

Letter

Spatiotemporal Variability of Influencing Factors on Tuberculosis Incidence in Jiangsu Province, China during 2011–2021



Yifan Tang^{1,2,&}, Tenglong Li^{1,&}, Cheng Chen³, Kai Wang^{1,2}, Mingming Chen¹, Junhui Lin¹, Sifan Wang¹, Limei Zhu³, Chengxiu Ling^{1,4,5,#}, and Qiao Liu^{3,#}

Tuberculosis (TB), one of the oldest infectious diseases caused by *Mycobacterium tuberculosis*, poses a considerable challenge to global public health. There are approximately 10 million new TB cases worldwide annually, and TB claims the lives of nearly 3 million people each year, making it one of the leading causes of death from a single infectious disease^[1]. China ranks third globally in terms of TB burden, with approximately 733,000 TB cases reported in 2023^[2]. Based on the ecological model of health determinants developed by Whitehead and Dahlgren, health determinants can be classified into direct causes, such as disease-specific causes or individual influencing factors, and indirect causes, such as social determinants or environmental factors. Jiangsu Province, located on the southeastern coast of China, has reported approximately 20,000 cases of TB annually in recent years, and significant spatiotemporal heterogeneity because of non-medical determinants has contributed to variation in TB incidence across counties. Previous studies examined the factors influencing TB and its temporal and spatial disparities to improve healthcare allocation and management^[3].

Bayesian statistical modeling has recently gained popularity for analyzing data incorporating temporal and spatial information. This method incorporates spatiotemporal random effects into regression models, allowing the simultaneous analysis of disease dynamics, spatial heterogeneity, and potential predictors, thereby improving the robustness of estimating the relation between influencing factors and disease incidence^[4]. The overall impact of influencing factors on TB incidence

has been assessed using the extreme gradient boosting (XGBoost) model, conditioned on the spatial and temporal distribution of TB, along with models, such as geographical detectors and empirical orthogonal functions^[5,6]. To the best of our knowledge, no study has quantitatively analyzed the contributions of temporal and spatial variability to external factors. To fill this gap, this study introduces a novel spatiotemporal decomposition method based on Bayesian spatiotemporal modeling. This approach utilizes XGBoost to rank the importance of all influencing factors, providing a comprehensive partitioned contribution of spatial, temporal, and residual components to variations in TB incidence. This can provide insights into effective TB control strategies in specific regions and seasons.

The TB data used in this study consisted of the total number of newly reported TB cases annually in 95 counties or districts of Jiangsu Province, an economically developed region in Eastern China^[7], from 2011 to 2021. These data were obtained by aggregating 353,931 individual records from the TB Information Management System maintained by the Jiangsu Provincial Center for Disease Control and Prevention. Our main outcome was the reported incidence of TB, defined as the number of reported cases divided by the resident population (per thousand individuals; yielding values between zero and one); for simplicity, it will be referred to as “TB incidence” henceforth. The incidence of TB in Jiangsu Province varied significantly across counties/districts and years (Supplementary Figure S1), with incidence in the northern region being higher than that in the southern and central regions. The northern city of Xuzhou had the highest TB incidence, with a

doi: [10.3967/bes2026.006](https://doi.org/10.3967/bes2026.006)

1. Academy of Pharmacy, Xi'an Jiaotong-Liverpool University, Suzhou 215123, Jiangsu, China; 2. Department of Mathematical Sciences, University of Liverpool, Liverpool L693BX, UK; 3. Jiangsu Provincial Center for Diseases Control and Prevention, Nanjing 210009, Jiangsu, China; 4. Jiangsu Provincial Higher Education Key Laboratory of Cell Therapy Nanoformulation (Construction), Suzhou 215123, Jiangsu, China; 5. Suzhou Industrial Park Center of Excellence for Single-Molecule and Single-Cell Analysis, Xi'an Jiaotong-Liverpool University, Suzhou 215123, Jiangsu, China

multiyear average of 42 cases per 100,000 individuals. The lowest incidences of TB were reported in Wuxi and Suzhou, with multiyear average incidence rates of 32 and 34 per 100,000, respectively. Over the study period, TB incidence across all 95 districts showed an overall decreasing trend from 2011 to 2021.

