RESEARCH ARTICLE

Integrating distributed lag non-linear models and deep learning reveals meteorological drivers of influenza A and B positivity in a 6-year subtropical surveillance study

Long Xie[1†], Meng-Jie Zhang[2†], Jin-Lin Tan[3,4†], Yi-Xin Ling[2], Zhe-Qiang Xue[3,4], Zhi-Shen Wu[3,4], Jun-Ju Huang[3,4], Jian-Ling Chen[2], Ze-Fan Ruan[3,4], Jing Qian[2], Hai-Yong Pan[2], Xiao Han[5,6]*, Sheng Xiong[7]*, Long-Mei Ling[2]*, Xi-Wen Jiang[3,4,8]*

[1]Translational & Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, United Kingdom; [2]Department of Microbiology Laboratory, Putian Centre for Disease Control and Prevention, Putian, China; [3]Research Institute, DAAN Gene Co., Ltd, Guangzhou, China; [4]The Medicine and Biological Engineering Technology Research Centre of the Ministry of Health, Guangzhou, China; [5]College of Biological Science and Engineering, Fuzhou University, Fuzhou, China; [6]Department of Infection, Mengchao Hepatobiliary Hospital of Fujian Medical University, Fuzhou, China; [7]College of Life Science and Technology, Jinan University, Guangzhou, China; [8]School of Life Sciences and Biopharmaceutics, Guangdong Pharmaceutical University, Guangzhou, China

***For correspondence:** hanxiao@fzu.edu.cn (XH); xiongsheng@jnu.edu.cn (SX); ptcdc001@163.com (L-ML); yuanyecat@vip.sina.com (X-WJ)

[†]These authors contributed equally to this work

Competing interest: See page 26

Funding: See page 26

Preprint posted
15 July 2025
Sent for Review
16 June 2025
Reviewed preprint v1
11 September 2025
Reviewed preprint v2
04 September 2026
Reviewed preprint v3
18 September 2026
Version of Record
20 November 2026

Reviewing Editor: George N Okoli, University of Hong Kong, Hong Kong

© Copyright Xie, Zhang, Tan *et al*. This article is distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use and redistribution provided that the original author and source are credited.

eLife Assessment

This revised submission remains a **valuable** contribution to influenza surveillance. After further addressing a previously highlighted methodological issue, the claims of providing an early warning tool aligns more with the reported study results and is a **solid** addition to the evidence base.

Abstract Subtropical influenza forecasting is complicated by non-linear meteorological dynamics and surveillance bias stemming from pandemic-era testing fluctuations. Existing approaches either characterize lagged environmental associations without forecasting capacity or apply deep learning without epidemiological interpretability. We analyzed multisite influenza surveillance data from Putian, southeastern China (2018–2023); test positivity was the primary outcome to mitigate testing-related bias. Distributed lag nonlinear models estimated type-specific associations between eight meteorological variables and positivity over a 15-day lag window. A Bayesian-optimized long short-term memory (LSTM) network incorporating meteorological, autoregressive, and socio-behavioral covariates was benchmarked against covariate-matched multivariate autoregressive integrated moving average (ARIMA) and extreme gradient boosting (XGBoost) models. For influenza A, the temperature-associated relative risk (RR) peaked at 15°C (3.01; 95% confidence interval [CI]: 1.03–8.78), and positivity was associated with narrow diurnal temperature ranges. Influenza B positivity was associated with colder conditions (8°C; RR = 26.12, 95% CI: 7.79–87.61), higher humidity (91%; RR = 1.52, 95% CI: 1.00–2.29), and lower solar radiation. The LSTM had lower mean absolute error for influenza A (0.009 vs 0.136 for ARIMA and 0.138 for XGBoost) and B (0.002 vs 0.049 and 0.070); symmetric mean absolute percentage error was 0.521 vs 1.212 and 1.081 for A and 0.484

vs 0.810 and 0.737 for B. Covariate ablation and Shapley additive explanations analyses suggested that historical positivity and testing-related variables contributed to forecast stability. External validation in Sanming provided preliminary evidence of regional portability. This framework combines interpretable meteorological associations with short-horizon forecasting under variable surveillance conditions.

Introduction

Influenza imposes a substantial global health burden, with an estimated 290,000–650,000 respiratory deaths and 3–5 million severe cases among one billion infections annually (***Iuliano et al., 2018***). In China alone, influenza is associated with an estimated 88,100 excess respiratory deaths per year (***Li et al., 2019***), with children, the elderly, and individuals with chronic comorbidities disproportionately affected. Effective preparedness therefore requires both a mechanistic understanding of the drivers of influenza activity and the operational capacity to forecast its short-term trajectory. Because influenza A and B viruses are transmitted primarily through respiratory droplets and contact (***Uyeki et al., 2022***), with transmission dynamics principally influenced by viral characteristics, host susceptibility, and environmental conditions (***Javanian et al., 2021***), meteorological variables, particularly temperature, humidity, and precipitation, represent a critical, potentially modifiable dimension of transmission (***Moriyama et al., 2020***).

