Journal of Clinical and Nursing Research, 2026, Volume 10, Issue 8
https://www.bbwpublisher.com/journal/JCNR
Online ISSN: 2208-3693
Print ISSN: 2208-3685

Multi-Factor-Driven Inter-City Forecasting of Viral and Bacterial Epidemics: A Comparative Modeling Study

Yunhui Xiang[1]†, Guokang Sun[2]†, Lvbo Tian[1], Jiangtao Hu[1], Jianfeng Li[3]*, Qin Zhang[1], Junxian Wang[1], Pinpin Xiang[4], Ping Chen[1], Chunbao Xie[5]*

[1]Department of Laboratory Medicine and Port Epidemic Disease Monitoring Key Laboratory of Sichuan Province, Sichuan International Travel Healthcare Center (Port Clinic of Chengdu Customs), Chengdu 610042, Sichuan, China
[2] Department of Laboratory Medicine, West China School of Public Health and West China Fourth Hospital of Sichuan University, Chengdu 610041, Sichuan, China
[3]Faculty of Resources and Environment, Chengdu University of Information Technology, Chengdu 610225, Sichuan, China
[4]Department of Laboratory Medicine, Xiping Community Healthcare Center of Longquanyi District, Chengdu 610107, Sichuan, China
[5]Department of Laboratory Medicine and Genetic Diseases Key Laboratory of Sichuan Province, Sichuan Provincial People's Hospital & University of Electronic Science and Technology of China, Chengdu 610072, Sichuan, China
† These authors contributed equally to this work and share first authorship.

***Corresponding authors*: Jianfeng Li, lijianfeng_ljf@126.com; Chunbao Xie, xiechunbao2008@163.com

Abstract: *Objective*: To compare the performance of prediction models incorporating multidimensional urban environmental factors in cross-city forecasting of viral and bacterial infectious diseases, and to explore how these multidimensional factors jointly shape epidemic dynamics, thereby providing a reference for optimizing urban epidemic forecasting strategies. *Methods*: Three representative cities in China were selected to characterize the correlation networks between multi-factor urban characteristics, including climate, air pollutants, economic activities, and control measures, and the incidence of viral infections (represented by hand, foot, and mouth disease and influenza) as well as bacterial infections (represented by pertussis and scarlatina). Using the seasonal auto-regressive integrated moving average (SARIMA) model as a benchmark, we systematically compared a long short-term memory (LSTM) model driven by these multidimensional factors with a SARIMAX model incorporating the same covariates. Through multi-disease forecasting experiments, the models were comprehensively evaluated across multiple dimensions, including predictive accuracy, robustness, and peak event forecasting capability. *Results*: The associations between multiple urban factors and infectious diseases exhibited marked spatial heterogeneity, and close statistical correlations were observed between climate variables and air pollutants, suggesting potential interactions among environmental covariates. Overall, both the multi-factor-driven LSTM and SARIMAX models significantly outperformed the traditional SARIMA model, with each demonstrating distinct advantages in predictive accuracy and peak event forecasting. For viral infections characterized by high incidence and rapid spread, the LSTM model showed the best performance, achieving the highest

correlation between predicted and observed values, the smallest error metrics, and greater consistency across cross-validation folds. For bacterial infections with relatively lower incidence, the SARIMAX model performed better. *Conclusions*: Multidimensional urban factors may exert synergistic effects on infectious disease transmission. Multi-factor-driven forecasting strategies considerably outperformed traditional time-series models, and the respective strengths of specific models were closely associated with the epidemiological characteristics of the diseases. These findings may offer practical insights into methodological selection for constructing multi-category epidemic forecasting models adapted to heterogeneous urban environments.

Keywords: Infectious disease; Prediction model; Long short-term memory network; SARIMA

Online publication: Aug 31, 2026

1. Introduction

Previous studies have shown that community transmission of major epidemics such as SARS and COVID-19 often precedes pathogen confirmation by several weeks, resulting in delays in isolation, contact tracing, and other control measures, and thereby underscoring the urgent need for sensitive and efficient surveillance and early-warning systems [1]. In recent years, prediction techniques based on multi-source data and artificial intelligence have demonstrated considerable potential [2]. For instance, LSTM models are well suited to capturing temporal dependencies, while approaches such as XGBoost have been applied to specific risk prediction tasks; nevertheless, these methods face challenges related to interpretability and data quality [3–5]. Traditional statistical models, such as ARIMA, offer relatively good interpretability; SARIMA can account for seasonal patterns, and ARIMAX allows the incorporation of exogenous covariates. However, their capacity to capture complex nonlinear relationships is generally more limited than that of machine learning models [6–9].

Notably, climate change and air pollution may act synergistically to increase population susceptibility to respiratory infectious diseases, which calls for the integration of environmental health risk assessment into public health systems [10]. However, most existing prediction models focus on a single city or region, and frameworks capable of simultaneously assessing epidemic risks across multiple cities remain scarce, which constrains the development of coordinated regional prevention and control strategies. To address this gap, this study developed a cross-city infectious disease prediction framework that integrates multi-dimensional factors, including economic indicators, passenger traffic, climate conditions, air pollution, and epidemic control measures. Using the SARIMA model as a benchmark, we systematically compared the predictive performance of an LSTM model enhanced with feature engineering and multi-dimensional city-level factors against a SARIMAX model incorporating the same covariates, across four representative viral and bacterial infectious diseases, hand, foot, and mouth disease (HFMD), influenza, pertussis, and scarlatina. This work aims to provide public health authorities with technical support and evidence for the simultaneous assessment of epidemic risks across multiple regions, as well as for optimizing resource allocation and coordinating prevention and control efforts.

2. Materials and methods

2.1. Data sources and curation procedures

This study used a cross-city dataset covering March 2019 to October 2025 across Chongqing (CQ), Shanghai (SH), and Hong Kong (HK). The dataset integrates indicators from multiple domains, including demography,

economy, infectious diseases, climate, and air pollutants. Economic and mobility data were obtained from publicly available statistics released by the municipal statistics bureaus and transportation authorities. Air pollutant data were collected from https://aqicn.org, covering the concentrations of $PM_{2.5}$, PM_{10}, O_3, NO_2, SO_2, and CO, which were averaged into monthly concentrations.

Climate data were retrieved from https://rp5.ru/, including temperature at 2 meters above ground (T), station-level atmospheric pressure (Po), relative humidity at 2 meters (U), mean wind speed at 10–12 meters (Ff), horizontal visibility (Vv), dew point temperature at 2 meters (Td), and precipitation (RRR). These variables were likewise aggregated into monthly values: precipitation was summed to obtain the monthly total (MRRR), while monthly means were computed for the remaining variables. To remove the influence of population size, monthly passenger traffic volumes and key economic indicators were normalized on a per capita basis.

