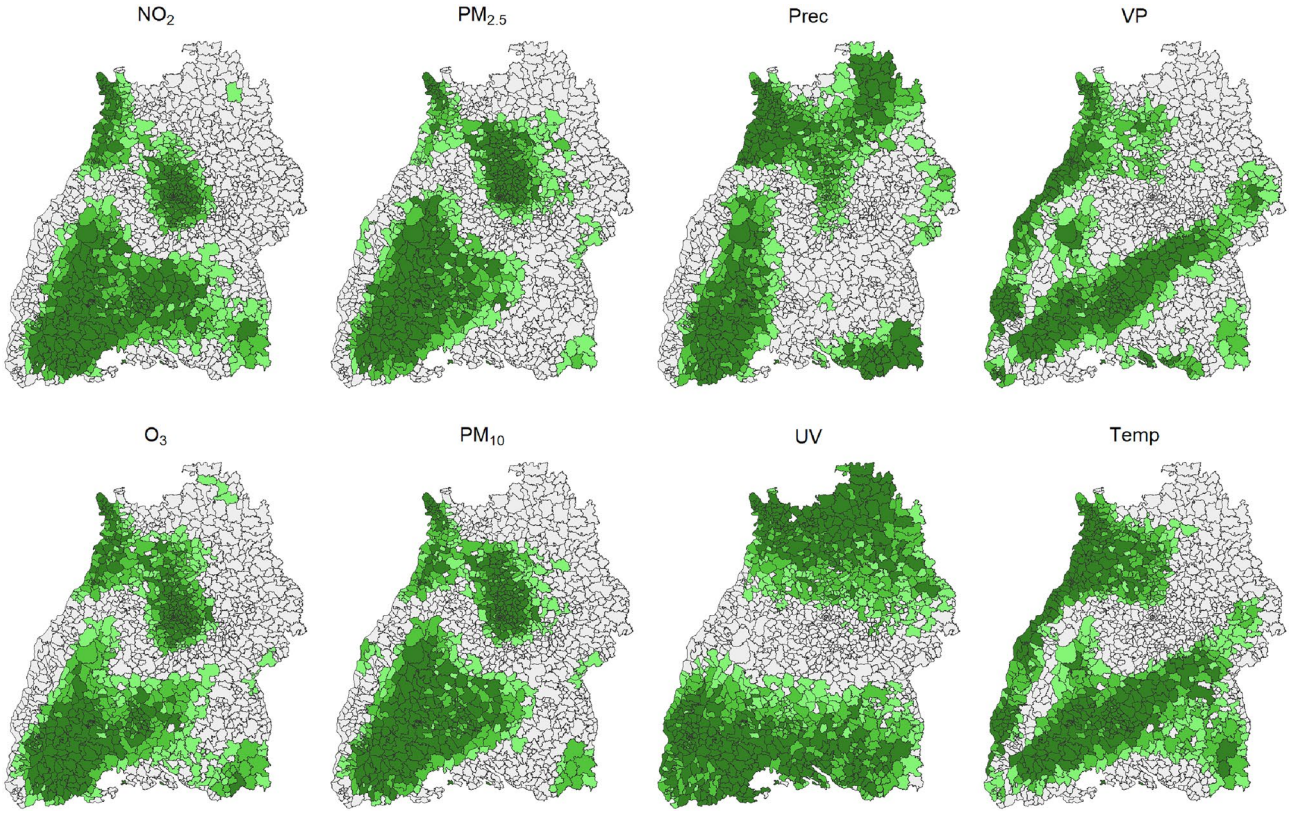


**Figure 6.** Local significance map of the individual environmental stressors in BW. The figure represented the results of the significance test of LISA analysis using Moran's I. Different shades of green indicated different thresholds: light green $p < 0.05$; medium green $p < 0.01$ and dark green $p < 0.001$.

The hot and cold spot resulted from the LISA analysis can be used to present a new level of spatial units in addition to the predefined categorical spatial units for population density in Table 2. When comparing the LISA hot and cold spots, it could be seen that they do not always corresponded to categories 1 through 4 of population densities. The parameters *Prec*, *UV*, *VP*, and *Temp*, in particular, differed from the distribution of population density in BW. Additionally, the LISA hot spots do not entirely matched the high population density in Table 2 for the $NO_2$, $PM_{2.5}$, $PM_{10}$, and $O_3$ parameters. Table 3 provided an overview of descriptive statistics for the new spatial units as obtained from the LISA analysis.

We introduced a new level of insight that was defined by the spatial variability of the stressors rather than relying on a predefined quantity such as population density. The strong discrepancy between the mean values was remarkable when looking at the hot and cold spots. These varied from 18.5 to 7.6 for $NO_2$, 59.7 to 45.1 for $O_3$, and 12.0 to 9.5 for $PM_{2.5}$.

Figure 9 displayed the Pearson correlation matrices for the hot and cold spots of the LISA spatial units, similar to the population density categories.

The correlation directions in both matrices were identical. However, the correlation coefficients for $NO_2$ with *Temp*, *VP*, and *UV* were higher for the cold spots. In addition, the $PM_{2.5}$ and $PM_{10}$ correlation values with all other stressors were overall slightly higher in the hot spots than in the cold spots.

We also generated cross-correlation plots for $NO_2$ and $O_3$ in 2018, categorized by both population density and hot and cold spots. However, the cross-correlations do not showed strong changes in the associations between environmental stressors.

## Discussion

Meteorological and air pollution variables were strongly correlated between and among themselves, with specific seasonal and spatial features. For example, $NO_2$ and $O_3$ were strongly interdependent, and the Pearson correlation varied with time. In January, there was a negative correlation of −0.84, whereas in July, the correlation coefficient was 0.45. Figure 10 illustrated that $NO_2$ and $O_3$ correlated not only with each other but also with other environmental stressors. It is particularly intriguing to note the contrasting values of the two months, as their correlation directions often differed.

The cross-correlation showed a noticeable change in the correlation direction for $O_3$ and $NO_2$. Spatially, $NO_2$, $PM_{2.5}$, and $PM_{10}$ concentrations were significantly higher in urban than in rural regions. For $O_3$, the effect is reversed. This result is also confirmed by LISA analysis, where distinct hot and cold spots of the different environmental stressors could be identified. In addition, the Pearson correlation coefficients suggest that $PM_{10}$ variation was almost entirely explained by $PM_{2.5}$ and vapor pressure by temperature.

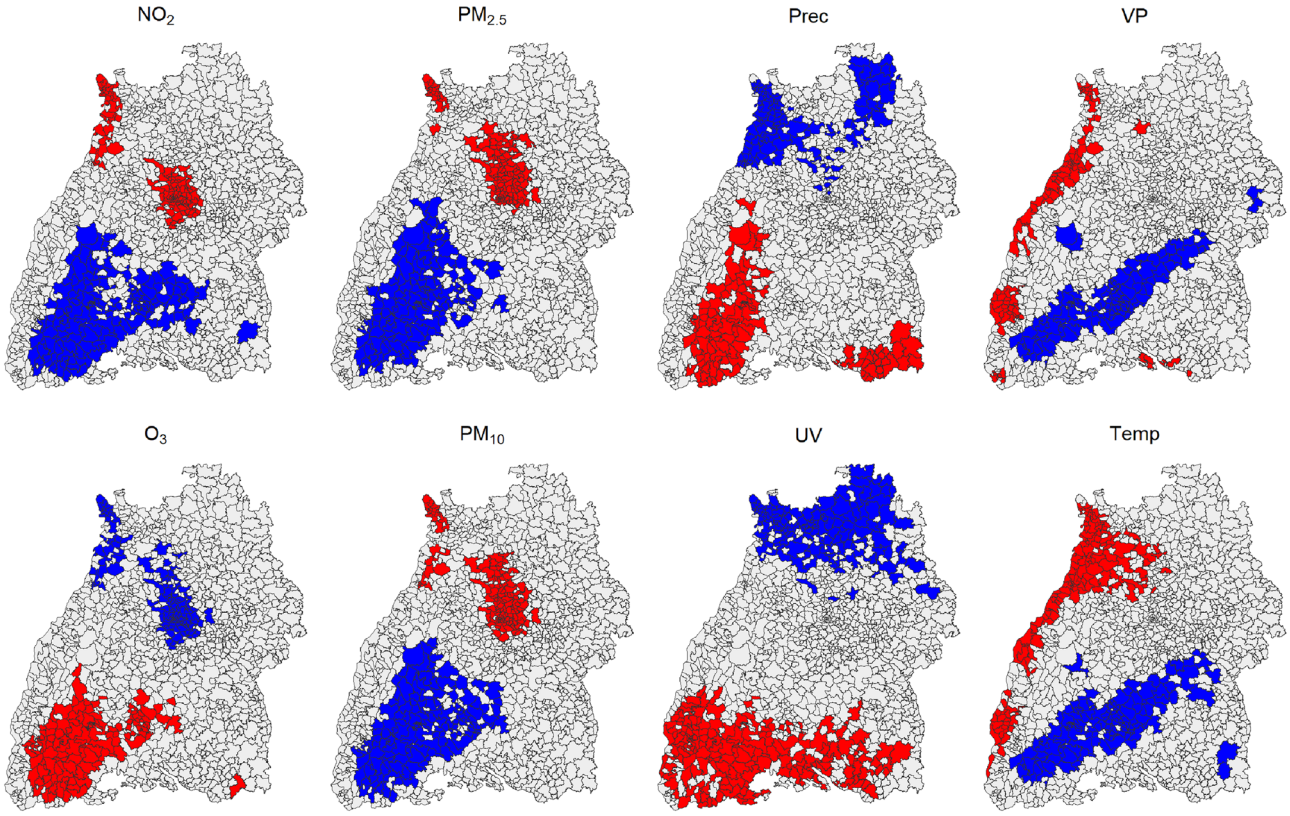


**Figure 7.** Local Cluster Map of the individual environmental stressors in BW. The significance level was set to $p \leq 0.001$. LISA hot spots were red, representing positive spatial autocorrelation with high values. LISA cold spots were blue, representing low values. BW had three postal code areas (78,266, 78,465, 78,479) that were isolated without neighboring postal code areas. We omitted these areas from any further consideration. The estimation was always relative to the mean, as constructed in formula (4). The combination of the red and blue areas results in the dark green areas in Fig. 6.

It is essential to emphasize that correlation does not necessarily imply causality. Correlation means that a change in one variable is related to another variable. However, correlation does not mean that one variable directly influences the other. Thus, there is not necessarily a cause-and-effect relationship. A third variable could affect the two variables that are supposed to be causally related.

Some atmospheric processes can explain the spatio-temporal variability of the correlation coefficients. From an atmospheric chemistry point of view $NO_2$ and $O_3$ need to be considered together since they are a function of each other which explains the opposite relationship[51]:

$$O_3 + NO \rightarrow O_2 + NO_2. \tag{6}$$

Their spatio-temporal variability is governed by superimposed emission-based and photochemistry-based regimes[52]. The predominantly anthropogenic nitrogen oxide ($NO_x$) emissions and the resulting $NO_2$ concentrations have a pronounced seasonal cycle, with higher values in winter than in summer[53]. This is due to the fact that in addition to the higher emissions also the lifetime of $NO_2$ is longer in winter (about 21 h) than in summer (about 6 h)[54,55]. Peak $NO_2$ concentrations in winter are resulting from superimposed atmospheric inversion conditions. The photochemically produced tropospheric $O_3$ exhibits higher levels in the summer months when there is more solar radiation. Nevertheless, the production depends strongly and non-linearly on precursors like $NO_x$ and volatile organic compound ($VOC$) concentrations as well as meteorological conditions. However, photolysis of $NO_2$ is the primary chemical source of tropospheric ozone[51,56].

Consequently, the relationship between $O3$ and $NO_2$ is complex, influenced by a variety of factors and thus needs to be distinguished between rural and urban areas. Since $NO_x$ are emitted from traffic, industrial processes, and other human activities the resulting $NO_2$ concentration are higher in urban areas and industrial agglomerations as can be inferred from Fig. 7. This is in line with the findings of[23] where the urban pollution island of the Stuttgart city region could be delineated from satellite.

Ozone in urban areas is primarily formed as a secondary pollutant through chemical reactions involving precursor pollutants, especially during sunny, warm weather conditions and can be transported to rural areas, affecting rural air quality. Paradoxically, locally high emissions of $NO_x$, such as from traffic, tend to favor ozone destruction in urban areas, resulting in $NO_2$ formation. As can be seen in Fig. 7 this results in ozone cold spots in urban areas. In rural areas, natural sources like vegetation (emitting biogenic volatile organic compounds) and soil contribute to ozone formation. $O_3$ levels tend to be higher in rural areas where there are fewer local emissions of $NOx$ to destroy any $O_3$ that was photochemically produced. As can be inferred from Fig. 7 the Black Forest

| Variable | Parameter | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| Temp (°C) | Mean (sd) | 9.0 (7.7) | 9.7 (7.7) | 10.0 (7.7) | 10.3 (7.6) |
| | Median [min, max] | 9.2 [−18.5, 29.9] | 9.9 [−18.9, 30.8] | **10.2** [−18.6 | **10.4** [−18.2 |
| Prec (mm/day) | Mean (sd) | 3.6 (6.3) | 3.4 (6.0) | 3.3 (5.9) | 3.1 (5.6) |
| | Median [min, max] | 1.0 [0, 99.3] | 0.9 [0, 98.1] | 0,9 [0, 98.0] | 0.8 [0, 85.4] |
| VP (hPa) | Mean (sd) | 9.7 (4.2) | 10.0 (4.2) | 10.1 (4.3) | 10.2 (4.3) |
| | median [min, max] | 9.1 [1.0, 24.0] | 9.4 [1.1, 24.6] | 9.5 [1.1, 24.6] | 9.6 [1.1, 24.3] |
| UV (W) | Mean (sd) | 15.5 (9.6) | 15.4 (9.7) | 15.3 (9.7) | 15.2 (9.7) |
| | Median [min, max] | 14.1 [0.4, 37.0] | 14.1 [0.4, 37.0] | 14.0 [0.4, 37.0] | 13.9 [0.4, 37.0] |
| $O_3$ ($\mu g/m^3$) | Mean (sd) | 54.4 (22.3) | 52.0 (22.8) | 50.0 (23.2) | 47.2 (23.5) |
| | Median [min, max] | 55.1 [0.4, 149] | 52.9 [0.6, 148] | 51.0 [0.3, 149] | 48.3 [0.5, 149] |
| $NO_2$ ($\mu g/m^3$) | Mean (sd) | 9.6 (5.8) | 11.4 (6.6) | 13.5 (7.4) | 16.5 (7.5) |
| | Median [min, max] | 8.1 [1.7, 31.3] | 9.7 [2.0, 38.3] | 11.9 [2.5, 45.5] | 14.8 [3.3, 53.0] |
| $PM_{2.5}$ ($\mu g/m^3$) | Mean (sd) | 10.5 (6.1) | 10.9 (6.3) | 11.2 (6.5) | 11.6 (8.7) |
| | Median [min, max] | 8.0 [1.1, 66.4] | 9.7 [1.1, 60.0] | 11.7 [1.1, 66.0] | 14.7 [1.4, 66.4] |
| $PM_{10}$ ($\mu g/m^3$) | Mean (sd) | 14.1 (7.9) | 14.7 (8.3) | 15.3 (8.7) | 16.0 (9.1) |
| | Median [min, max] | 12.6 [0.9, 79.6] | 13.1 [1.0, 80.0] | 13.6 [1.1, 83.5] | 14.2 [1.3, 82.5] |

**Table 2.** Overview of environmental stressors split by population density categories. Overview of environmental stressors split by population density category 1: 0–150 inhabitants/km$^2$, category 2: 151–300 inhabitants/km$^2$, category 3: 301–1000 inhabitants/km$^2$, category 4: > 1000 inhabitants/km$^2$ including information on standard deviation (sd), median, minimum (min) and maximum (max) value. The variables covered all postal code areas in BW and were based on daily data from 2010 to 2018.

mountain range depicts an $O_3$ hot spot in BW due to the high solar irradiation and the abundance of biogenic volatile organic substances *BVOC* ozone precursors. This is also substantiated in Table 2: the more inhabitants there are, the less $O_3$ and the more $NO_2$ occur.

The anticorrelation of $O_3$ and $NO_2$, for the reasons outlined above, can be confirmed for all temporal and spatial aggregation levels in the study: Fig. 2 (entire BW, entire period), Fig. 3 (entire BW; Jan, Apr, Oct), Fig. 8 (population density classes, entire period) and Fig. 9 (hot and cold spots; entire period) except for Fig. 3 (entire BW, Jul) where a positive correlation is reported. The negative correlations agree with in-situ measurements of $O_3$ and $NO_2$ for Munich, a similar city region to Stuttgart at almost the same latitude, with −0.58 and −0.64 for the period January to July 2019 and 2020, respectively[57]. The same authors find a positive correlation of $O_3$ and *Temp* of 0.67 and 0.49 for the time periods given and addresses also the interannual variability. This agrees with r = 0.62 in our study for the period 2010 to 2018. Further process-oriented regimes can be identified in the results of the study. Temperature-dominated effects can best be seen in the Pearson correlation matrix for July (Fig. 3C). In the first column the variables *VP*, *UV*, $O_3$, $NO_2$, $PM_{2.5}$ and $PM_{10}$ are all positively correlated with *Temp*, partly because they are produced by photochemical processes and linked high solar irradiance ($O_3$, $NO_2$, *UV*) or increased by dry weather conditions ($PM_{2.5}$ and $PM_{10}$)[1]. In the second column the effect of wet deposition and cleaning effect is evident by the negative correlation of all air pollutants with precipitation. *UV* shows the strongest negative correlation, due to the presence of clouds. Emission-dominated variables and effects can best be seen in winter (Fig. 3A). It can be inferred that low temperatures, low precipitation and low water pressure favor high concentrations of $NO_2$, $PM_{2.5}$ and $PM_{10}$, partly due to increased heating, longer photochemical lifetimes and accumulation under inversion conditions or low windspeeds[1,55]. More previous studies have discussed the effect of meteorological conditions on the concentration of atmospheric pollutants and meteorological variables such as wind direction, wind speed and precipitation that have a constraining effect on atmospheric pollutant concentrations, but not a simple linear relationship[20,21,58]. This study is limited to daily data. For further studies, hourly observations could be considered as applied in[10]. The different stressors usually interact differently during the day and at night since e.g. anthropogenic emissions exhibit a pronounced daily cycle and photochemical reactions are confined to sunlit conditions[26]. Another potential extension of the current analysis is to expand the study area to encompass all of Europe. The study uses air pollution and meteorological data that represents background conditions and mesoscale variability. As such air pollution from point sources or along roads cannot be resolved. However, such data is not yet available to our knowledge for the entire BW and the time period under investigation. A possible addition to the variables considered could be wind speed, wind direction and boundary layer height since these parameters have shown a large impact on the variability of particulate matter[59,60]. Furthermore, boundary layer height and $O_3$ showed the strongest positive correlation among all the analyzed variables in[57]. The Pearson correlation assumes a linear relationship between two continuous variables. Linearity was deemed sufficient for our initial analysis of the internal dependencies of environmental stressors, although there are other correlation coefficients like Spearman correlation that deal with nonlinear associations. We recommend using nonlinear statistical methods such as generalized additive models with splines for advanced studies of air pollution and health factors.

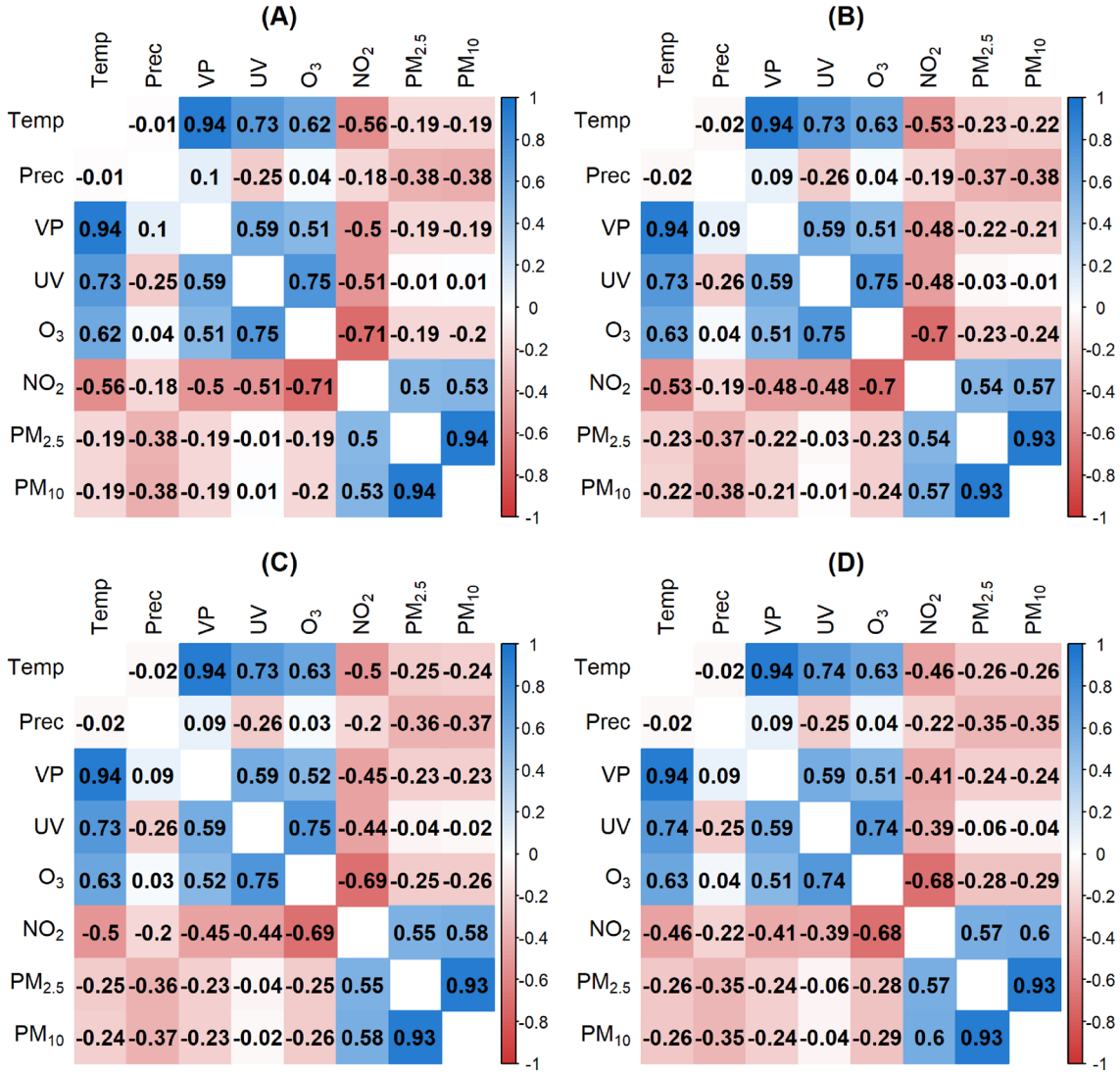


**Figure 8.** Pearson correlation matrices based on daily data aggregated by population density category 1 (**A**): 0–150 inhabitants/km$^2$, category 2 (**B**): 151–300 inhabitants/km$^2$, category 3 (**C**): 301–1000 inhabitants/km$^2$, category 4 (**D**): > 1000 inhabitants/km$^2$.

| Variable | Parameter | LISA hot spots | LISA cold spots | Isolated | Non significant |
|---|---|---|---|---|---|
| $NO_2$ ($\mu g/m^3$) | Mean (sd) | 18.5 (8.7) | 7.6 (4.6) | 11.2 (6.5) | 12.0 (6.9) |
| | Median [min, max] | 16.7 [2.5, 66.4] | 6.24 [1.1, 42.9] | 9.2 [1.7, 40.3] | 10.3 [1.1, 66.4] |
| $O_3$ ($\mu g/m^3$) | Mean (sd) | 59.7 (21.8) | 45.1 (23.2) | 53.7 (23.5)) | 51.5 (22.8) |
| | Median [min, max] | 59.5 [1.1, 149] | 46.4 [0.5, 136] | 55.2 [3.3, 125] | 52.4 [0.3, 149] |
| $PM_{2.5}$ ($\mu g/m^3$) | Mean (sd) | 12.0 (7.1) | 9.5 (5.9) | 11.3 (6.9) | 11.0 (6.7) |
| | Median [min, max] | 10.5 [1.2, 63.1] | 8.3 [0.8,50.4] | 9.8 [1.1, 57.3] | 9.6 [0.7, 72.3] |

**Table 3.** Overview of environmental stressors split by LISA spatial units. Overview of $NO_2$, $O_3$ and $PM_{2.5}$ split by LISA spatial units hot spot, cold spot, isolated and other non significant postal code areas including information on standard deviation (sd), median, minimum (min) and maximum (max) value. The variables covered all postal code areas in BW and included all days from 2010 to 2018. Note that the LISA spatial units differed based on the environmental stressor, as illustrated in Fig. 7. As a result, the quantity of values in each category varied.

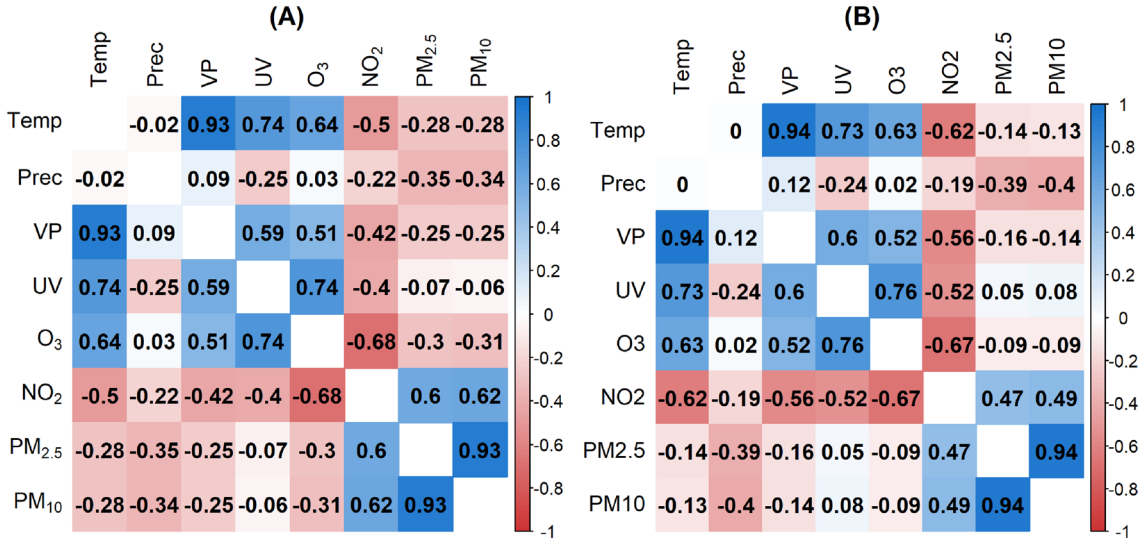


**Figure 9.** Pearson correlation matrices based on daily measurements aggregated by postal code areas in hot (**A**) and cold (**B**) spots as obtained from the LISA analysis.

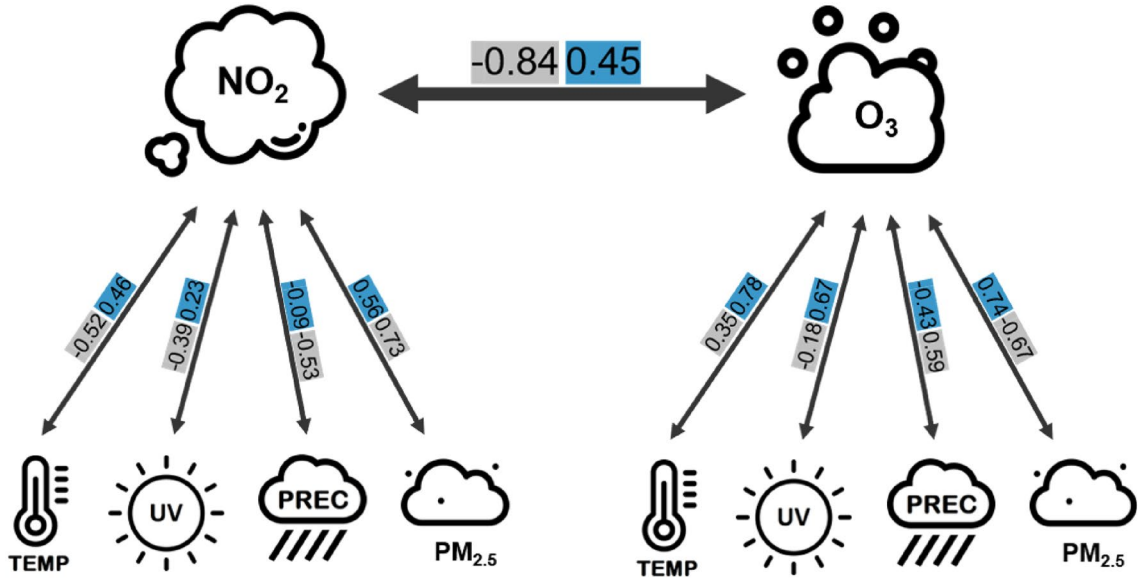


**Figure 10.** Illustration of the correlations between $NO_2$ and $O_3$ in January (gray) and July (blue). The correlation values originated from Fig. 3. This diagram had been designed using images from flaticon.com.

## Conclusion

Selecting the appropriate variables for a statistical model can be a challenging task. This paper offers decision-making assistance for upcoming analyses describing the health effects of environmental stressors. Including a single environmental variable in the model may result in information loss, while including too many variables may lead to correlations and biases. Finding the right balance is important. The optimal choice of variables relies on the specific research question and the given data. However, this paper provides recommendations regarding the variable selection that can be considered. In this work, it turned out to be sufficient to consider $PM_{2.5}$. $PM_{10}$ has larger particles but almost identical temporal and spatial characteristics. The only possible deviation would be for Saharan dust[61,62]. The variables *VP* and *Temp* show strong similarities by design so that future investigations can be limited to the temperature.

The opposite relationship between $NO_2$ and $O_3$ was confirmed both temporally and spatially. $NO_2$ is more often observed in metropolitan areas and $O_3$ in rural areas. How can this knowledge be addressed in a future model describing the health effects of environmental stressors? For future analyses, we propose incorporating interaction terms to effectively illustrate the relationships between the two variables, $NO_2$ and $O_3$, and their impact on the dependent variable. Considering only one environmental stressor in a future model may lead to loss of information and confusion in interpreting. Moreover, including both variables in the model without

an interaction term is not advisable, as high Pearson correlation coefficients may cause bias. Based on this, we recommend using an interaction term between $NO_2$ and $O_3$.

It is important to note that different stressors have different health effects. Although the previous analysis suggests that $NO_2$ and $O_3$ are opposite, they have different impacts on human health[40]. This fact makes the choice of model and the interpretation of the relationships more complex and must be considered in future analyses.