Meteorological drivers of influenza transmission have been extensively characterized in temperate settings, where a pronounced winter peak (***Ryu and Cowling, 2021***) is associated with low ambient temperatures (5–8°C) that stabilize aerosolized influenza virions and impair host mucociliary clearance (***Moriyama et al., 2020***), and with low absolute humidity (<40%) that prolongs airborne aerosol residence while compromising respiratory epithelial defenses (***Peci et al., 2019***; ***Lowen and Steel, 2014***; ***Eccles, 2002***; ***Tamerius et al., 2013***). However, these temperate-zone patterns do not generalize to subtropical regions, where influenza circulates year-round, exhibits multiple annual peaks, and demonstrates meteorological coupling that differs substantially in directionality and lag structures from temperate observations (***Deyle et al., 2016***; ***Soebiyanto et al., 2015***). Despite the demographic and climatic significance of subtropical regions for global influenza dynamics (***Bloom-Feshbach et al., 2013***; ***Tamerius et al., 2013***), dedicated quantitative studies in subtropical Chinese urban areas, particularly in non-megacity settings that nonetheless harbor substantial populations and represent prevailing regional urbanization patterns, remain sparse. This geographic and methodological gap leaves subtropical influenza preparedness without the region-specific evidence required for precision-oriented public health planning.

A key reason subtropical weather–influenza relationships remain difficult to characterize is their non-linear, threshold-dependent, and lagged nature. Conventional linear analyses, such as logistic regression, assume simple contemporaneous relationships and thus fail to capture these dynamics with the precision required for epidemiological inference (***Ng et al., 2022***). Moreover, meteorological effects on influenza transmission also unfold over days to weeks, further complicating association analysis (***Ng et al., 2022***; ***Si et al., 2024***). Distributed lag non-linear models (DLNM) provide a framework for jointly analyzing exposure-response and lag-response relationships (***Gasparrini et al., 2010***; ***Gasparrini, 2011***). DLNM has been applied in empirical analyses of influenza transmission patterns and in lagged risk analyses between influenza and meteorological factors (***Liu et al., 2020***; ***Guo et al., 2019***). However, DLNM is fundamentally an inferential tool designed for retrospective association characterization and does not generate operational forecasts capable of supporting prospective public health decision-making.

On the forecasting side, traditional statistical models, such as autoregressive integrated moving average (ARIMA) approaches, despite their widespread use for influenza prediction (***Wang et al., 2017***), struggle to capture non-linear, lagged interactions among multiple meteorological drivers (***Shoji et al., 2011***). Existing studies often overlook spatiotemporal heterogeneity intrinsic to subtropical contexts (***Zhang et al., 2022***). Although time series analysis approaches and deep neural learning architectures, such as recurrent neural networks (RNNs; ***Venna et al., 2018***), can accommodate non-linear temporal patterns (***Venna et al., 2018***), their performance is hampered when baseline algorithms are not deeply customized, due to gradient vanishing or explosion and memory decay over longer horizons (***Liu et al., 2021***). Long short-term memory (LSTM) algorithms, an advanced RNN

variant, address these issues through gating mechanisms that enable learning of long-term temporal dependencies (***Absar et al., 2022***; ***Liu et al., 2021***), and have shown potential in influenza forecasting (***Absar et al., 2022***), although prior applications have often struggled with architectural complexity, high computational cost, and typically lower interpretability than regression-based methods (***Liu et al., 2021***). For instance, LSTM algorithms have demonstrated good predictive performance in the Chinese subtropical Fujian region but remained limited in identifying subtype-specific virological differences (***Zhu et al., 2022***). Another study proposed a complex seasonal autoregressive integrated moving average-LSTM (SARIMA-LSTM) hybrid utilizing singular spectrum analysis to predict influenza in Shanxi. However, its analytical complexity and computational intensity limited the explicit incorporation of confounding behavioral variables (***Zhao et al., 2023***).

Reviewing these literatures, three substantive gaps persist at the intersection of meteorological characterization and forecasting for subtropical influenza. First, most existing studies apply DLNM or LSTM in isolation, foregoing the potential synergy between interpretable association modeling and flexible non-linear forecasting. Second, forecasting models trained on data spanning the Coronavirus Disease 2019 (COVID-19) period rarely account explicitly for the epidemiological non-stationarity introduced by non-pharmaceutical interventions (NPIs) and fluctuating testing intensity, which can distort case-based signals in subtype-specific ways. Third, the choice of modeling endpoint, typically raw case counts rather than positivity, renders them highly vulnerable to surveillance intensity biases, as counts confound underlying transmission with testing behavior, particularly during the pandemic and post-pandemic period.