Infectious disease data were obtained from the official websites of the municipal health authorities, comprising monthly reported cases of four notifiable infectious diseases: HFMD, scarlatina, pertussis, and influenza. All cases were diagnosed according to the unified criteria issued by the National Health Commission. It should be noted that the notifiable disease reporting system in HK does not include confirmed cases of HFMD or influenza. To quantify the intensity of epidemic control measures, we constructed a containment intensity index, in which a value of 0 was assigned to months of normal activity; for months with large-scale lockdowns, the value equaled the number of lockdown days in that month; months without large-scale lockdowns but under strict public health requirements were assigned a value of 15; for other situations, the value was adjusted according to the duration of partial lockdowns and the stringency of the measures.

2.2. Spearman correlation networks linking infectious diseases and urban factors

The Spearman's rank correlation matrix was computed between the incidence of infectious diseases and each environmental and socioeconomic variable. Significant associations with $|\rho| \geq 0.3$ and $p < 0.05$ were retained as network edges. Nodes were classified into two categories, diseases (HFMD, influenza, scarlatina, and pertussis) and factors (including climate, air pollution, and socioeconomic activity variables), represented by blue rectangles and yellow circles, respectively. In the network, disease–factor associations are depicted as solid lines, whereas factor–factor correlations are shown as dashed lines. Edges are color-coded according to the direction of the association: red for positive and green for negative correlations, with line width proportional to the absolute value of the correlation coefficient ρ.

2.3. Model setup: Benchmark and candidates

To systematically evaluate the predictive utility of multi-dimensional factors for the monthly incidence of specific infectious diseases, we developed a comparative time series forecasting framework. The framework integrates data preprocessing, feature engineering, modeling, and validation, all implemented in R (version 4.2.2), with the deep learning components built using TensorFlow (version 2.15.0) and Keras (version 2.15.0). To ensure strict temporal generalization and prevent data leakage, we adopted a rolling-block cross-validation scheme. For each city, the time series were arranged in chronological order and partitioned using a fixed training length (48 months), validation period (12 months), and step size (1 month). This scheme ensures that the training set contains only historical observations, whereas the validation set strictly corresponds to future data points. The model look-back windows were determined according to the sample size to maintain robustness along the temporal dimension.

All factors were robustly scaled based on the median and interquartile range prior to training to mitigate the influence of outliers. Key time-series features were then engineered from the case count series: the average rate of change over a three-month window was computed as Slope_3m to capture medium-term trends; the maximum value within a six-month sliding window was extracted as Max_6m to identify local outbreaks and extreme events; and a binary indicator, Is_peak, was created to flag periods in which the incidence exceeded the 95th percentile of the training set.

To construct a robust and low-redundancy set of predictive inputs, we employed a two-step factor and engineered feature selection procedure. First, mutual information (MI) scores between each candidate factor or engineered feature and the target incidence were computed to rank them by their predictive relevance. A sequential selection algorithm was then applied, in which a factor or engineered feature was included only if the ρ between the factor and those already selected was below a threshold of 0.5, thereby effectively reducing multicollinearity.

A SARIMA model was constructed as the benchmark, with its optimal parameters automatically identified using the *auto.arima* function in the R package forecast (v8.23.0). Next, an LSTM network taking the selected inputs was built to capture complex nonlinear temporal dependencies. The network architecture consists of an optimized LSTM layer, followed by fully connected layers and an output layer with a Softsign activation function. To emphasize the accurate prediction of epidemic peaks, a customized peak-weighted loss function was employed.

$$Peak\_weighted\_loss=\frac{1}{n}\sum_{i=1}^{n}(1+a\times\Pi(y_{ture,i}>b\times\bar{y}_{ture}))\times(y_{true,i}-y_{pred,i})^2$$

Here, n denotes the batch size; $y_{ture,i}$ and $y_{pred,i}$ represent the observed and predicted values of the i-th sample, respectively; and $\bar{y}_{ture}$ denotes the mean of the observed values. The indicator function $\Pi(y_{ture,i}>b\times\bar{y}_{ture})$ equals 1 when an observed value exceeds a dynamic threshold defined as b times the mean of the observed values, in which case the loss weight is amplified by a factor of $(1+a)$.

The hyperparameters of the LSTM model were tuned via Bayesian optimization, and a customized callback was used to implement a curriculum learning strategy that progressively increased the temporal complexity of training, thereby enhancing model generalization. Training was then performed using the Adam optimizer combined with gradient clipping to ensure stable convergence. Finally, the SARIMA model was extended to a SARIMAX model. To ensure comparability with the LSTM model, exactly the same inputs were incorporated, allowing the contribution of the city-level factors and engineered features to be evaluated directly within a traditional statistical framework.

2.4. Model evaluation

All models were trained and validated on the same data splits generated by the rolling cross-validation scheme. Predicted and observed values were back-transformed to the original scale, and predictive performance was comprehensively assessed using a set of metrics: the strength of association was measured by ρ; error magnitude was quantified by the root mean square error (RMSE), mean absolute error (MAE), and symmetric mean absolute percentage error (SMAPE); and relative accuracy was evaluated using the mean absolute scaled error (MASE), benchmarked against seasonal naïve forecasts. In addition, the Diebold–Mariano test was applied to determine whether differences in predictive accuracy between models were statistically significant, with

$p < 0.05$ considered significant. Given that public health decision-making places particular emphasis on identifying epidemic peaks, we further assessed model performance on this critical event: the area under the receiver operating characteristic curve (AUC-ROC) was used, complemented by the precision–recall curve, which is better suited to imbalanced data. Furthermore, using the 85th percentile of case counts as the threshold, the F1 score, harmonizing precision and recall, was computed as a threshold-sensitive summary metric.

3. Results

Analyses of the disease–factor correlation networks for each city revealed marked spatial heterogeneity in the association patterns between diseases and factors **(Figure 1)**. Nevertheless, several associations were relatively consistent across cities: the containment intensity index was negatively correlated with the incidence of all four infectious diseases in every city, whereas the number of international visitors was positively correlated with disease incidence in most contexts (except for HFMD in CQ, which was significantly associated only with MRRR). In SH, the associations between HFMD and climatic indicators, including temperature and humidity, were also noticeably stronger than those observed for other diseases. In addition, the network analysis further revealed close statistical associations among the multi-dimensional factors, particularly between climatic variables and air pollutants, suggesting potential collinearity or synergistic effects among environmental covariates.