Based on individual TB records from the Jiangsu Provincial Center for Disease Control and Prevention, we aggregated case-level information to the county-year level and computed six clinical indicators (age, delay time, delay rate, confirmed rate, and admission rate) to reflect the efficiency and management capacity of the local healthcare system in screening, diagnosing, and treating TB. Additionally, indirect influencing factors related to socioeconomic status [Gross Domestic Product (GDP), health technicians], air pollution (PM_1 , $PM_{2.5}$, PM_{10} , and O_3), meteorological conditions (wind speed, precipitation, temperature, and relative humidity), and geographical information [Normalized Difference Vegetation Index (NDVI)] were also included. For factors beyond socioeconomic status, we clipped the raster layers to county administrative boundaries and computed the annual mean values within each county to derive the corresponding variables. A brief description and summary of the explanatory variables are provided in Supplementary Table S1.

We established a spatiotemporal model using logit-linked Bayesian Beta regression to analyze TB incidence across 95 districts and counties in Jiangsu Province from 2011 to 2021 (Model S1). Clinical factors were incorporated into the Bayesian Beta regression model, and an innovative variation partitioning procedure was proposed to analyze the proportion of temporal and spatial effects of socioeconomic status, air pollution, meteorological conditions, and geographical information on TB incidence. Each factor was decomposed into temporal, spatial, and residual components (Model S2). First, each factor was decomposed into temporal, spatial, and residual parts to isolate the time-related, location-related, and unexplained variations in TB incidence. Second, the partitioned contributions of these components were quantified based on the relative reductions in the marginal log-likelihoods of the nested Bayesian Beta regression models, representing the variations in TB incidence in terms of temporal, spatial, and residual components (cf. Eq.(S.1)).

We subsequently applied the XGBoost machine learning algorithm to quantify the relative

importance of all factors influencing TB incidence (Model S3).

We constructed a baseline Bayesian Beta regression model using only the clinical factors (Supplementary Table S2). Our analysis revealed significant age-related differences, with a 3% increase in TB incidence for each year of age (Odds Ratio (*OR*): 1.030, 95% Confidence Interval (*CI*): 1.022–1.037), highlighting the greater susceptibility of the elderly to TB, likely because of a weakened immune system and limited knowledge of prevention. Additionally, diagnostic delay had a significant impact, with a 0.3% decrease in TB incidence for every 1% increase in the delay rate (*OR*: 0.997; 95% *CI*: 0.995–0.999). The confirmed diagnosis rate was inversely associated with TB incidence (*OR*: 0.985, 95% *CI*: 0.983–0.988). The observed association suggests that areas with better diagnostic capabilities may be more effective for the early detection and treatment of TB, leading to reduced transmission rates. However, this relation might also be confounded by underlying factors, such as the quality of the local health system, including case detection practices and diagnostic capacity, which could influence the observed TB incidence^[8]. This underscores the challenges of making causal inferences in ecological studies, as the results may be influenced by health-system factors rather than solely by the epidemiological factors under investigation.

Figure 1 presents the partitioned contributions of the temporal, spatial, and residual components for each influencing factor to the variations in TB incidence. Spatial variations in NDVI (48%), temperature (42%), wind speed (40%), and humidity (36%) contributed significantly to the heterogeneity in TB incidence, whereas temporal variations in GDP, PM_{10} , O_3 , and PM_1 exerted a substantial impact (40%–50%). Overall, the impact of socioeconomic status and air pollution on TB incidence was less pronounced spatially (14.81% and 28.31%, respectively) than temporally (40.71% and 40.58%, respectively). By contrast, meteorological conditions and geographical information exhibited the opposite trends, with their impact on TB incidence being more significant for spatial variations.

In particular, air pollution factors significantly affect TB incidence over time, likely due to the long-term effects of pollutant exposure on human health. Prolonged exposure to high pollution levels leads to the accumulation of harmful substances that cause chronic damage to the respiratory and immune systems, thereby increasing the risk of developing

TB. Therefore, the effects of air pollution on TB incidence were more persistent and cumulative over time. By contrast, the spatial impact of meteorological conditions on TB incidence was more pronounced, likely because of significant variations in precipitation, temperature, and humidity across regions. Different factors may be highly correlated in both space and time, such that their patterns are partially driven by shared underlying influences. Notably, our current approach does not account for intervariable interactions or potential confounding effects, which are partially absorbed into the residual component and thus fall outside the scope of the present analysis.