To address these gaps, we developed a dual-stage analytical framework leveraging 6 years (2018–2023) of multi-site influenza surveillance data and detailed meteorological records from Putian, a representative subtropical coastal city in southeastern China that exemplifies the under-investigated non-megacity urban context described above. The 2018–2023 window deliberately spans a structural epidemiological transition: two pre-pandemic baseline years (2018–2019), three NPI-modulated years (2020–2022), and one post-suppression rebound year (2023). In the first stage, DLNMs characterize subtype-specific, non-linear, and lag-distributed associations between meteorological variables and influenza A and B positivity. In the second stage, a Bayesian-optimized LSTM network integrating meteorological, autoregressive, and socio-behavioral covariates is used for short-horizon forecasting, benchmarked against two covariate-matched baselines representing distinct methodological paradigms, a multivariate ARIMA model (canonical linear statistical reference) and an XGBoost gradient-boosting ensemble (leading non-deep-learning machine-learning framework), and evaluated in an independent subtropical city (Sanming) as a preliminary test of model portability. Influenza positivity rate is used as the primary modeling endpoint to mitigate testing-related surveillance bias. Our aim is to provide an interpretable, climate-informed forecasting approach for subtropical influenza that can serve as a methodological foundation for future operational surveillance developments.

Methods

Study area

Putian city, on China's southeast coast (latitudes 24°59′N–25°46′N; longitudes 118°27′E–119°39′E), is an exemplary site for examining the impact of meteorological factors on influenza transmission in subtropical regions. It experiences a subtropical monsoon climate with warm, humid summers and mild, dry winters. The annual mean temperature ranges from 18 to 21°C, with summer highs up to 38 °C and winter lows at 0.4 °C. Annual precipitation averages 1538.8 mm, with 72% occurring from April to September. Covering 4200 km² and comprising five counties/districts with a population exceeding 3.2 million, Putian is a commercial/manufacturing hub with high population density and robust transport infrastructure, promoting mobility and social interactions that drive influenza transmission. Since 2004, Putian CDC has operated a comprehensive influenza surveillance network, meeting national standards, employing standardized RT-PCR testing for influenza A/B since late 2017, supported by a trained workforce, ensuring high-quality longitudinal data for robust analyses. Together, Putian's climatic variability, socio-economic characteristics, and surveillance capacity position it as a representative site for this study, with insights potentially generalizable to other subtropical settings with similar environmental and epidemiological traits. Sanming city, between latitudes