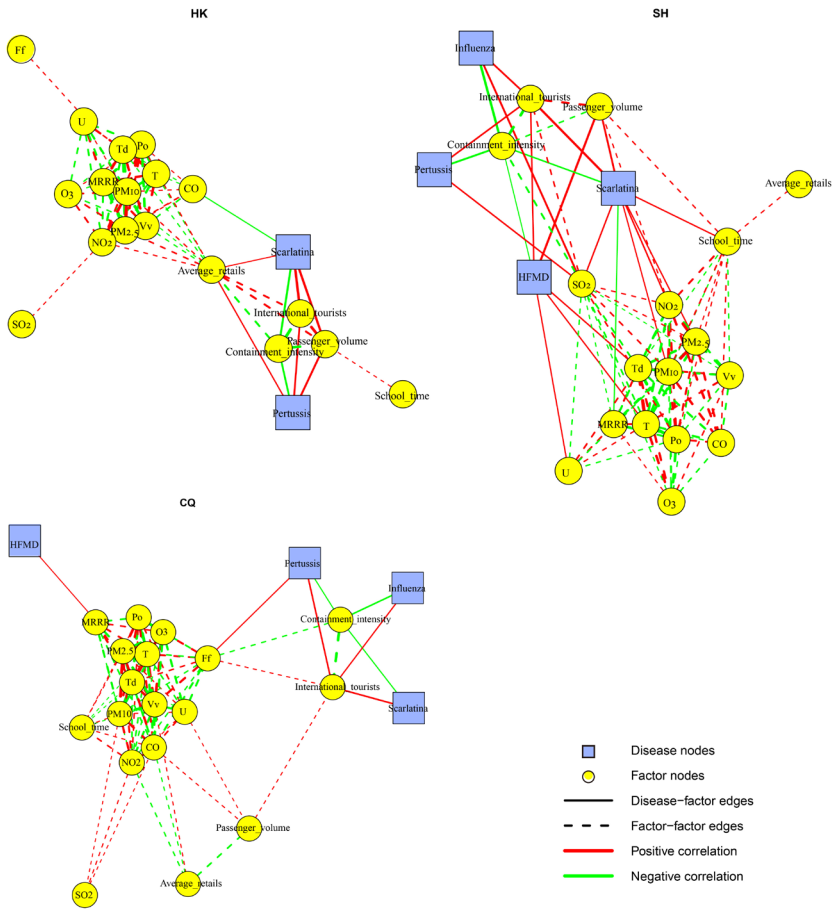

Figure 1. Network of correlations between infectious diseases and multidimensional factors across cities.

After constructing the rolling cross-validation function, completing factor standardization, and generating the engineered features, we screened the inputs based on MI and Spearman correlation analyses. The MI score rankings of all factors with respect to the four diseases are presented separately in **Figure 2**. Apart from certain hyperparameters tuned via Bayesian optimization, the optimal settings of the key parameters, including the look-back period (Lookback = 3 months), the number of inputs, the sample weight configuration, and the loss function parameters, were determined through systematic repeated experiments, with detailed results summarized in **Table 1**.

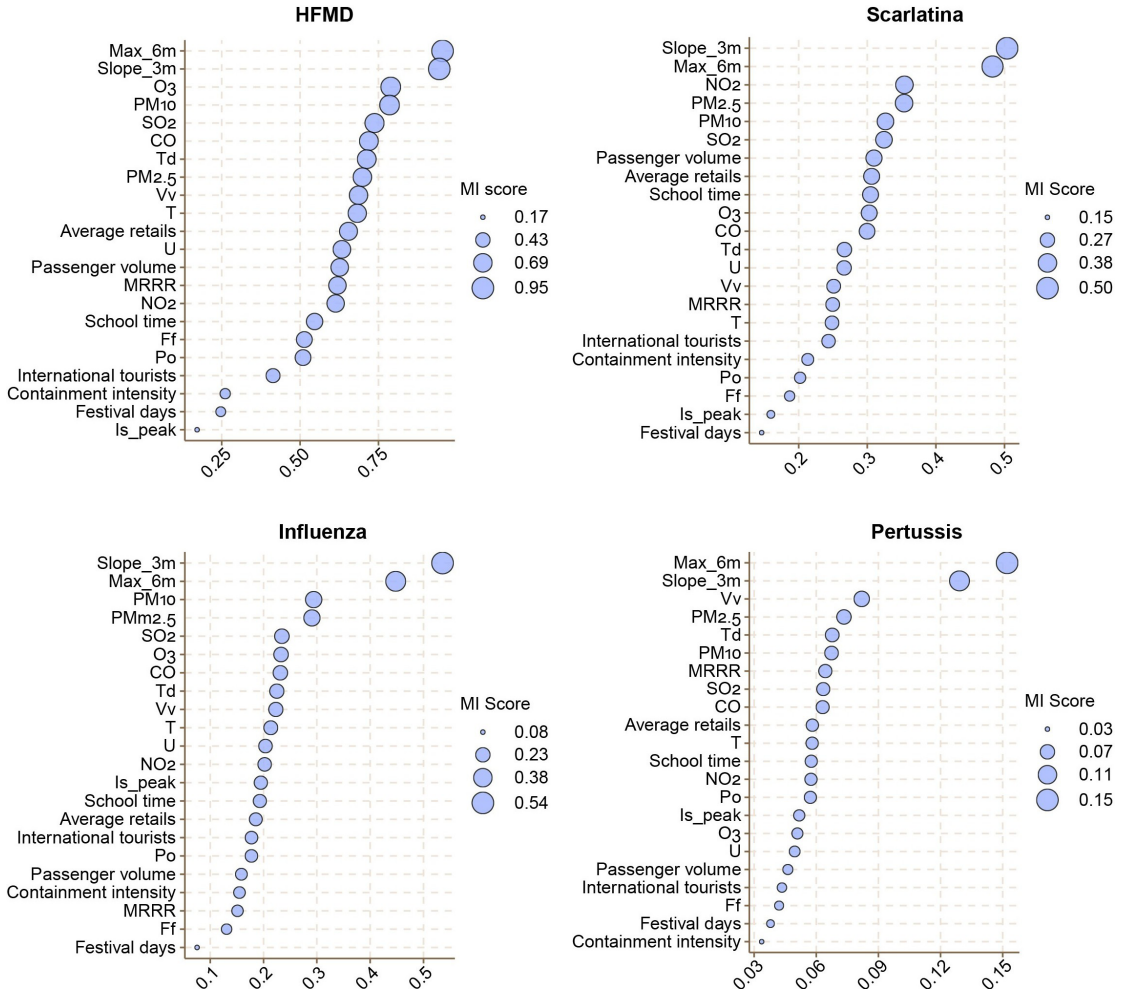

Figure 2. Mutual information scores between multi-dimensional factors and the four infectious diseases.

Table 1. LSTM model inputs and characteristics for each disease

Diseases	Data	Number of inputs	Sample weight	Parameters of Loss function
HFMD	Cities: 2 Folds: 21	4 (Max_6m; Slope_3m; O_3; SO_2)	Is_peak = 1–6, Containment intensity < 10–1.2	a = 2 b = 4
Influenza	Cities: 2 Folds: 21	7(Slope_3m; Max_6m; PM_{10} SO_2; U; Is_peak; School time)	Is_peak = 1–2, Containment intensity < 10–1.4	a = 2 b = 5
Scarlatina	Cities: 3 Folds: 21	13 (Max_6m; Slope_3m; NO_2; SO_2; Passenger volume; Average retails; School time; O_3; U; MRRR; Containment intensity Ff; Is_peak)	Is_peak = 1–2, Containment intensity < 10–1.3	a = 2 b = 4
Pertussis	Cities: 3 Folds: 21	9 (Max_6m; Slope_3m; Vv; MRRR; SO_2, Average retails School time; Is_peak; O_3)	Is_peak = 1–2 Containment intensity < 10–1.2	a = 3 b = 5

After completing the final model training, we aggregated the predictions of each model across the cross-validation folds and visualized them using box plots overlaid with lines of the observed values, while model performance was further assessed using Spearman's rank correlation analysis. In the prediction of HFMD, the LSTM model achieved ρ values of 0.56 and 0.92 in CQ and SH, respectively. In contrast, the SARIMA model showed the weakest performance in both cities, with ρ values of only 0.45 and 0.41, and a limited ability to capture incidence peaks. When the same inputs used by the LSTM model were incorporated as covariates into the SARIMAX model, its ρ values increased to 0.65 and 0.62 in CQ and SH, respectively (**Figure 3**).