When applying the LISA model, we found substantial spatial differences in some variables (e.g. $PM_{2.5}$), but not in others (*UV*) between urban and rural areas. This indicates that this spatial variation can be statistically exploited for epidemiological studies. Notably, a large fraction of postal code regions show lack of coherence with their neighbors, as can seen from the high proportion of uncolored areas in Fig. 7. In addition, we also identified clear patterns of LISA hot and cold spots, particularly in urban areas, mountainous regions Schwarzwald, and Schwäbische Alb. All of these identified patterns show positive autocorrelations, and no negative autocorrelation was observed.

We spatially categorized the state of Baden-Württemberg in two ways: first, by population density (Fig. 2 and table 2) and second, by LISA hot and cold spots (Fig. 7 and table 3). The two categorizations matched spatially well for some air pollution variables (e.g. $PM_{2.5}$ and $NO_2$) and less well for some meteorological variables (e.g. *UV*).

To conclude, it will be straightforward to implement the principle findings of our study, namely (a) the temporal coherence of stressor patterns, and (b) the spatial clustering into statistical models for the epidemiological study of stressor effects on human health, e.g. by affording the necessary spatial categorical variables and the opportune interaction terms into the statistical models.

## Data availability

The data that support the findings of this study are available on request from the corresponding author LH.

Received: 2 November 2023; Accepted: 7 March 2024
Published online: 12 March 2024

## References

1. Jacob, D. J. & Winner, D. A. Effect of climate change on air quality. *Atmos. Environ.*https://doi.org/10.1016/j.atmosenv.2008.09.051 *(2009).*
2. Doherty, R. M., Heal, M. R. & O'Connor, F. M. Climate change impacts on human health over Europe through its effect on air quality. *Environ. Health.* **16**(1), 33–44. https://doi.org/10.1186/s12940-017-0325-2 (2017).
3. World Health Organization (WHO). *WHO Global Air Quality Guidelines: Particulate Matter (PM2.5 and PM10), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide*. (World Health Organization, 2021). https://www.who.int/publications/i/item/9789240034228.
4. World Health Organization (WHO). *Air Quality Guidelines: Global Update 2005: Particulate Matter, Ozone, Nitrogen Dioxide, and Sulfur Dioxide*. (World Health Organization, 2006). WHO/SDE/PHE/OEH/06.02. https://www.who.int/publications/i/item/WHO-SDE-PHE-OEH-06.02.
5. The European Parliament and the Council of the European Union. Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe. *J. Eur. Union.* (2008).
6. World Health Organization (WHO). *Review of Evidence on Health Aspects of Air Pollution: REVIHAAP Project: Technical Report*. (World Health Organization, Regional Office for Europe, 2021). WHO/EURO: 2013-4101-43860-61757. https://apps.who.int/iris/handle/10665/341712.
7. Manisalidis, I., Stavropoulou, E., Stavropoulos, A. & Bezirtzoglou, E. Environmental and health impacts of air pollution: A review. *Front. Public Health*https://doi.org/10.3389/fpubh.2020.00014 *(2020).*
8. Kampa, M. & Castanas, E. Human health effects of air pollution. *Environ. Pollut.* **151**(2), 362–367. https://doi.org/10.1016/j.envpol.2007.06.012 (2008).
9. Schwartz, J. Air-pollution and daily mortality: A review and meta analysis. *Environ. Res.* **64**(1), 36–52. https://doi.org/10.1006/enrs.1994.1005 (1994).
10. Gilardi, L., Marconcini, M., Metz-Marconcini, A., Esch, T. & Erbertseder, T. Long-term exposure and health risk assessment from air pollution: Impact of regional scale mobility. *Int. J. Health Geogr.* **22**(1), 11. https://doi.org/10.1186/s12942-023-00333-8 (2023).
11. Chen, J. & Hoek, G. Long-term exposure to PM and all-cause and cause-specific mortality: A systematic review and meta-analysis. *Environ. Int.*https://doi.org/10.1016/j.envint.2020.105974 *(2020).*
12. Burnett, R. *et al.* Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. *Proc. Natl. Acad. Sci.* **115**(38), 9592–9597. https://doi.org/10.1073/pnas.1803222115 (2018).
13. European Environment Agency (EEA). *Health Impacts of Air Pollution in Europe, 2022 [Web Page]*. European Environment Agency (EEA). https://www.eea.europa.eu/publications/air-quality-in-europe-2022/health-impacts-of-air-pollution.
14. Liu, X. *et al.* Assessment of German population exposure levels to PM10 based on multiple spatial-temporal data. *Environ. Sci. Pollut. Res.* **27**, 6637–6648. https://doi.org/10.1007/s11356-019-07071-0 (2020).
15. Rittweger, J. *et al.* Temperature and particulate matter as environmental factors associated with seasonality of influenza incidence: An approach using Earth observation-based modeling in a health insurance cohort study from Baden-Wurttemberg (Germany). *Environ. Health.*https://doi.org/10.1186/s12940-022-00927-y *(2022).*
16. Zhang, Y., Wang, S. J., Feng, Z. X. & Song, Y. Influenza incidence and air pollution: Findings from a four-year surveillance study of prefecture-level cities in China. *Front. Public Health.*https://doi.org/10.3389/fpubh.2022.1071229 *(2022).*
17. Villeneuve, P. J. & Goldberg, M. S. Methodological considerations for epidemiological studies of air pollution and the SARS and COVID-19 coronavirus outbreaks. *Environ. Health Perspect.*https://doi.org/10.1289/Ehp7411 *(2020).*
18. Katoto, P. D. M. C. *et al.* Acute and chronic exposure to air pollution in relation with incidence, prevalence, severity and mortality of COVID-19: A rapid systematic review. *Environ. Health.* https://doi.org/10.1186/s12940-021-00714-1 (2021).
19. Marques, M. & Domingo, J. L. Positive association between outdoor air pollution and the incidence and severity of COVID-19: A review of the recent scientific evidences. *Environ. Res.*https://doi.org/10.1016/j.envres.2021.111930 *(2022).*
20. Elminir, H. K. Dependence of urban air pollutants on meteorology. *Sci. Total Environ.* **350**(1–3), 225–237. https://doi.org/10.1016/j.scitotenv.2005.01.043 (2005).
21. Liu, Y. S., Zhou, Y. & Lu, J. X. Exploring the relationship between air pollution and meteorological conditions in China under environmental governance. *Sci. Rep.* https://doi.org/10.1038/s41598-020-71338-7 (2020).
22. Battista, G. & Vollaro, R. D. Correlation between air pollution and weather data in urban areas: Assessment of the city of Rome (Italy) as spatially and temporally independent regarding pollutants. *Atmos. Environ.* **165**, 240–247. https://doi.org/10.1016/j.atmosenv.2017.06.050 (2017).

23. Samad, A. *et al.* Meteorological and air quality measurements in a city region with complex terrain: Influence of meteorological phenomena on urban climate. *Meteorol. Z.*https://doi.org/10.1127/metz/2023/1124 *(2023).*
24. Varotsos, C., Christodoulakis, J., Tzanis, C. & Cracknell, A. P. Signature of tropospheric ozone and nitrogen dioxide from space: A case study for Athens, Greece. *Atmos. Environ.* **89**, 721–730. https://doi.org/10.1016/j.atmosenv.2014.02.059 (2014).
25. Venkitaswamy, S. & Bhaskar, V. Relationship between ozone with nitrogen dioxide and climatic impacts over major cities in India. *Sustain. Environ. Res.* **25**(6), 295–304 (2015).
26. Nguyen, D. H. *et al.* Tropospheric ozone and NOx: A review of worldwide variation and meteorological influences. *Environ. Technol. Innov.* https://doi.org/10.1016/j.eti.2022.102809 (2022).
27. Souza, J. B. *et al.* Generalized additive models with principal component analysis: An application to time series of respiratory disease and air pollution data. *J. R. Stat. Soc. C* **67**(2), 453–480. https://doi.org/10.1111/rssc.12239 (2018).
28. Statheropoulos, M., Vassiliadis, N. & Pappa, A. Principal component and canonical correlation analysis for examining air pollution and meteorological. *Atmos. Environ.* **32**(6), 1087–1095. https://doi.org/10.1016/S1352-2310(97)00377-4 (1998).
29. Stafoggia, M., Breitner, S., Hampel, R. & Basagana, X. Statistical approaches to address multi-pollutant mixtures and multiple exposures: The state of the science. *Curr. Environ. Health Rep.* **4**, 481–490. https://doi.org/10.1007/s40572-017-0162-z (2017).
30. Sun, Z. C. *et al.* Statistical strategies for constructing health risk models with multiple pollutants and their interactions: Possible choices and comparisons. *Environ. Health.* **12**(1), 1–19. https://doi.org/10.1186/1476-069X-12-85 (2013).
31. Höffner, J. Geografie. Staatsministerium Baden-Württemberg. https://www.baden-wuerttemberg.de/de/unser-land/land-und-leute/geografie.
32. Esri, D. Postleitzahlgebiete in Deutschland. Esri Deutschland Content. https://opendata-esri-de.opendata.arcgis.com/datasets/5b203df4357844c8a6715d7d411a8341_0.
33. Klüsener, S. Bevölkerungsdichte in Deutschland (Kreisebene, 2020). Bundesinstitut für Bevölkerungsdichte. https://www.bib.bund.de/DE/Fakten/Fakt/B77-Bevoelkerungsdichte-Kreise.html.
34. European Centre for Medium-Range Weather Forecasts (ECMWF). CAMS European air quality reanalyses. European Centre for Medium-Range Weather Forecasts (ECMWF). https://ads.atmosphere.copernicus.eu/cdsapp#!/dataset/cams-europe-air-quality-reanalyses?tab=overview.
35. Muñoz Sabater, J. ERA5-Land hourly data from 1950 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/cds.e2161bac.
36. Gilardi, L. NO2, O3, PM10 and PM2.5 concentrations - Daily geographical aggregates at ZIP-code level from CAMS European Air Quality Re-analyses. Zenodo. https://doi.org/10.5281/zenodo.8325533 (2023).
37. Lavers, D. A., Simmons, A., Vamborg, F. & Rodwell, M. J. An evaluation of ERA5 precipitation for climate monitoring. *Q. J. R. Meteorol. Soc.* **148**(748), 3152–3165. https://doi.org/10.1002/qj.4351 (2022).
38. Marécal, V. *et al.* A regional air quality forecasting system over Europe: The MACC-II daily ensemble production. *Geosci. Model Dev.* **8**(9), 2777–2813. https://doi.org/10.5194/gmd-8-2777-2015 (2015).
39. Akritidis, D. *et al.* A complex aerosol transport event over Europe during the 2017 Storm Ophelia in CAMS forecast systems: Analysis and evaluation. *Atmos. Chem. Phys.* **20**(21), 13557–13578. https://doi.org/10.5194/acp-20-13557-2020 (2020).
40. World Health Organization (WHO). Ambient (outdoor) air pollution. World Health Organization (WHO). https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health.
41. Minkos, A. *et al.* Luftqualität 2015: Vorläufige Auswertung. Umweltbundesamt; 2016. https://www.umweltbundesamt.de/publikationen/luftqualitaet-2015.
42. Bolton, D. The computation of equivalent potential temperature. *Mon. Weather Rev.* **108**(7), 1046–1053. https://doi.org/10.1175/1520-0493(1980)108<1046:TCOEPT>2.0.CO;2 (1980).
43. Hinkle, D. E., Wiersma, W. & Jurs, S. G. *Applied Statistics for the Behavioral Sciences* (Houghton Mifflin, 2003).
44. Bourke, P. *Cross Correlation, Autocorrelation, 2d Pattern Identification*, vol. 2019 (1996). http://paulbourke.net/miscellaneous/correlate/.
45. Brockwell, P. J. & Davis, R. A. *Introduction to Time Series and Forecasting* (Springer, 2016).
46. Anselin, L. Local indicators of spatial association: Lisa. *Geogr. Anal.* **27**(2), 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x (1995).
47. Müller, I., Erbertseder, T. & Taubenbock, H. Tropospheric $NO_2$: Explorative analyses of spatial variability and impact factors. *Remote Sens. Environ.*https://doi.org/10.1016/j.rse.2021.112839 *(2022).*
48. Getis, A. & Ord, J. K. The analysis of spatial association by use of distance statistics. *Geogr. Anal.* **24**(3), 189–206. https://doi.org/10.1111/j.1538-4632.1992.tb00261.x (1992).
49. Anselin, L. A local indicator of multivariate spatial association: Extending Geary's c. *Geogr. Anal.* **51**(2), 133–150. https://doi.org/10.1111/gean.12164 (2019).
50. R Core Team. *R: A Language and Environment for Statistical Computing*. https://www.R-project.org/.
51. Seinfeld, J. H. & Pandis, S. N. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change* (Wiley, 2016).
52. Han, S. *et al.* Analysis of the relationship between $O_3$, NO and $NO_2$ in Tianjin, China. *Aerosol Air Qual. Res.* **11**(2), 128–139. https://doi.org/10.4209/aaqr.2010.07.0055 (2011).
53. Beirle, S., Boersma, K. F., Platt, U., Lawrence, M. G. & Wagner, T. Megacity emissions and lifetimes of nitrogen oxides probed from space. *Science.* **333**(6050), 1737–1739. https://doi.org/10.1126/science.1207824 (2011).
54. Pommier, M. Estimations of NOx emissions, $NO_2$ lifetime and their temporal variation over three British urbanised regions in 2019 using TROPOMI $NO_2$ observations. *Environ. Sci. Atmos.* **3**(2), 408–421. https://doi.org/10.1039/D2EA00086E (2023).
55. Shah, V. *et al.* Effect of changing NOX lifetime on the seasonality and long-term trends of satellite-observed tropospheric $NO_2$ columns over China. *Atmos. Chem. Phys.* **20**(3), 1483–1495. https://doi.org/10.5194/acp-20-1483-2020 (2020).
56. Kleinman, L. I. *et al.* Dependence of ozone production on NO and hydrocarbons in the troposphere. *Geophys. Res. Lett.* **24**(18), 2299–2302. https://doi.org/10.1029/97GL02279 (1997).
57. Balamurugan, V. *et al.* Tropospheric $NO_2$ and $O_3$ response to COVID-19 lockdown restrictions at the national and urban scales in Germany. *J. Geophys. Res. Atmos.*https://doi.org/10.1029/2021JD035440 *(2021).*
58. Gong, W., Reich, B. J. & Chang, H. H. Multivariate spatial prediction of air pollutant concentrations with INLA. *Environ. Res. Commun.* **3**(10), 101002. https://doi.org/10.1088/2515-7620/ac2f92 (2021).
59. Handschuh, J., Erbertseder, T., Schaap, M. & Baier, F. Estimating PM2.5 surface concentrations from AOD: A combination of SLSTR and MODIS. *Remote Sens. Appl. Soc. Environ.* **26**, 100716. https://doi.org/10.1016/j.rsase.2022.100716 (2022).
60. Handschuh, J., Erbertseder, T. & Baier, F. Systematic evaluation of four satellite AOD datasets for estimating PM2.5 using a random forest approach. *Remote Sens.* **15**(8), 2064. https://doi.org/10.3390/rs15082064 (2023).
61. Klüser, L., Kleiber, P., Holzer-Popp, T. & Grassian, V. H. Desert dust observation from space: Application of measured mineral component infrared extinction spectra. *Atmos. Environ.* **54**, 419–427. https://doi.org/10.1016/j.atmosenv.2012.02.0 (2012).
62. Capraz, O. & Deniz, A. Particulate matter (PM10 and PM2.5) concentrations during a Saharan dust episode in Istanbul. *Air Qual. Atmos. Health.* **14**, 109–116. https://doi.org/10.1007/s11869-020-00917-4 (2021).

## Acknowledgements
We kindly acknowledge the Copernicus Climate Change Service (C3S) and the Copernicus Atmosphere Monitoring Service (CAMS) for providing the environmental datasets.

## Author contributions
LG provided the re-elaborated and pre-processed environmental data. JR, LH and MTS conceived and designed the analysis. LH performed the analysis and wrote the paper. TE wrote parts of the discussion. MTS and MS verified the analytical methods. TE, LG, SW and MB helped interpret the analysis results within physico-chemical atmospheric contexts. All authors reviewed and approved the manuscript.

## Funding
Open Access funding enabled and organized by Projekt DEAL. This research has been supported by the Deutsche Forschungsgesellschaft (DFG) funded project "Influence of air quality on the expected burden on the health care system in the event of pandemics" (Project Number 458531714) and the German Aerospace Center (DLR) funded project "Environmental stressors and health costs".

## Competing interests
The authors declare no competing interests.

## Additional information
**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-56513-4.

**Correspondence** and requests for materials should be addressed to L.H.

# Investigating the spatiotemporal associations between meteorological conditions and air pollution in the federal state Baden-Württemberg (Germany)

Leona Hoffmann[1*], Lorenza Gilardi[2], Marie-Therese Schmitz[3], Thilo Erbertseder[2], Michael Bittner[2], Sabine Wüst[2], Matthias Schmid[3], Jörn Rittweger[1,4]

[1]Institute of Aerospace Medicine, German Aerospace Center (DLR), Cologne, Germany.
[2]German Remote Sensing Data Center, German Aerospace Center (DLR), Weßling, Germany.
[3]Institute of Medical Biometry, Informatics and Epidemiology, University Hospital Bonn, Bonn, Germany.
[4]Department of Pediatrics and Adolescent Medicine, University Hospital Cologne, Cologne, Germany.

*Corresponding author. E-mail: leona.hoffmann@dlr.de;

## Supplementary Information

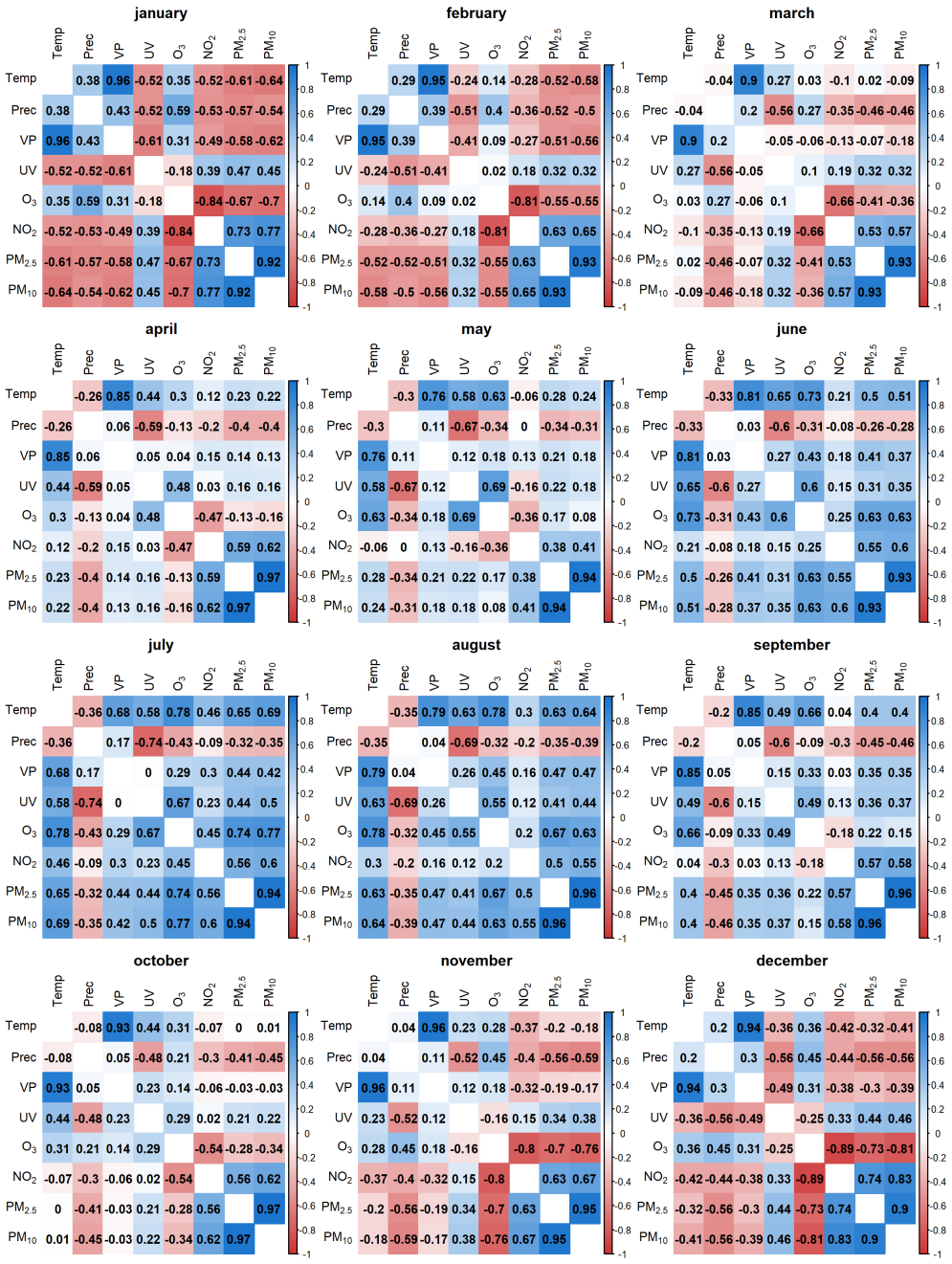


**Fig. S1** Pearson correlation matrices based on daily measurements from 2010 to 2018 across BW split by all months.

## 3.2 Publication 2: Modulation of COVID-19 incidence by environmental stressors is variant between pre-Omicron and Omicron periods

DOI: https://doi.org/10.1038/s41598-025-13521-2

www.nature.com/scientificreports

scientific reports

Check for updates

OPEN

# Modulation of COVID-19 incidence by environmental stressors is variant between pre-Omicron and Omicron periods

Leona Hoffmann[1]✉, Lorenza Gilardi[2], Tobias Antoni[3], Maxana Baltruweit[3], Michael Bittner[2], Susanne Breitner[4,5], Simon Dally[3], Thilo Erbertseder[2], Sabine Hawighorst-Knapstein[3], Marie-Therese Schmitz[6], Rochelle Schneider[7], Sabine Wüst[2] & Jörn Rittweger[1,8,9]

**COVID-19 had a devastating impact on humanity. We investigated how residential air pollution (ozone ($O_3$), nitrogen dioxide ($NO_2$), fine particulate matter ($PM_{2.5}$)) and meteorological factors (temperature (Temp), precipitation (Prec)) are associated with COVID-19 incidence in Baden-Württemberg (BW), Germany. We utilized data from the Copernicus Atmosphere Monitoring Service and the Copernicus Climate Change Service to model environmental exposure from 2020 to 2022 in postal code areas in BW. Health insurance data on SARS-CoV-2 infections were provided from the health insurance AOK BW on a quarterly level covering approximately 12 million person-years. We examined the spatiotemporal variability with a generalized additive model including various stressors, demographic factors, and area-wide data, offering a comprehensive analysis of the environmental stressor- COVI-10 incidence associations. In 2022, during the prevalence of the Omicron variant, the number of COVID-19 cases tripled compared to 2020. During the pre-Omicron period, COVID-19 incidence showed a positive association with $PM_{2.5}$ (relative risk [RR] 2.41; 95% confidence interval [CI] (2.31, 2.52)), a negative association with Temp (RR 0.39 (0.32, 0.48)), and no clear or slight associations with $O_3$, Prec, and $NO_2$. During the Omicron period, there were either no clear or slight negative associations with Temp (RR 0.92 (0.74, 1.30)), $PM_{2.5}$ (RR 0.70 (0.64, 0.79)), $NO_2$, and Prec and a negative association with $O_3$ (RR 0.46 (0.40, 0.53)). The analysis found clear links between environmental stressors and COVID-19 incidence, which strongly differed between pre-Omicron and Omicron periods. Consideration of environmental stressor concentration could be relevant in the management of the pandemic.**

**Abbreviations**

| | |
|---|---|
| AOK | Allgemeine Ortskrankenkasse |
| BW | Baden-Württemberg |
| C3S | Copernicus climate change service |
| CAMS | Copernicus atmosphere monitoring service |
| CI | Confidence interval |
| COVID-19 | Coronavirus disease of 2019 |
| DLR | German aerospace center |
| ECMWF | European centre of medium-range weather forecasts |
| ESA | European space agency |
| GAM | Generalized additive model |
| ICD-10 | International classification of diseases 10th revision |
| IMBIE | Institute of medical biometry, informatics and epidemiology |

[1]Institute of Aerospace Medicine, German Aerospace Center (DLR), Cologne, Germany. [2]German Remote Sensing Data Center, German Aerospace Center (DLR), Weßling, Germany. [3]Allgemeine Ortskrankenkasse Baden-Württemberg (AOK-BW), Stuttgart, Germany. [4]Institute of Epidemiology, Helmholtz Zentrum München - German Research Centre for Environmental Health, Neuherberg, Germany. [5]IBE-Chair of Epidemiology, LMU Munich, Munich, Germany. [6]Institute of Medical Biometry, Informatics and Epidemiology (IMBIE), University Hospital Bonn, Bonn, Germany. [7]Φ-Lab, European Space Agency (ESA), Frascati, Italy. [8]Department of Pediatrics and Adolescent Medicine, University Hospital Cologne, Cologne, Germany. [9]Professor JörnRittweger, passed away. ✉email: .de

| | |
|---|---|
| $NO_2$ | Nitrogene dioxide |
| NPI | Non-pharmaceutical intervention |
| $O_3$ | Ozone |
| $PM_{2.5}$ | Fine particulate matter with a diameter of 2.5 µm or smaller |
| Prec | Precipitation |
| RKI | Robert Koch Institute |
| RR | Relative risk |
| SARS-CoV-2 | Severe acute respiratory syndrome coronavirus 2 |
| Temp | Temperature |

The COVID-19 pandemic has been a fatal disruption at the start of the twenty-first century, with long-reaching consequences. From 2020 to 2022 there were 732 million COVID-19 cases globally, 269 million in Europe, and 37 million in Germany (status: January 1, 2023)[1]. Ambient air pollution is a major risk to human health, with fine particulate matter of diameter ≤ 2.5 µm ($PM_{2.5}$) and nitrogen dioxide ($NO_2$) being prominent contributors. Several previous studies have reported associations between environmental stressors and COVID-19 infections[2,3]. However, most scientific papers on this topic were published in the early days of the pandemic, thus with relatively short observation interval, and with a focus on densely populated areas[4–6]. Review articles have summarized the environmental impact on COVID-19 peaks in 2020 and 2021 for various regions worldwide, including Asia, Europe, America, and the Middle East[7], concluding that exposure to air pollution can facilitate COVID-19 transmission[2,8–12]. More specifically, positive associations with COVID-19 incidence were repeatedly reported for $PM_{2.5}$ and $NO_2$[2,3,7–9], and a negative association with temperature[13–15] . For a comprehensive assessment of environmental effect modulations, it is therefore necessary to analyze a more extended time span, covering the entire 2020–2022 period.

Moreover, the emergence of COVID-19 variants has introduced major antigenic changes to the SARS-CoV-2 virus, with implications for transmissibility, and infection dynamics[16]. These changes have likely influenced not only viral transmissibility and clinical severity, but also the interaction between the virus and external factors[17], including environmental exposures. The Omicron variant, in particular, spread rapidly despite high levels of population immunity[18], suggesting altered infection dynamics and transmission patterns. We therefore hypothesized that the associations between environmental exposures, such as $PM_{2.5}$, $NO_2$, and temperature, and COVID-19 incidence may differ between the pre-Omicron and Omicron periods and conducted a stratified analysis accordingly.

Previous scientific studies have often concentrated on particular urban areas, such as Milan (Italy)[5], Wuhan (China)[19] or Vienna (Austria)[4], or on particular environmental stressors, like $NO_2$[20,21], and particulate matter[22,23]. Urban and rural areas differ in pollution levels, and also in the spread and course of diseases. Our study covered the entire federal state of Baden-Württemberg (Germany), thus including both urban and rural regions and thereby offering a better basis for generalizable findings than spatially selective studies. Moreover, air pollutions have been shown to interact with meteorological factors. For example, ambient temperature and $PM_{2.5}$ can jointly explain seasonal modulation of influenza incidence[24]. The relationships between air pollution parameters and meteorological conditions are complex and not always obvious[25], offering potential for misinterpretation when environmental stressors such as $PM_{2.5}$ are studied stand-alone. A comprehensive understanding also requires a comprehensive set of observations[11]. There is a need to include confounding factors such as age, sex assigned at birth, and population density in the analysis[8,9,26].

Therefore, the present paper presents a comprehensive analysis of the association between COVID-19 incidence and environmental stressors in Baden-Württemberg, Germany, between 2020 and 2022, stratified by the pre-Omicron and Omicron periods.

## The pandemic in Baden-Württemberg

In March 2020, the World Health Organization declared Europe an epicenter of the pandemic, leading to border closures and strict restrictions on public life. Member states implemented contact and exit restrictions to control the spread of the virus. Baden-Württemberg is a federal state in southwestern Germany with a population of 11.28 million, making it more populous than some European countries like Austria or Finland[27]. It is a diverse state with urban, rural, and mountainous regions. The COVID-19 pandemic significantly impacted Baden-Württemberg due to its dynamic pandemic activity. The government implemented several restrictive measures to control the pandemic's spread, including mandatory facial masks, social distancing, school closures, travel and restaurant restrictions, and public and private gatherings limits.

The German Federal Ministry for Economic Affairs and Climate Action has commissioned the creation of a non-pharmaceutical intervention (NPI) to evaluate COVID-19 measures implemented continuously since March 1, 2020[28]. This index, which is provided by infas 360[29], is based on the Oxford COVID-19 Government Response Tracker[30], and amalgamates the various restriction types at any given time into a single score value. The NPI is based on 21 key areas of public life, including gatherings, schools, childcare, events, cultural institutions, retail, nightlife, accommodations, sports, travel restrictions, mask mandates, workplace regulations, curfews, public transport capacity, distancing rules, and testing policies. Each area was coded with specific measures ranked on a scale from 0 ("no restriction") to higher values for more stringent interventions. To make it compatible with our analysis, we aggregated the monthly information as quarterly mean values. The NPI index peaks in Q1 of 2021 (Fig. 1).