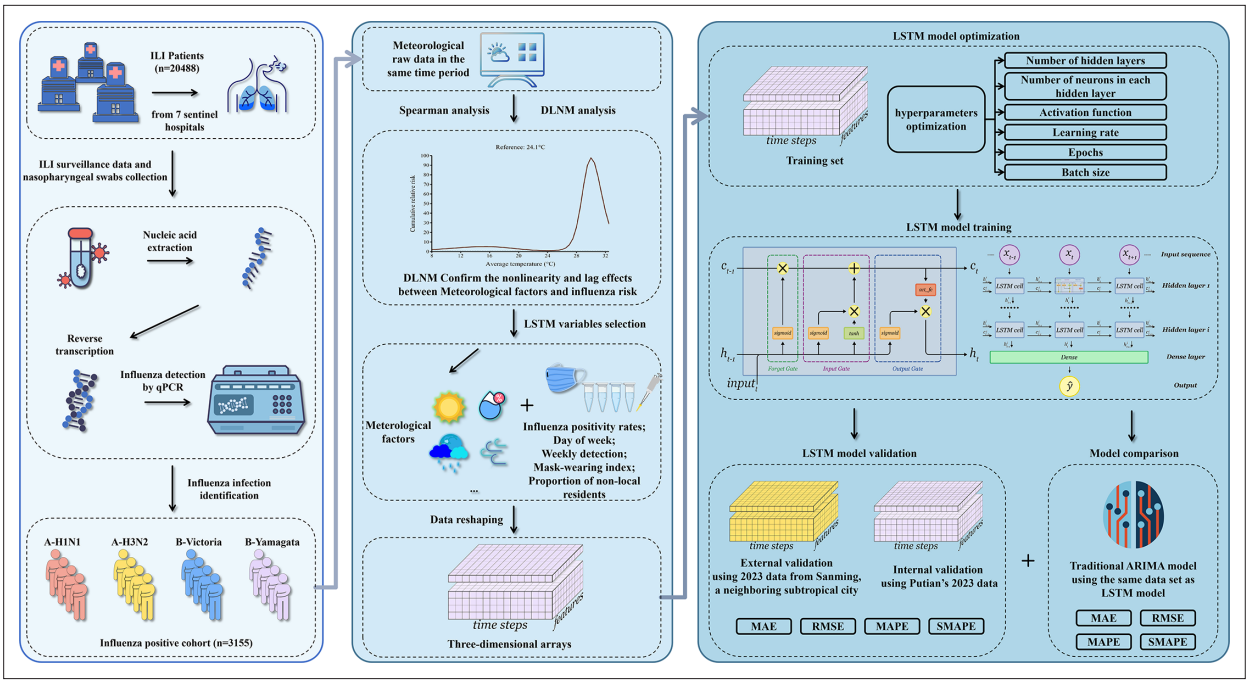


Figure 1. Schematic of the integrated analytical framework. Sentinel surveillance data, daily meteorological records, and contextual covariates (weekly testing volume, weekday/weekend status, proportion of non-local residents, mask-wearing stringency indices) were jointly analyzed using two complementary approaches: (i) distributed lag non-linear models (DLNM) to characterize non-linear, lagged exposure–response relationships between meteorological variables and subtype-specific influenza positivity, providing mechanistic interpretability; and (ii) a Bayesian-optimized long short-term memory (LSTM) neural network for short-term forecasting of daily positivity rates, benchmarked against a seasonal ARIMA model. The primary outcome variable for both analytical streams is the influenza positivity rate (laboratory-confirmed cases ÷ ILI specimens tested), selected to mitigate surveillance bias arising from temporal variation in testing intensity.

The online version of this article includes the following figure supplement(s) for figure 1:

Figure supplement 1. Detailed structure of the long short-term memory (LSTM) cell.

25°30′N and 27°97′N and longitudes 116°22′E and 118°39′E, is a mountainous subtropical area about 330 km from Putian and was selected as an external validation data source.

Study design

This collaborative study by Putian CDC and MBETRC utilized a prospective cohort design, with retrospective analysis of historical data to systematically assess non-linear quantitative relationships between multiple meteorological factors and both influenza A and B infection risks, ultimately constructing a predictive network for outbreaks. The technical workflow is illustrated in ***Figure 1***.

The study spanned January 1, 2018, to December 31, 2023, integrating four categories of data sources: (i) ILI and laboratory-confirmed cases from seven influenza sentinel hospitals across Putian's urban and rural areas, ensuring representative coverage of diverse healthcare-seeking populations; (ii) daily meteorological data; (iii) COVID-19-associated public health intervention indicators (mask-wearing stringency indices) recorded from January 2020 onwards; and (iv) demographic mobility indicators (non-local population proportion) obtained from ILI consultation records. To ensure consistency between meteorological measurements and influenza incidence records across all data sources, we applied rigorous quality control and pre-processing procedures, including: temporal alignment of all data streams to a unified daily resolution; missing value imputation using temporally adjacent observations for sporadic gaps (<5% of records); cross-validation of laboratory-confirmed cases against ILI consultation records to identify and resolve coding inconsistencies; and standardization of meteorological measurements against the regional monitoring network's established calibration protocols. Final datasets underwent independent verification by two co-investigators to ensure analytical reliability.