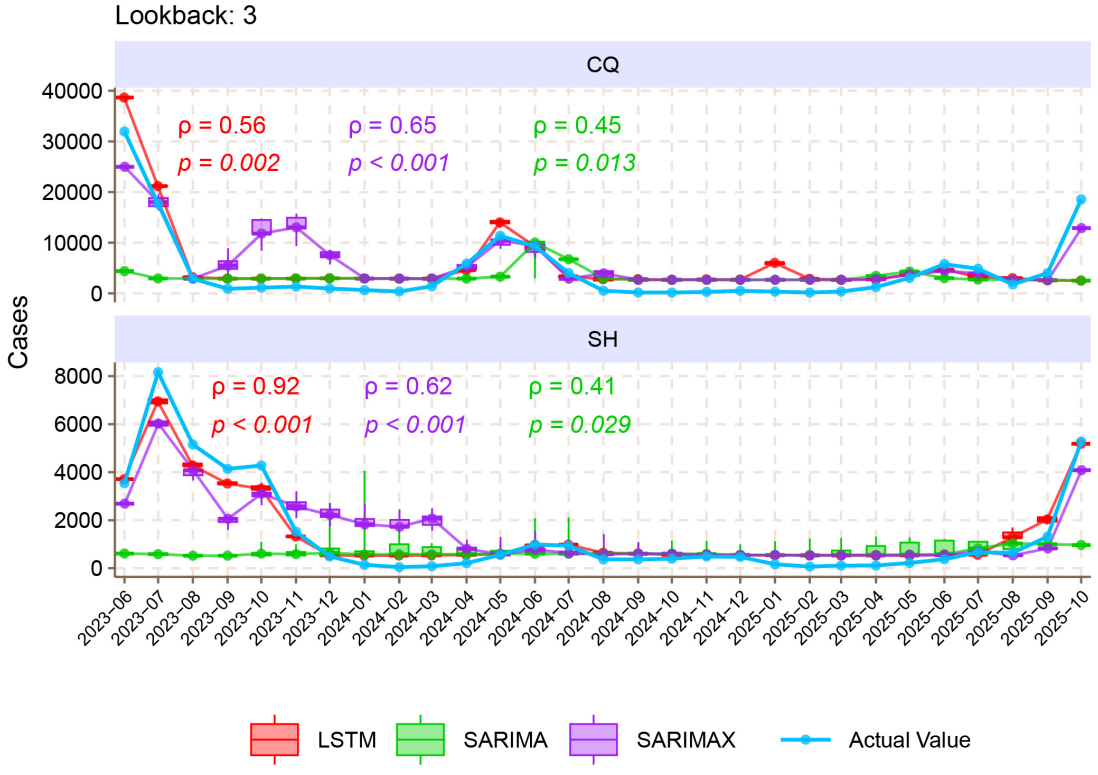

Figure 3. Reported versus predicted HFMD incidence for the three forecasting models.

In influenza prediction, the LSTM model again exhibited the highest correlations, with ρ values of 0.84 and 0.73 in CQ and SH, respectively. After incorporating the same covariates, the SARIMAX model yielded ρ values of 0.29 and 0.49 in CQ and SH, respectively. Similarly, the SARIMA model remained the weakest performer, with ρ values of −0.09 and −0.31, neither of which reached statistical significance (**Figure 4**).

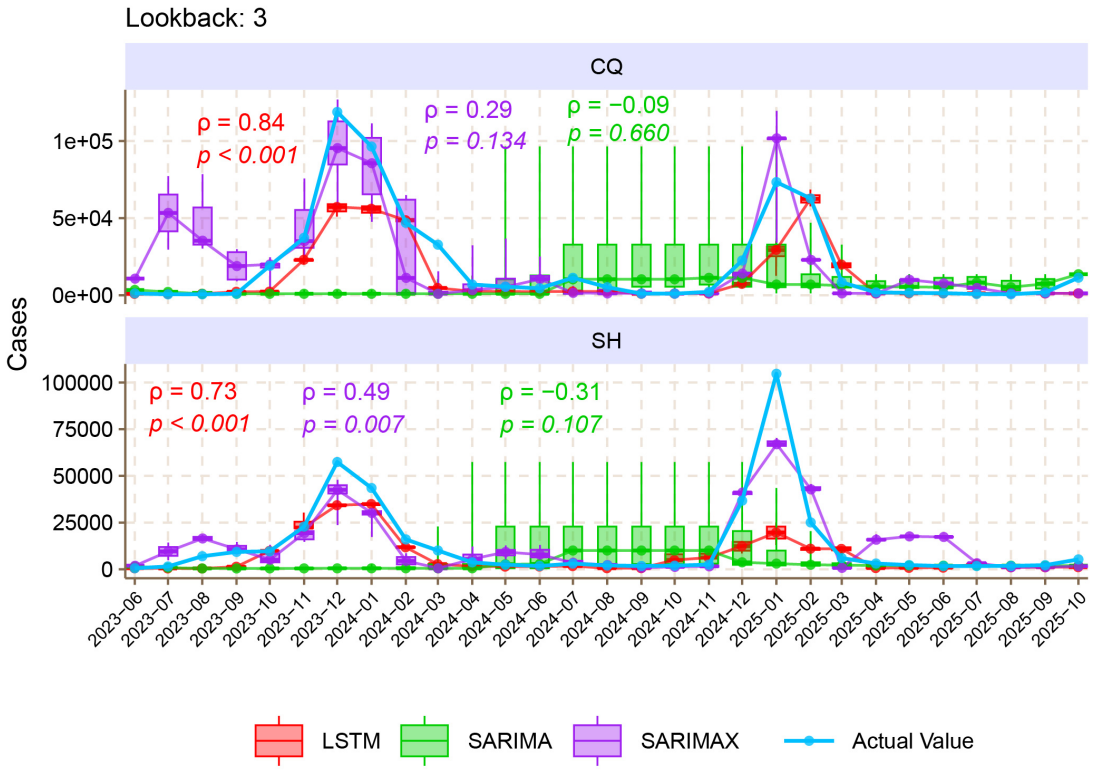

Figure 4. Reported versus predicted influenza incidence for the three forecasting models.