The COVID-19 vaccination campaign in Germany began on December 27, 2020. Initially, high-risk individuals and those with occupational exposure were prioritized, until the vaccine eventually became available to everybody in June 2022[31]. The basic immunization rate reached 71% of the BW population by the end of 2022, with the largest increase observed in Q2 and Q3 of 2021 (Fig. 1). Although it was initially assumed that two

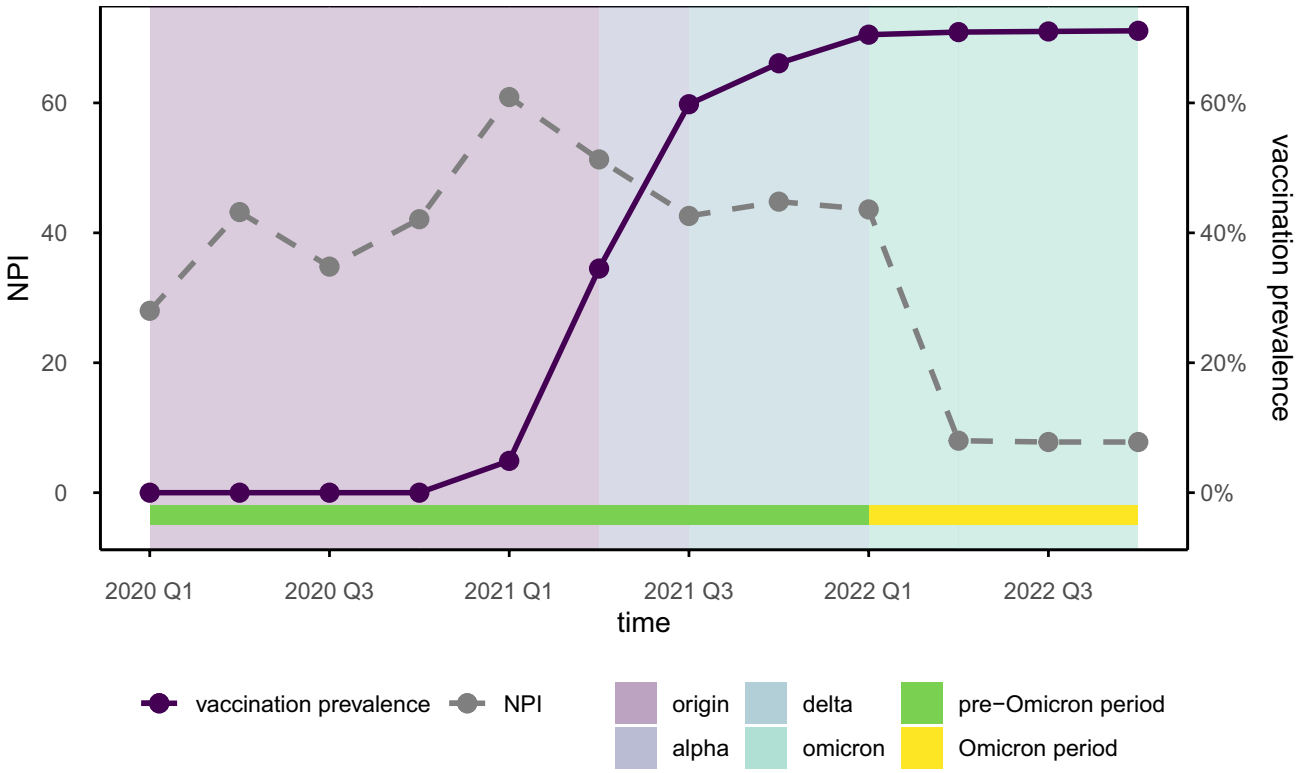


**Fig. 1**. Illustration of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variants, non-pharmaceutical intervention (NPI) and vaccination prevalence in Baden-Württemberg. Vaccination data are aggregated as quarterly means. Reference population for BW as of Dec 31, 2022, is 11.28 M people[34]. SARS-CoV-2 variants are classified quarterly based on the prevailing virus strain. As no specific SARS-CoV-2 variants data are available for BW, it is assumed that the variants behaved similarly as in the rest of Germany. Using the NPI, we could track and observe the government's responses to COVID-19 over time in Baden-Württemberg.

vaccinations suffice for basic immunization, infections occurred frequently despite vaccination[18]. Information on German COVID-19 vaccination is available from Zenodo[32], while information on SARS-CoV-2 variants can be retrieved from the weekly dashboard of the Robert Koch Institute (RKI)[33]. The original SARS-CoV-2 variant dominated up to and including Q1/2021, to be replaced by the Alpha variant in Q2/2021, the Delta variant in Q3/2021, and the Omicron variant in Q1/2022 (Fig. 1).

## Materials and methods

### Health data

The Allgemeine Ortskrankenkasse Baden-Württemberg (AOK BW) is one of Germany's largest health insurance companies, operating since more than 130 years, and covering over 4.6 million people. Our data include the number of COVID-19 infections and the total number of insured persons from 2020 to 2022, aggregated by quarter, postal code area, sex assigned at birth, and the age group in 10-year increments where the numerical value corresponds to the age of the oldest person in the group. As in our previous study[24], the data sources within the AOK BW electronic system were outpatient and inpatient hospital data, sick-leave notes and outpatient diagnoses within the framework of home and specialist centered care and standard care. COVID-19 cases were identified on a quarterly basis via the International Classification of Diseases 10$^{th}$ Revision (ICD-10) codes U07.1 (COVID-19, virus detected) and U07.2 (COVID-19, virus not detected). Using these data sources ensures maximum sensitivity when identifying infected person groups. All methods were conducted by relevant guidelines and regulations. A positive ethics vote and a waiver of informed consent was obtained from the North Rhine Medical Association (number: 2020092), and data processing was approved by AOK BW and DLR. The incidence estimates are based on data from AOK Baden-Württemberg, which covers a large and demographically representative segment of the regional population. Internal analyses by AOK indicate similarity in key factors such as age, sex, and morbidity between the insured and the general population in BW. Nonetheless, incidence estimates apply specifically to the insured group, and caution is advised when generalizing to the entire population.

### Environmental data

Surface-level data on outdoor air pollution, including $NO_2$ (in μg/m$^3$), ozone ($O_3$, in μg/m$^3$), and $PM_{2.5}$ (in μg/m$^3$)[35], based on the Copernicus Atmosphere Monitoring Service (CAMS) Air Quality Reanalysis dataset[36]. The meteorological data precipitation (Prec, in mm/day) and temperature (Temp, in °C) is sourced from the ERA5-

Land, a reanalysis dataset provided by the Copernicus Climate Change Service (C3S) of the European Centre of Medium-Range Weather Forecasts (ECMWF)[37]. Both initial datasets have a 0.1°×0.1° spatial and hourly temporal resolutions. To address the spatial mismatch between the coarser environmental datasets and the finer scale of postal code areas, we applied bilinear interpolation to oversample the spatial resolution of all variables. Based on the interpolated data, daily mean values were computed for each grid cell, and spatial aggregation was then performed by calculating the mean across all grid cells intersecting each postal code area defined by a polygon, following established recommendations[38,39]. For ozone ($O_3$), the daily maximum of the 8-h rolling mean was calculated in line with WHO air quality guidelines[40]. Finally, all daily postal code-level estimates were averaged to the quarterly level for the years 2020 to 2022 to align with the temporal resolution of the health data.

## Data processing

The health and environmental data sets were processed and analyzed using R version 4.3.0[41] and combined using the `merge()` function from R's base package. This merging process was based on three key variables: PLACE, which serves as the geographic identifier for postal code areas; YEAR, representing the calendar year; and QUARTER, indicating the quarter of the year. This approach ensured accurate temporal and geographic alignment of the two data sets. We utilized a shapefile containing five-digit postal codes sourced from the Esri Germany database to manage postal code areas effectively[42].

## Statistical methods

We generated descriptive statistics, categorizing COVID-19 incidence by year, sex assigned at birth, and age (Fig. 2), and with incidence maps for each year (Fig. 3). The overall infection rates by sex were calculated, along with 95% confidence intervals, using R's binom.test() function. In addition, we utilized a generalized additive model (GAM)[43], as previously[24], to investigate the associations between environmental stressors and COVID-19 incidence. The model was implemented using the bam() function from the mgcv package in R[44], employing penalized maximum likelihood estimation (REML) for smoothing parameter selection. The dependent variable was the number of new COVID-19 cases, modeled with a negative binomial distribution. The explanatory variables included both linear and nonlinear terms. We fitted smooth functions with penalized splines for the environmental stressors Temp, $PM_{2.5}$, $NO_2$, $O_3$, and Prec[45]. The categorical variables age group, gender, quarter and year were included as fixed effects to address temporal effects and demographic information. To account for spatial variation, we incorporated a Markov random field (MRF) smooth term for postal code areas (bs = "mrf" in bam()), where the neighborhood structure was automatically constructed from shared polygon boundaries provided via the polygon list[43,44]. The smoothing basis for this term was restricted to 20 dimensions. We included an offset for the number of AOK-insured persons per five-digit postal code to adjust for population density and other features linked to post code (e.g. socio-economic status). A stratified subgroup analysis was performed

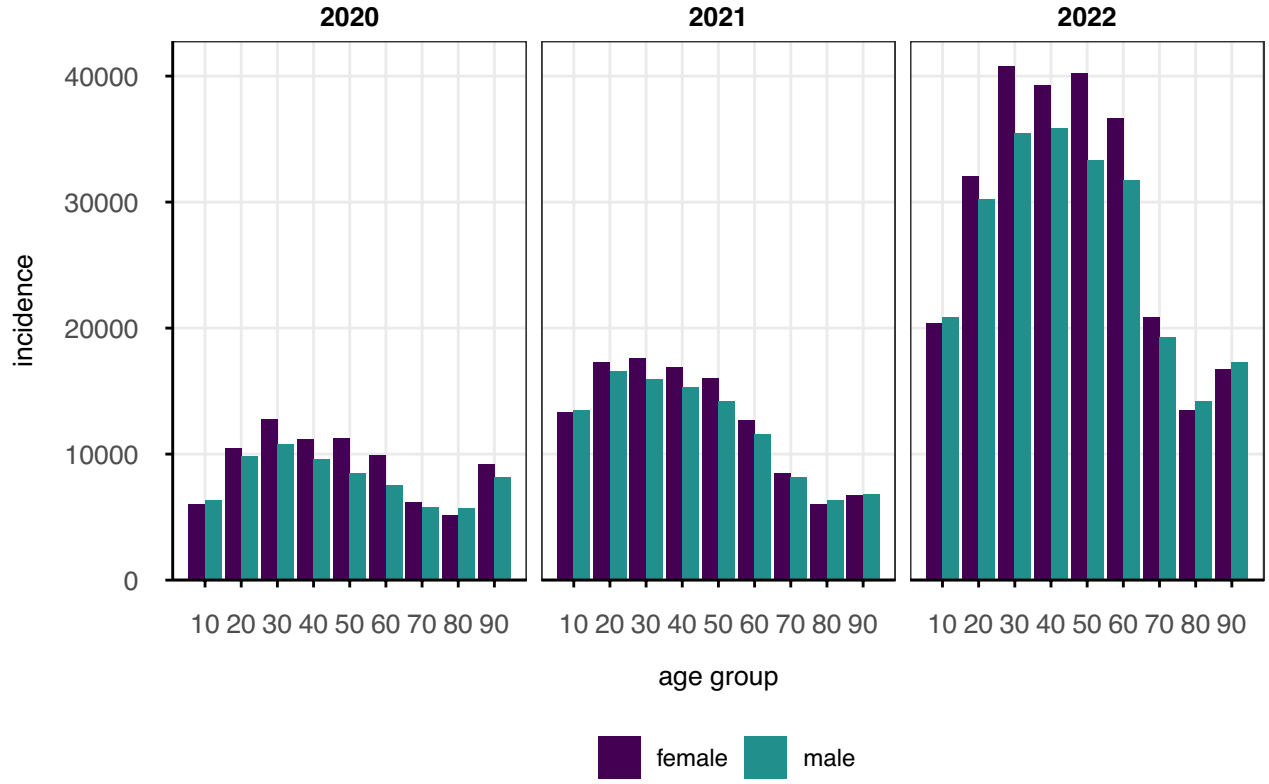


**Fig. 2**. COVID-19 incidence categorized by year, age, and sex assigned at birth. The incidence was defined as the number of COVID-19 cases per 100 thousand AOK-insured individuals per year.

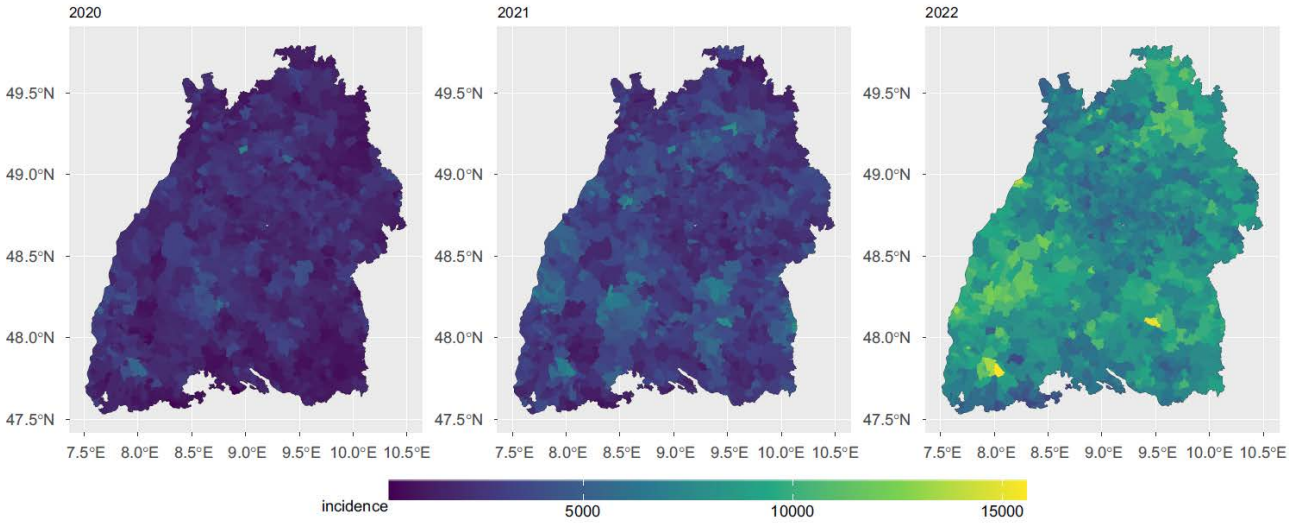


**Fig. 3**. Spatial distribution of COVID-19 incidence by postal code area faceted by year. The incidence was defined as the number of COVID-19 cases per 100 thousand AOK-insured individuals per year.

for the pre-Omicron and Omicron periods by run separate GAMs for each period. We defined the incidence as the number of COVID-19 cases per 100 thousand AOK-insured individuals per year. per year. Based on the GAMs we estimated the incidence depending on the individual environmental stressors and used R's predict() function. The datasets were split into a training set (70%) and a test set (30%) to evaluate model performance. The split was balanced across postal code area and year to preserve the temporal and spatial structure of the data. Model performance was assessed by comparing observed and predicted values[46]. To quantify the impact of the environmental stressors, the relative risk (RR) was calculated as the 95%-to-5% ratio of the environmental stressor predictions, following our previous approach[24]. The 95% confidence intervals for the RR values were calculated using the percentile bootstrap method with 1000 repetitions for each model[46]. To test the robustness of results, we conducted sensitivity analyses excluding the top 1% of extreme values for each air pollutant.

## Results

### Descriptive results

Between 2020 and 2022, the AOK BW reported more than 2 million COVID-19 infections. The incidence rose over the years, with about 10,000 new infections per age group per 100,000 persons reported in 2020, about 15,000 in 2021, and over 30,000 in 2022, indicating that COVID-19 incidence has tripled in 2022 compared to 2020 (Fig. 2). The overall infection rate for females was 53.24% [95% confidence interval (CI): 53.17%, 53.30%] and for males 49.76% [95% CI: 49.69%, 49.83%].

As illustrated in Fig. 3, the geographical distribution of the year-wise increase in incidence showed no clear distinction between urban and rural areas. In 2022, two outstandingly high incidences (yellow spots) were reported in Bernau im Schwarzwald (postal code 79872) and Ertingen (postal code 88521).

The inter-relationships among the various environmental stressors in Baden-Württemberg were discussed previously[25]. The Supplementary Material provides a detailed table of the concentration levels of the environmental stressors (Table S1) and correlation matrices (Figure S1) specifically for the years 2020 to 2022.

### GAM results

Whilst $PM_{2.5}$ (RR 6.27, 95% confidence interval CI (6.05, 6.50), Table 1) and $O_3$ (RR 2.04 (1.94, 2.16), Table 1) were positively associated with COVID-19 incidence, ambient Temp (RR 0.003 (0.003, 0.004), Table 1) was negatively associated (Fig. 4). No clear-cut or weak associations were found for $NO_2$ (RR 1.05 (1.01, 1.10), Table 1) and Prec (RR 1.29 (1.26, 1.33), Table 1).

When comparing pre-Omicron and Omicron periods, the relationships revealed a different picture (Fig. 5). Whilst associations for the pre-Omicron period were very similar to the overall model (Fig. 4), the Omicron period showed a negative association of $O_3$ (RR 0.46 (0.40, 0.53), Table 1) and a slight negative association of $PM_{2.5}$ (RR 0.7 (0.64, 0.79), Table 1) and $NO_2$ (RR 0.78 (0.70, 0.83), Table 1) with COVID-19 incidence and no clear-cut associations for Temp (RR 0.92 (0.74, 1.30), Table 1) and Prec (RR 0.85 (0.83, 0.88), Table 1) concentrations.

According to our GAM models, the expected number of COVID-19 cases was 8–9% lower for males than for females in all three COVID-19 models (RR Overall: 0.92 (0.91, 0.92); RR Omicron: 0.91 (0.91,0.92); RR pre-Omicron: 0.92 (0.91, 0.92)). Comparing 2022 to 2020, the expected number of cases increased by a factor of 3.97 for the COVID-19 overall model. Compared to the 50–60 age group, the expected number of new COVID-19 cases was lower for the 70–80 age group (RR Overall: 0.49 (0.48, 0.50); RR Omicron: 0.40 (0.39, 0.41); RR pre-Omicron: 0.57 (0.65, 0.58)) and higher for the 20–30 age group (RR Overall: 1.23 (1.21,1.24); RR Omicron: 1.04 (1.03, 1.06); RR pre-Omicron: 1.38 (1.36, 1.40)).

| | Output | COVID-19 Overall | pre-Omicron period | Omicron period |
|---|---|---|---|---|
| $PM_{2.5}$ | Prediction<br>5-Percentile<br>Prediction<br>95-Percentile<br>**Relative Risk** | 6072.98<br>(5745.25, 6419.41)<br>38,085.95<br>(36,176.79, 40,095.87)<br>**6.27 (6.05, 6.50)** | 2548.32<br>(2322.88, 2795.63)<br>6136.83<br>(5550.48, 6785.12)<br>**2.41 (2.31, 2.52)** | 17,571.66<br>(12,571.68, 19,723.30)<br>12,304.79<br>(10,147.77, 14,920.32)<br>**0.70 (0.64, 0.79)** |
| $NO_2$ | Prediction<br>5-Percentile<br>Prediction<br>95-Percentile<br>**Relative Risk** | 8892.68<br>(8417.80, 9394.36)<br>9351.00<br>(8816.68, 9917.69)<br>**1.05 (1.01, 1.10)** | 3121.89<br>(2835.20, 3437.58)<br>2188.84<br>(1982.89, 2416.17)<br>**0.70 (0.66, 0.74)** | 16,372.97<br>(13,591.75, 19,723.30)<br>12,701.18<br>(10,524.07, 15,328.68)<br>**0.78 (0.70, 0.83)** |
| $O_3$ | Prediction<br>5-Percentile<br>Prediction<br>95-Percentile<br>**Relative Risk** | 5653.11<br>(5287.54, 6043.96)<br>11,538.9<br>(10,982.98, 12,122.97)<br>**2.04 (1.94, 2.16)** | 4255.69<br>(3840.01, 4716.37)<br>4263.45<br>(3873.63, 4692.50)<br>**1.00 (0.93, 1.08)** | 27,664.95<br>(22,033.03, 34,736.47)<br>12,696.59<br>(10,628.35, 15,167.29)<br>**0.46 (0.40, 0.53)** |
| Temp | Prediction<br>5-Percentile<br>Prediction<br>95-Percentile<br>**Relative Risk** | 361,166.50<br>(332,729.00, 392,034.40)<br>1232.60<br>(1154.46, 1316.03)<br>**0.003 (0.003, 0.004)** | 12,184.78<br>(10,552.70, 14,069.29)<br>4785.45<br>(4326.25, 5293.39)<br>**0.39 (0.32, 0.48)** | 15,975.74<br>(12,630.65, 20,206.74)<br>14,769.59<br>(13,639.19, 15,993.68)<br>**0.92 (0.74, 1.30)** |
| Prec | Prediction<br>5-Percentile<br>Prediction<br>95-Percentile<br>**Relative Risk** | 7653.14<br>(7270.31, 8056.12)<br>9896.35<br>(9441.55, 10,373.05)<br>**1.29 (1.26, 1.33)** | 2925.48<br>(2658.21, 3219.62)<br>3927.30<br>(3601.46, 4282.63)<br>**1.34 (1.29, 1.39)** | 16,543.51<br>(13,761.47, 19,887.97)<br>14,109.23<br>(10,628.35, 16,905.05)<br>**0.85 (0.83, 0.88)** |

**Table 1**. Estimated COVID-19 incidence and associations to environmental stressors during different periods of the pandemic. The table enlisted the estimated number of cases per 100,000 persons per year for 5% and 95% percentiles of the environmental stressors split by the overall model (Fig. 4) and the pre-Omicron and Omicron period (Fig. 5). The values of the confidence interval are given in brackets. Green cells indicated RR > 1.3 (strong positive association), yellow cells RR < 0.7 (strong negative association), and uncolored cells $0.7 \leq RR \leq 1.3$ (weak or no association). The incidence was defined as the number of COVID-19 cases per 100 thousand AOK-insured individuals per year.

After splitting the datasets to training and test sets, the Root Mean Squared Error (RMSE) values and scatter plots of actual versus predicted values and predicted values versus residuals (see Fig. S2) indicated that it performed consistently well: the RMSE was 5.03 and 9.69 for the pre-Omicron and the Omicron periods, respectively, and 9.87 for the overall COVID-19 data.

Removing the top 1% of exposure values in separate sensitivity analyses for the models Overall, Omicron period and pre-Omicron period did not substantially change the shape or strength of the exposure–response functions (see Figs. S3 and S4). This supports the robustness of the main findings to potential outliers or extreme exposure events.

## Discussion

Our analysis aimed to comprehensively investigate the associations between environmental stressors and COVID-19 over time, encompassing multiple stressors and demographic factors. The model incorporated temporal effects as independent variables. While non-pharmaceutical interventions (NPIs), immunization rates, population density, and circulating SARS-CoV-2 variants were not explicitly modeled, quarterly aggregation allows for partial adjustment for these changing factors. Postal code areas are included as MRF in the GAMs. The models included age group and sex assigned at birth as fixed effects to adjust for their potential confounding effect. Three models with the same structure but different time periods were fitted, the Overall period (2020–2022), pre-Omicron period (2020–2021) and Omicron period (2022). For the pre-Omicron period, the directions of association were broadly consistent with findings for influenza incidence before the COVID-19 pandemic[24]. In the Omicron period, these associations were entirely lost (Temp), or even reverted ($PM_{2.5}$) (Table 1). It is an attractive hypothesis to ascribe the extensive loss of the association between environmental stressor and COVID-19 to the exaggerated virulence of the Omicron variant.

The Omicron variant is considered to be a particularly contagious and vaccine-resistant mutation[47–49]. During the period dominated by the Omicron variant, the main driver of the epidemic changed from contact rates to contagiousness[50]. Models suggest that the Omicron variant of SARS-CoV-2 may be up to 10 times more transmissible than the original strain and 2.8 times more transmissible than the Delta variant. Furthermore, it has been estimated that there is an 88% probability of Omicron evading the current vaccines[48]. On the other hand, the Omicron variant seems to result in less severe outcomes in terms of hospitalization, ventilation therapy, and death compared to previous variants[51–53].

Most studies established a positive association between air pollutants, such as $PM_{2.5}$, $NO_2$, and $O_3$, and respiratory diseases in general[54], and COVID-19 in particular[2,3,7–9]. However, two review articles also indicated negative associations, as shown by Monoson[55] and Carballo[3].

A global meta-analysis of short-term exposure to air pollution discovered that COVID-19 incidence was positively associated with $PM_{2.5}$[12]. A second meta-analysis found that 10 μg/m$^3$ of $PM_{2.5}$ increased odds of infection by 66%[56]. A review highlighted that most studies reviewed demonstrated a positive connection

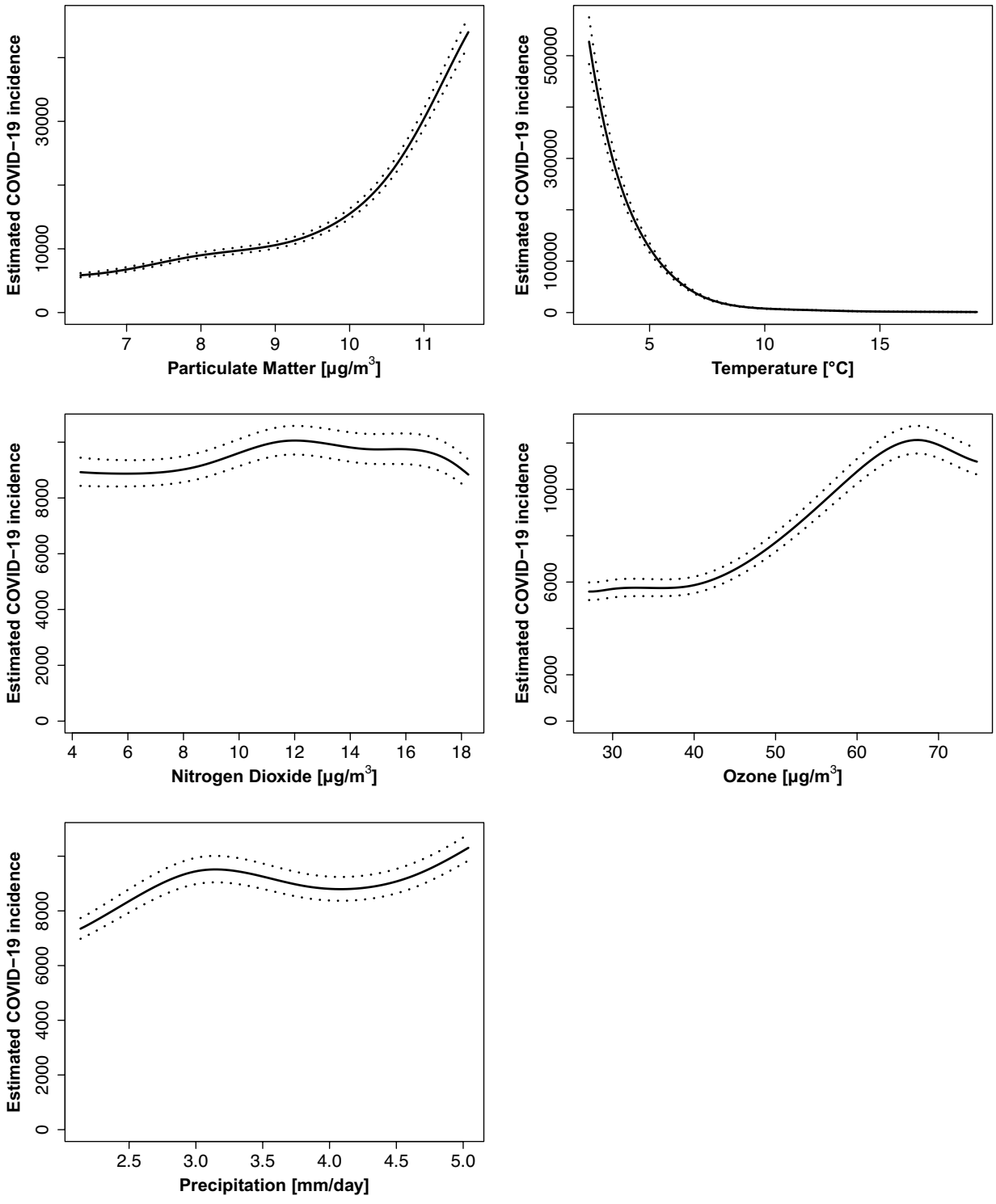


**Fig. 4**. Estimated COVID-19 incidence per 100,000 persons per year in response to each environmental stressor. All other parameters were kept constant, and were set to the following values: Age group: 50–60 years, sex assigned at birth: female, postal code: 70376, quarter: Q2, year: 2021 and environmental stressors: median. Estimates were obtained using the R function predict for Temp, Prec, $NO_2$, $O_3$, and $PM_{2.5}$, including their 5th and 95th percentiles.

between COVID-19 infection and $PM_{2.5}$ exposure to air pollution over both short and long term[9]. The positive association is also detected in the United Kingdom[57,58] and Italy[13,59,60]. Additionally, a study in Germany revealed that each one-unit increase in $PM_{2.5}$ resulted in almost 200 more cases of COVID-19 per 100,000 inhabitants by February 2021[22]. There is a biologically plausible mechanism that supports the high infection rate associated with $PM_{2.5}$ exposure[61]. A positive correlation between $NO_2$ and COVID-19 incidence was often found in the

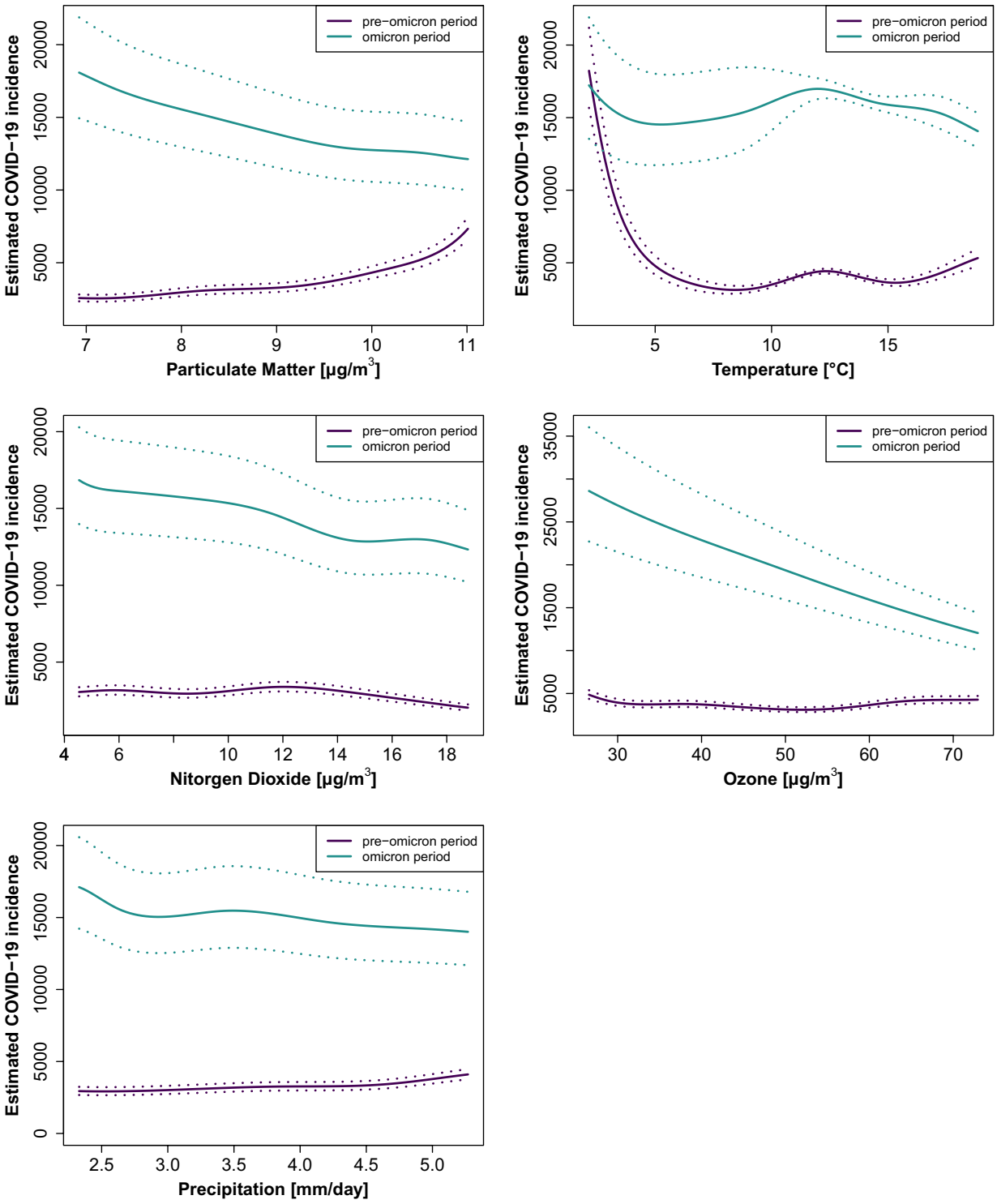


**Fig. 5**. Estimated COVID-19 incidence per 100,000 persons per year in response to each environmental stressor based on pre-Omicron (2020, 2021, purple) and Omicron (2022, teal) model. All other parameters were kept constant and were set to the following values: quarter: Q2, year (optional): 2021, postal code: 70376, sex assigned at birth: female, age group: 50 to 60 years, and environmental stressor: median. Estimates were obtained using the R function predict for Temp, Prec, $NO_2$, $O_3$, and $PM_{2.5}$, including their 5th and 95th percentiles.

literature[4,9,12,59], especially in Germany during the observation period of COVID-19 cases until September 2020[62]. On the other hand, two studies conducted in Italy indicated a negative correlation between $NO_2$ levels and the incidence of COVID-19, at the beginning of the pandemic[20,63]. There is no clear association direction for $O_3$ in the literature. Positive as well as negative associations between $O_3$ exposure and COVID-19 incidence

have been found[63–65]. A study across 409 cities in 26 countries also found little evidence of meteorological conditions like precipitation influencing COVID-19 transmission[66]. The literature has shown a link between higher temperatures and lower COVID-19 incidence rates, indicating a negative association[13–15].