DLNM was first constructed to screen meteorological factors and other covariates with substantial influence on influenza seasonality. Subsequently, an LSTM neural network was developed within the same covariate system. The integrated DLNM–LSTM framework was employed as methodologically complementary rather than redundant components, leveraging the distinctive strengths of each approach to address different analytical objectives within a unified investigation. Specifically, DLNM models characterize the non-linear exposure–lag–response relationships between meteorological factors and influenza risk, providing biologically interpretable insights into the temporal structure of weather-influenza associations and identifying meteorological factors with statistically and clinically significant effects on influenza dynamics. Building upon this DLNM-derived foundation, the LSTM network constructs a time-series forecasting tool within an identical covariate framework, evaluating predictive capability for influenza transmission trends. Beyond meteorological factors, our DLNM–LSTM framework systematically incorporated influenza positivity rates as the primary outcome variable, weekly detection volumes, mask-wearing stringency indices (indicator of NPIs during the COVID-19 pandemic), day of the week (DOW, distinguishing weekdays from weekends), and non-local population proportion (defined as the ratio of the number of individuals whose reported residential district at the time of testing lies outside Putian city to the total number of tests) as covariates within both DLNM and LSTM modeling pipelines. This comprehensive covariate framework ensures that observed meteorological associations are estimated after controlling for surveillance intensity, public health intervention status, behavioral healthcare-seeking cycles, and population mobility patterns. The LSTM network was trained on data from Putian corresponding to the four categories described above, with the time period 2018–2022, and Putian's 2023 data serving as the internal validation set. To rigorously evaluate model transferability beyond the training context, data from Sanming city, a mountainous subtropical city exhibiting comparable climatic characteristics, influenza seasonality, and public health intervention frameworks to Putian, were acquired for the period January 1, 2023, to December 31, 2023, temporally aligned with the Putian internal validation set. This Sanming dataset constituted our external validation set, facilitating a geographically independent assessment of model generalization within subtropical Chinese environments. The predictive performance of the LSTM algorithm for influenza A/B prevalence in 2023 Putian data was benchmarked against two parallel baseline models: (1) the multivariate ARIMA model with exogenous variables (ARIMAX), representing the canonical linear statistical reference in public-health surveillance, and (2) the XGBoost gradient-boosting regression model, representing the leading non-deep-learning machine-learning framework for structured predictive tasks. To ensure paradigm-agnostic methodological comparisons, both baseline models were supplied with the exact same meteorological and epidemiological covariate matrix as the LSTM, without any additional hand-engineered lag features, sliding-window statistics, or autoregressive terms in the XGBoost input.

Study cohort

The study systematically screened outpatients and emergency departments across seven influenza surveillance sentinel hospitals within Putian's prefecture, with an emphasis on high-incidence departments such as fever clinics, internal medicine, and pediatrics. The screening encompassed ILI cases from January 1, 2018, to December 31, 2023. According to earlier mentioned Diagnostic Criteria for Influenza, ILI cases are defined as acute onset with high fever (≥38 °C) accompanied by respiratory symptoms such as cough or sore throat yet lacking a definitive causative diagnosis. Respiratory samples, together with onset dates, ages, genders, and residence information, were collected from all ILI patients. Unified diagnostic standards, laboratory testing methods, and quality control measures were employed throughout the research process to ensure the reliability and comparability of results.

After conducting nucleic acid screening for influenza viruses from the collected respiratory samples of ILI cases, the criteria for inclusion in this study prescribed the following: (1) the patient presented ILI symptoms; (2) the nucleic acid test for the influenza virus yielded a positive result; (3) the patient is a permanent resident of Putian city; (4) the duration of symptoms from onset to consultation did not exceed three days. No restrictions were placed on age or gender. The exclusion criteria were: (1) patients with clear alternative diagnoses, such as bacterial pneumonia or tuberculosis; (2) patients with unstable underlying conditions, such as acute exacerbations of chronic obstructive pulmonary disease (COPD); (3) patients with missing data. This focused the analysis on acute ILI presentations.

Sample collection and pathogen typing

Respiratory specimens (nasopharyngeal swabs) were collected from ILI patients during their initial visit, prior to treatment, and stored at 4 °C in viral transport medium. All samples were delivered to the Putian CDC laboratory within 24 hr of collection. Nucleic acid extraction and one-step real-time fluorescent RT-PCR for influenza A/B subtyping (H1N1, H3N2, Victoria, and Yamagata lineages) were executed within 24 hr upon sample arrival. All assays utilized commercial diagnostic kits (supplied by Da'an Gene Co., Ltd., Guangzhou, China) and were processed in a Biosafety Level 2 (BSL-2) laboratory, strictly adhering to the manufacturer's instructions and China CDC's standardized protocols. Detailed laboratory procedures, including RNA integrity parameters and PCR amplification thresholds, are comprehensively documented in Appendix 1.

Meteorological and mask-wearing stringency data sources

Daily meteorological observation data were sourced from the authoritative public database Visual Crossing (https://www.visualcrossing.com/), which provides extensive historical time-series data across multiple meteorological variables on a global scale. We extracted meteorological variables for Putian city, including daily average temperature (°C), relative humidity (%), precipitation (mm), diurnal temperature range (°C), wind speed (km/h), wind direction, solar radiation (W/m^2), and the ultraviolet (UV) index, wherein the diurnal temperature range was defined as the difference between the highest and lowest temperatures of the day. All indicators were synchronized with the influenza data. To ensure data quality, we conducted missing value and anomaly checks on the raw datasets.