In scarlatina prediction, the LSTM model achieved ρ values of 0.82, 0.84, and 0.60 in CQ, HK, and SH, respectively. When the same inputs were used, the SARIMAX model yielded ρ values of 0.82, 0.93, and 0.72 in these cities, while the SARIMA model continued to rank last, with ρ values of 0.50, 0.45, and 0.34 (**Figure 5**). For pertussis prediction, the LSTM model obtained ρ values of 0.71, −0.03, and 0.82 in CQ, HK, and SH, respectively; the SARIMAX model yielded ρ values of 0.51, 0.64, and 0.70 in the corresponding cities; and the SARIMA model again performed worst, with ρ values of −0.34, −0.26, and −0.06, none of which were statistically significant (**Figure 6**).

Overall, for HFMD and influenza, diseases with relatively high incidence, the correlations between the LSTM predictions and observed values were higher than those of the SARIMAX model, whereas for scarlatina, a disease with lower incidence, the SARIMAX model performed better. In pertussis prediction, the SARIMAX model achieved its best correlation in HK, the city with the lowest incidence. These results suggest that the SARIMAX model may be better suited to forecasting infectious diseases with relatively low incidence, whereas the LSTM model may offer stronger predictive capability for diseases capable of widespread transmission over short periods. In addition, at the same time points across different cross-validation folds, the LSTM model exhibited greater reproducibility in its predictions than the SARIMAX

model, with the SARIMA model performing worst, particularly in pertussis prediction, reflecting the more robust predictive behavior of the LSTM model.

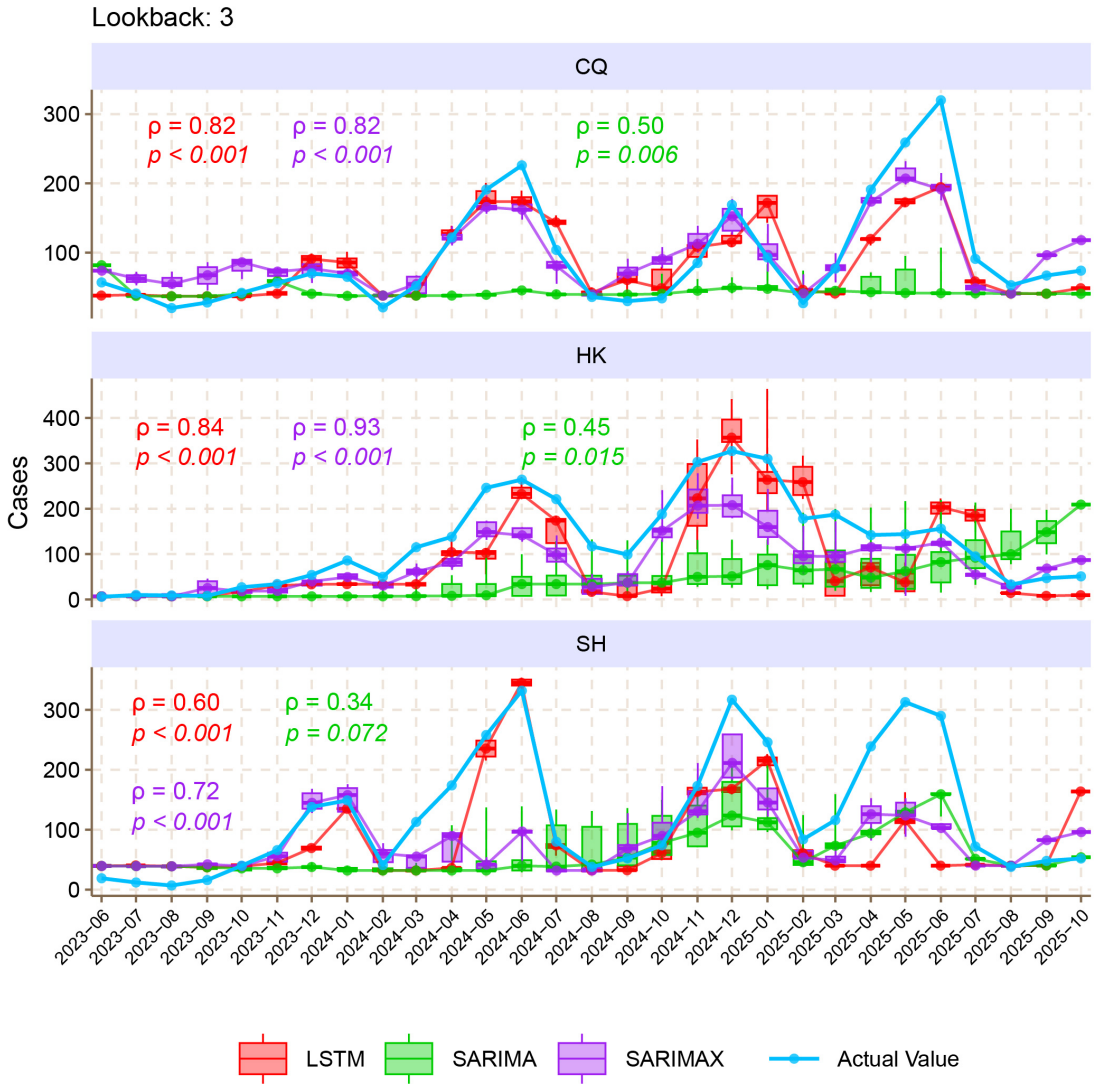

Figure 5. Reported versus predicted scarlatina incidence for the three forecasting models.

Table 2 further presents the error magnitude and relative accuracy of the three models in terms of RMSE, MAE, SMAPE, and MASE: the LSTM model achieved the smallest errors in the prediction of HFMD, influenza, and pertussis, whereas the SARIMAX model showed the smallest errors for scarlatina.

Table 2. Performance metrics for model prediction values

Metrics	HFMD models			Influenza models			Scarlatina models			Pertussis models		
	LSTM	SARIMA	SARIMAX	LSTM	SARIMA	SARIMAX	LSTM	SARIMA	SARIMAX	LSTM	SARIMA	SARIMAX
RMSE	2761	5165	2939	17670	30620	15326	69.3	111	68.3	1031	2384	1082
MAE	1477	2603	1789	8358	16351	9835	47.1	77.9	46.9	389	1310	465

SMAPE	70.9	96.8	79.2	69.6	118.6	84.7	52.7	80.1	47.4	98	144	87.5
MASE	1.40	2.48	1.70	1.47	2.88	1.73	2.45	4.04	2.43	4.84	16.3	5.79
SD	4007	4007	4007	27063	27063	27063	97.3	97.3	97.3	1071	1071	1071

MAE, mean absolute error; MASE, mean absolute scaled error; RMSE, root mean square error; SMAPE, symmetric mean absolute percentage error; SD, standard deviation for disease cases.