According to our COVID-19 models, the relative risk of contracting COVID-19 was lower for men than women, with a coefficient between 0.91 and 0.92 across all models. Northern Italy showed a similar RR value of 0.96 for men[60]. However, in general, men are thought to be more vulnerable to COVID-19 infections and more likely to experience severe symptoms of the disease[67,68].

The above findings during the pre-Omicron period align with our model. COVID-19 incidence is positively associated with $PM_{2.5}$ and negatively associated with Temp. Our model shows weak or no links with $NO_2$, $O_3$, and Prec, which are not clearly or distinctly described in the literature.

As far as we know, no study has been conducted on the relationship between the incidence of COVID-19 and environmental stressors covering the entire period from 2020 to 2022. Renard's study examined the same time frame for Western Europe but with regard to COVID-19 mortality, reporting a positive correlation between short term exposure to $PM_{2.5}$ and mortality[69]. However, our analysis focused on COVID-19 incidence. The overall model generally has similar trends to influenza incidence found from 2010 to 2018[24]. The literature classification is similar to the pre-Omicron period. The deviating positive association between incidence and $O_3$ aligns with the literature mentioned.

To our knowledge, almost no studies have isolated the Omicron period from the preceding periods in terms of the relationship between COVID-19 incidence and environmental stressors. This is an area that requires further investigation in future studies. Only one study on the relationship between new infections and temperature in Verona up to and including July 2022 explicitly considered the Omicron period[70]. In this study, the previously observed negative correlation between temperature and the number of new infections is no longer observed due to the dominance of the Omicron variant. Instead, a positive correlation is observed. The paper discussed that the virus and its biological properties may have continuously changed over time. It was also suggested that the extremely high transmissibility of the virus could diminish the influence of environmental stressors.

Our study has high value due to its broad period, area-wide analysis, multiple stressors, demographic factors, high acquisition rate (notifiable disease in Germany)[71], and differentiation between COVID-19 overall, Omicron, and pre-Omicron variants. To the best of our ability, our analysis incorporated Villeneuve's commentary on methodological considerations for epidemiological studies of air pollution and COVID-19[72]. We introduced confounding variables such as age and sex and used high spatial resolution to ensure accuracy. Using an MRF smoother captures constant spatial structures, but not recorded time-varying, location-specific factors in postcode areas. As a result, potential spatial confounding factors may be unconsidered. While the GAM offers valuable insights, like any modeling approach, it approximates reality and cannot fully capture the complexity of real-world dynamics. A restriction of our analysis is the limited temporal resolution in quarters. This prevents us from examining potential short-term lag effects of environmental stressors on COVID-19 incidence; future studies with data of finer temporal resolution—when they become available—could explore this aspect in more detail. In addition, COVID-19 cases tend to be underestimated due to the dynamic situation regarding tests, doctor's visits, and specific diagnoses. We excluded the NPI, the ratio of basic immunization, and SARS-CoV-2 variants due to their link to the quarter. Their exclusion avoided multicollinearities and improved model interpretation while indirectly retaining information via quarter. This study relies on observational data, which means that the identified associations should not be interpreted as causal relationships. Although we have adjusted for relevant covariates, factors that were not measured and potential reverse causality—such as pollution reductions resulting from mobility restrictions—could affect the observed associations. Our analysis was limited to COVID-19 incidence and cases. Only outdoor air pollution was considered. Future analysis may benefit from including the year 2023, which was also dominated by Omicron.

To study the impact of outdoor air pollution, we used the validated regional reanalysis from the Copernicus Atmosphere Monitoring Service (CAMS)[73]. Validation studies and continuous evaluation ensure its data quality. An evaluation study for Germany showed an underestimation of $PM_{2.5}$ exposure compared to in-situ measurements[74]. Its spatial resolution represents background conditions but has proven reliable for air pollution and health risk assessment studies on small geographical areas[38]. At the time of our study, the CAMS regional reanalysis was found to be the most comprehensive data set available for capturing air pollution variability.

## Conclusion

The results obtained from our models highlight the complex association between environmental stressors and COVID-19 incidence. In the pre-Omicron period, the impact of these environmental stressors appears comparable to their known influence on influenza incidence. Higher levels of $PM_{2.5}$ and $O_3$ are linked to increased COVID-19 incidence, while temperature shows a negative association. In contrast, associations with Prec and $NO_2$ are weak. During the Omicron period however, these associations change, particularly for Temp, $PM_{2.5}$ and $O_3$. All stressors, with the exception of $O_3$, show weak or no clear cut-off association with COVID-19 incidence. Notably, the effects of $O_3$ reverses to a negative association. This change may reflect Omicron's increased infectiousness and transmissibility, potentially reducing the response to environmental stressors. These findings nourish the hypothesis that Omicron is sufficiently contagious that environmental factors no longer have a measurable impact on COVID-19 incidence. In contrast, during the pre-Omicron period, reductions in air pollution might have contributed to lowering COVID-19 transmission, as has been observed for other respiratory infections.

## Data availability

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

Received: 19 December 2024; Accepted: 24 July 2025
Published online: 29 July 2025

## References

1. World Health Organization (WHO). WHO COVID-19 dashboard. https://data.who.int/dashboards/covid19/data (2023).
2. Copat, C. et al. The role of air pollution (PM and NO2) in COVID-19 spread and lethality: A systematic review. *Environ. Res.* **191**, 110129 (2020).
3. Carballo, I. H., Bakola, M. & Stuckler, D. The impact of air pollution on COVID-19 incidence, severity, and mortality: A systematic review of studies in Europe and North America. *Environ. Res.* **215**, 114155 (2022).
4. Hutter, H.-P. et al. Air pollution is associated with COVID-19 incidence and mortality in Vienna, Austria. *Int. J. Environ. Res. Public Health* **17**, 9275 (2020).
5. Zoran, M. A., Savastru, R. S., Savastru, D. M. & Tautan, M. N. Assessing the relationship between surface levels of PM2.5 and PM10 particulate matter impact on COVID-19 in Milan Italy. *Sci. Total Environ.* **738**, 139825 (2020).
6. Li HaX, X., Dai, D. W., Huang, Z., Ma, Z. & Guan, Y. J. Air pollution and temperature are associated with increased COVID-19 incidence: A time series study. *Int. J. Infect. Dis.* **97**, 278–282 (2020).
7. Marques, M. & Domingo, J. L. Positive association between outdoor air pollution and the incidence and severity of COVID-19. A review of the recent scientific evidences. *Environ. Res.* **203**, 111930 (2022).
8. Ali, N. et al. Exposure to air pollution and COVID-19 severity: A review of current insights, management, and challenges. *Integr. Environ. Assess. Manag.* **17**, 1114–1122 (2021).
9. Ali, N. & Islam, F. The effects of air pollution on COVID-19 infection and mortality—A review on recent evidence. *Front. Public Health* **8**, 580057 (2020).
10. Comunian, S., Dongo, D., Milani, C. & Palestini, P. Air pollution and COVID-19: The role of particulate matter in the spread and increase of COVID-19's morbidity and mortality. *Int. J. Environ. Res. Public Health* **17**, 4487 (2020).
11. Maleki, M., Anvari, E., Hopke, P. K., Noorimotlagh, Z. & Mirzaee, S. A. An updated systematic review on the association between atmospheric particulate matter pollution and prevalence of SARS-CoV-2. *Environ. Res.* **195**, 110898 (2021).
12. Zang, S.-T. et al. Ambient air pollution and COVID-19 risk: Evidence from 35 observational studies. *Environ. Res.* **204**, 112065 (2022).
13. De Angelis, E. et al. COVID-19 incidence and mortality in Lombardy, Italy: An ecological study on the role of air pollution, meteorological factors, demographic and socioeconomic variables. *Environ. Res.* **195**, 110777 (2021).
14. Guo, C. et al. Meteorological factors and COVID-19 incidence in 190 countries: An observational study. *Sci Total Environ* **757**, 143783 (2021).
15. Yuan, J. et al. Association between meteorological factors and daily new cases of COVID-19 in 188 countries: A time series analysis. *Sci. Total Environ.* **780**, 146538 (2021).
16. Harvey, W. T. et al. SARS-CoV-2 variants, spike mutations and immune escape. *Nat. Rev. Microbiol.* **19**, 409–424 (2021).
17. Vasireddy, D., Vanaparthy, R., Mohan, G., Malayala, S. V. & Atluri, P. Review of COVID-19 variants and COVID-19 vaccine efficacy: What the clinician should know?. *J. Clin. Med. Res.* **13**, 317 (2021).
18. Telenti, A., Hodcroft, E. B. & Robertson, D. L. The evolution and biology of SARS-CoV-2 variants. *Cold Spring Harb. Perspect. Med.* **12**, a041390 (2022).
19. Jiang, Y., Wu, X.-J. & Guan, Y.-J. Effect of ambient air pollutants and meteorological variables on COVID-19 incidence. *Infect. Control Hosp. Epidemiol.* **41**, 1011–1015 (2020).
20. Ogen, Y. Assessing nitrogen dioxide (NO2) levels as a contributing factor to coronavirus (COVID-19) fatality. *Sci. Total Environ.* **726**, 138605 (2020).
21. Huang, G. & Brown, P. E. Population-weighted exposure to air pollution and COVID-19 incidence in Germany. *Spatial Stat.* **41**, 100480 (2021).
22. Prinz, A. L. & Richter, D. J. Long-term exposure to fine particulate matter air pollution: An ecological study of its effect on COVID-19 cases and fatality in Germany. *Environ. Res.* **204**, 111948 (2022).
23. Czwojdzińska, M., Terpińska, M., Kuźniarski, A., Płaczkowska, S. & Piwowar, A. Exposure to PM2.5 and PM10 and COVID-19 infection rates and mortality: A one-year observational study in Poland. *Biomed. J.* **44**, S25–S36 (2021).
24. Rittweger, J. et al. Temperature and particulate matter as environmental factors associated with seasonality of influenza incidence—An approach using Earth observation-based modeling in a health insurance cohort study from Baden-Wurttemberg (Germany). *Environ. Health-Glob.* **21**, 131 (2022).
25. Hoffmann, L. et al. Investigating the spatiotemporal associations between meteorological conditions and air pollution in the federal state Baden-Württemberg (Germany). *Sci. Rep.-UK* **14**, 5997 (2024).
26. Katoto, P. D. M. C. et al. Acute and chronic exposure to air pollution in relation with incidence, prevalence, severity and mortality of COVID-19: A rapid systematic review. *Environ. Health-Glob.* **20**, 1–21 (2021).
27. Baden-Württemberg. Baden-Württemberg in Zahlen: Bevölkerung. [07/15/2024]. https://www.baden-wuerttemberg.de/de/unser-land/land-und-leute/bevoelkerung/.
28. Bundesministerium für Wirtschaft und Klimaschutz. Die Corona-Datenplattform: Grundlage für eine effiziente und effektive Politik in Zeiten der Pandemie. Bundesministerium für Wirtschaft und Klimaschutz; [07/15/2024]. https://www.bmwk.de/Redaktion/DE/Coronavirus/corona-datenplattform.html.
29. infas360. Massnahmenindex Bundesländer pro Monat. healthcare daten plattform; 2023 [07/15/2024]. https://www.healthcare-datenplattform.de/dataset/massnahmenindex_bundeslaender_pro_monat.
30. Hale T, Angrist N, Kira B, Petherick A, Phillips T, Webster S. Variation in government responses to COVID-19. *Blavatnik Centre for Government Working Paper* (2020).
31. Waize, M. et al. Die Impfung gegen COVID-19 in Deutschland zeigt eine hohe Wirksamkeit gegen SARS-CoV-2-Infektionen, Krankheitslast und Sterbefälle. *Epid. Bull.* **35**, 3–10 (2021).
32. Wünsche H. COVID-19-Impfungen in Deutschland. Robert Koch Institut. https://zenodo.org/records/10874856 (2024).
33. Robert Koch Institut. SARS-CoV-2 Varianten in Deutschland: Daten aus der integrierten genomischen Surveillance von SARS-CoV-2. Robert Koch Institut; [07/15/2024]. https://public.data.rki.de/t/public/views/IGS_Dashboard/DashboardVOC?%3Aembed=y%26%3AisGuestRedirectFromVizportal=y.
34. Baden-Württemberg Statistisches Landesamt. Pressemitteilung 129/2023: 11,28 Millionen: Noch nie gab es so viele Menschen im Südwesten - Enorme Zuwanderung im Jahr 2022, aber deutlich weniger Geburten und mehr Sterbefälle als 2021. Baden-Württemberg statistisches Landesamt; 2023 [07/15/2024]. https://www.statistik-bw.de/Presse/Pressemitteilungen/2023129.
35. Gilardi L. NO2, O3, PM10 and PM2.5 concentrations - Daily geographical aggregates at ZIP-code level from CAMS European Air Quality Re-analyses. Zenodo. https://doi.org/10.5281/zenodo.8325533 (2023).

36. European Centre for Medium-Range Weather Forecasts (ECMWF). CAMS European air quality reanalyses. European Centre for Medium-Range Weather Forecasts (ECMWF); [07/15/2024]. https://ads.atmosphere.copernicus.eu/cdsapp#!/dataset/cams-europe-air-quality-reanalyses?tab=overview.
37. Muñoz Sabater J. ERA5-Land hourly data from 1981 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/cds.e2161bac (2019).
38. Gilardi, L., Marconcini, M., Metz-Marconcini, A., Esch, T. & Erbertseder, T. Long-term exposure and health risk assessment from air pollution: Impact of regional scale mobility. *Int. J. Health Geogr.* **22**, 11 (2023).
39. Stroh, E., Harrie, L. & Gustafsson, S. A study of spatial resolution in pollution exposure modelling. *Int. J. Health Geogr.* **6**, 1–13 (2007).
40. World Health Organization (WHO). WHO global air quality guidelines. Particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. 2021 [06/16/2025]. https://www.ncbi.nlm.nih.gov/books/NBK574594/.
41. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. (2023).
42. Esri Deutschland. Postleitzahlgebiete in Deutschland. [07/15/2024]. https://opendata-esri-de.opendata.arcgis.com/datasets/5b203df4357844c8a6715d7d411a8341_0.
43. Wood, S. *Generalized Additive Models: An Introduction with R* 2nd edn. (Chapman and Hall/CRC, 2017).
44. Wood, S. Package 'mgcv'. *R Package Version* **1**, 729 (2025).
45. Eilers, P. H. & Marx, B. D. Flexible smoothing with B-splines and penalties. *Stat. Sci.* **11**, 89–121 (1996).
46. Hastie, T., Tibshirani, R., Friedman, J. H. & Friedman, J. H. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* 2nd edn. (Springer, 2009).
47. Ren, S.-Y., Wang, W.-B., Gao, R.-D. & Zhou, A.-M. Omicron variant (B. 1.1. 529) of SARS-CoV-2: Mutation, infectivity, transmission, and vaccine resistance. *World J. Clin. Cases* **10**, 1 (2022).
48. Chen, J., Wang, R., Gilby, N. B. & Wei, G.-W. Omicron variant (B. 1.1. 529): Infectivity, vaccine breakthrough, and antibody resistance. *J. Chem. Inf. Model.* **62**, 412–22 (2022).
49. Hoffmann, M. et al. The Omicron variant is highly resistant against antibody-mediated neutralization: Implications for control of the COVID-19 pandemic. *Cell* **185**, 447–56 e11 (2022).
50. Kendall, M., Ferretti, L., Wymant, C., Tsallis, D., Petrie, J., Di Francia, A. et al. Drivers of epidemic dynamics in real time from daily digital COVID-19 measurements. *Science*, eadm8103 (2024).
51. Relan, P. et al. Severity and outcomes of Omicron variant of SARS-CoV-2 compared to Delta variant and severity of Omicron sublineages: A systematic review and metanalysis. *BMJ Glob. Health* **8**, e012328 (2023).
52. Kim, K. et al. The case fatality rate of COVID-19 during the delta and the omicron epidemic phase: A meta-analysis. *J. Med. Virol.* **95**, e28522 (2023).
53. Wang L, Berger NA, Kaelber DC, Davis PB, Volkow ND, Xu R. COVID infection rates, clinical outcomes, and racial/ethnic and gender disparities before and after Omicron emerged in the US. *MedRxiv* (2022).
54. Thurston GD, Kipen, H., Annesi-Maesano, I., Balmes, J., Brook, R. D., Cromar, K. A joint ERS/ATS policy statement: What constitutes an adverse health effect of air pollution? An analytical framework. *Eur. Respir. J.* 49 (2017).
55. Monoson, A. et al. Air pollution and respiratory infections: The past, present, and future. *Toxicol. Sci.* **192**, 3–14 (2023).
56. Sheppard, N., Carroll, M., Gao, C. & Lane, T. Particulate matter air pollution and COVID-19 infection, severity, and mortality: A systematic review and meta-analysis. *Sci. Total Environ.* **880**, 163272 (2023).
57. Travaglio, M. et al. Links between air pollution and COVID-19 in England. *Environ. Pollut.* **268**, 115859 (2021).
58. Scalsky, R. J., Chen, Y.-J., Ying, Z., Perry, J. A. & Hong, C. C. The social and natural environment's impact on SARS-CoV-2 infections in the UK Biobank. *Int. J. Environ. Res. Public Health* **19**, 533 (2022).
59. Ranzi A, Stafoggia M, Giannini S, Ancona C, Bella A, Cattani G, et al. Long-term exposure to ambient air pollution and the incidence of SARS-CoV-2 infections in Italy: The EpiCovAir study. *Epidemiologia e Prevenzione* 47 (2023).
60. Veronesi, G., De Matteis, S., Calori, G., Pepe, N. & Ferrario, M. M. Long-term exposure to air pollution and COVID-19 incidence: A prospective study of residents in the city of Varese, Northern Italy. *Occup. Environ. Med.* **79**, 192–199 (2022).
61. Marchetti, S. et al. On fine particulate matter and COVID-19 spread and severity: An in vitro toxicological plausible mechanism. *Environ. Int.* **179**, 108131 (2023).
62. Huang G, Brown PE. Population-weighted exposure to air pollution and COVID-19 incidence in Germany. *Spat. Stat.* 41 (2021).
63. Zoran, M. A., Savastru, R. S., Savastru, D. M. & Tautan, M. N. Assessing the relationship between ground levels of ozone (O3) and nitrogen dioxide (NO2) with coronavirus (COVID-19) in Milan, Italy. *Sci. Total Environ.* **740**, 140005 (2020).
64. Bilal, Bashir MF. et al. Environmental pollution and COVID-19 outbreak: Insights from Germany. *Air Qual. Atmos. Health* **13**, 1385–94 (2020).
65. Isphording, I. E. & Pestel, N. Pandemic meets pollution: Poor air quality increases deaths by COVID-19. *J. Environ. Econ. Manag.* **108**, 102448 (2021).
66. Sera, F. et al. A cross-sectional analysis of meteorological factors and SARS-CoV-2 transmission in 409 cities across 26 countries. *Nat. Commun.* **12**, 5968 (2021).
67. Gebhard, C., Regitz-Zagrosek, V., Neuhauser, H. K., Morgan, R. & Klein, S. L. Impact of sex and gender on COVID-19 outcomes in Europe. *Biol. Sex Differ.* **11**, 1–13 (2020).
68. Pijls BG, Jolani S, Atherley A, Dijkstra JI, Franssen GH, Hendriks S, et al. Temporal trends of sex differences for COVID-19 infection, hospitalisation, severe disease, intensive care unit (ICU) admission and death: a meta-analysis of 229 studies covering over 10M patients. *F1000Research* 11 (2022).
69. Renard, J.-B. et al. Relation between PM2.5 pollution and Covid-19 mortality in Western Europe for the 2020–2022 period. *Sci. Total Environ.* **848**, 157579 (2022).
70. Mattiuzzi, C., Henry, B. M. & Lippi, G. Regional association between mean air temperature and case numbers of multiple SARS-CoV-2 lineages throughout the pandemic. *Viruses* **14**, 1913 (2022).
71. Diercke, M., Claus, H., Rexroth, U. & Hamouda, O. Anpassung des Meldesystems gemäß Infektionsschutzgesetz im Jahr 2020 aufgrund von COVID-19. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* **64**, 388 (2021).
72. Villeneuve, P. J. & Goldberg, M. S. Methodological considerations for epidemiological studies of air pollution and the SARS and COVID-19 coronavirus outbreaks. *Environ. Health Persp.* **128**, 095001 (2020).
73. Marécal, V. et al. A regional air quality forecasting system over Europe: The MACC-II daily ensemble production. *Geosci. Model Dev.* **8**, 2777–2813 (2015).
74. Handschuh, J., Erbertseder, T. & Baier, F. On the added value of satellite AOD for the investigation of ground-level PM2.5 variability. *Atmos. Environ.* **331**, 120601 (2024).

## Author contributions

T.A., M.Ba., M.Bi., S.B., S.W., and J.R. made the funding acquisition. L.H., T.A., M.Ba., and J.R. conceptualized the project. L.H., MT.S., and J.R. developed the methodology. L.H., L.G., S.D., and T.E. were responsible for data curation. L.H. did the formal analysis, made the visualization, and wrote the software. T.A., M.Ba., S.D., and S.HK. provided resources. L.H. wrote the original draft. All authors reviewed and approved the manuscript.

## Funding

Open Access funding enabled and organized by Projekt DEAL. This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) [Grant no 458531714] and internal funding at DLR and AOK-BW.

## Declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-13521-2.

**Correspondence** and requests for materials should be addressed to L.H.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

# Supplemented Material

# Modulation of COVID-19 incidence by environmental stressors is variant between pre-Omicron and Omicron periods

Leona Hoffmann[1*], Lorenza Gilardi[2], Tobias Antoni[3], Maxana Baltruweit[3], Michael Bittner[2], Susanne Breitner[4,5], Simon Dally[3], Thilo Erbertseder[2], Sabine Hawighorst-Knapstein[3], Marie-Therese Schmitz[6], Rochelle Schneider[7], Sabine Wüst[2], Jörn Rittweger[1,8]

[1]Institute of Aerospace Medicine, German Aerospace Center (DLR), Cologne, Germany
[2]German Remote Sensing Data Center, German Aerospace Center (DLR), Weßling, Germany
[3]Allgemeine Ortskrankenkasse Baden-Württemberg (AOK-BW), Stuttgart, Germany
[4] Institute of Epidemiology, Helmholtz Zentrum München - German Research Centre for Environmental Health, Neuherberg, Germany
[5] IBE-Chair of Epidemiology, LMU Munich, Munich, Germany.
[6]Institute of Medical Biometry, Informatics and Epidemiology (IMBIE), University Hospital Bonn, Bonn, Germany
[7]Φ-lab, European Space Agency (ESA), Frascati, Italy
[8]Department of Pediatrics and Adolescent Medicine, University Hospital Cologne, Cologne, Germany

[*]Corresponding author. Email: leona.hoffmann@dlr.de

| | Statistical Parameter | 2020 | 2021 | 2022 |
|---|---|---|---|---|
| **$PM_{2.5}$** | Mean (SD) | 8.43 (1.07) | 8.89 (1.34) | 8.51 (2.08) |
| | Median [Min, Max] | 8.39 [5.25, 12.1] | 8.49 [6.21, 12.0] | 7.84 [5.30, 14.3] |
| **$NO_2$** | Mean (SD) | 11.1 (4.48) | 10.5 (4.53) | 9.34 (3.97) |
| | Median [Min, Max] | 10.6 [2.59, 23.7] | 9.73 [2.77, 22.9] | 8.75 [2.78, 20.5] |
| **$O_3$** | Mean (SD) | 52.8 (16.3) | 52.6 (13.5) | 58.1 (16.5) |
| | Median [Min, Max] | 58.3 [20.3, 81.2] | 53.1 [24.9, 78.6] | 65.4 [23.6, 87.1] |
| **Temp** | Mean (SD) | 10.6 (5.89) | 9.02 (5.61) | 10.9 (5.91) |
| | Median [Min, Max] | 9.74 [2.09, 21.1] | 8.10 [0.36, 18.5] | 10.7 [1.56, 20.9] |
| **Prec** | Mean (SD) | 3.24 (0.76) | 3.85 (0.84) | 2.98 (0.66) |
| | Median [Min, Max] | 3.17 [1.45, 6.16] | 3.63 [2.36, 6.98] | 2.74 [1.68, 5.65] |

**Table S1: Overview of environmental stressors.** Overview of meteorological data and air pollutants, including statistical parameters mean, standard deviation (SD), median, minimum (min), and maximum (max) values. The data cover all postal code areas in BW and are based on quarterly measurements taken from 2020 to 2022.

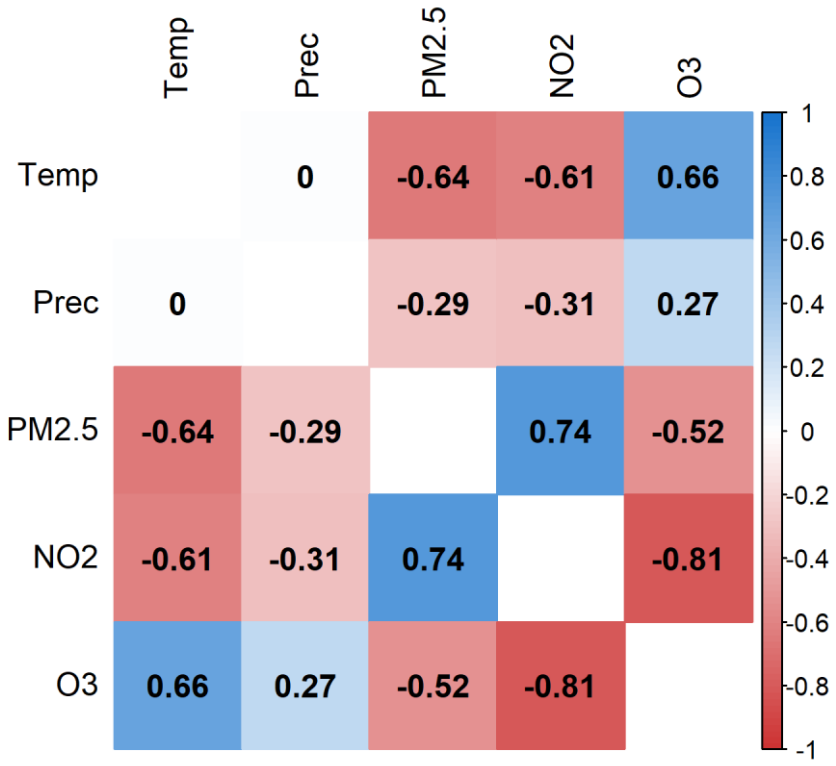


**Figure S1: Pearson correlation matrix based on quarterly measurements from 2020 to 2022 across the German federal state Baden-Württemberg (BW).** The intensity of the color indicates the strength of the correlation between the two variables, with blue indicating positive correlations and red indicating negative correlations.