Early COVID-19 studies revealed a notable decline in positivity rates of seasonal respiratory viruses relative to pre-pandemic levels, likely due to containment measures such as masking and behavioral changes during surges (***Groves et al., 2021***; ***Xiao et al., 2022***). To evaluate the potential impact of NPIs on influenza incidence, we obtained policies on face-covering stringency indices for China during the same period from the Our World in Data (OWID) database (https://ourworldindata.org/coronavirus) as a reference indicator for social prevention measures. This database categorizes stringency indices into four levels: 0 represents 'No policy'; 1 represents 'Recommended'; 2 represents 'Required in some specified shared/public spaces outside the home with other people present, or some situations when social distancing is not possible'; 3 represents 'Required in all shared/public spaces outside the home with other people present or all situations when social distancing is not possible'; 4 represents 'Required outside the home at all times regardless of location or presence of other people'. While OWID stringency indices represent national-level policy, publicly accessible and standardized daily NPI datasets at the municipal level are currently unavailable in China. To ensure the spatial validity of these indices, co-authors from Putian CDC, who actively managed the local epidemic response, conducted a rigorous cross-validation. Our local public health experts confirmed that temporal fluctuations of OWID indices (ranging from Level 4 strict mandates in 2020 to Level 1 recommendations in 2023) exhibited high fidelity with the actual, on-the-ground enforcement of NPIs in both Putian and Sanming. Thus, OWID stringency indices serve as a highly reliable contextual proxy for local social contact restrictions. The datasets from these public databases, along with influenza surveillance data, are devoid of missing values, eliminating the need for imputation. For influenza and COVID-19 data, days with no reported cases were explicitly recorded as zero per the sentinel hospitals' and government's daily reporting protocols, respectively.

DLNM analysis

In the DLNM stage, we prioritized epidemiological interpretability by focusing on the isolated, nonlinear, and lag-distributed main effects of individual meteorological variables. We deliberately avoided incorporating high-dimensional interaction cross-bases (e.g. temperature ×humidity interactions) within the DLNM to prevent the 'curse of dimensionality', which typically induces severe multicollinearity, variance inflation, and loss of interpretability. Instead, the computational task of capturing complex, synergistic interactions among concurrent meteorological factors and epidemiological covariates was entirely delegated to the subsequent LSTM network. Separate DLNM models were constructed for influenza A and B incidence risks. To rigorously address the inherent confounding effects of 'surveillance intensity bias', where fluctuations in raw case counts may merely reflect transient surges in clinical testing capacity rather than true community transmission. The primary outcome metric for all DLNM modeling was mathematically defined as influenza positive rates, with meteorological factors

serving and weekly detection volumes as independent variables. The formula for calculating the daily influenza positivity rate is as follows:

$$R\left(\%\right)=\frac{I}{N}\times 100\%$$

where R represents the daily influenza positivity rate, I is the number of daily influenza positive cases, and N denotes the total number of daily influenza tests performed. By adopting this WHO-recommended surveillance metric, our analytical framework explicitly standardizes the epidemiological denominator. This mathematical normalization effectively neutralizes the severe surveillance intensity bias caused by dramatic testing volume surges during the COVID-19 pandemic, ensuring that our models capture intrinsic viral transmissibility rather than artificial fluctuations in healthcare-seeking behavior or diagnostic capacity.

The maximum lag time was set to 15 days for all models, optimizing biological validity and analytical precision to capture typical influenza incubation and transmission cycles while minimizing model complexity and statistical noise from extended periods, aligning with previous research. A quasi-Poisson function was used to account for overdispersion in the incidence data. The model formula is as follows:

$$\log\left(E\left(y_{time}\right)\right)=\alpha+cb\left(x,k\right)+ns\left(time,df\right)$$

where α is the intercept, cb is the cross-basis function of the meteorological factor x, which includes the exposure-response relationship $ns\left(x,\ df_x\right)$ defined using natural splines and the lag-response relationship $g_{(k)}$ defined using polynomial functions, with k as the lag time; $\mathrm{ns}\left(\mathrm{time},\mathrm{df}_{\mathrm{time}}\right)$ is a natural spline function reflecting long-term trends, used to control for long-term trends and seasonality. The degrees of freedom df_x for the exposure-response relationship ns spline were selected from the range {2, 3, 4, 5}, while the degrees of freedom df_{time} per year for the long-term trend ns spline were selected from the range {6, 7, 8, 9, 10}. To determine the optimal df, including df_x for the natural spline modeling the exposure-response relationship and df_{time} per year for the natural spline modeling the long-term trend, the combinations of above ranges across selected meteorological factors were systematically evaluated using the quasi-Akaike information criterion (QAIC), with the combination yielding the smallest QAIC being selected. This indicates the best balance between goodness-of-fit and model complexity for our influenza-specific dataset. The QAIC formula is:

$$QAIC=-2\lambda+2\varphi k$$

where λ represents the log-likelihood of the model, φ is the dispersion parameter, and k indicates the number of model parameters. When plotting prediction results, the median was selected as the reference value for its robustness against outliers, offering a consistent depiction of typical meteorological conditions in Putian and enabling the evaluation of bidirectional risk deviations, especially during extreme weather events. To further explore the impact of extreme weather conditions, we predicted and displayed the effects of extreme scenarios across the meteorological factors: except for precipitation extremes set at 5 mm and 25 mm and wind direction extremes at the most common eastern and rarest northwestern winds, all other meteorological factors were assessed at their 5th and 95th percentiles to estimate the lagged risk curves relative to the median. All results underwent seasonality and long-term trend corrections to eliminate the inherent temporal variation effects between meteorological factors and influenza incidence. Furthermore, to evaluate the robustness of our findings against testing intensity, we conducted a targeted sensitivity analysis by reconstructing the DLNM models without the 'weekly detection volumes' covariate. We then compared the non-linear cumulative risks and extreme weather lag effects between these ablated models and the original fully adjusted models. This allowed us to determine whether the identified climate-transmission associations were stable and biologically intrinsic or merely artifacts of weather-correlated testing behaviors (***Figure 4—figure supplement 2***; ***Figure 5—figure supplement 1***; Appendix 1).

LSTM modeling and analysis

Building on the significant non-linear and lagged effects through DLNM analysis of real-world environmental exposures, this study further constructed multi-factor influenza A and B prediction LSTM networks. Leveraging a recurrent architecture with non-linear gating mechanisms, these networks are

able to automatically capture and represent complex, high-dimensional interactions among meteorological variables without the need for manual pre-specification. This functional division of labor between the DLNM and LSTM models enhances predictive performance while preserving the interpretability and inferential stability established in the DLNM stage.

Data preprocessing

To respect the temporal dependence inherent in LSTM architectures and avoid data leakage, we adopted a strict chronological out-of-time (OOT) validation strategy. Time-series data from January 1, 2018, to December 31, 2022 (88.26% of the Putian dataset) were used for model training, with 10% reserved during Bayesian optimization as an internal validation subset for convergence monitoring and hyperparameter tuning only. Data from January 1, 2023, to December 31, 2023 (11.74%) were held out as a chronologically internal validation set for final performance evaluation, covering a complete annual cycle. External validation was conducted using concurrent 2023 data from Sanming city to assess spatial generalizability. To prevent distributional leakage, all normalization parameters were derived exclusively from the training set and consistently applied to the validation sets, with predictions subsequently transformed back to the original scale. First, we performed data normalization, a crucial step to ensure that training and test set data are compared on a unified scale. We normalized the training and test set data separately within the range [0, 1]. For the test set normalization, we used the maximum and minimum values from the training set as boundaries. This approach ensured consistency between the normalized test set data and the training set data. After the algorithm conducted predictions on the test set data, we performed denormalization to convert the predicted results back to the original data scale and rounded them to integers. These steps ensured that the final prediction results accurately and objectively reflected the LSTM's performance in real-world scenarios and provided reliable data for subsequent calculation of evaluation metrics.

The normalization formula is as follows:

$$y_{\text{normalization}} = \frac{y - min_Y}{max_Y - min_Y}.$$

The denormalization formula is as follows:

$$y_{denormalization} = \hat{y} \times \left(max_Y - min_Y\right) + min_Y$$

where $y_{normalization}$ is the normalized result, y is the original data, max_Y is the maximum value of the variable in the training set, min_Y is the minimum value of the variable in the training set, y is the predicted value output by the network, and $y_{denormalization}$ is the denormalized result.

Finally, we used the above data while accounting for potential confounders by incorporating features such as DOW (1 for Monday–Friday, 0 for Saturday–Sunday), weekly detection volumes, non-local population proportion, and mask-wearing stringency indices, as input variables. The time step was set to 15 (i.e. using data from the previous 15 days to predict the number of influenza cases for the following day), and the two-dimensional data were organized into three-dimensional arrays to serve as input for the LSTM network.

Network construction

Our algorithm is a unidirectional LSTM, using Adam as the optimizer and Mean Squared Error (MSE) as the loss function. The number of hidden layers was determined through hyperparameter optimization to ensure optimal network performance. Each LSTM layer consists of multiple LSTM units, including input gates, output gates, and forget gates. The basic framework of LSTM and the structure of LSTM units are shown in the LSTM model training section of ***Figure 1***. Detailed principles and formulas are provided in the LSTM cell description section of Appendix 1 and illustrated in ***Figure 1—figure supplement 1***.

In the LSTM network, we used Hyperopt for hyperparameter tuning, with MSE established as the reference metric for evaluating and selecting optimal parameters. A detailed description of the Hyperopt Bayesian optimization strategy, including the specific search spaces for learning rates, network architecture (layers and neurons), and activation functions, along with their associated theoretical justifications, is provided in Appendix 1.