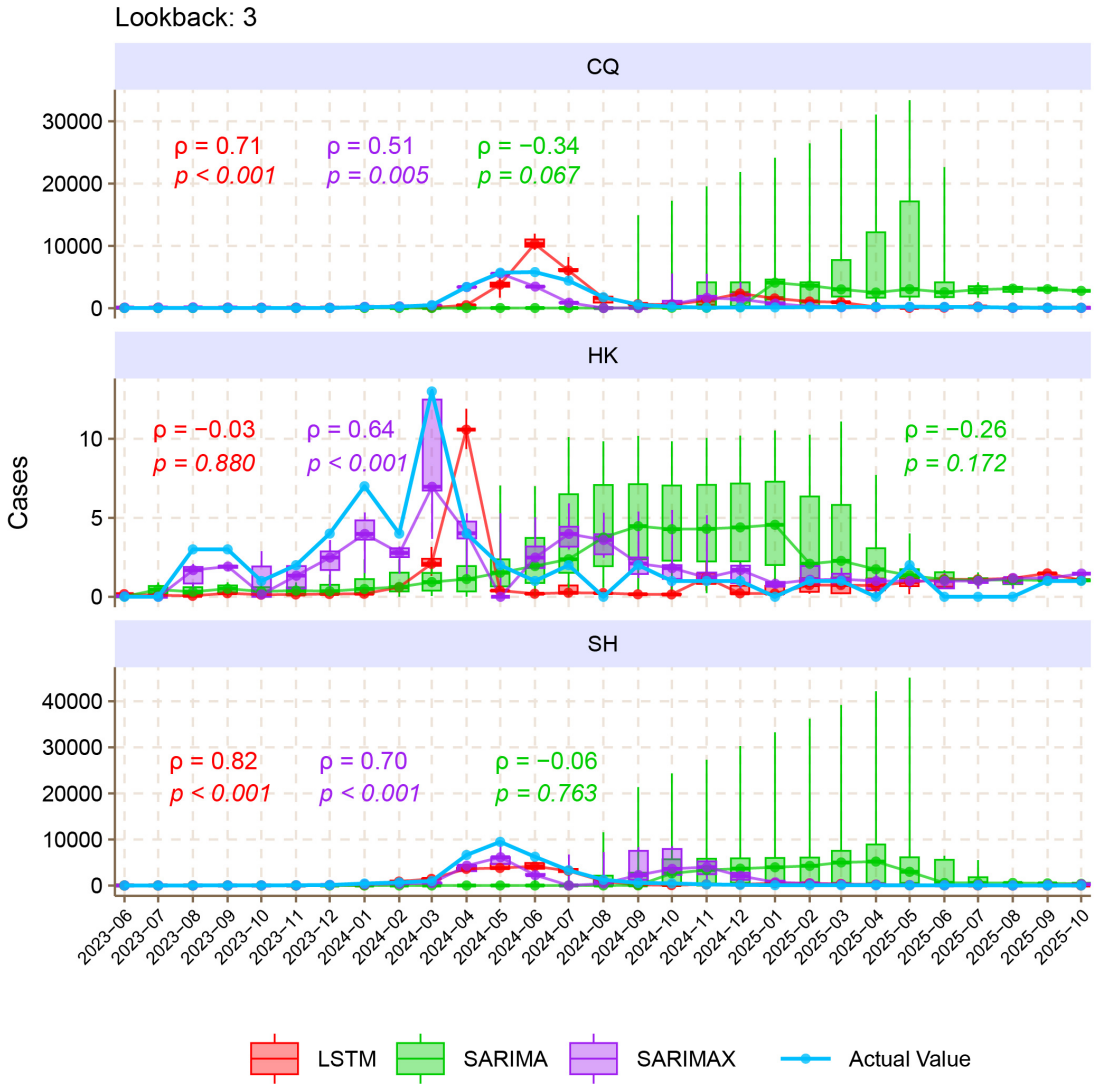

Figure 6. Reported versus predicted pertussis incidence for the three forecasting models.

Subsequently, Diebold–Mariano tests on the forecasting results of the three models for the four diseases indicated that both the LSTM and SARIMAX models were significantly more accurate than the SARIMA model. The LSTM and SARIMAX models, however, performed comparably overall, with the LSTM model showing an advantage in HFMD prediction in SH and the SARIMAX model performing better in pertussis prediction in CQ. Regarding the prediction of peak events, ROC curve analyses revealed that the LSTM and SARIMAX models each excelled in different diseases: the LSTM model performed best in predicting HFMD and pertussis, with AUC values of 0.863 and 0.872, respectively, whereas the SARIMAX model showed

better performance for influenza and scarlatina, with AUC values of 0.859 and 0.930, respectively. Across all diseases, the SARIMA model yielded the lowest AUC values (**Figure 7**).