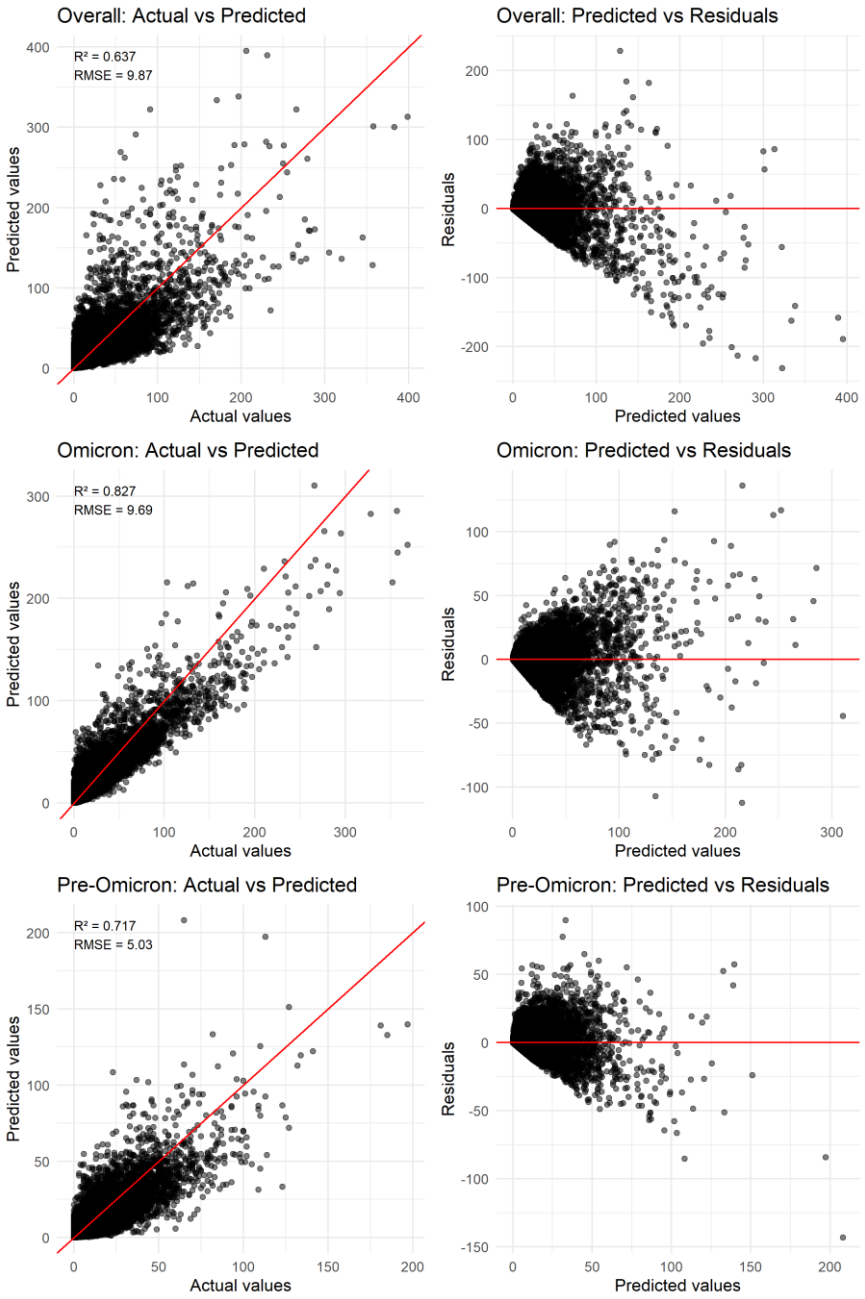


**Figure S2: Scatter plots of actual versus predicted values and predicted values versus residuals for all three models: Overall, Omicron and pre-Omicron period.** Additional, $R^2$ and RMSE are provided.

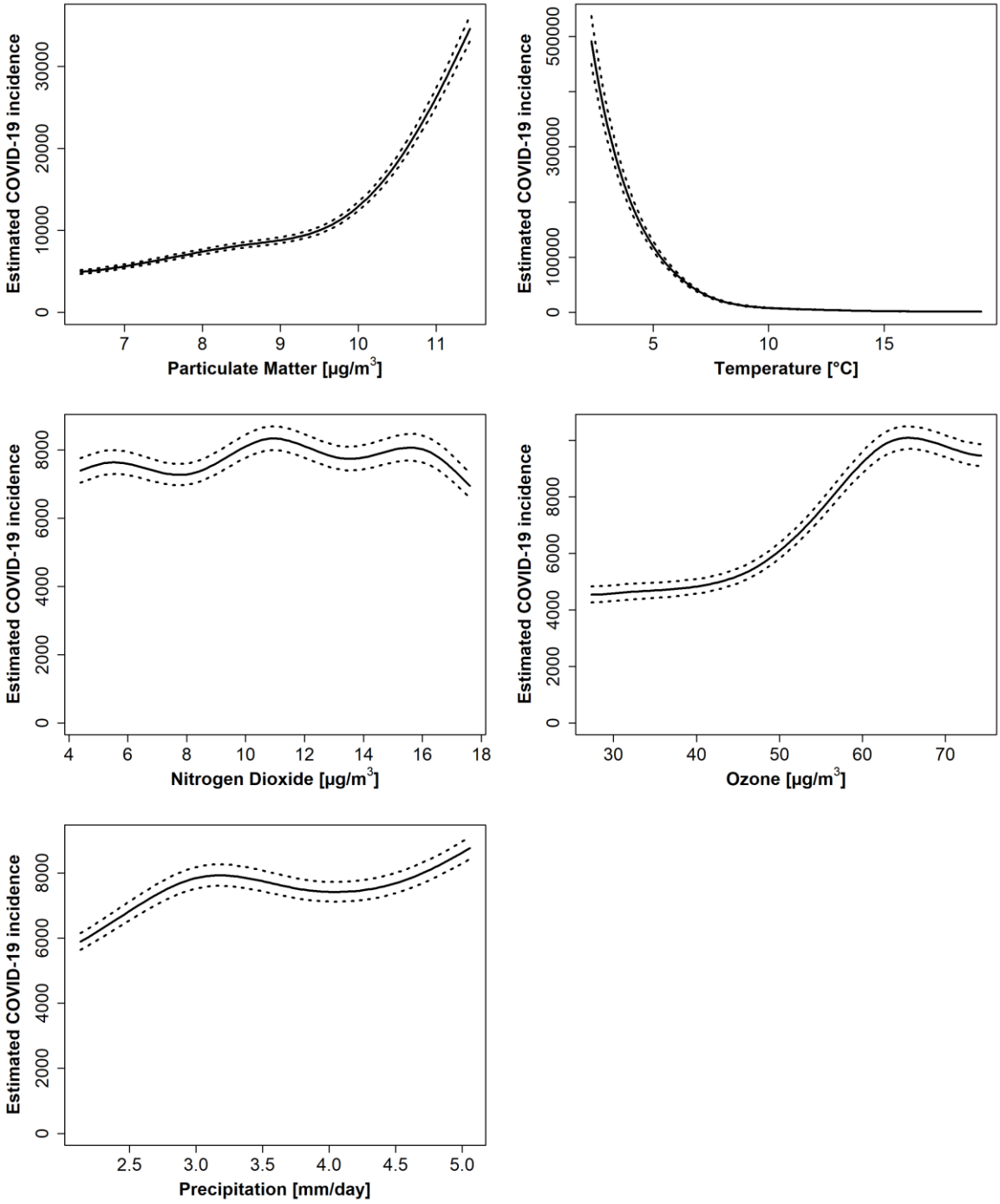


**Figure S3: Sensitivity analyses excluding extreme pollution values (top 1%): Estimated COVID-19 incidence per 100,000 persons per year in response to each environmental stressor.** All other parameters were kept constant, and were set to the following values: Age group: 50-60 years, sex assigned at birth: female, postal code: 70376, quarter: Q2, year: 2021 and environmental stressors: median. Estimates were obtained using the R function predict for Temp, Prec, $NO_2$, $O_3$, and $PM_{2·5}$, including their 5th and 95th percentiles.

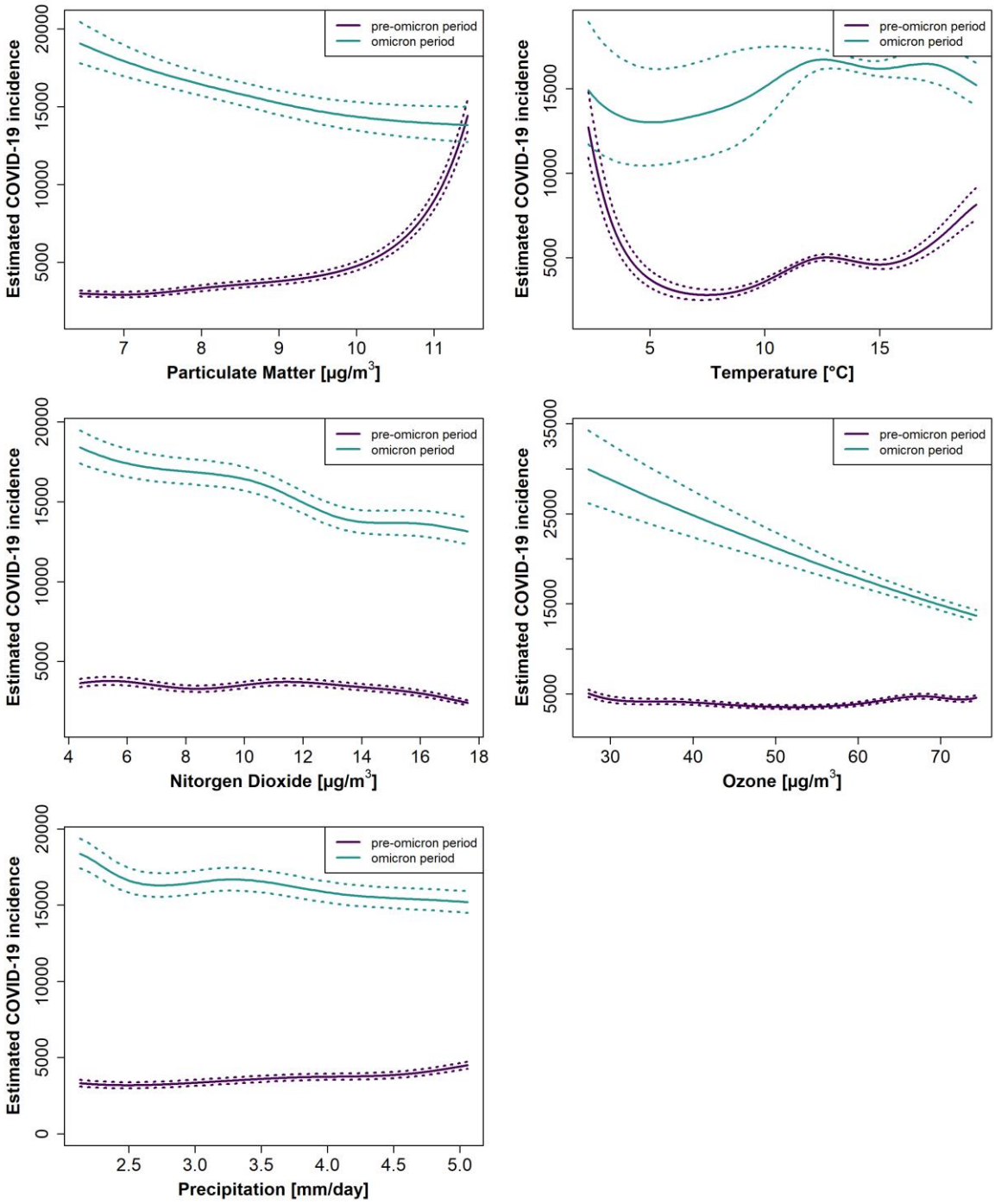


**Figure S4: Sensitivity analyses excluding extreme pollution values (top 1%): Estimated COVID-19 incidence per 100,000 persons per year in response to each environmental stressor based on pre-Omicron (2020, 2021, purple) and Omicron (2022, teal) model.** All other parameters were kept constant and were set to the following values: quarter: Q2, year (optional): 2021, postal code: 70376, sex assigned at birth: female, age group: 50 to 60 years, and environmental stressor: median. Estimates were obtained using the R function predict for Temp, Prec, $NO_2$, $O_3$, and $PM_{2\cdot5}$, including their 5th and 95th percentiles.

## 3.3 Publication 3: Associations between COVID-19 incidence and environmental stressors in Brazil: a nationwide study from 2020 to 2022

DOI: https://doi.org/10.3389/fenvh.2025.1635503

frontiers | Frontiers in Environmental Health

TYPE Original Research
PUBLISHED 03 October 2025
DOI 10.3389/fenvh.2025.1635503



OPEN ACCESS

EDITED BY
Mehdi Hazari,
United States Environmental Protection Agency (EPA), United States

REVIEWED BY
Sanjay Kumar Mehta,
SRM Institute of Science and Technology (Deemed to be University) Research Kattankulathur, India
Elif Yıldırım,
Konya Technical University, Türkiye

*CORRESPONDENCE
Leona Hoffmann
✉ leona.hoffmann@dlr.de

†Deceased

RECEIVED 26 May 2025
ACCEPTED 17 September 2025
PUBLISHED 03 October 2025

CITATION
Hoffmann L, Viana Jacobson L, Erbertseder T, Berger M, Rittweger J and Schneider R (2025) Associations between COVID-19 incidence and environmental stressors in Brazil: a nationwide study from 2020 to 2022. Front. Environ. Health 4:1635503. doi: 10.3389/fenvh.2025.1635503



# Associations between COVID-19 incidence and environmental stressors in Brazil: a nationwide study from 2020 to 2022

Leona Hoffmann[1]*, Ludmilla Viana Jacobson[2], Thilo Erbertseder[3], Moritz Berger[4,5], Jörn Rittweger[1,6†] and Rochelle Schneider[7,8]

[1]Institute of Aerospace Medicine, German Aerospace Center (DLR), Cologne, Germany, [2]Department of Statistics, Institute of Mathematics and Statistics, Fluminense Federal University, Niterói, Brazil, [3]German Remote Sensing Data Center, German Aerospace Center (DLR), Weßling, Germany, [4]Institute of Medical Biometry, Informatics and Epidemiology (IMBIE), University Hospital Bonn, Bonn, Germany, [5]Core Facility Biostatistics, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany, [6]Department of Pediatrics and Adolescent Medicine, University Hospital Cologne, Cologne, Germany, [7]Φ-lab, European Space Agency (ESA), Frascati, Italy, [8]Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom

**Background:** Environmental stressors such as temperature (Temp), relative humidity (RHumid), and fine particulate matter ($PM_{2.5}$) may influence the incidence of COVID-19. While many studies have examined these associations in Europe and Asia, research in Brazil—a country with diverse climatic zones and a high burden of COVID-19—remains limited.

**Objective:** This study aimed to assess the associations between environmental stressors and COVID-19 incidence at the municipality level across Brazil over a three-year period, differentiating between climate zones and pre-Omicron/Omicron periods.

**Methods:** We utilized a generalized additive model (GAM) framework to analyze monthly COVID-19 incidence while adjusting for population size, spatial structure, and temporal trends. Distributed lag nonlinear models (DLNM) were used to evaluate lagged exposure-response associations. Separate models were fitted for five climate zones to assess regional variations.

**Results:** In the overall analysis, Temp was positively associated with COVID-19 incidence [relative risk RR 2.47, 95% confidence interval (2.04, 2.91)], while $PM_{2.5}$ [RR 1.03 (0.95, 1.11)] and RHumid [RR 1.02 (0.91, 1.13)] did not demonstrate clear effects. Climate zone-specific analyses revealed diverse effects: Temp had a positive association with COVID-19 in Temperado (TE) [RR 17.9 (15.26, 22.19)] and Tropical Brazil Central [RR 1.87 (1.57, 2.10)], but a negative association in Tropical Nordeste Oriental [RR 0.008 (0.004, 0.012)] and Tropical Zona Equatorial (TZE) [RR 0.12 (0.08, 0.15)] climate zones. RHumid showed varying positive and negative associations depending on the climate zone, while high levels of $PM_{2.5}$ are positive associated with COVID-19 incidence in zones TE [RR 2.10 (1.93, 2.28)] and TZE [RR 1.87 (1.54, 2.31)]. DLNM results revealed parabolic lag response curves, with extreme values of Temp and RHumid raising risks in certain zones.

**Significance:** Our study provides a comprehensive, long-term analysis of environmental stressors and COVID-19 incidence across diverse climate zones in Brazil. The results reveal considerably spatial and temporal variations in how Temp, RHumid, and $PM_{2.5}$ influence COVID-19 incidence. These findings emphasize the importance of considering regional climatic conditions when assessing environmental risk factors for COVID-19. Understanding these associations can inform targeted public health interventions and preparedness strategies for future respiratory disease outbreaks.

KEYWORDS

COVID-19, temperature, $PM_{2.5}$, relative humidity, DLNM, Brazil

## 1 Introduction

The SARS-CoV-2 virus and the associated COVID-19 disease have had a deep impact on global health and society. Even years after the initial outbreak, continued analysis of the pandemic remains important to deepen our understanding of the virus and to inform strategies for managing potential future pandemics.

Various studies have shown an association between environmental stressors and new COVID-19 cases, consistently demonstrating a link across different timeframes and geographical areas (1–3). Air pollution as well as meteorological factors like temperature and relative humidity, significantly impacts on human health. Especially particulate matter which originates from natural and anthropogenic sources is known to be associated with respiratory diseases (4), including COVID-19 (5–7). In addition to the influence of environmental stressors, several studies have also shown the combined effects with socio-economic and demographic parameters like income, government health expenditure or population density on COVID-19 (8–10). A study in China highlighted the importance of considering socio-economic factors, such as population density, when analyzing particulate matter concentrations and their potential impacts (11).

Research has further demonstrated that the associations between environmental stressors and health outcomes differ considerably between tropical and non-tropical regions (12, 13). Brazil's extensive and diverse landscape encompasses a wide range of climatic zones and meteorological conditions, making it a particularly suitable setting for examining differential effects of the COVID-19 pandemic across varying ecological contexts.

On February 26, 2020, Brazil reported the first case of a SARS-CoV-2 infection in South America. Following the World Health Organization's declaration of SARS-CoV-2 as a pandemic on March 11, 2020, Brazil's government implemented various measures to curb the spread of the virus. These measures included social isolation, school closures, mandatory mask use, and event cancellations. Despite these interventions, travel between Brazilian federal states remained largely unrestricted (14). Like many countries, Brazil experienced various virus mutations called variants of concern. The Gamma variant was the dominant strain from March 2021 until the Delta variant replaced it in October 2021. As of January 2022, the Omicron variant became the dominant strain (15). Vaccinations in Brazil began on January 17, 2021, with a limited supply of doses. Several factors contributed to delays in vaccinating the Brazilian population, including the lack of a consistent national vaccination campaign, hesitancy among the population to receive the vaccine, distrust in the healthcare system, and political rivalries (16). In January 2022, the vaccination rate for the first dose was 75.6%, while the rate for the second dose was 67.4% (17).

Previous studies investigating exposure-response effects in Brazil primarily focused on the early stages of the pandemic and were mostly limited to urban areas (12, 13, 18). These studies utilized various methods, including generalized additive models (12), principal component analysis (13), and linear regression models (18). Another study analyzed the indirect response of temperature, humidity, and rainfall on the spread of COVID-19 across five cities in the Indian Monsoon region between April 26 and December 5, 2020, emphasizing humidity and temperature as key factors influencing the transmission of the virus (19). Sarkodie and Owusu (10) conducted a global assessment of the influence of air pollution, climatic variables, and socio-economic factors on the spread of the COVID-19 pandemic for 615 cities, but confined to the early period from Janurary 1 to June 11, 2020. A study of Turkey and all 12 NUTS level 1 regions highlighted the importance of considering environmental factors in pandemic management and investigated the relationship between air pollution ($PM_{10}$ and $SO_2$) and the number of COVID-19 cases during the early pandemic phase (June – November 2020) (20).

However, to achieve a more comprehensive understanding, it is essential to analyze data over a longer period and across a broader geographic scope. Our study provides a long-term (three-year), countrywide, municipality-level analysis of the association between environmental stressors and COVID-19 incidence in Brazil from 2020 to 2022. By systematically analyzing data from all municipalities, rather than from selected single cities, we gained a better understanding of how environmental factors link to COVID-19 incidence, including

Abbreviations

AOK, Allgemeine Ortskrankenkasse; BW, Baden-Württemberg; COVID-19, Corona Virus Disease 2019; DLNM, Distributed Lag Nonlinear Model; EQ, Equatorial; GAM, Generalized Additive Model; $PM_{2.5}$, Fine Particulate Matter with a diameter of 2.5 µm or smaller; RHumid, Relative Humidity; Temp, Temperature; TBC, Tropical Brasil Central; TE, Temperado; TNO, Tropical Nordeste Oriental; TZE, Tropical Zona Equatorial; WHO, World Health Organization.

the effects of virus mutations. Our research utilizes monthly aggregated data and combines generalized additive models (GAM) and distributed lag nonlinear models (DLNM), providing a robust analysis of longer-term trends, lagged associations, and regional variability.

The specific research questions of this study are as follows:

1. How are meteorological factors (temperature and relative humidity) and air pollutants ($PM_{2.5}$) associated with COVID-19 incidence across Brazilian municipalities and climate zones?
2. Do these associations differ between pre-Omicron and Omicron periods?
3. How do the lagged effects of environmental stressors contribute to trends in COVID-19 incidence?

This article is organized as follows: Section 2 describes the Materials and Methods, including health and environmental data, data processing and the statistical modeling framework. Section 3 presents the results of the analyses, divided into descriptive results, generalized additive model (GAM) findings and distributed lag nonlinear model (DLNM) results. Section 4 gives a short summary of the results, discusses the findings in relation to existing literature, addresses political implications and points out limitations. Finally, Section 5 concludes with the main contributions of the study.

# 2 Materials and methods

## 2.1 Health data

Notification records for COVID-19 cases for all Brazilian municipalities (from March 2020 to December 2022), as reported by State and Municipal Health Departments, were obtained at https://covid.saude.gov.br (last access March 2025). The Brazilian Ministry of Health provides an interactive dashboard with the number of new confirmed cases and deaths by epidemiological week of notification. For this study, we aggregated data by month.

## 2.2 Environmental data

To analyze the monthly mean concentration of particulate matter with a diameter of 2.5 μm or smaller ($PM_{2.5}$, in $\mu g/m^3$) over the three-year period, we utilized regional estimates from the Atmospheric Composition Analysis Group at the University of Wahington (21) and van Donkelaar (22). This high-resolution dataset ($0.01° \times 0.01°$) is well-established in the literature for Brazilian data (23–25), and incorporates information from satellite, simulation, and monitor-based sources. The air temperature variable at two meters above ground level (Temp, in °C) was obtained from the ERA5-Land global reanalysis dataset provided by the Copernicus Climate Change Service of the European Centre of Medium-Range Weather Forecasts (26). The same data source was utilized to calculate relative humidity (RHumid, in %) using the Magnus equation (27), which incorporated dew point temperature and air temperature at two meters above ground level. The data was accessed through Google Earth Engine (28).

## 2.3 Data processing

All data processing was done in Python (29). The health dataset contained infection cases within month-municipality aggregates, and we added zero counts for those aggregate windows without cases to ensure completeness for subsequent analysis. In order to ensure compatibility between the health and environmental datasets, municipality codes were converted from 6-digit format to 7-digit format. For this matching, the official matching table provided by the Brazilian Statistics Institute was used, and code mismatches and exceptional cases were corrected manually (30). Three very small and sparsely populated municipalities were excluded due to missing environmental data; their omission is unlikely to affect the overall analysis because of their small population sizes and geographic isolation. Details on code conversions and exclusions are provided in Supplementary Material S1. To ensure precise alignment, the merging of the datasets was based on three key variables: the seven-digit municipality code, month, and year. The municipality shapefile based on 2020 and the national climate zone information are sourced from the Brazilian Institute of Geography and Statistics (30). The climate zone data showed three small climate zones near the coast, which we relabeled them according to their neighboring climate zones, resulting in five main climate zones: Equatorial (EQ), Tropical Zona Equatorial (TZE), Tropical Brasil Central (TBC), Tropical Nordeste Oriental (TNO), and Temperado (TE). In municipalities with multiple climate zones, we identified the largest and set it as the dominant climate zone. We converted the temperature from Kelvin to degrees Celsius.

Plausibility checks on the health dataset were conducted, revealing an anomalous high number of infections in June 2020. We decided that this particular month significantly deviated from expected trends and exceeded other records by an unrealistic margin. Therefore, given the lack of corroborating evidence and the high likelihood of a data entry error, we removed this record from the analysis to maintain data integrity.

## 2.4 Statistical methods

Statistical analysis was carried out using R version 4.4.2 (31). To examine the distribution of average Temp, RHumid, and $PM_{2.5}$ concentrations across all municipalities, choropleth maps were created with the ggplot2 package in R as a prerequisite for the further statistical analyses. In addition, climate zones and population density were mapped as choropleth maps to provide a better context for understanding spatial patterns and regional variations. We created line plots for the temporal descriptive analysis with months on the $x$-axis and incidence, RHumid, Temp, and $PM_{2.5}$ levels on the $y$-axis. The years 2020–2022 were color-coded. Descriptive metrics for the environmental

stressors were summarized by means with standard deviations and median with ranges.

We used two different statistical methods to investigate the research question about the association between environmental stressors and COVID-19 incidence: A Generalized Additive Model (GAM) and a Distributed Lag Non-linear Model (DLNM). The GAM was mainly applied to determine the effect of the three aforementioned environmental stressors; the DLNM was additionally applied to determine a possible (non-linear) lagged functional association.

The GAM model was implemented with the *"bam()"* function from the mgcv package in R (32). We used the number of new COVID-19 cases per municipality as the dependent variable and applied the log-transformed population as an offset to express cases per 100,000 people. The exposure of interest included mean Temp, $PM_{2.5}$, and RHumid per municipality. For each exposure we specified natural cubic splines with 3 degrees of freedom to capture potential nonlinear links. Additionally, we incorporated categorical variables for both month and year to adjust for temporal variations. To adjust for spatial dependencies between municipalities, we utilized a Markov Random Field based on geographic identifiers specific to each municipality (33). A stratified subgroup analysis was performed for the five climate zones. We fitted the models assuming a negative binomial distribution for the dependent variable, given substantial overdispersion in the outcome (variance-to-mean ratio > 10,000) and applied Restricted Maximum Likelihood for unbiased estimation of variance components, enhancing the model's reliability. The GAM model assumptions were confirmed using gam.check() plots of the mgvc package in R (32). We used the R's *"predict()"* function with the fitted GAM to make predictions and calculate confidence intervals based on standard error. Additionally, we stratified for the pre-Omicron and Omicron periods for each climate zone and the overall area to assess differences in environmental stressors. While the model structure remained the same, we reduced the number of kernels in the Markov Random Field from 100 to 75 for the stratified analysis of the climate zones over different periods. To enhance interpretation, we expressed predicted cases as incidence per 100,000 people per year. The GAM model for each analysis—stratified by climate zone, pre-Omicron and Omicron periods, and overall—can be expressed as:

$$\begin{aligned}\log\left(E[Y_{it}|X_{it}]\right) = \alpha &+ \mathrm{ns}_{\mathrm{Temp}}(\mathrm{Temp}_{\mathrm{it}}) + \mathrm{ns}_{\mathrm{PM}_{2.5}}(\mathrm{PM2.5}_{\mathrm{it}}) \\ &+ \mathrm{ns}_{\mathrm{Humid}}(\mathrm{Humid}_{\mathrm{it}}) + u_i + \beta_1\,\mathrm{Year}_t \\ &+ \beta_2\,\mathrm{Month}_t + \log\left(\mathrm{PopTotal}_{\mathrm{it}}\right)\end{aligned}$$

Where $Y_{\mathrm{it}}$ denotes the COVID-19 incidence cases in month $t$ for municipality $i$, $X_{\mathrm{it}}$ denotes the set of independent variables, ns() are natural cubic splines, $u_i$ is the Markov Random Field term, $\beta_1\mathrm{Year}_t + \beta_2\mathrm{Month}_t$ represents temporal effects, and $\log(\mathrm{PopTotal}_{\mathrm{it}})$ is an offset for population size.

As a sensitivity analysis, the overall model was extended by including interaction terms between the period (Omicron/pre-Omicron) and the environmental stressors to formally assess potential differences in exposure–response relationships between the two periods.

As previously (4), we calculated relative risks (RR) for each environmental factor, defined as the ratio of predicted cases at the 95th percentile to those at the 5th percentile of the environmental variable. The percentile bootstrap method was applied with 1,000 repetitions to calculate 95% confidence intervals (CI) for the RR values, stratified by year and municipality (34).

We used a DLNM time series approach to analyze the lagged associations between environmental stressors and COVID-19 incidence. The model was implemented with the *"crossbasis()"* and *"crosspred()"* functions from the dlnm package in R (35) considering lag effects for up to 12 months. We applied natural splines to the exposure-response associations for temperature (knots at the 10th, 25th, 75th, and 90th percentiles), $PM_{2.5}$ (10th, 50th, and 90th percentiles), and RHumid (50th and 90th percentiles). Lag response associations used natural splines with knots placed at equally-spaced values on the log scale. The crosspred function elements were optimized using the Akaike Information Criterion. We fitted the model utilizing the *"bam()"* function, similar to the GAM model, with the three aforementioned stressors considered as the crossbasis function elements. Separate models were fitted for each climate zone to capture region-specific associations between environmental stressors and COVID-19 incidence, as well as an overall model. The models can be expressed as:

$$\begin{aligned}\log\left(E[Y_{\mathrm{it}}|X_{\mathrm{it}}]\right) = \alpha &+ cb_{\mathrm{Temp}}(\mathrm{Temp}_{\mathrm{it}}, \mathrm{lag}) + cb_{\mathrm{PM}_{2.5}}(\mathrm{PM}_{2.5_{\mathrm{it}}}, \mathrm{lag}) \\ &+ cb_{\mathrm{Humid}}(\mathrm{Humid}_{\mathrm{it}}, \mathrm{lag}) + u_i + \beta_1\mathrm{Year}_t \\ &+ \beta_2\mathrm{Month}_t + \log\left(\mathrm{PopTotal}_{\mathrm{it}}\right)\end{aligned}$$

Where $Y_{\mathrm{it}}$ is the COVID-19 incidence in month $t$ for municipality $i$, $E[Y_{it}|X_{it}]$ denotes the expected number of cases given the independent variables, $\alpha$ is the intercept, $\mathrm{cb}_{\mathrm{Temp}}()$, $\mathrm{cb}_{\mathrm{PM}_{2.5}}()$, $\mathrm{cb}_{\mathrm{Humid}}()$ are $c$ ross-basis functions modeling nonlinear and lagged effects, $u_i$ is the Markov Random Field term, $\beta_1\mathrm{Year}_t + \beta_2\mathrm{Month}_t$ represents temporal effects, and $\log(\mathrm{PopTotal}_{\mathrm{it}})$ is an offset for population size.

# 3 Results

## 3.1 Descriptive analysis

The descriptive analysis examined both spatial and temporal aspects. Figure 1 illustrates the spatial distribution of Brazilian data averaged over all years. Temp (Figure 1A) were lowest in the south, around 25°C in the north and west, and highest in the country's interior with a maximum of 28.55°C. The RHumid distribution (Figure 1B) indicated that northern Brazil and the east and southeast coast experienced an increased average monthly relative humidity, peaking at 88.59%. In contrast, RHumid was minimal in the central region with a minimum of 53.33%. The $PM_{2.5}$ distribution (Figure 1C) indicated lower concentrations in the east and south of the country.