We investigated the relative importance of all factors in the change in TB incidence using the XGBoost method; the hyperparameters and model performance are shown in Supplementary Table S3. As shown in Figure 2, air pollution and clinical factors were the dominant factors influencing TB incidence, with cumulative proportions of 42.59% and 25.84%, respectively. Among these, O_3 showed the strongest association with TB incidence (37.96%), followed by health technicians (12.13%), confirmation rate (11.03%), age (8.24%), and temperature (6.01%). By contrast, the admission rate, PM_{10} , $PM_{2.5}$, and precipitation had relatively lower effects on TB

incidence (all less than 2%). Clearly, O_3 outweighs other factors, including clinical factors. Tropospheric O_3 is formed through photochemical reactions between ozone precursors under sunlight and is influenced by both anthropogenic and natural factors, such as vehicular emissions and vegetation release^[9]. The observed association between O_3 and TB incidence may be attributed to its role as a proxy for a complex mixture of photochemical pollutants and other unmeasured environmental and social factors, rather than a direct biological effect^[10], warranting further investigation in future studies.

We propose a joint probability framework that first attributes each factor's share of the explained variance to its XGBoost gain, which is computed independently for every predictor. Subsequently, the total factor-specific contribution is further decomposed into three terms: the contributions of the temporal, spatial, and residual components of the specific factor to the variations in TB incidence (Model S3). Our analyses suggest that air pollution, in addition to clinical factors, has a major influence on the spatial and temporal variation in TB incidence. By contrast, meteorological conditions contributed relatively little to TB incidence, which is consistent with the findings of previous studies. The government needs to strengthen collaboration

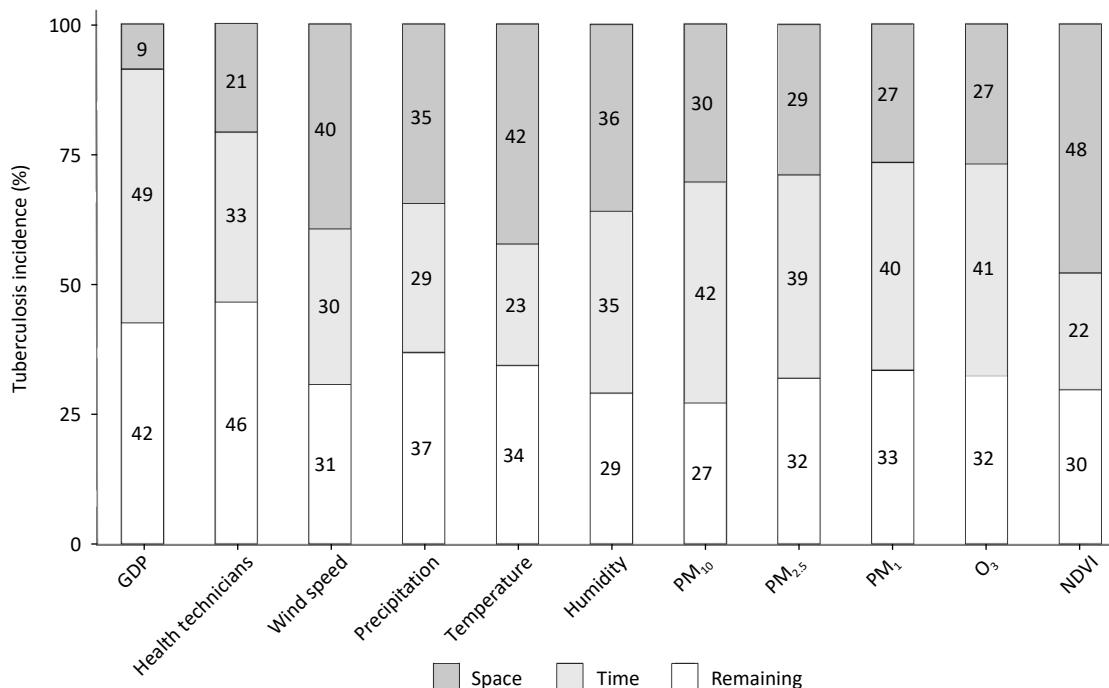


Figure 1. Partitioned contribution of socioeconomic status, meteorological conditions, air pollution, and geographical information to variations in tuberculosis incidence in terms of temporal, spatial, and residual components.