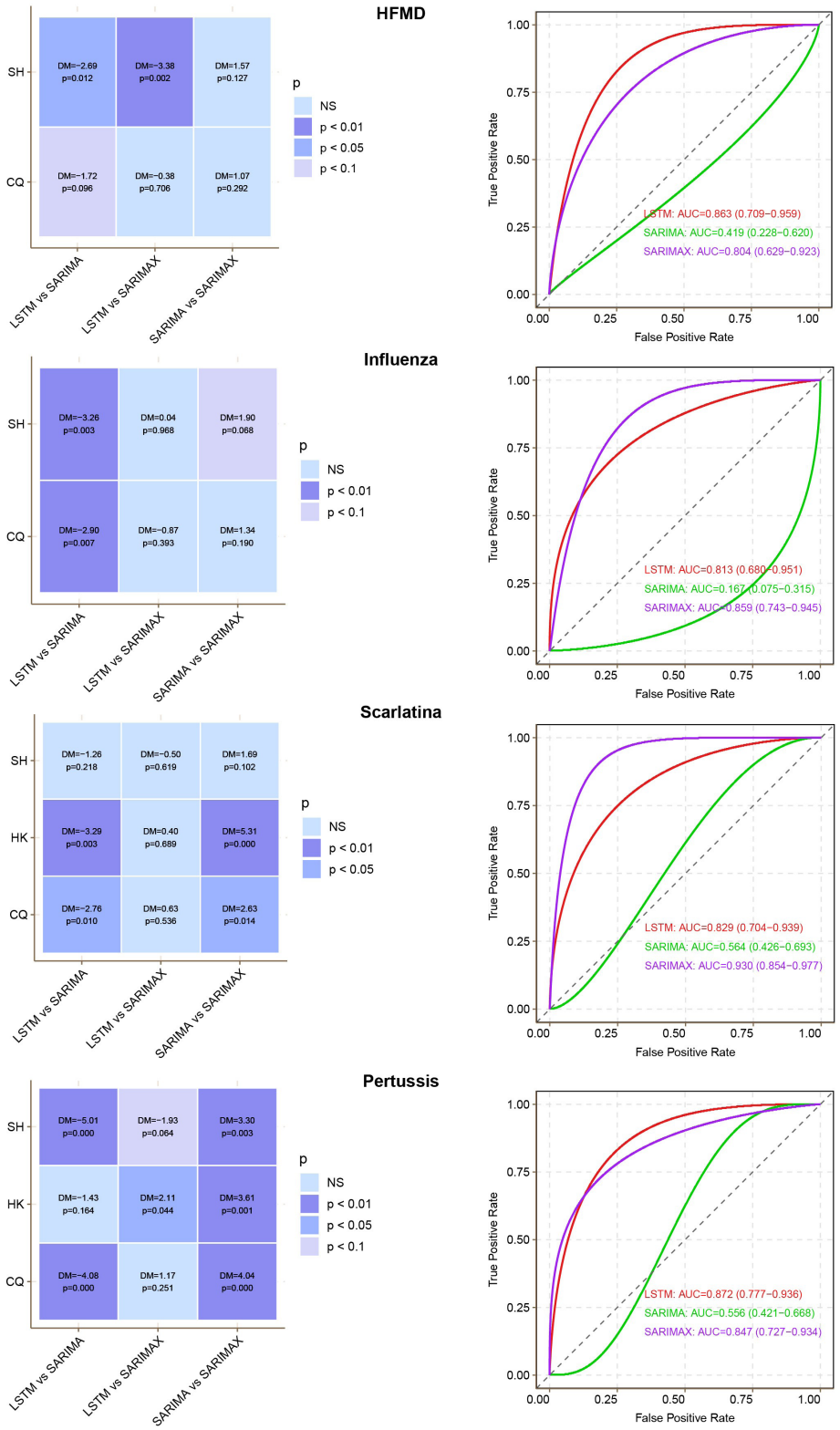

Figure 7. Diebold-Mariano tests for models and ROC curves for evaluating peak prediction performance.

This study further adopted the precision–recall (PR) curve and the F1 score as complementary metrics for evaluating peak prediction accuracy in imbalanced scenarios (**Table 3**). Overall, the LSTM and SARIMAX models performed similarly in terms of both the F1 score and AUPRC: the LSTM model showed a slight advantage in predicting HFMD and influenza, whereas the SARIMAX model performed marginally better for scarlatina and pertussis, which is consistent with the findings of the aforementioned correlation analyses.

Table 3. Precision, recall, and F1-score evaluation of the models

Metrics	HFMD models			Influenza models			Scarlatina models			Pertussis models		
	LSTM	SAR IMA	SARI MAX	LSTM	SAR IMA	SARI MAX	LSTM	SAR IMA	SARI MAX	**LSTM**	**SAR IMA**	**SARI MAX**
Precision	0.50	0.00	0.36	0.77	0.00	0.55	0.53	0.00	0.89	0.56	0.13	0.55
Recall	0.33	0.00	0.44	0.67	0.00	0.73	0.63	0.00	0.50	0.71	0.14	0.76
F1-score	0.40	0.00	0.40	0.71	0.00	0.63	0.57	0.00	0.64	0.63	0.13	0.64
AUPRC	0.59	0.13	0.53	0.63	0.15	0.60	0.57	0.19	069	0.73	0.23	0.77

AUPRC, area under the precision-recall curve.

4. Discussion

In recent years, the epidemiological characteristics of infectious diseases have become increasingly complex, as reflected in the dynamic evolution of pathogen spectra and the nonlinear effects of environmental factors, such as climate and air pollution, on disease risk [11,12]. Influenza surveillance data have revealed continuous antigenic variation in the virus, with different strains alternately dominating summer and winter epidemics [13]. Meanwhile, non-pharmaceutical interventions have substantially altered disease transmission patterns, and ongoing urbanization may also attenuate the effectiveness of traditional measures such as school closures; together, these factors have added to the complexity of infectious disease forecasting [14,15]. Using Spearman correlation networks between infectious diseases and multi-dimensional factors across cities, we identified marked spatial heterogeneity in the association patterns between diseases and factors, as well as potential collinearity or synergistic effects among environmental covariates. These findings suggest that effective disease control still requires interventions tailored to local circumstances, with full consideration of differences in climate, air pollutants, socioeconomic conditions, and public health policies, rather than a "one-size-fits-all" strategy [16].

The cross-city LSTM and SARIMAX models built on multi-dimensional factors performed well overall in predicting the four infectious diseases studied, HFMD, influenza, scarlatina, and pertussis, with ρ and RMSE comparable to or better than those reported in previous studies[17,18]. In contrast, the SARIMA model yielded unsatisfactory predictions for three of the four diseases, which may be attributable to the pronounced heterogeneity in epidemiological contexts across cities, as well as the substantial disruption of natural disease transmission patterns by containment policies during 2020–2022 [19].

Furthermore, by encompassing diseases caused by different types of pathogens, both viral and bacterial, this study has enhanced our understanding of the respective application scenarios of deep learning models and traditional time series models. The marked differences in the performance of the LSTM and SARIMAX models across diseases indicate that urban heterogeneity in epidemiological data constrains the general

applicability of forecasting models. Specifically, the LSTM model performed slightly less well than the SARIMAX model in predicting scarlatina and pertussis, but outperformed the latter for HFMD and influenza, which have higher incidence. This contrast highlights the potential advantage of LSTM models in forecasting diseases with high epidemic intensity, such as "Disease X" [20]. Nevertheless, the LSTM model employed a fixed look-back window of three historical time steps. Although this design takes into account the delays commonly encountered in actual epidemic reporting and enhances the feasibility of deployment during evolving epidemics, it remains more dependent on training sample size than SARIMAX, which requires no look-back window. This suggests that in the early stages of emerging infectious disease outbreaks, when only limited data are available, SARIMAX may adapt more rapidly.

This study has several limitations. (1) Data availability and methodological discrepancies: the number of cities for which multi-dimensional public data can currently be obtained for model training is quite limited, which may affect generalizability. Data on vaccine effectiveness and coverage are often difficult to obtain, underscoring the need for cross-sector data sharing. In addition, methodological differences in statistical practices across cities and the approach used to estimate containment intensity may introduce systematic bias when integrating data across regions. (2) Benchmarking gaps: due to computational constraints, this study did not include comparative evaluations against zero-shot forecasting foundation models such as TimeGPT, which have demonstrated strong generalization capabilities in forecasting biomedical or infectious disease time series driven by exogenous variables. (3) Region-specific threshold effects: the threshold effects of air pollution and climatic variables may vary across regions, potentially weakening model generalizability [21]. (4) Geographic equity: although intelligent technologies and data-driven tools hold considerable promise for infectious disease surveillance and sustainable urban development, their benefits and risk burdens are often unevenly distributed. Without equity-focused design and implementation, such innovations may exacerbate existing socioeconomic disparities rather than foster universal resilience [22].

Based on the above discussion, future research could proceed in the following directions: (1) exploring hybrid architectures that combine SARIMAX and LSTM to complementarily offset their respective limitations; (2) applying transfer learning to improve model performance in data-scarce settings, given that LSTM models typically require substantial training data [23]; (3) introducing causal inference frameworks, such as structural equation modeling, in combination with LSTM-based factor attribution methods, to build forecasting frameworks that are interpretable and translatable into interventions [24]; (4) incorporating attention mechanisms to further improve both predictive accuracy and model transparency [25]; and (5) developing interval forecasting. This study focused on point forecasts, but quantifying uncertainty is essential for public health decision-making, and interval forecasts could provide more reliable support for this purpose [26]. Nevertheless, the complexity of artificial intelligence models dictates their sample size requirements; collecting longer time-series data from more cities remains fundamental to improving predictive accuracy and generalizability.