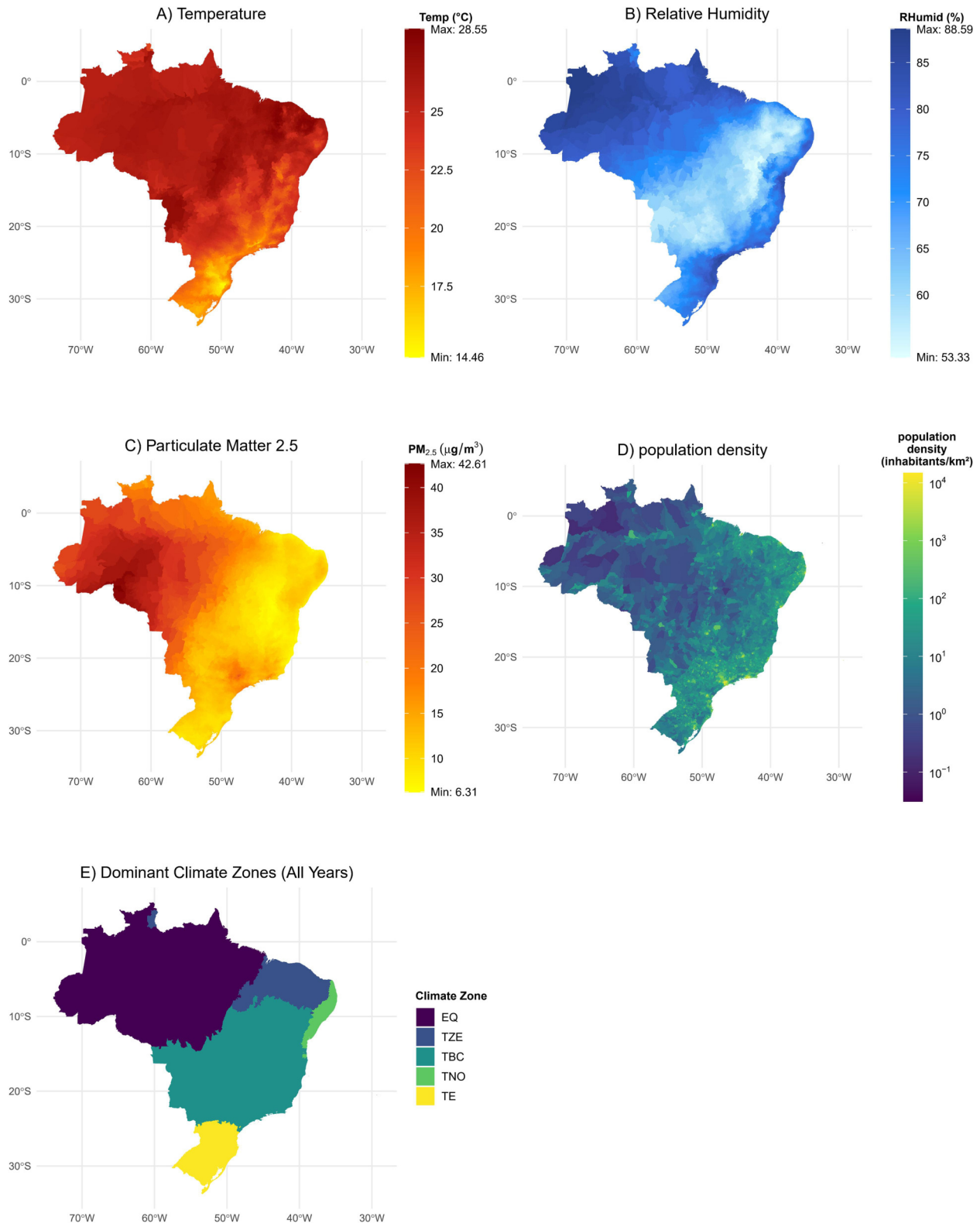


FIGURE 1
Spatial descriptive analysis maps of Brazil. The spatial distribution of the environmental stressors across all years is displayed for Temp **(A)**, RHumid **(B)**, and $PM_{2.5}$ **(C)** The population density map is calculated by the number of inhabitants per square kilometer **(D)** Additionally, the five dominant climate zones Equatorial (EQ), Tropical Zona Equatorial (TZE), Tropical Brazil Central (TBC), Tropical Nordeste Oriental (TNO) and Temperado (TE) are color-coded by municipality in Brazil **(E)**.

A remarkably high concentration was recorded in the southwest, with a maximum level of 42.61 μg/m$^3$. The population density map (Figure 1D) showed that the northwest of Brazil had the lowest population density. In contrast, population densities were higher in the eastern region, peaking in the southeast. Brazil was divided into five dominant climate zones (Figure 1E): EQ (433 municipalities), TZE (870), TBC (2,688), TNO (564), and TE (1,012). Population density varies across climate zones. The lowest population density was observed in EQ with 4.7 inhabitant/km$^2$ (total population 19,662,557), followed by TZE with 31.3 inhabitant/km$^2$ (total population 22,896,582), TBC with 40.3 inhabitant/km$^2$ (total population 117,654,874), TE with 53.6 inhabitant/km$^2$ (total population 26,498,324), and TNO with the highest population density with 151.7 inhabitant/km$^2$ (total population 24,996,485). All population figures refer to the year 2020.

The temporal descriptive analysis for Brazil (Figure 2) illustrates the incidence of COVID-19 cases over the months (Figure 2A). The incidence rose until June 2021, then declined towards December 2021. In January 2022, it increased again, peaking in February. After a decrease, there was a spike in July 2022, followed by another low in October, with a slight increase in the last two months of the year. The incidence by climate zone (Figure 2B) was similar to the general trend (Figure 2A). TNO tended to have lower incidence rates, whereas TE generally showed higher rates than the average. The monthly average RHumid (Figure 2C) revealed that the EQ, TBC, and TZE climate zones reached lower RHumid levels in the first and fourth quarter (Q1 and Q4) and higher levels in the second and third quarter (Q2 and Q3). In contrast, the TNO and TE zones exhibited relatively consistent RHumid throughout the year, with only slight increases observed in Q2 and Q3. Average

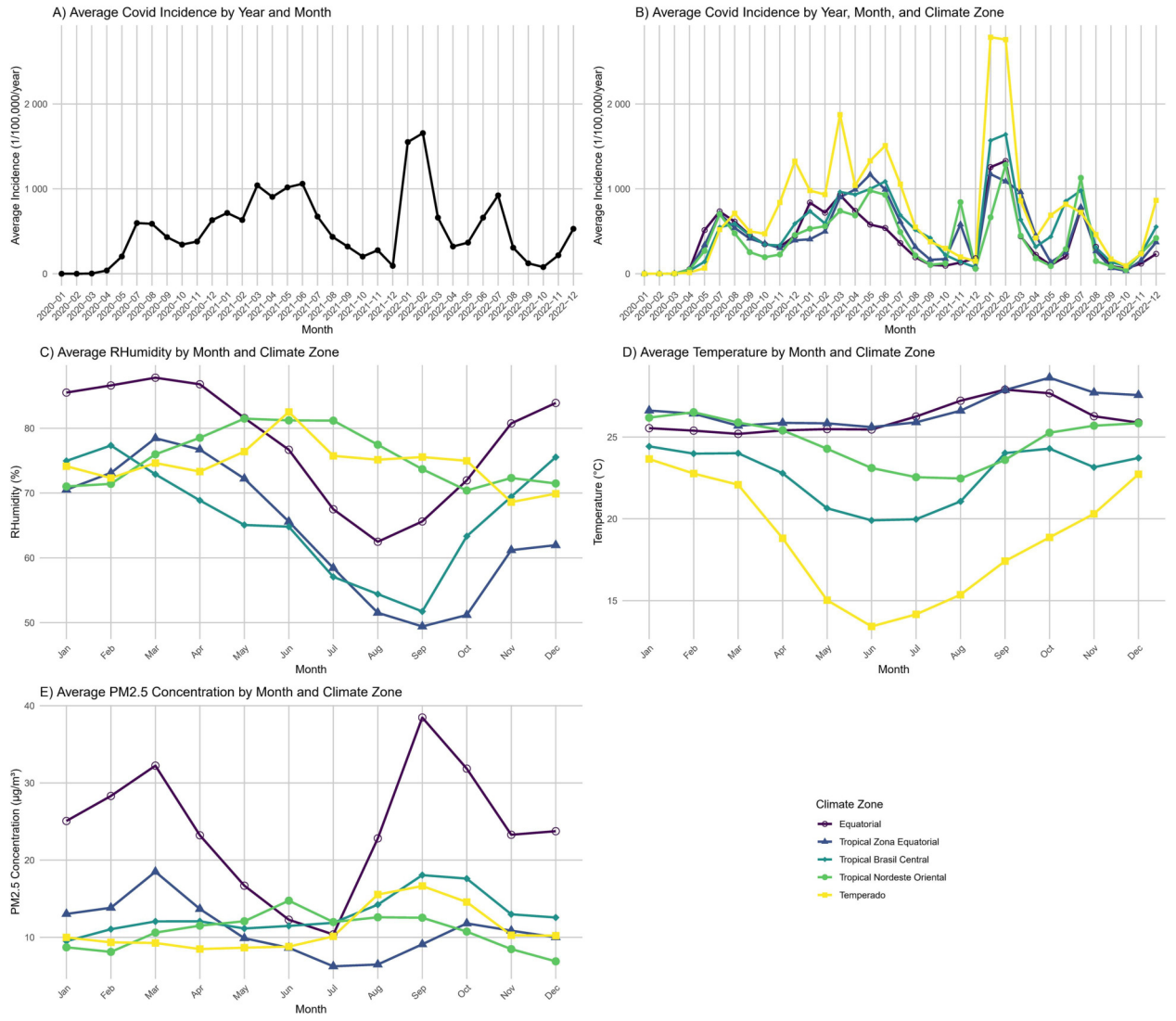


FIGURE 2
Temporal descriptive analysis of COVID-19 incidence and environmental stressors in Brazil. **(A)** Temporal distribution of COVID-19 incidence by year and month, defined as the number of cases per 100,000 people per year; **(B)** Incidence distribution stratified by climate zone; **(C–E)** Monthly averages of environmental stressors over the three-year study period: relative humidity (RHumid) **(C)**, temperature (Temp) **(D)**, and particulate matter ($PM_{2.5}$) **(E)**, with climate zones represented by distinct colors.

monthly Temp (Figure 2D) varied most in climate zone TE, followed by TBC and TNO, with the coldest temperatures typically in mid-year. In contrast, TZE and EQ had consistently high temperatures year-round, peaking in September and October across all years. Average monthly $PM_{2.5}$ concentrations (Figure 2E) were highest in the EQ climate zone, particularly at the year's start and end, with a dip in the middle. Other climate zones were more stable: TBC and TNO showed slight increases from August to October, while TE had a minor peak in June. Meanwhile, TZE experienced higher levels in Q1, lower levels in Q2 and Q3, and an increase in the Q4.

$PM_{2.5}$ exhibited the highest mean concentration and considerable variation throughout the observation period, with an average of 24.4 $\mu g/m^3$ in climate zone EQ (Table 1). The other climate zones had mean concentrations ranging from 10.6 to 12.9 $\mu g/m^3$. The average Temp in the EQ, TNO, and TZE zones exceeded 24 °C, while TE recorded the lowest mean Temp at 18.9 °C. Regarding RHumid, TZE had the lowest average at 64.2%, followed by TBC at 66.3%, TE at 74.2%, TNO at 75.3%, and EQ at 78.1%. Although the mean temperatures in EQ (26.2 °C) and TZE (26.7 °C) were similar, the relative humidity in TZE (64.2%) was significantly lower than in EQ (78.1%). Conversely, while the relative humidity was comparable between TE (74.2%) and TNO (75.3%), TE had a lower mean temperature (18.9 °C) compared to TNO (24.8 °C).

## 3.2 GAM

The exposure-response curves between environmental stressors and COVID-19 incidence (Figure 3) and the related relative risk (RR) values (Table 2) varied by climate zone. Temp showed a strong positive association with estimated COVID-19 incidence in the TBC [RR 1.87, 95% confidence interval (1.57, 2.10)] and TE [RR 17.9 (15.26, 22.19)] zones, a strong negative association in the TNO [RR 0.008 (0.004, 0.012)] and TZE [RR 0.12 (0.08, 0.15)] zones, and no cut-off association in EQ zone. For $PM_{2.5}$ levels, there was a positive trend in association with estimated COVID-19 incidence in the TE [RR 2.10 (1.93, 2.28)] and TZE [RR 1.87 (1.54, 2.31)] zones, while the EQ, TBC, and TNO zones did not show a clear association. Regarding RHumid, the estimated incidence of COVID-19 showed a negative association with the TZE [RR 0.01 (0.006, 0.014)], TBC [RR 0.66 (0.57, 0.73)] and TNO [RR 0.05 (0.03, 0.06)] zones, a positive relation with the TE [RR 3.69 (3.38, 4.12)] zone, and no significant association in the EQ zone. The overall model showed a positive link between COVID-19 incidence and Temp [RR 2.47 (2.04, 2.91)], but no statistically significant association with RHumid [RR 1.02 (0.91, 1.13)] and $PM_{2.5}$ [RR 1.03 (0.95, 1.11)].

The stratified models for the pre-Omicron (2020–2021) and Omicron periods (2022) showed some differences in the exposure-response curves (Supplementary Figure S1) and the RR values (Supplementary Table S1). During the Omicron period, the exposure-response curves for all stressors were higher than in the pre-Omicron period across all climate zones, while the trend in TZE was reversed. Most associations seen in the overall period were also present in the pre-Omicron period, except for the $PM_{2.5}$ effect in TZE, which was minimal in that period [RR 0.84 (0.83, 0.85)]. Additionally, strong new associations identified in the pre-Omicron period include a negative correlation between COVID-19 incidence and $PM_{2.5}$ levels in TNO [RR 0.27 (0.26, 0.27)] as well as with Temp [RR 0.64 (0.62, 0.65)] and RHumid [RR 0.47 (0.46, 0.47)] in EQ zone. During the pre-Omicron period, Temp showed in the overall model a positive association [RR 1.79 (1.65, 1.92)], while $PM_{2.5}$ [RR 0.72 (0.69, 0.76)] and RHumid [RR 0.84 (0.80, 0.89)] had no specific links. In contrast, the RR values in the overall model showed no statistically significant threshold effect for all environmental stressors. TE showed no statistically significant link with the three environmental stressors during the Omicron period. The only consistent associations observed between the two periods were a negative association between COVID-19 incidence and RHumid in the EQ climate zone [pre-Omicron period RR 0.47 (0.46, 0.47), Omicron period RR 0.67 (0.66, 0.68)] and a positive association with Temp in the TBC climate zone [pre-Omicron period RR 1.42 (1.33, 1.54), Omicron period RR 2.21 (1.37, 2.89)]. New positive associations emerged during the Omicron period, specifically between Temp and the incidence of COVID-19 in the TZE [RR 2.73 (2.66, 2.79)], between RHumid and COVID-19 in the TZE [RR 2.26 (2.24, 2.27)] and TBC [RR 2.24 (1.35, 2.95)] zones, and between $PM_{2.5}$ and COVID-19 in the TNO zone [RR 1.33 (1.29, 1.36)]. Supplementary Figure S2 displayed the exposure-response curves for all climate zones in one graph, with separate plots for the

TABLE 1 Descriptive statistics [mean, standard deviation (SD), median, minimum (Min), and maximum (Max)] of the environmental stressors $PM_{2.5}$, Temp, and RHumid across the five Brazilian climate zones.

| | | $PM_{2.5}$ ($\mu g/m^3$) | Temp (°C) | RHumid (%) |
|---|---|---|---|---|
| Equatorial (EQ) | Mean (SD) | 24.4 (14.1) | 26.2 (1.23) | 78.1 (12.6) |
| | Median [Min, Max] | 22.5 [2.12., 103] | 25.9 [21.5, 30.7] | 83.0 [32.5, 95.2] |
| Tropical Zona Equatorial (TZE) | Mean (SD) | 11.1 (5.30) | 26.7 (1.71) | 64.2 (13.9) |
| | Median [Min, Max] | 9.80 [1.70, 37.4] | 26.6 [20.9, 32.2] | 65.0 [20.9, 32.2] |
| Tropical Brasil Central (TBC) | Mean (SD) | 12.9 (6.16) | 22.7 (2.99) | 66.3 (13.1) |
| | Median [Min, Max] | 12.0 [1.86, 113] | 23.0 [11.4, 32.1] | 69.2 [26.7, 91.2] |
| Tropical Nordeste Oriental (TNO) | Mean (SD) | 10.6 (3.17) | 24.8 (1.81) | 75.3 (6.88) |
| | Median [Min, Max] | 10.4 [2.76, 23.4] | 25.0 [18.9, 29.6] | 76.0 [48.5, 90.0] |
| Temperado (TE) | Mean (SD) | 11.1 (3.98) | 18.9 (3.91) | 74.2 (7.99) |
| | Median [Min, Max] | 10.1 [4.14, 39.6] | 19.0 [8.49, 29.9] | 75.4 [45.4, 90.1] |

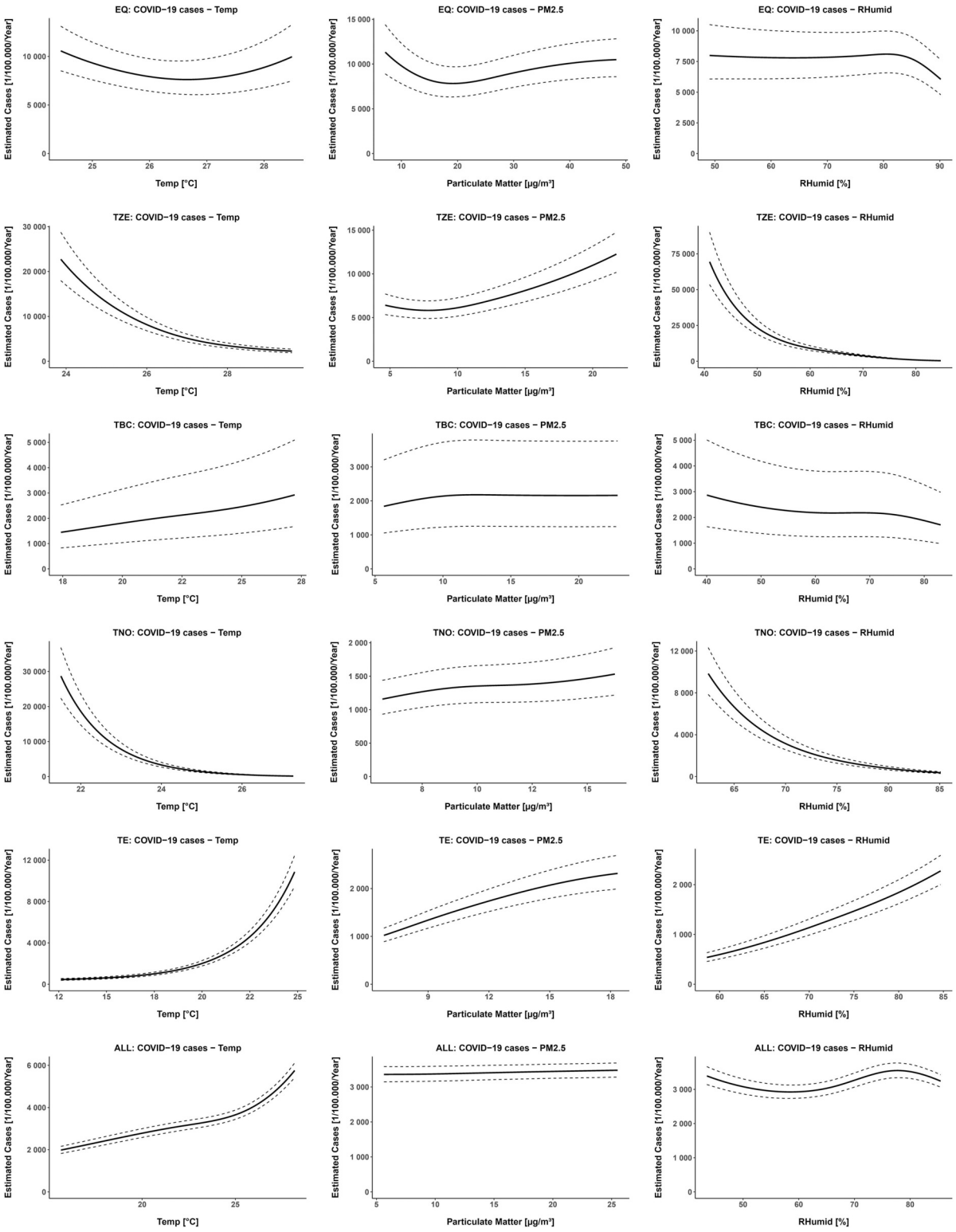


FIGURE 3
Estimated COVID-19 incidence per 100,000 people per year in response to each environmental stressor and climate zone. The five climate zones are: Equatorial (EQ), Tropical Zona Equatorial (TZE), Tropical Brasil Central (TBC), Tropical Nordeste Oriental (TNO), and Temperado (TE). All other parameters were kept constant, and were set to the following values: municipality code: 1100114 (EQ, ALL), 2107803 (TZE), 3500600 (TBC), 2610202 (TNO), 4122206 (TE), month: February, year: 2021 and environmental stressors: median. Estimates were obtained using the R function predict for Temp, RHumid, and $PM_{2.5}$, including their 5th and 95th percentiles.

TABLE 2 Estimated COVID-19 incidence and associations with environmental stressors by climate zone.

| | Output | EQ | TZE | TBC | TNO | TE | ALL |
|---|---|---|---|---|---|---|---|
| $PM_{2.5}$ | Prediction 5-Percentile | 10,388.39 (10,279.31, 10,457.69) | 6,190.40 (5,165.46, 8,215.02) | 1,902.07 (1,238.85, 2,624.00) | 1,186.25 (917.61, 1,502.26) | 1,088.02 (938.47, 1,252.74) | 3,359.68 (2,885.94, 3,982.01) |
| | Prediction 95-Percentile | 10,428.85 (10,391.85, 10,457.69) | 11,603.62 (9,531.89, 15,662.22) | 2,160.37 (1,428.25, 2,935.60) | 1,502.59 (1,193.12, 1,960.26) | 2,283.93 (1,980.77, 2,600.35) | 3,472.08 (2,924.59, 4,134.17) |
| | **Relative Risk** | **1.00 (0.99, 1.02)** | **1.87 (1.54, 2.31)** | **1.14 (1.07, 1.21)** | **1.27 (1.07, 1.55)** | **2.10 (1.93, 2.28)** | **1.03 (0.95, 1.11)** |
| Temp | Prediction 5-Percentile | 10,164.41 (10,107.74, 10,226.26) | 20,340.52 (16,078.16, 32,799.20) | 1,501.46 (979.12, 2,078.02) | 23,391.29 (19,045.96, 35,526.53) | 480.11 (389.47, 574.84) | 2,075.08 (1,685.75, 2,602.67) |
| | Prediction 95-Percentile | 9,460.34 (9,359.05, 9,579.74) | 2,402.54 (2,008.08, 3,005.25) | 2,803.41 (1,772.19, 3,768.15) | 392.84 (113.95, 255.81) | 8,605.50 (7,537.20, 9,853.36) | 5,124.08 (4,484.77, 5,814.32) |
| | **Relative Risk** | **0.93 (0.92, 0.94)** | **0.12 (0.08, 0.15)** | **1.87 (1.57, 2.10)** | **0.008 (0.004, 0.012)** | **17.9 (15.26, 22.19)** | **2.47 (2.04, 2.91)** |
| RHumid | Prediction 5-Percentile | 7,953.63 (7,924.05, 7,986.63) | 55,340.72 (43,371.65, 89,964.79) | 2,766.24 (1,833.89, 3,930.14) | 8,527.89 (7,048.94, 12,404.80) | 581.42 (482.86, 691.91) | 3,295.67 (2,745.15, 4,080.95) |
| | Prediction 95-Percentile | 6,835.95 (6,804.29, 6,869.81) | 570.66 (468.96, 742.94) | 1,831.50 (1,191.17, 2,477.41) | 438.13 (323.65, 560.73) | 2,144.80 (1,874.23, 2,448.49) | 3,377.40 (2,989.51, 3,930.18) |
| | **Relative Risk** | **0.86 (0.85, 0.86)** | **0.01 (0.006, 0.014)** | **0.66 (0.57, 0.73)** | **0.05 (0.03, 0.06)** | **3.69 (3.38, 4.12)** | **1.02 (0.91, 1.13)** |

This table shows the estimated cases per 100,000 people per year for the 5th and 95th percentiles of environmental stressors. Confidence intervals are in brackets. Green cells indicate a strong positive association (RR > 1.3), yellow cells indicate a strong negative association (RR < 0.7), and unshaded cells represent a weak or no association ($0.7 \leq RR \leq 1.3$).
All relative risk values are written in bold for clarity.

Omicron and pre-Omicron periods. The x-scaling and range of environmental stressors differed considerably by climate zone, making direct comparisons between these zones challenging.

In the sensitivity analysis based on the overall model, the interaction terms suggested that both the magnitude and shape of the associations differed between the pre-Omicron and Omicron periods, broadly consistent with the trends observed in the stratified analyses.

## 3.3 DLNM

Figure 4 shows distinct lag-response curves for Temp, $PM_{2.5}$ and RHumid at the 1% and 99% quantiles within each climate zone. In TZE and TNO, relative risk (RR) was higher for extreme cold and below 1 for extreme heat compared to the median (for TZE 26.62°C, for TNO 25.05°C), peaking at lags 0 and 12 for hot temperatures and reaching the lowest values at the same lags for cold temperatures. In TE, the RR for hot temperatures peaked at 2.41 [95% confidence interval: (2.15,2.71)] at lag 0, while cold temperatures had an RR of 0.66 (0.59, 0.73). In TBC, RR was 1.10 (1.01, 1.19)) for hot and 0.91 (0.86, 0.98)) for cold temperatures, whereas in EQ, RR remained above 1, initially higher for cold than hot temperatures, reversing by lag 3.

For RHumid, 1% quantile values in EQ, TZE, and TNO started with an RR above 1 at lag 0, dipping at lags 5–7 and following a parabolic curve. At the 99% quantile, extremely high RHumid resulted in RR <1 in EQ, TZE, and TNO, with minor fluctuations. In TBC, RR at lag 0 was >1 for extremely low and <1 for extremely high RHumid, with both curves converged at 1 by lag 6 and remained stable. In contrast, RR was >1 for high and <1 for low RHumid in TE.

Concering $PM_{2.5}$, RR values were calculated relative to the minimum concentration in each zone. In TE, TZE, TBC, and EQ, RR at lag 0 was >1 and decreased over time. The 99% quantile curves exceeded the 15 $\mu g/m^3$ threshold in TZE and EQ, with EQ reaching the highest value (80 $\mu g/m^3$) and TNO the lowest (19 $\mu g/m^3$). In TBC and TE, RR values were nearly identical at both levels. Unlike other zones, TNO had RR <1 at lag 0, approached 1 over time, and its 19-curve remained above 15 $\mu g/m^3$.

The overall model indicated low Temp (12.71°C) below 1 and high Temp (29.65°C) above 1, both following parabolic curves. For RHumid, the low value (34.70%) is above the high value (89.07%). In $PM_{2.5}$, the curves are nearly aligned with 1. At lag 0, the threshold (15 $\mu g/m^3$) exceeds the 99th percentile (40 $\mu g/m^3$), while an inverted structure appears at lag 12.

# 4 Discussion

## 4.1 Summary

Our analysis aimed to investigate the association between environmental stressors and COVID-19 incidence over a three-year period, using monthly data at municipality level across Brazil. The GAM modeling approach provides a robust framework for assessing the association between environmental factors and COVID-19 incidence, effectively accounting for population size, spatial structure, and temporal trends. To address regional variability, we developed an overall model and further adjusted for climatic differences by fitting separate models for each climate zone while maintaining consistent model parameters. We further disentangled two distinct periods, pre-Omicron and Omicron, to assess how environmental stressors link to COVID-19 incidences. In addition, we examined lag effects in the exposure-response association using DLNM. Our findings revealed a positive association between COVID-19 incidence and Temp in the overall model, while no statistically significant associations could be observed with $PM_{2.5}$ or RHumid. In climate zone-specific models, only TE and TZE showed a positive association with $PM_{2.5}$. RHumid effects varied: negative in TZE, TBC and TNO; positive in TE; and

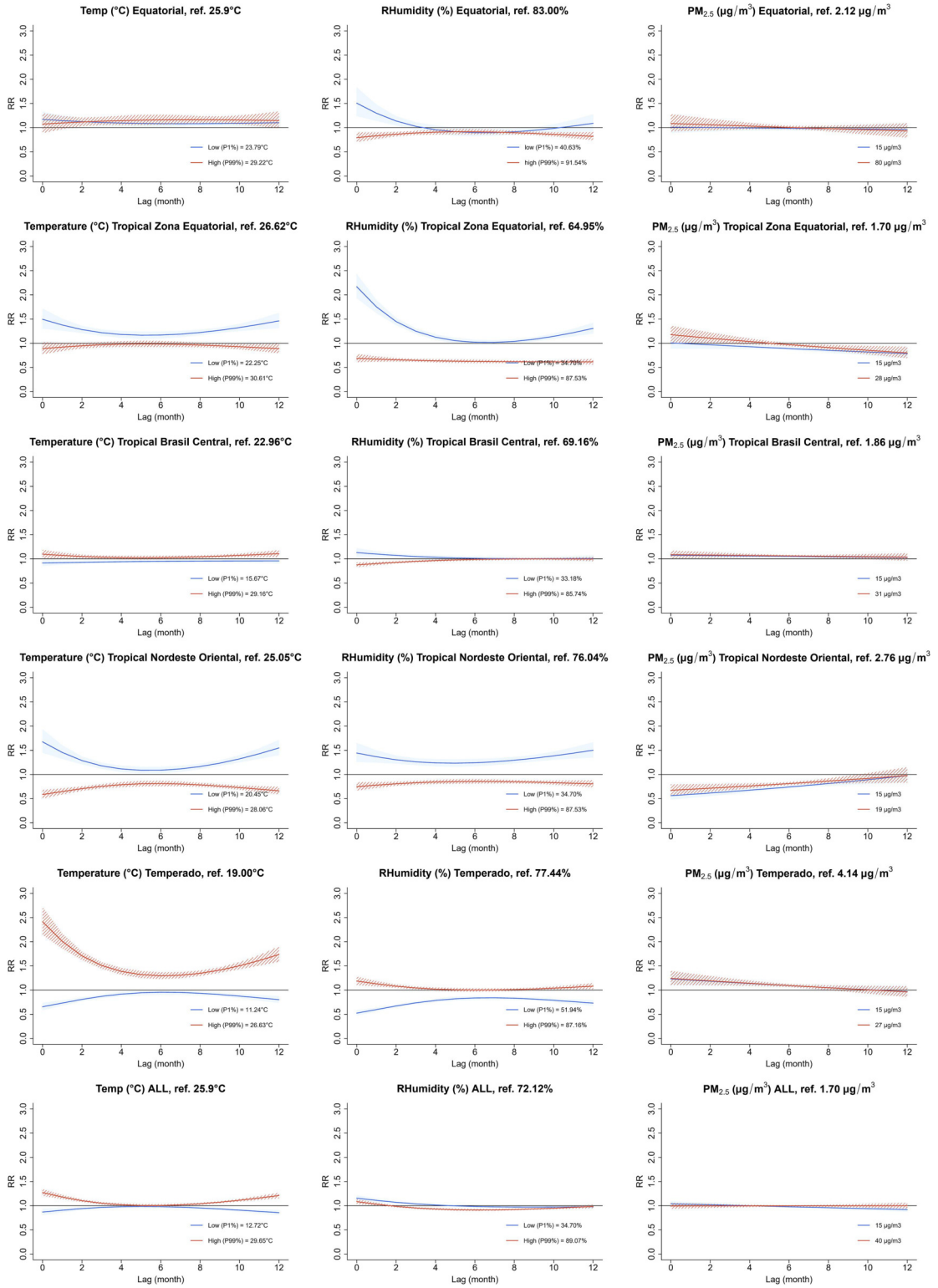


FIGURE 4
Lag-response plots of the distributed Lag nonlinear model (DLNM) for temp, $PM_{2.5,}$ and RHumid across different climate zones. Each plot displays the lagged effects (in months) of the 1% and 99% quantiles for Temp and RHumid, as well as the threshold of 15 µg/m$^3$ and the 99% quantile for $PM_{2.5}$. These effects are presented as relative risks (RR), which indicate the incidence risk in comparison to the median exposure for Temp and RHumid, and the minimum exposure for $PM_{2.5}$ across the time lags. Exposure-response associations were modeled with natural splines (Temp knots: 10th, 25th, 75th, 90th percentiles; $PM_{2.5}$: 10th, 50th, 90th percentiles; RHumid: 50th, 90th percentiles), and lag-response associations used natural splines with knots equally spaced on the log scale.

neutral in TBC. Temperature associations were positive in TE and TBC, negative in TNO and TZE, and neutral in EQ. The analysis revealed no consistent concentration-response pattern between the pre-Omicron and Omicron periods. The lag effects were modeled using a DLNM time series approach. In the TZE and TNO zones, high Temp and RHumid levels (99th percentiles) were linked to increased risk of COVID-19 incidence, while low values (1st percentile) were associated with decreased risk of incidence, with parabolic lag response curves. In the TE zone, the associations were reversed. The TBC and EQ zones showed minimal and consistent effects over lag periods for Temp and RHumid, except low RHumid in EQ reaching up to 1.5 at lag 0. For $PM_{2.5}$, extreme exposures and the threshold 15 $\mu g/m^3$ at lag 0 in the TE, TZE, TBC, and EQ zones were associated with higher incidence, decreasing with subsequent lags, while TNO showed the opposite trend.