among multiple sectors, including health, environmental protection, and education, to achieve information sharing and resource integration and to establish a working mechanism for joint prevention and control. By implementing comprehensive prevention and control measures, the level of TB prevention and control can be significantly enhanced, effectively reducing both the incidence of the disease and risk of its spread, thereby safeguarding public health and social stability. Using this framework, we generated a cross-tabulation (Table 1) showing the partitioning of temporal, spatial, and residual contributions to variations in TB

incidence. Supplementary Figure S2 illustrates the XGBoost-derived relative importance of the 11 external indirect factors, excluding clinical variables, and Supplementary Table S4 presents the corresponding cross-tabulations.

This study has some limitations. First, owing to data limitations, certain unmeasured factors, such as occupational categories, population migration, and host-related susceptibility, were excluded from the analysis of TB incidence. Second, the coarse spatial resolution of the ecological data at the county/district level, together with the use of annual data on TB incidence and associated factors, may

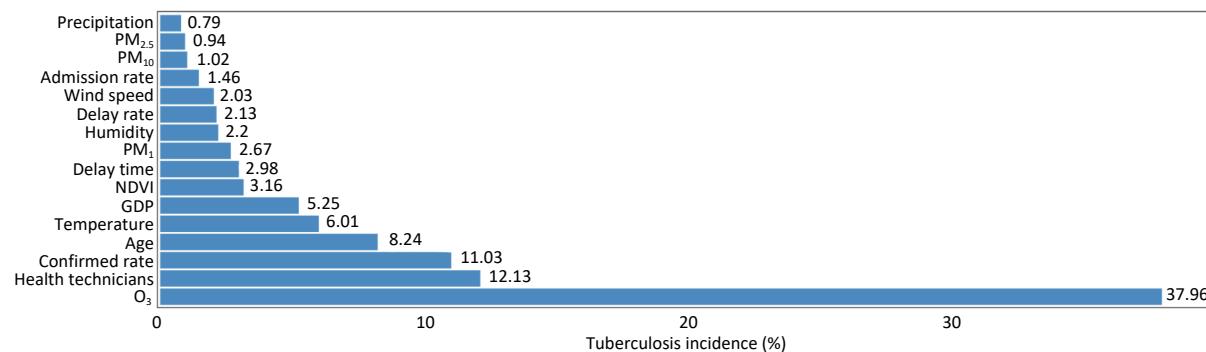


Figure 2. Relative importance of exposure to clinical factors, socioeconomic status, meteorological conditions, air pollution, and geographical information in explaining variations in tuberculosis incidence using the XGBoost model.

Table 1. Relative importance of influencing factors based on the partitioned contributions of temporal, spatial, and residual components to variations in tuberculosis incidence

Factors	Variables	Time (%)	Space (%)	Remaining (%)	Relative Importance (%)
Socioeconomic	GDP	2.56	0.45	2.23	5.24
	Health Technicians	3.95	2.54	5.64	12.13
	Total	6.52	3.00	7.85	17.37
Meteorological	Wind Speed	0.61	0.80	0.62	2.03
	Precipitation	0.23	0.27	0.29	0.79
	Temperature	1.40	2.55	2.06	6.01
	Humidity	0.77	0.79	0.64	2.20
	Total	3.00	4.42	3.61	11.03
Air Pollution	PM ₁	0.43	0.31	0.28	1.02
	PM _{2.5}	0.37	0.27	0.30	0.94
	PM ₁₀	1.07	0.71	0.89	2.67
	O ₃	15.49	10.26	12.21	37.96
Geographical	Total	17.36	11.56	13.68	42.60
	NDVI	0.71	1.52	0.93	3.16

Note. Clinical factors (age: 8.24%, delay time: 2.98%, delay rate: 2.13%, confirmed rate: 11.03%, admission rate: 1.46%) contribute a total of 25.84%.

have led to exposure misclassification and an ecological fallacy. Therefore, our results may not be directly applicable to individual-level conclusions. Finally, using reported cases as proxies for actual prevalence may introduce bias, as potential underreporting and reporting inconsistencies may undermine the reliability of the results.