5. Conclusion

In summary, the multi-dimensional forecasting models generally outperformed traditional time series models, and the strengths of different models were associated with disease type. Therefore, flexibly selecting models according to disease characteristics and epidemic phases may constitute a more effective strategy. The multi-

city, multi-disease forecasting framework developed in this study improves model generalizability and can provide technical support for public health authorities in simultaneously assessing epidemic risks across multiple regions and guiding coordinated regional prevention and control.

Funding

General Administration of Customs of the People's Republic of China Science and Technology Program (Project No.: 2024HK161); Chengdu Science and Technology Program (Project No.: 2025-YF08-00230-SN); Science and Technology Research and Development Program of China Railway Chengdu Group Co., Ltd (Project No.: CX25014); Sichuan Science and Technology Program (Project No.: 2024JDRC0032)

Disclosure statement

The authors declare no conflict of interest.

References

[1] Oeschger T, McCloskey D, Buchmann R, et al., 2021, Early Warning Diagnostics for Emerging Infectious Diseases in Developing Into Late-Stage Pandemics. Accounts of Chemical Research, 54(19): 3656–3666.

[2] Guo K, Shen C, Zhou X, et al., 2024, Traffic Data-Empowered XGBoost-LSTM Framework for Infectious Disease Prediction. IEEE Transactions on Intelligent Transportation Systems, 25(1): 774–785.

[3] Wu H, Cao Y, Wei H, et al., 2024, Prediction of COVID-19 Infections Based on LSTM With Self-Attention Mechanism. Computer Applications and Software, 41(9): 106–113.

[4] Casillas N, Torres A, Moret M, et al., 2022, Mortality Predictors in Patients With COVID-19 Pneumonia: A Machine Learning Approach Using Extreme Gradient Boosting Model. Internal and Emergency Medicine, 17(7): 1929–1939.

[5] Kraemer M, Tsui J, Chang S, et al., 2025, Artificial Intelligence for Modelling Infectious Disease Epidemics. Nature, 638(8051): 623–635.

[6] Lv Z, Sun R, Liu X, et al., 2024, Evaluating the Effectiveness of Self-Attention Mechanism in Tuberculosis Time Series Forecasting. BMC Infectious Diseases, 24(1): 1377.

[7] Prendin F, Díez J, Del Favero S, et al., 2022, Assessment of Seasonal Stochastic Local Models for Glucose Prediction Without Meal Size Information Under Free-Living Conditions. Sensors (Basel, Switzerland), 22(22): 8682.

[8] Rancati S, Bosoni P, Schiaffini R, et al., 2024, Exploration of Foundational Models for Blood Glucose Forecasting in Type-1 Diabetes Pediatric Patients. Diabetology, 5(6): 584–599.

[9] Yang H, Ren Z, Li C, et al., 2025, Prediction and Analysis of Monthly Brucellosis Incidence in Linyi City Based on the SARIMA Model. Chinese Journal of Health Statistics, 42(6): 882–884.

[10] Zhang C, 2024, Modification Effects of Long-Term Air Pollution on the Associations of Short-Term Temperature, Humidity, and Their Combined Exposure With the Incidence of Hand, Foot, and Mouth Disease, thesis, Shandong University.

[11] Yang J, 2024, Association Between Air Pollutants and the Transmissibility of Influenza, thesis, Peking Union Medical College.

[12] Shao T, Li H, Xue N, et al., 2025, Epidemic Trends and Clinical Characteristics of SARS-CoV-2 Omicron Variants. Chinese Journal of Nosocomiology, 2025(3): 337–342.

[13] Chinese Center for Disease Control and Prevention, 2024, Technical Guidelines for Seasonal Influenza Vaccination in China (2023–2024). Chinese Journal of Viral Diseases, 14(1): 1–19.

[14] Zhang X, 2023, Impact of COVID-19 Non-Pharmaceutical Interventions on Influenza Activity and Development of Influenza Forecasting Models, thesis, Peking Union Medical College.

[15] Zhang Y, Zhang J, Zhang X, et al., 2025, How Urbanization Shapes the Effectiveness of School Closures on Hand, Foot, and Mouth Disease. International Journal of Infectious Diseases: IJID: Official Publication of the International Society for Infectious Diseases, 154: 107845.

[16] Xiang Y, Sun G, Xiang P, et al., 2026, Multidimensional Drivers of Urban Infectious Diseases: A Comparative Longitudinal Study Informing Planetary Health Governance. Npj Urban Sustainability.

[17] Chen X, Ba J, Liu Y, et al., 2025, Spatiotemporal Filtering Modeling of Hand, Foot, and Mouth Disease: A Case Study From East China, 2009–2015. Epidemiology and Infection, 153: e61.

[18] Geng X, Shi Y, Ou Y, et al., 2025, Probabilistic Forecasting of Hand, Foot and Mouth Disease in Mainland China Using Bayesian Additive Regression Tree Model. Scientific Reports, 15: 33745.

[19] Li Y, 2024, Causal Effects of Non-Pharmaceutical Interventions During the COVID-19 Pandemic on Tuberculosis and Prediction of Its Epidemic Trends, thesis, Xinxiang Medical University.

[20] Haque S, Mengersen K, Barr I, et al., 2024, Towards Development of Functional Climate-Driven Early Warning Systems for Climate-Sensitive Infectious Diseases: Statistical Models and Recommendations. Environmental Research, 249: 118568.

[21] Zhang G, 2024, Lagged Exposure–Response Relationships of Meteorological, Pollution, and Industrial Factors With Tuberculosis in Northeast China, thesis, China Medical University.

[22] Sengupta U, Sengupta U, 2022, SDG-11 and Smart Cities: Contradictions and Overlaps Between Social and Environmental Justice Research Agendas. Frontiers in Sociology, 7: 995603.

[23] Guo W, Dong Y, Hao G, 2024, Transfer Learning Empowers Accurate Pharmacokinetics Prediction of Small Samples. Drug Discovery Today, 29(4): 103946.

[24] Liu Y, Liu Y, Pan H, et al., 2025, Unveiling New Insights Into Migraine Risk Stratification Using Machine Learning Models of Adjustable Risk Factors. The Journal of Headache and Pain, 26(1): 103.

[25] Xu N, Yang G, Ming L, et al., 2025, PSOA-LSTM: A Hybrid Attention-Based LSTM Model Optimized by Particle Swarm Optimization for Accurate Lung Cancer Incidence Forecasting in China (1990–2021). Frontiers in Medicine, 12: 1620257.

[26] Abbas W, Lee S, Kim S, 2025, Joint Estimation of Hand-Foot-Mouth Disease Model and Prediction in Korea Using the Ensemble Kalman Filter. PLoS Computational Biology, 21(4): e1012996.