## 4.2 Discussion of the results

The incidence of COVID-19 cases in Brazil was consistent with data from other sources (36). Many studies have explored the impact of environmental stressors on COVID-19 infections, primarily in Europe and Asia. However, Brazil, a large country with various climate types, was significantly impacted by the pandemic and has been less studied. To the best of our knowledge, our research is unique in both the duration of the observation period and the level of spatial resolution. We are not aware of other studies that have conducted similar analyses at municipality level over a three-year period in Brazil. While some studies have explored the association between COVID-19 cases and factors such as $PM_{2.5}$ and Temp concentrations in Brazil, they often faced limitations regarding spatial and/or temporal scale. These studies typically analyzed shorter observation periods, focused on specific cities, or smaller geographic areas. Nevertheless, we will discuss the similarities and differences between the findings of these studies and our analysis, which examined Brazil as a whole, split into five climate zones.

Köppen's climate classification maps for Brazil indicate comparable patterns for high and low temperatures (37). However, our analysis revealed that highest temperatures were concentrated in central Brazil, specifically, between the northeastern and southwestern region. This discrepancy may be attributed to differences in data periods, as the literature-based climate data ended in 2013, while our study data began in 2020 (38). Two previous studies investigated the association between temperature and COVID-19 incidence in Brazilian state capitals, with one study including Brasília, focusing on the initial phase of the pandemic. These studies found an inverse association between the number of COVID-19 cases and mean temperatures below 25.8°C. However, this association was not observed in cities where the mean temperature was equal to or greater than 25.8°C (39, 40). In our climate zone-specific analysis, EQ and TZE, which had mean temperatures exceeding 25.8°C, showed no significant association with estimated COVID-19 cases. In contrast, TE and TBC, where mean temperatures were below 25.8°C, demonstrated a positive association with COVID-19 incidence. These findings are consistent with the results reported in the previously mentioned studies. The TNO climate zone, with a mean temperature of 24.7°C, is an exception, as it showed a negative exposure-response association. The literature indicated a difference between tropical and non-tropical regions in relation to temperature and COVID-19 associations. In tropical areas, high temperatures and relative humidity were risk factors, while in subtropical regions, both high and low temperatures were risk factors (13, 41). Other early pandemic research of capital cities showed a positive correlation between mean temperature and COVID-19 cases in tropical regions but an inverse link in subtropical areas (12). Köppen's climate classification indicates that northern and central Brazil are primarily tropical, while the south is subtropical (37). We found a positive association between temperature and COVID-19 cases in the TE (subtropical) region, a negative association in the TNO (tropical) region, and no statistically significant association in the EQ (tropical) region, which contrasts with previous findings. Research conducted from April 2022 to July 2023 in a temperate subtropical region supported our results, finding a positive correlation between temperature and SARS-CoV-2 (42). It is essential to examine the temperature range in different areas. The TZE, TNO, and EQ regions had a slight temperature range and predominantly high temperatures, while the TBC and TE zones exhibited both high temperatures and a wider range. Our results suggested that consistently high temperatures may not correlate positively with COVID-19 incidence, whereas high temperatures in areas with lower temperatures were positively associated with COVID-19 cases.

The study on the initial phase of the Covid-19 pandemic in Brasília found additionally that lower mean relative humidity was associated with lower case numbers, especially when relative humidity was 65% or lower (39). In the TZE climate zone, our GAM indicated that lower relative humidity was linked to increased COVID-19 incidence, which contradicts the findings for the same humidity range reported in the Brasília study. Our analysis of the TBC climate zone (including Brasília) revealed similar results. Additionally, our DLNM showed that extreme RHumid levels below 65% raised infection risk across most climate zones, which contradicted the papers results with focus on Brasília. A study conducted in 2020, focusing on the municipalities of Mato Grosso (in TBC and EQ) and Cuiaba Varzea (in TBC), found that RHumid had a significant negative correlation with the number of confirmed daily COVID-19 cases (43). We also observed a negative association in TBC and no consistent pattern across the range of relative humidity in EQ.

$PM_{2.5}$ concentrations in Brazil were highest in the western and northern regions (Figure 1C), which aligns with existing research (44). This pollution was in general caused by human activities such as traffic, industry, and agriculture, along with wildfires. The central-western region of Brazil experienced the highest annual average of $PM_{2.5}$ related to wildfires (44).

A study using DLNM found that COVID-19 morbidity is nearly twice as high among individuals exposed to high levels of

wildfire-generated $PM_{2.5}$ (45). Our DLNM model aligned with that finding and showed a higher incidence risk for high $PM_{2.5}$ levels across all climate zones except TNO. This Research over two years in five Brazilian states—Acre (AC), Rondônia (RO), Mato Grosso (MT), Pará (PA), and Amazonas (AM)—found that AC and PR had a higher risk of COVID-19 morbidity linked to $PM_{2.5}$ concentrations, while MT, RO, and PA showed a decreasing trend (45). AC, PA, and RO were classified in the EQ climate zone, PR in the TE zone, and MT was both EQ and TBC zones. Our GAM indicated a positive exposure-response association between $PM_{2.5}$ and COVID-19 incidence for TE, aligning with the literature, while showing a weak or unclear association for EQ and TBC. A study examined the link between COVID-19 and $PM_{2.5}$ levels in the cities Curitiba and Araucaria using sensor data. From March to November 2021, a positive correlation was observed between $PM_{2.5}$ and COVID-19 incidence (46). The areas studied are in climate zone TE, where our GAM aligned with the literature and showed a RR of 2.03.

Brazil's climatic conditions differed considerably from those in Europe, affecting the influence of environmental stressors on COVID-19 incidence. A study in Germany showed a strong negative association between temperature and COVID-19 incidence: as temperatures dropped, COVID-19 incidence rose (6). A global multi-city study found that the risk of COVID-19 infection was 1.22 times higher at 5°C compared to 17°C (47). In contrast, for temperatures above 25.8°C, a strong positive correlation was observed, indicating that higher temperatures lead to increased COVID-19 incidence (39, 40). Our models found these patterns in the TE, TZE, and TBC climate zones and the overall data. For TNO, the association began negatively, but weakened at higher temperatures, unlike the consistent positive correlation seen in other zones. The EQ zone showed no statistically significant temperature effect, differing from other climate zones.

These findings suggest that both low and high temperatures may boost COVID-19 infections. While this may appear counterintuitive, there are plausible pathophysiological explanations. Exposure to both cold and hot ambient temperatures can cause physiological stress (48, 49), and the negative impact of chronic stress on immune function is well-documented (50). Additionally, behavioral adaptations, such as increased indoor time and the use of heating or air conditioning, may alter exposure to indoor air pollution, further influencing infection resk. The delayed effects observed in the DLNM lag-response curves likely reflect a combination of direct environmental exposures, stress-mediated immune responses, and behavioral patterns. Future research should continue to investigate the underlying mechanisms to gain a better understanding of the causal implications of these findings.

## 4.3 Policy implications

The findings of this study have implications for public health and environmental policy, particularly in tropical and subtropical regions like Brazil. First, the dual role of high and low temperatures in exacerbating infection risk supports the need for temperature-sensitive public health strategies and risk communication. Second, the findings reinforce the need of regional environmental surveillance systems that monitor pollution and weather in real-time, allowing for adaptive response measures during high-risk periods.

Moreover, policymakers should consider integrating environmental data into epidemic and pandemic preparedness frameworks. For instance, incorporating temperature and air quality thresholds into early warning systems may improve the timing and targeting of public health interventions. In urban planning and housing policies, strategies to reduce indoor air pollution exposure—such as promoting natural ventilation, regulating air conditioning use, and improving insulation—could mitigate the compounding effects of environmental stressors on health.

Finally, public health policies must acknowledge the regional variability in environmental risks and avoid one-size-fits-all approaches. Tailored interventions considering local climatic realities, particularly in vulnerable populations and under-resourced regions, are critical for achieving equitable health outcomes in the face of both infectious disease threats and ongoing climate change.

## 4.4 Limitations

The availability of PCR tests fluctuated significantly over the weeks, peaking during the second wave (2021), followed by the first (2020) and third (2022) waves (51). This inconsistency and the lack of tests have led to a substantial number of unreported COVID-19 infections, causing a significant underestimation of incidence rates (16). Moreover, socioeconomic inequality among Brazilian municipalities impacted COVID-19 detection and the occurrence of negative outcomes. In poorer areas, access to health services is often precarious and fewer COVID-19 tests are performed, so there might be a detection bias that could reduce the number of cases in poorer areas (52). Some limitations of the $PM_{2.5}$ dataset include uncertainties in Aerosol Optical Depth (AOD) measurements, limited validation due to sparse ground monitoring in Brazil, and a reduced temporal resolution that affects short-term analyses (18, 22, 53). On the other hand, the van Donkelaar $PM_{2.5}$ dataset offers high resolution, effectively filling gaps in $PM_{2.5}$ coverage when ground stations cannot ensure continuous reporting (22, 54). Well-documented in Brazilian studies (23–25), it combines satellite-derived AOD data with chemical transport models, providing reliable coverage even in areas without monitoring infrastructure (21). The reanalysis data from the ERA5-land products serve as a reliable source for climate data (55). Our analysis used a monthly aggregation level due to the availability of our $PM_{2.5}$ dataset. While imposing some temporal limitations, this approach is meaningful for examining the longer-term health impacts of $PM_{2.5}$, Temp, and RHumid on COVID-19 incidence rates. Monthly data helps to reduce short-term fluctuations caused by reporting delays, allowing for a more robust analysis of these associations. The month June 2020 has been excluded from the data due to plausibility checks revealing an unrealistically high number of infections. We focused on environmental stressors without considering age group and

gender information, which is a limitation of our work. Literature showed that from March 2020 to September 2021 in Brazil, the disease was more common in men (55.6%), with the highest prevalence in the 50–59 age group (20.2%) (56). A study from February to November 2020 examined Brazil's 27 state capitals and discovered that meteorological conditions influenced COVID-19 in periodic ways, with a positive effect from March to May and a negative effect from June to August (57). A possible extension and future work of our analysis would be to examine interaction terms between months and environmental stressors to reach similar analysis.

## 5 Conclusion

Our study provides a comprehensive and systematic analysis of the association between selected environmental stressors and COVID-19 incidence in Brazil from 2020 to 2022. The findings reinforce existing evidence that air pollution, particularly particulate matter, as well as meteorological factors such as temperature and relative humidity, play a significant role in influencing respiratory health and the spread of COVID-19. By examining data across Brazil's diverse climatic regions, we observe that the impact of these environmental variables differs between climate zones, in particular, tropical and non-tropical areas, highlighting the importance of considering regional differences in epidemiological assessments. Furthermore, our study underscores the necessity of long-term analyses beyond the early stages of the pandemic to fully capture the evolving relationships between environmental exposures and COVID-19 outcomes. The influence of virus mutations, shifting public health measures, and vaccination efforts further emphasize the complex interplay between environmental and epidemiological factors. These insights contribute to a broader understanding of the environmental determinants of infectious diseases and can inform future public health policies aimed at mitigating the effects of pandemics in varying ecological contexts. Specifically, the findings emphasize the importance of implementing continuous environmental surveillance systems, which guide interventions in areas with high pollution or extreme climate conditions. Additionally, integrating environmental data into pandemic preparedness strategies is essential to reducing health risks.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

## Author contributions

LH: Writing – original draft, Formal analysis, Visualization, Data curation, Methodology, Conceptualization, Writing – review & editing, Software, Investigation. LV: Data curation, Writing – review & editing, Formal analysis, Methodology. TE: Writing – review & editing. MB: Formal analysis, Writing – review & editing. JR: Funding acquisition, Conceptualization, Supervision, Writing – review & editing, Methodology. RS: Supervision, Conceptualization, Writing – review & editing, Methodology.

## Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) [Grant no 458531714].

## Acknowledgments

COVID-19 data is freely available at https://covid.saude.gov.br. $PM_{2.5}$ data is freely available from the Atmospheric Composition Analysis Group at the Washington University in St. Louis (21). Meteorological data is freely available from the ERA5-Land global reanalysis dataset provided by the Copernicus Climate Change Service of the European Centre of Medium-Range Weather Forecasts (26).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## Publisher's note



## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fenvh.2025.1635503/full#supplementary-material

# References

1. Marques M, Domingo JL. Positive association between outdoor air pollution and the incidence and severity of COVID-19. A review of the recent scientific evidences. *Environ Res*. (2022) 203:111930. doi: 10.1016/j.envres.2021.111930

2. Zang S-T, Luan J, Li L, Yu H-X, Wu Q-J, Chang Q, et al. Ambient air pollution and COVID-19 risk: evidence from 35 observational studies. *Environ Res*. (2022) 204:112065. doi: 10.1016/j.envres.2021.112065

3. Copat C, Cristaldi A, Fiore M, Grasso A, Zuccarello P, Santo Signorelli S, et al. The role of air pollution (PM and NO2) in COVID-19 spread and lethality: a systematic review. *Environ Res*. (2020) 191:110129. doi: 10.1016/j.envres.2020.110129

4. Rittweger J, Gilardi L, Baltruweit M, Dally S, Erbertseder T, Mittag U, et al. Temperature and particulate matter as environmental factors associated with seasonality of influenza incidence - an approach using earth observation-based modeling in a health insurance cohort study from Baden-Wurttemberg (Germany). *Environ Health-Glob*. (2022) 21(1):131. doi: 10.1186/s12940-022-00927-y

5. Carballo IH, Bakola M, Stuckler D. The impact of air pollution on COVID-19 incidence, severity, and mortality: a systematic review of studies in Europe and North America. *Environ Res*. (2022) 215:114155. doi: 10.1016/j.envres.2022.114155

6. Hoffmann L, Gilardi L, Antoni T, Baltruweit M, Bittner M, Breitner S, et al. Modulation of COVID-19 incidence by environmental stressors is variant between pre-omicron and omicron periods. *Sci Rep*. (2025) 15(1):27636. doi: 10.1038/s41598-025-13521-2

7. Bianconi V, Bronzo P, Banach M, Sahebkar A, Mannarino M, Pirro M. Particulate matter pollution and the COVID-19 outbreak: results from Italian regions and provinces. *Arch Med Sci*. (2020) 16(1):985–92. doi: 10.5114/aoms.2020.95336

8. Han Y, Zhao W, Pereira P. Global COVID-19 pandemic trends and their relationship with meteorological variables, air pollutants and socioeconomic aspects. *Environ Res*. (2022) 204:112249. doi: 10.1016/j.envres.2021.112249

9. Mehmood K, Bao Y, Abrar MM, Petropoulos GP, Soban A, Saud S, et al. Spatiotemporal variability of COVID-19 pandemic in relation to air pollution, climate and socioeconomic factors in Pakistan. *Chemosphere*. (2021) 271:129584. doi: 10.1016/j.chemosphere.2021.129584

10. Sarkodie SA, Owusu PA. Global effect of city-to-city air pollution, health conditions, climatic & socio-economic factors on COVID-19 pandemic. *Sci Total Environ*. (2021) 778:146394. doi: 10.1016/j.scitotenv.2021.146394

11. Bhatti UA, Marjan S, Wahid A, Syam M, Huang M, Tang H, et al. The effects of socioeconomic factors on particulate matter concentration in China's: new evidence from spatial econometric model. *J Cleaner Prod*. (2023) 417:137969. doi: 10.1016/j.jclepro.2023.137969

12. Prata D, Rodrigues W, Bermejo PHDS, Moreira M, Camargo W, Lisboa M, et al. The relationship between (sub) tropical climates and the incidence of COVID-19. *PeerJ*. (2021) 9:e10655. doi: 10.7717/peerj.10655

13. Auler A, Cássaro F, Da Silva V, Pires L. Evidence that high temperatures and intermediate relative humidity might favor the spread of COVID-19 in tropical climate: a case study for the most affected Brazilian cities. *Sci Total Environ*. (2020) 729:139090. doi: 10.1016/j.scitotenv.2020.139090

14. Giovanetti M, Slavov SN, Fonseca V, Wilkinson E, Tegally H, Patané JSL, et al. Genomic epidemiology of the SARS-CoV-2 epidemic in Brazil. *Nature Microbiology*. (2022) 7(9):1490–500. doi: 10.1038/s41564-022-01191-z

15. Alcantara LCJ, Nogueira E, Shuab G, Tosta S, Fristch H, Pimentel V, et al. SARS-CoV-2 epidemic in Brazil: how the displacement of variants has driven distinct epidemic waves. *Virus Res*. (2022) 315:198785. doi: 10.1016/j.virusres.2022.198785

16. Martins JP, Siqueira BA, Sansone NMS, Marson FAL. COVID-19 in Brazil: a three-year update. *Diagn Microbiol Infect Dis*. (2023) 107(4):116074. doi: 10.1016/j.diagmicrobio.2023.116074

17. Yang J, Cordeiro G, Longato M, Vaghela S, Kyaw MH, Mendoza CF, et al. Burden of COVID-19 during the omicron predominance in Brazil: a nationwide retrospective database study. *J Med Econ*. (2023) 26(1):1201–11. doi: 10.1080/13696998.2023.2262323

18. Leirião LFL, Debone D, Miraglia SGEK. Does air pollution explain COVID-19 fatality and mortality rates? A multi-city study in São Paulo state, Brazil. *Environ Monit Assess*. (2022) 194(4):275. doi: 10.1007/s10661-022-09924-7

19. Mehta SK, Ananthavel A, Reddy TR, Ali S, Mehta SB, Kakkanattu SP, et al. Indirect response of the temperature, humidity, and rainfall on the spread of COVID-19 over the Indian monsoon region. *Pure Appl Geophys*. (2023) 180(1):383–404. doi: 10.1007/s00024-022-03205-7

20. Yıldırım E. The relationship between PM10 and SO2 exposure and COVID-19 infection rates in Turkey using nomenclature of territorial units for statistics level 1 regions. *Heliyon*. (2023) 9(11):e21795. doi: 10.1016/j.heliyon.2023.e21795

21. Atmospheric Composition Analysis Group. Satellite-derived PM2.5 (2024). Available online at: https://sites.wustl.edu/acag/datasets/surface-pm2-5/ (Accessed 06/01/2024).

22. Van Donkelaar A, Hammer MS, Bindle L, Brauer M, Brook JR, Garay MJ, et al. Monthly global estimates of fine particulate matter and their uncertainty. *Environ Sci Technol*. (2021) 55:15287–300. doi: 10.1021/acs.est.1c05309

23. Couto LO, Jacobson LSV, Périssé ARS, Hacon SS. Identifying high occurrence areas of hospitalization and mortality from respiratory diseases in the Brazilian legal Amazon: a space-time analysis. *Cad Saúde Pública*. (2024) 40:e00148023. doi: 10.1590/0102-311XEN148023

24. Fleischer NL, Merialdi M, van Donkelaar A, Vadillo-Ortega F, Martin RV, Betran AP, et al. Outdoor air pollution, preterm birth, and low birth weight: analysis of the world health organization global survey on maternal and perinatal health. *Environ Health Persp*. (2014) 122(4):425–30. doi: 10.1289/ehp.1306837

25. Gouveia N, Rodriguez-Hernandez JL, Kephart JL, Ortigoza A, Betancourt RM, Sangrador JLT, et al. Short-term associations between fine particulate air pollution and cardiovascular and respiratory mortality in 337 cities in Latin America. *Sci Total Environ*. (2024) 920:171073. doi: 10.1016/j.scitotenv.2024.171073

26. Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, et al. The ERA5 global reanalysis. *Q J R Metereol Soc*. (2020) 146:1999–2049. doi: 10.1002/qj.3803

27. Alduchov OA, Eskridge RE. Improved Magnus form approximation of saturation vapor pressure. *J Appl Meteorol*. (1996) 35(4):601–9. doi: 10.1175/1520-0450(1996)035<0601:IMFAOS>2.0.CO;2

28. Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, Moore R. Google earth engine: planetary-scale geospatial analysis for everyone. *Remote Sens Environ*. (2017) 202:18–27. doi: 10.1016/j.rse.2017.06.031

29. Python Software Foundation. Python (Version 3.12.1, packaged by conda-forge) (2023). Available online at: https://www.python.org/ (Accessed 05/10/2025).

30. Brazilian Institute of Geography and Statistics. Instituto Brasileiro de Geografia e Estatística (2024). Available online at: https://www.ibge.gov.br/en/home-eng.html (Accessed 07/05/2024).

31. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing (2023).

32. Wood S. Package 'mgcv'. R package version. (2025) 1(29):729. doi: 10.32614/CRAN.package.mgcv

33. Wood S. *Generalized Additive Models: An introduction with R*. 2nd ed. New York: Chapman and Hall/CRC (2017).

34. Hastie T, Tibshirani R, Friedman JH, Friedman JH. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. New York: Springer (2009).

35. Gasparrini A. Distributed lag linear and non-linear models in R: the package dlnm. *J Stat Softw*. (2011) 43(8):1. doi: 10.18637/jss.v043.i08

36. Mathieu E, Ritchie H, Lucas R-G, Appel C, Gavrilov D, Giattino C, et al. Coronavirus (COVID-19) Cases (2020). Available online at: https://ourworldindata.org/covid-cases (Accessed 20/01/2025).

37. Alvares CA, Stape JL, Sentelhas PC, Gonçalves JM, Sparovek G. Köppen's climate classification map for Brazil. *Meteorol Zeitschrift*. (2013) 22(6):711–28. doi: 10.1127/0941-2948/2013/0507

38. Curado LFA, de Paulo SR, de Paulo IJC, de Oliveira Maionchi D, da Silva HJA, de Oliveira Costa R, et al. Trends and patterns of daily maximum, minimum and mean temperature in Brazil from 2000 to 2020. *Climate*. (2023) 11(8):168. doi: 10.3390/cli11080168

39. Olinto MTA, Garcêz AS, Brunelli G, Olinto FA, Fanton M, Canuto R. Relationship between temperature and relative humidity on initial spread of COVID-19 cases and related deaths in Brazil. *J Infect Dev Ctries*. (2022) 16(5):759–67. doi: 10.3855/jidc.15324

40. Prata DN, Rodrigues W, Bermejo PH. Temperature significantly changes COVID-19 transmission in (sub) tropical cities of Brazil. *Sci Total Environ*. (2020) 729:138862. doi: 10.1016/j.scitotenv.2020.138862

41. Martins LD, da Silva I, Batista WV, de Fátima Andrade M, de Freitas ED, Martins JA. How socio-economic and atmospheric variables impact COVID-19 and influenza outbreaks in tropical and subtropical regions of Brazil. *Environ Res*. (2020) 191:110184. doi: 10.1016/j.envres.2020.110184

42. Decker SRR, Wolf JM, Pille A, Freese L, Petek H, de Oliveira Rocha B, et al. Temporal trends in respiratory pathogens following the COVID-19 pandemic and climate variables: a unicentric retrospective evaluation of 24 pathogens in a temperate subtropical region. *J Med Virol*. (2024) 96(7):e29797. doi: 10.1002/jmv.29797

43. Brunelli TC, Paiva S, Siqueira AY, Santana CE, Curvo LO, Marques JB, et al. Environmental parameters and relationships with COVID-19 cases in Central South America. *Química Nova*. (2021) 44(10):1236–44. doi: 10.21577/0100-4042.20170786

44. Ye T, Xu R, Yue X, Chen G, Yu P, Coêlho MS, et al. Short-term exposure to wildfire-related PM2. 5 increases mortality risks and burdens in Brazil. *Nat Commun*. (2022) 13(1):7651. doi: 10.1038/s41467-022-35326-x

45. Gonçalves KS, Cirino GG, Costa MO, Couto LO, Tortelote GG, Hacon SS. The potential impact of PM2. 5 on the COVID-19 crisis in the Brazilian Amazon region. *Rev Saúd Pública*. (2023) 57:67. doi: 10.11606/s1518-8787.2023057005134

46. da Costa G, Pauliquevis T, Heise EFJ, Potgieter-Vermaak S, Godoi AFL, Yamamoto CI, et al. Spatialized PM2. 5 during COVID-19 pandemic in Brazil's most populous southern city: implications for post-pandemic era. *Environ Geochem Hlth*. (2024) 46(1):29. doi: 10.1007/s10653-023-01809-z

47. Feurer D, Riffe T, Kniffka MS, Acosta E, Armstrong B, Mistry M, et al. Meteorological factors, population immunity, and COVID-19 incidence: a global multi-city analysis. *Environ Epidemiol*. (2024) 8(6):e338. doi: 10.1097/EE9.0000000000000338

48. McMorris T, Swain J, Smith M, Corbett J, Delves S, Sale C, et al. Heat stress, plasma concentrations of Adrenaline, noradrenaline, 5-hydroxytryptamine and cortisol, mood state and cognitive performance. *Int J Psychophysiol*. (2006) 61(2):204–15. doi: 10.1016/j.ijpsycho.2005.10.002

49. Leppäluoto J, Westerlund T, Huttunen P, Oksa J, Smolander J, Dugué B, et al. Effects of long-term whole-body cold exposures on plasma concentrations of ACTH, beta-endorphin, cortisol, catecholamines and cytokines in healthy females. *Scand J Clin Lab Invest*. (2008) 68(2):145–53. doi: 10.1080/00365510701516350

50. Zefferino R, Di Gioia S, Conese M. Molecular links between endocrine, nervous and immune system during chronic stress. *Brain Behav*. (2021) 11(2):e01960. doi: 10.1002/brb3.1960

51. Barberia LG, Boing A, Gusmão J, Miyajima F, Abud A, Kemp B, et al. An assessment of the public health surveillance strategy based on molecular testing during three major pandemic waves of COVID-19 in Brazil. *PLoS Glob Public Health*. (2023) 3(8):e0002164. doi: 10.1371/journal.pgph.0002164

52. Boing AF, Boing AC, Veras MA, de Lacerda JT, da Silva RLP, Barbato PR, et al. Area-level inequalities in COVID-19 outcomes in Brazil in 2020 and 2021: an analysis of 1,894,165 severe COVID-19 cases. *Prev Med*. (2022) 164:107298. doi: 10.1016/j.ypmed.2022.107298

53. Van Donkelaar A, Martin RV, Brauer M, Kahn R, Levy R, Verduzco C, et al. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application. *Environ Health Persp*. (2010) 118(6):847–55. doi: 10.1289/ehp.0901623

54. Wei J, Li Z, Lyapustin A, Wang J, Dubovik O, Schwartz J, et al. First close insight into global daily gapless 1 km PM2. 5 pollution, variability, and health impact. *Nat Commun*. (2023) 14(1):8349. doi: 10.1038/s41467-023-43862-3

55. Mistry MN, Schneider R, Masselot P, Royé D, Armstrong B, Kyselý J, et al. Comparison of weather station and climate reanalysis data for modelling temperature-related mortality. *Sci Rep-Uk*. (2022) 12(1):5178. doi: 10.1038/s41598-022-09049-4

56. Raymundo CE, Oliveira MC, de Araujo Eleuterio T, de Arruda Santos Junior ÉC, da Silva MG, André SR, et al. Spatial–temporal distribution of incidence, mortality, and case-fatality ratios of coronavirus disease 2019 and its social determinants in Brazilian municipalities. *Sci Rep*. (2023) 13(1):4139. doi: 10.1038/s41598-023-31046-4

57. Yin C, Zhao W, Pereira P. Meteorological factors' effects on COVID-19 show seasonality and spatiality in Brazil. *Environ Res*. (2022) 208:112690. doi: 10.1016/j.envres.2022.112690

# Associations between COVID-19 Incidence and Environmental Stressors in Brazil: A Nationwide Study from 2020 to 2022

Leona Hoffmann[1*], Ludmilla Viana Jacobson[2], Thilo Erbertseder[3], Moritz Berger[4,5], Jörn Rittweger[1,6], Rochelle Schneider[7,8]

[1]Institute of Aerospace Medicine, German Aerospace Center (DLR), Cologne, Germany
[2] Fluminense Federal University. Institute of Mathematics and Statistics. Department of Statistics, Niterói, RJ, Brazil
[3]German Remote Sensing Data Center, German Aerospace Center (DLR), Weßling, Germany
[4]Institute of Medical Biometry, Informatics and Epidemiology (IMBIE), University Hospital Bonn, Bonn, Germany
[5]Core Facility Biostatistics, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany
[6]Department of Pediatrics and Adolescent Medicine, University Hospital Cologne, Cologne, Germany
[7]Φ-lab, European Space Agency (ESA), Frascati, Italy
[8]Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom

[*]Corresponding author. Email: leona.hoffmann@dlr.de

**Supplementary Material S1 – Data Harmonization Details**

We converted the six-digit code in the municipality variable in the health data to a seven-digit code to align it with the environmental dataset using a mapping table that ensured accurate correspondence between the two codes (21). We noted that all municipalities with six-digit codes starting with 53 are administrative regions of the Federal District Brasília. As a result, all of these regions were recoded under Brasília. The outdated seven-digit code 4314530 (Pinto Bandeira city in Rio Grande do Sul State) has been renamed to 4314548. Three seven-digit codes (2605459, 2916104, and 2919926) from sparsely populated municipalities (3101, 22337, 21432 inhabitants) were excluded from the dataset due to a lack of coverage by ERA5-land data.