Funding This research was partially supported by the National Natural Science Foundation of China (82574173, 82003516), Jiangsu Provincial Natural Science Foundation (BK20251958), Jiangsu Provincial Medical Key Discipline (ZDXK202250), Top Talent Awards Project Fund (RDF-TP-0023, RDF-TP-0030), and Postgraduate Research Fund (PGRS2112022) at Xi'an Jiaotong–Liverpool University.

Competing Interests The authors declare that they have no competing interests.

Ethics This work did not require ethical approval from human participants or animal welfare committees, as it did not involve personal information.

Patient and Public Involvement Not applicable.

Authors' Contributions Q.L. and C.L. contributed to the research design and conceptualization. C.L., Y.T., and T.L. developed the methodology and conceptualization. Y.T. prepared the initial draft of the manuscript and programmed the code. C.L., Y.T., and T.L. analyzed the main results. C.C., L.Z., and Q.L. conducted the surveys and collected the data. Y.T., K.W., J.L., and S.W. processed the data. C.C., L.Z., Q.L., M.C., C.L., and T.L. commented on and revised the manuscript. All the authors have read and approved the final version of the manuscript.

Acknowledgements We thank the editor and reviewers for their valuable comments and suggestions, which greatly improved the article.

Data Sharing The data that support the findings of this study are available from the Jiangsu Provincial Center for Disease Control and Prevention, but restrictions apply to the availability of these data, which were used under license for the current study and so are not publicly available. However, the data are available from the authors upon reasonable request and with permission from the Jiangsu Provincial Center for Disease Control and Prevention.

The supplementary materials will be available in www.besjournal.com.

⁸These authors contributed equally to this work.

⁹Correspondence should be addressed to Chengxiu Ling, E-mail: Chengxiu.Ling@xjtu.edu.cn; Qiao Liu, E-mail: liuqiaonjmu@163.com

Biographical notes of the first authors: Yifan Tang, PhD, Candidate, majoring in spatiotemporal modeling of infectious diseases, E-mail: Yifan.Tang21@student.xjtu.edu.cn; Tenglong Li, Assistant Professor, majoring in statistical computation in infectious disease, E-mail: Tenglong.Li@xjtu.edu.cn

Received: September 15, 2025;

Accepted: November 27, 2025

REFERENCES

1. Bagcchi S. WHO's global tuberculosis report 2022. *Lancet Microbe*, 2023; 4, e20.
2. World Health Organization. Global tuberculosis report 2024. <https://www.who.int/teams/global-programme-on-tuberculosis-and-lung-health/tb-reports/global-tuberculosis-report-2024/>. [2026-01-04]
3. Liu K, Chen SH, Zhang Y, et al. Tuberculosis burden caused by migrant population in Eastern China: evidence from notification records in Zhejiang Province during 2013–2017. *BMC Infect Dis*, 2022; 22, 109.
4. Liu FQ, Zhang ZX, Chen HY, et al. Associations of ambient air pollutants with regional pulmonary tuberculosis incidence in the central Chinese province of Hubei: a Bayesian spatial-temporal analysis. *Environ Health*, 2020; 19, 51.
5. Li JX, Luan QY, Li BB, et al. Outdoor environmental exposome and the burden of tuberculosis: findings from nearly two million adults in northwestern China. *J Hazard Mater*, 2023; 459, 132222.
6. Wei X, Fu T, Chen D, et al. Spatial-temporal patterns and influencing factors for pulmonary tuberculosis transmission in China: an analysis based on 15 years of surveillance data. *Environ Sci Pollut Res*, 2023; 30, 96647–59.
7. Liang SY, Li ZF, Zhang N, et al. Epidemiological and spatiotemporal analysis of severe fever with thrombocytopenia syndrome in Eastern China, 2011–2021. *BMC Public Health*, 2023; 23, 508.
8. Preston L, Thompson A, Baxter S, et al. A rapid review of the causes of diagnostic and treatment delays for tuberculosis in low-burden countries. *J Public Health*, 2025; 47, e530–9.
9. Elshorbagy Y, Ziemke JR, Strode S, et al. Tropospheric ozone precursors: global and regional distributions, trends, and variability. *Atmos Chem Phys*, 2024; 24, 12225–57.
10. Pang Z, Jiang XN, Liu K, et al. Beyond additive effects: examining the combined impact of air pollutant interactions on pulmonary tuberculosis in China. *BMC Public Health*, 2025; 25, 3102.