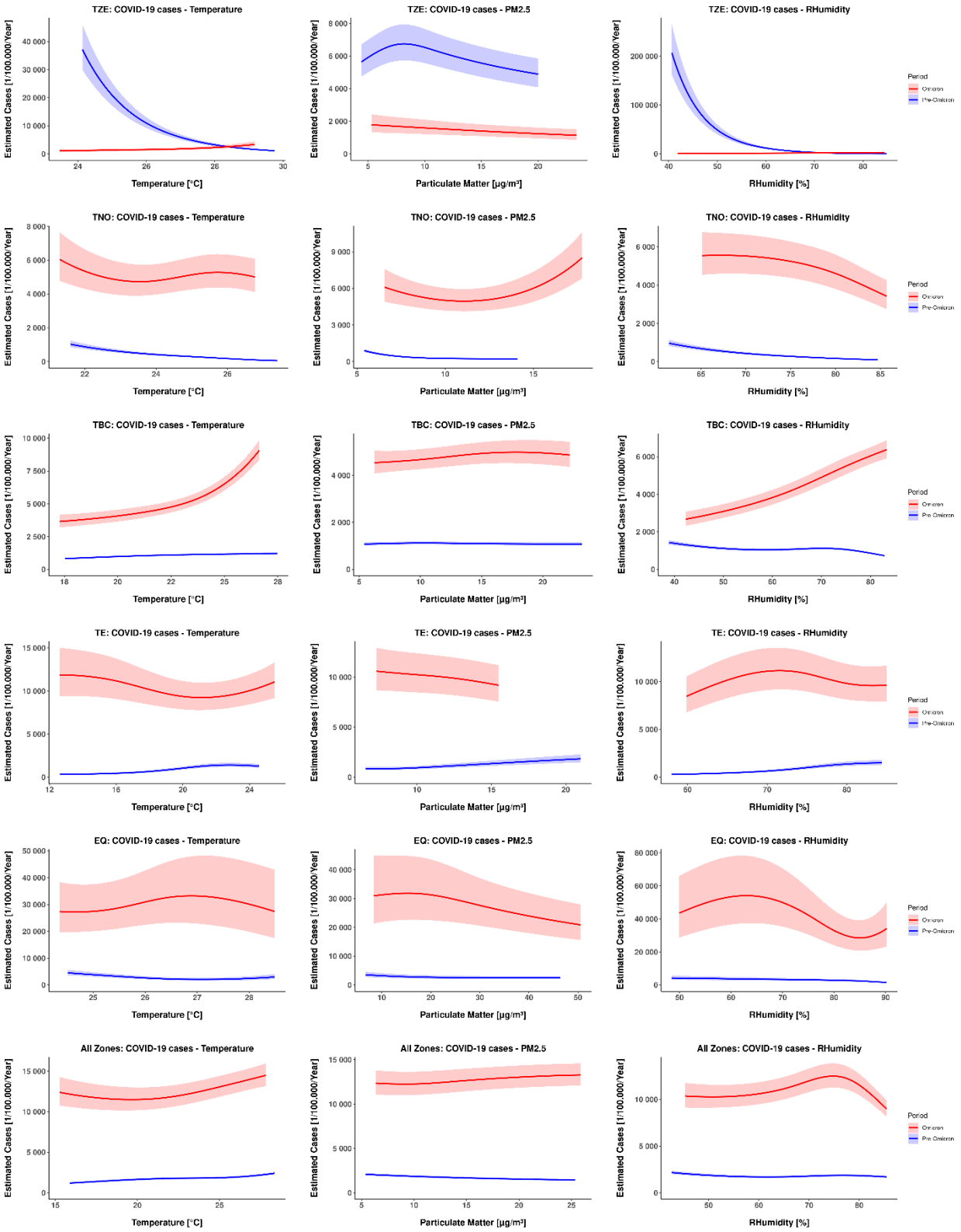


**Figure S1: Estimated COVID-19 incidence per 100,000 people per year in response to each environmental stressor, climate zone, and pre- Omicron and Omicron period.** The five climate zones are: Equatorial (EQ), Tropical Zona Equatorial (TZE), Tropical Brasil Central (TBC), Tropical Nordeste Oriental (TNO), and Temperado (TE). All other parameters were kept constant, and were set to the following values: municipality code: 1100114 (EQ, ALL), 4122206 (TE), 2107803 (TZE), 2610202 (TNO), 3500600 (TBC), month: February,

year: 2021 (in Omicron period model skipped) and environmental stressors: median. Estimates were obtained using the R function predict for Temp, RHumid, and $PM_{2·5}$, including their 5th and 95th percentiles.

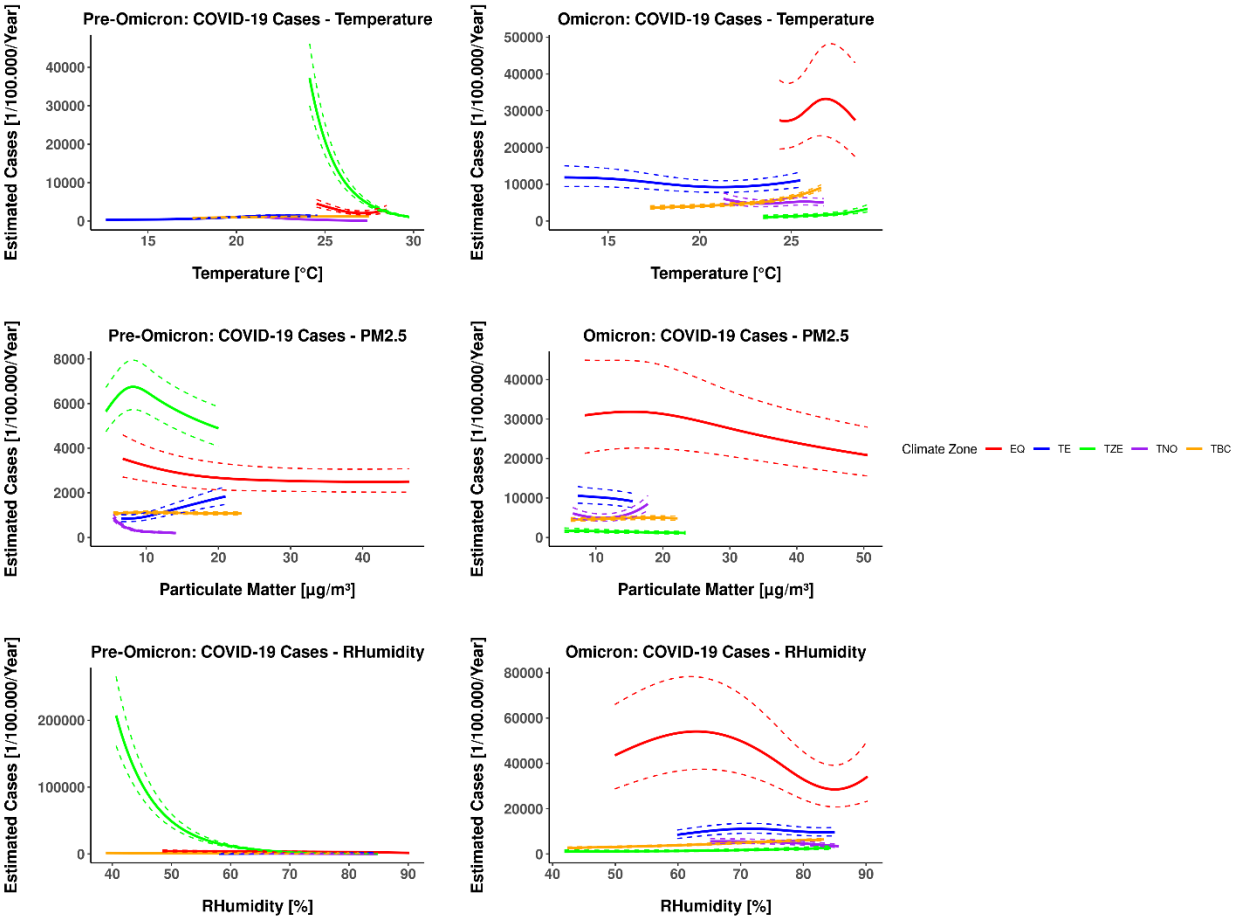


**Figure S2: Estimated COVID-19 incidence per 100,000 people per year across three environmental stressors, with incidence trends for five climate zones shown in one plot, split by pre-Omicron and Omicron periods.** The five climate zones are: Equatorial (EQ), Tropical Zona Equatorial (TZE), Tropical Brasil Central (TBC), Tropical Nordeste Oriental (TNO), and Temperado (TE). All other parameters were kept constant, and were set to the following values: municipality code: 1100114 (EQ, ALL), 4122206 (TE), 2107803 (TZE), 2610202 (TNO), 3500600 (TBC), month: February, year: 2021 (in Omicron period model skipped) and environmental stressors: median. Estimates were obtained using the R function predict for Temp, RHumid, and $PM_{2·5}$, including their 5th and 95th percentiles.

| | | Output | EQ | TZE | TBC | TNO | TE | ALL |
|---|---|---|---|---|---|---|---|---|
| Pre-Omicron period | $PM_{2.5}$ | Prediction 5-Percentile | 3367.97 (3332.35, 3396.41) | 5935.50 (4810.10, 5131.81) | 1086.37 (1030.02, 1156.57) | 768.91 (747.20, 787.89) | 849.93 (835.25, 865.21) | 2033.52 (1854.44, 2219.35) |
| | | Prediction 95-Percentile | 2497.95 (2479.35, 2513.98) | 5981.65 (5748.51, 6110.33) | 1076.18 (1010.44, 1156.57) | 205.54 (201.68, 208.94) | 1785.20 (1746.38, 1823.45) | 1466.35 (1358.82, 1576.66) |
| | | **Relative Risk** | **0.74 (0.73, 0.75)** | **0.84 (0.83, 0.85)** | **0.99 (0.94, 1.04)** | **0.27 (0.26, 0.27)** | **2.10 (2.07, 2.13)** | **0.72 (0.69, 0.76)** |
| | **Temp** | Prediction 5-Percentile | 4280.62 (4231.31, 4332.25) | 31939.14 (30292.09, 33458.10) | 847.77 (794.89, 906.67) | 934.39 (908.96, 952.51) | 344.03 (340.77, 347.50) | 1254.18 (1108.04, 1415.88) |
| | | Prediction 95-Percentile | 2721.70 (2655.65, 2798.96) | 1292.71 (1241.84, 1337.89) | 1206.54 (1144.25, 1290.04) | 72.51 (70.45, 74.53) | 1351.22 (1338.40, 1365.38) | 2241.90 (2056.98, 2425.56) |
| | | **Relative Risk** | **0.64 (0.62, 0.65)** | **0.04 (0.038, 0.042)** | **1.42 (1.33, 1.54)** | **0.08 (0.075, 0.080)** | **3.93 (3.90, 3.95)** | **1.79 (1.65, 1.92)** |
| | **RHumid** | Prediction 5-Percentile | 4105.17 (4055.56, 4140.67) | 156317.20 (150245.60, 162986.80) | 1362.99 (1261.32, 1472.70) | 874.15 (845.98, 901.10) | 309.65 (304.36, 315.07) | 2084.66 (1895.16, 2276.21) |
| | | Prediction 95-Percentile | 1911.53 (1889.72, 1932.41) | 244.56 (239.55, 251.38) | 833.32 (791.81, 882.79) | 108.33 (106.18, 110.79) | 1486.82 (1462.64, 1511.31) | 1765.25 (1622.45, 1891.06) |
| | | **Relative Risk** | **0.47 (0.46, 0.47)** | **0.002 (0.001, 0.002)** | **0.61 (0.58, 0.65)** | **0.12 (0.12, 0.13)** | **4.80 (4.74, 4.86)** | **0.85 (0.80, 0.89)** |
| Omicron period | $PM_{2.5}$ | Prediction 5-Percentile | 31299.93 (30866.07, 31806.06) | 1749.86 (1735.51, 1767.60) | 4556.86 (3567.23, 5800.36) | 5889.33 (5807.12, 5807.18) | 10544.78 (10501.43, 10594.54) | 12310.37 (8579.40, 16922.38) |
| | | Prediction 95-Percentile | 21425.30 (20981.21, 21931.37) | 1161.40 (1149.91, 1173.23) | 4904.63 (3907.02, 5971.08) | 7837.45 (7630.87, 7987.54) | 9301.16 (9256.31, 9349.79) | 13254.36 (8847.12, 18557.81) |
| | | **Relative Risk** | **0.68 (0.67, 0.70)** | **0.66 (0.66, 0.67)** | **1.08 (0.95, 1.21)** | **1.33 (1.29, 1.36)** | **0.88 (0.88, 0.88)** | **1.07 (0.90, 1.23)** |
| | **Temp** | Prediction 5-Percentile | 27245.33 (26846.29, 27700.59) | 1097.95 (1079.60, 1114,77) | 3698.52 (2761.67, 5131.11) | 5798.46 (5697.54, 5890.52) | 11851.48 (11825.91, 11888.24) | 12207.76 (7846.77, 18209.60) |
| | | Prediction 95-Percentile | 28620.93 (28092.80, 29103.60) | 2992.50 (2935.86, 3054.09) | 8178.38 (6770.67, 8498.98) | 5128.27 (5105.85, 5148.52) | 10573.98 (10491.42, 10672.14) | 14164.12 (9797.47, 18519.97) |
| | | **Relative Risk** | **1.05 (1.03, 1.06)** | **2.73 (2.66, 2.79)** | **2.21 (1.37, 2.89)** | **0.88 (0.87, 0.90)** | **0.89 (0.89, 0.90)** | **1.16 (0.83, 1.50)** |
| | **RHumid** | Prediction 5-Percentile | 45630.58 (44700.99, 46299.88) | 1097.77 (1092.11, 1104.77) | 2754.15 (1911.46, 4582.41) | 5553.40 (5528.79, 5578.36) | 8829.22 (8710.67, 8955.81) | 10320.15 (6957.95, 14832.00) |
| | | Prediction 95-Percentile | 30644.24 (30379.57, 30862.90) | 2476.09 (2461.13, 2492.48) | 6172.33 (5427.90, 6466.08) | 3642.10 (3594.76, 3680.00) | 9590.23 (9548.10, 9636.56) | 10024.22 (7287.82, 13017.42) |
| | | **Relative Risk** | **0.67 (0.66, 0.68)** | **2.26 (2.24, 2.27)** | **2.24 (1.35, 2.95)** | **0.66 (0.65, 0.66)** | **1.09 (1.07,1.10)** | **0.97 (0.74, 1.18)** |

**Table S1: Estimated COVID-19 Incidence and Associations with Environmental Stressors by Climate Zone and pre-Omicron and Omicron period.** This table shows the estimated cases per 100,000 people per year for the 5th and 95th percentiles of environmental stressors. Confidence intervals are in brackets. Green cells indicate a strong positive association (RR > 1.3), yellow cells indicate a strong negative association (RR < 0.7), and unshaded cells represent a weak or no association ($0.7 \leq RR \leq 1.3$). The five climate zones are: Equatorial (EQ), Tropical Zona Equatorial (TZE), Tropical Brasil Central (TBC), Tropical Nordeste Oriental (TNO), and Temperado (TE).

# 4. Discussion

While detailed findings were presented and discussed in the three individual papers, this discussion emphasizes the overarching patterns in the associations between COVID-19 incidence and environmental stressors across different climatic contexts.

## 4.1 Key results

The analyses across the three papers reveal that Temp and $PM_{2.5}$ have the strongest associations with COVID-19 incidence. Both variables were examined in all three manuscripts. The associations with COVID-19 incidence varied depending on the pandemic phase, specifically during the Omicron and pre-Omicron periods (Hoffmann et al., 2025a), with a relative risk (RR) value of 0.39 for pre-Omicron and 0.92 for Omicron period for Temp and a RR value of 2.41 for pre-Omicron and 0.70 for Omicron period for $PM_{2.5}$ (Hoffmann et al., 2025a). Additionally, the levels of these variables influenced the estimated COVID-19 incidence differently.

Climate is inherently complex, and different climate zones exhibited differing associations among these variables. In Brazil, the pattern observed in the models for the climate zones Tropical Brazil Central, Temperado and in the Overall model, showed that higher temperatures were associated with increased incidence rates (Hoffmann et al., 2025b). Conversely, in BW, the data indicated that lower temperatures are associated with higher estimated COVID-19 incidence rates (Hoffmann et al., 2025a). While this may initially seem contradictory, it is not when considering that the temperature ranges in these regions are quite different. This suggests a strong association with COVID-19 incidence for both extremely hot and extremely cold temperatures.

It is crucial to understand that individual factors do not operate in isolation but are influenced by other variables. Hoffmann et al. (2024) indicated that particular variables can be affected by additional environmental stressors. Other factors, such as vaccination status and socio-economic conditions, also played significant roles. In the absence of laboratory data, it must be assumed that additional influences exist that may not be fully captured in the analyses. Moreover, distinguishing between causality and correlation is

vital. The direction of the association cannot be determined with 100% accuracy, requiring caution when interpreting the results.

## 4.2 Interpretation

A considerable number of research articles have confirmed that COVID-19 transmission is affected by weather conditions and air pollution (Faruk et al., 2022; Han et al., 2023). In the literature, most research has identified temperature as an essential variable influencing COVID-19 transmission (Faruk et al., 2022).

COVID-19 is a respiratory disease, and environmental stressors such as temperature and PM2.5 have also been linked to other respiratory viral infections, including influenza (Rittweger et al., 2022). Higher incidence rates in colder temperatures may be due to spending more time indoors, where heated air, limited ventilation, and low humidity prevail. According to von Hahn et al. (2024) and Mannan et al. (2021), people in Europe and around the world spend an average of 90% of their lives indoors. The transmission of COVID-19 is particularly prevalent in poorly ventilated or poorly circulated indoor spaces (Yao et al., 2020). Additionally, Bulfone et al. (2021) found that the odds of transmission indoors were 18.7 times higher than outdoors. In Europe, the virus was found to spread rapidly with the mixture of the lower outdoor temperature and reduced relative humidity indoors (Faruk et al., 2022). Cold, dry air can trigger inflammatory processes in the nasal mucosa, compromising its integrity and weakening its barrier against viral pathogens (Togias et al., 1985). Cold temperatures and low humidity are associated with increased occurrence of respiratory tract infections (Mäkinen et al., 2009).

The analyses focused on outdoor air pollution and weather conditions, which were significantly associated with COVID-19 incidence. These associations may stem from the direct effects of environmental factors and the indirect influence of human behavior, such as heated indoor environments and limited ventilation. As Fu et al. (2022) concluded, indoor levels of particulate matter correlate exponentially with outdoor concentrations, highlighting the role of occupant behavior in the infiltration of outdoor pollution in indoor environments.

In temperate climates like Germany, distinct seasons and colder winter temperatures are common. In contrast, tropical and subtropical climates, such as those in Brazil, lack similar seasonal variations (see Hoffmann et al. (2025b), Figure 2). While respiratory viral diseases in cool, dry conditions have been extensively studied in temperate regions, their associations with environmental stressors in tropical climates is less understood (Tamerius et al., 2011; Yuan et al., 2021). Paynter (2015) found that influenza incidence rises during the rainy season, likely due to increased contact transmission, as high humidity enhanced droplet survival and surface deposition, potentially favoring indirect transmission. Additionally, Tamerius et al. (2011) listed several studies linking increased influenza activity to the rainy season in various tropical populations, where humidity is typically at its highest.

The association between environmental stressors and COVID-19 is complex, particularly in Brazil's diverse climate zones (Hoffmann et al., 2025b). It is not contradictory that both high and low temperatures can lead to increased COVID-19 infections, as climatic conditions, including the range and concentration levels of various factors like relative humidity, can vary significantly. Physiological stress responses to extreme temperatures may play a role (Hoffmann et al., 2025b). Further analysis, particularly in subtropical and tropical regions, is needed to understand the mechanisms behind the links between COVID-19 and environmental stressors.

### 4.3 Limitations

The strengths of this dissertation include the long observation period, broad spatial coverage, multiple stressors, and distinction between pandemic phases, which have already been outlined in section 2.3 Added value of this work. Potential limitations include monthly and quarterly temporal data aggregation, data source constraints, and the evolving nature of the COVID-19 pandemic, which affects testing, vaccination, and non-pharmaceutical interventions. The extent of underreporting in COVID-19 data is also uncertain. These limitations were addressed in detail in the three publications (Hoffmann et al., 2025a; Hoffmann et al., 2024; Hoffmann et al., 2025b).

There are some additional limitations to consider. Several confounding variables, such as socio-economic status, household density, access to healthcare, occupational exposure, mobility patterns, and behaviors such as mask use and testing, could not be included in the analysis. In some cases, this was due to limited data availability, while in others, variables were excluded because they were highly correlated with the temporal variable, potentially introducing multicollinearity and compromising the stability and interpretability of the model estimates. Future research would benefit from incorporating such data, where possible, to reduce residual bias and gain a more comprehensive understanding of the underlying associations.

As with any model, it is essential to emphasize that the results' generalizability is limited by strong contextual dependence and the specific focus on BW and Brazil. It would be valuable to expand the analysis to more countries to further investigate different or similar patterns and links between various climatic conditions, environmental stressors, and COVID-19 incidence.

This dissertation focused exclusively on COVID-19 incidence. However, the severity of COVID-19 and its potential associations with environmental stressors are of considerable scientific and epidemiological interest. Expanding the scope of this research to include disease severity, such as hospitalization and ventilation rates or mortality, represents a meaningful and logical extension of this work. Moreover, incorporating comorbidities could provide important insights into the link between environmental stressors and COVID-19 incidence or severity. While this would be a valuable extension of the current work, it requires individual-level data that was not available for this study design. Future analyses should consider integrating such data where feasible.

## 4.4 Conclusion

This dissertation has demonstrated that environmental stressors, particularly Temp and $PM_{2.5}$, show significant and context-dependent associations with COVID-19 incidence across different climate zones. The findings suggest that both extremely high and low temperatures may contribute to increased infection rates - likely due to a combination of direct physiological stress responses and indirect behavioral adaptations, such as time

spent indoors. In temperate regions like Germany, colder conditions are associated with higher incidence rates, while in tropical zones like Brazil, elevated temperatures and humidity may also increase the risk of infection.

These results emphasize the need to consider local climatic and environmental conditions when assessing the dynamics of respiratory disease transmission, particularly in relation to COVID-19. They highlight that climate-health associations are non-linear and region-specific, and that simplistic one-size-fits-all assumptions may overlook important mechanisms.

Additional research is essential to clarify the causal pathways that connect environmental stressors to disease incidence. This should ideally involve interdisciplinary approaches that bring together epidemiology, climatology, and behavioral sciences. Understanding these interactions is vital not only for responding to future outbreaks of COVID-19 but also for preparing for emerging public health challenges in a changing climate.

4.5 References

Bulfone, TC, Malekinejad, M, Rutherford, GW, & Razani, N. Outdoor transmission of SARS-CoV-2 and other respiratory viruses: a systematic review. In: The Journal of infectious diseases 2021; 223(4): 550-561

Faruk, MO, Rahman, MS, Jannat, SN, Arafat, Y, Islam, K, & Akhter, S. A review of the impact of environmental factors and pollutants on covid-19 transmission. In: Aerobiologia 2022; 38(3): 277-286

Fu, N, Kim, MK, Huang, L, Liu, J, Chen, B, & Sharples, S. Experimental and numerical analysis of indoor air quality affected by outdoor air particulate levels (PM1. 0, PM2. 5 and PM10), room infiltration rate, and occupants' behaviour. In: Science of the Total Environment 2022; 851: 158026

Han, J, Yin, J, Wu, X, Wang, D, & Li, C. Environment and COVID-19 incidence: A critical review. In: journal of environmental sciences 2023; 124: 933-951

Hoffmann, L, Gilardi, L, Antoni, T, Baltruweit, M, Bittner, M, Breitner, S, Dally, S, Erbertseder, T, Hawighorst-Knapstein, S, Schmitz, MT, Schneider, R, Wüst, S, & Rittweger, J. Modulation of COVID-19 incidence by environmental stressors is variant between pre-Omicron and Omicron periods. In: Scientific Reports 2025a; 15(1): 27636

Hoffmann, L, Gilardi, L, Schmitz, M-T, Erbertseder, T, Bittner, M, Wüst, S, Schmid, M, & Rittweger, J. Investigating the spatiotemporal associations between meteorological conditions and air pollution in the federal state Baden-Württemberg (Germany). In: Scientific Reports 2024; 14(1): 5997

Hoffmann, L, Jacobson, L, Erbertseder, T, Berger, M, Rittweger, J, & Schneider, R. Associations between COVID-19 Incidence and Environmental Stressors in Brazil: A Nationwide Study from 2020 to 2022. In: Frontiers in Environmental Health 2025b; 4: 1635503

Mäkinen, TM, Juvonen, R, Jokelainen, J, Harju, TH, Peitso, A, Bloigu, A, Silvennoinen-Kassinen, S, Leinonen, M, & Hassi, J. Cold temperature and low humidity are associated with increased occurrence of respiratory tract infections. In: Respiratory medicine 2009; 103(3): 456-462

Mannan, M, & Al-Ghamdi, SG. Indoor air quality in buildings: a comprehensive review on the factors influencing air pollution in residential and commercial structure. In:

International journal of environmental research and public health 2021; 18(6): 3276

Paynter, S. Humidity and respiratory virus transmission in tropical and temperate settings. In: Epidemiology & Infection 2015; 143(6): 1110-1118

Rittweger, J, Gilardi, L, Baltruweit, M, Dally, S, Erbertseder, T, Mittag, U, Naeem, M, Schmid, M, Schmitz, M-T, & Wüst, S. Temperature and particulate matter as environmental factors associated with seasonality of influenza incidence–an approach using Earth observation-based modeling in a health insurance cohort study from Baden-Württemberg (Germany). In: Environmental Health 2022; 21(1): 131

Tamerius, J, Nelson, MI, Zhou, SZ, Viboud, C, Miller, MA, & Alonso, WJ. Global influenza seasonality: reconciling patterns across temperate and tropical regions. In: Environmental Health Perspectives 2011; 119(4): 439-445

Togias, AG, Naclerio, RM, Proud, D, Fish, JE, Adkinson, N, Kagey-Sobotka, A, Norman, PS, & Lichtenstein, LM. Nasal challenge with cold, dry air results in release of inflammatory mediators. Possible mast cell involvement. In: The Journal of clinical investigation 1985; 76(4): 1375-1381

von Hahn, N, Graveling, R, Kuhl, K, & Van den Broek, K, 2024: Indoor air quality (IAQ). https://oshwiki.osha.europa.eu/en/themes/indoor-air-quality-iaq (Access date: 2025/04/29)

Yao, M, Zhang, L, Ma, J, & Zhou, L. On airborne transmission and control of SARS-Cov-2. In: Science of the Total Environment 2020; 731: 139178

Yuan, H, Kramer, SC, Lau, EH, Cowling, BJ, & Yang, W. Modeling influenza seasonality in the tropics and subtropics. In: PLoS computational biology 2021; 17(6): e1009050

## 5. Statement on own contribution

As part of the first publication entitled “Investigating the Spatiotemporal Associations Between Meteorological Conditions and Air Pollution in the Federal State of Baden-Württemberg (Germany)”, I was substantially involved in the planning of the scientific work, under the supervision of Professor Jörn Rittweger and with support from Marie-Therese Schmitz. I was responsible for merging and consolidating data from multiple sources. Lorenza Gilardi provided the preprocessed environmental data. I independently performed the formal analyses, created the visualizations, and developed the software used in the study. Marie-Therese Schmitz and Professor Matthias Schmid reviewed the analytical methods. The interpretation of the results was primarily carried out by me. I was responsible for placing the findings within the scientific context, deriving the main conclusions, and formulating the discussion section. Thilo Erbertseder, Lorenza Gilardi, Sabine Wüst, and Michael Bittner contributed to the interpretation of the analytical results within the physical-chemical and atmospheric context.

In my second publication, entitled “Modulation of COVID-19 Incidence by Environmental Stressors is Variant Between Pre-Omicron and Omicron Periods”, I was substantially involved in the conception of the study and the planning of the methodology. The project was developed under the guidance of Professor Jörn Rittweger, with support from Tobias Antoni and Maxana Baltruweit. I coordinated the data collection and merged and consolidated the data sources. Lorenza Gilardi provided preprocessed environmental data, and Simon Dally supplied health data. I conducted the analyses, created the visualizations, and developed the software. Marie-Therese Schmitz reviewed the analytical methods. I primarily interpreted the results, contextualized the findings, and drafted the discussion. I incorporated specific input from co-authors.

In my third publication, entitled “Associations Between COVID-19 Incidence and Environmental Stressors in Brazil: A Nationwide Study from 2020 to 2022”, I was substantially involved in the conception of the study and the planning of the methodology. Rochelle Schneider and Professor Jörn Rittweger contributed to the study design, while Rochelle Schneider, Professor Jörn Rittweger, and Moritz Berger supported the methodological planning. I collected the environmental data, merged and consolidated data from various sources, and performed the formal analyses. Ludmilla Viana Jacobson collected the health data. I created the visualizations and developed the software. Moritz Berger reviewed the analytical methods. I primarily interpreted the results, contextualized the findings, and drafted the discussion. Ludmilla Viana Jacobson contributed to the interpretation within the Brazilian context, and Thilo Erbertseder supported the section on policy implications.

I confirm that I have written this thesis independently and have not used any sources or aids other than those specified by me.

I hereby confirm that my thesis complies with the Statement by the Executive Committee of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) on the

Influence of Generative Models of Text and Image Creation on Science and the Humanities and on the DFG’s Funding Activities.

To specify, I used ChatGPT (OpenAI) to refine the language of certain English passages. The original content and ideas are entirely my own; ChatGPT was only used to improve phrasing and linguistic clarity. After using this tool, I reviewed and edited the relevant passages and take full responsibility for the content of the published dissertation.

## 6. Acknowledgements

I want to express my heartfelt gratitude to my dissertation committee members: Matthias Schmid, Jörn Rittweger, Sabine Wüst, Daniel Aeschbach, and Rochelle Schneider. Their thoughtful feedback and support were crucial to the success of my project. I especially thank Rochelle Schneider for making my time at the ESA phi lab possible, which greatly enriched my research and personal growth. I'm thankful to Sabine Wüst for stepping in as a second reviewer after Jörn Rittweger's passing and for her invaluable guidance.

I also want to thank my colleagues for many inspiring discussions, particularly my office mate Jonas. I’m grateful for all the interesting conversations, the joint lunches, and the sense of community we built in our shared space.

My deepest thanks go to my friends Carina, Carola, and Lea. Even though they are not part of the academic world, they were always there to celebrate small victories and to listen with patience and kindness during tough times. Their support always uplifted and encouraged me.

To my family – especially my parents, my sister Emilie, and my grandparents – thank you for reminding me of what truly matters in life. You have taught me the value of taking a step back and keeping perspective. You have also shown me the quiet strength that comes from unconditional support. Knowing I could always rely on you made all the difference.

Finally, thank you to my partner and fiancé Robin for being my constant through all the ups and downs. Your calm, your care, and your love carried me through more than you know.