LSTM performance evaluation

To comprehensively assess the predictive performance of the LSTM network, this study employed multiple commonly used error metrics, including Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Symmetric Mean Absolute Percentage Error (SMAPE; ***Chicco et al., 2021***).

MAE reflects the actual magnitude of the error between predicted and true values. A smaller MAE indicates more accurate network predictions. The calculation formula is:

$$MAE=\frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i\text{-}y_i|$$

RMSE is the arithmetic square root of the mean of the sum of squared errors. A smaller RMSE indicates better network prediction performance. The calculation formula is:

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i-y_i|^2}$$

A smaller MAPE indicates better network fitting. Since the denominator cannot be zero, this study only calculated MAPE for data where the true value was not zero. The calculation formula is:

$$MAPE=\frac{1}{n}\times\sum_{i=1}^{n}\frac{|\hat{y}_i-y_i|}{|\hat{y}_i|}\times 100\%$$

A smaller SMAPE indicates better network fitting. The calculation formula is:

$$SMAPE=\frac{1}{n}\times\sum_{i=1}^{n}\frac{|\hat{y}_i-y_i|}{\left(|y_i|+|\hat{y}_i|\right)/2}\times 100\%$$

In the above formulas, $\hat{y}_i$ is the denormalized predicted value, y_i is the observed value, and n is the number of observations.

Additionally, we plotted the loss function curves for the network on the training and validation sets to monitor LSTM convergence and the risk of overfitting. By comparing the trends and differences between the two curves, we could intuitively judge the training quality and generalization ability of the LSTM. Furthermore, the Bayesian-optimized LSTM architecture was directly applied, without re-training or hyperparameter adjustment, to both the Putian internal test set and the Sanming external validation set. This unified modeling framework ensures that performance differences between the two evaluation contexts primarily reflect geographic transferability rather than model re-optimization.

To further explain the prediction logic of the LSTM, this study employed the SHapley Additive exPlanations (SHAP) method for interpretability analysis. Based on Shapley values from game theory, SHAP can quantify the contribution of each feature to the LSTM's predicted values and generate intuitive feature importance rankings and dependency plots. Positive values indicate that the feature increased the predicted result, while negative values indicate a decrease in the predicted result, with the absolute value representing the intensity of the influence. SHAP values were computed by treating each meteorological variable's lagged sequence as a unified input feature, preserving the temporal dependencies within each sequence rather than decomposing lags into independent features. This approach not only aligns with the sequential nature of our LSTM inputs but also supports our focus on population-level trends rather than individual-level causal effects.

To compare the predictive performance between the LSTM and alternative forecasting paradigms, we concurrently developed two baseline models, the ARIMA model and the XGBoost machine-learning model, employing the same training and validation datasets. Detailed methodological descriptions for ARIMA and XGBoost are provided in Appendix 1. To assess the contribution of pandemic-related confounding variables, we performed targeted sensitivity analyses on the LSTM networks. Specifically, we sequentially removed covariates of mask-wearing stringency indices and weekly detection volumes from the input features while maintaining identical Bayesian-optimized hyperparameters. The performance of these ablated models was evaluated using MAE, RMSE, MAPE, and SMAPE, allowing us to

quantify the exact necessity of incorporating testing and behavioral covariates in forecasting models during periods of epidemiological non-stationarity.

Statistical analysis

Statistical analyses were primarily conducted using R software version 4.3.1 and Python version 3.8.3. DLNM modeling was implemented using the DLNM package in R. LSTM construction used Python 3.8.3, with the main program packages including TensorFlow 2.2.0, Keras 2.3.1, and SHAP 0.44.1. The ARIMA model was developed using the forecast package in R. The XGBoost model was implemented using the xgboost package (version 3.2.1.1) in R, with hyperparameter tuning conducted via grid search over tree depth, learning rate, and L2 regularization parameters; detailed specifications are provided in Appendix 1. All statistical tests were two-sided, with $p<0.05$ considered statistically significant.

Results

Descriptive epidemiological characteristics of influenza

Between 2018 and 2023, influenza surveillance conducted at seven sentinel hospitals in Putian city documented a total of 20,488 ILI cases, with the majority being influenza-negative (17,333 cases, 84.60%). Of these, 3155 cases were confirmed positive for influenza through PCR testing, representing a positivity rate of 15.40%. By viral type, 1685 cases were attributed to influenza A (8.23%), and 1470 cases were influenza B (7.17%). Further subtype analysis of influenza A revealed 790 cases of H1N1 (46.88%) and 895 cases of H3N2 (53.12%), while for influenza B, 1270 cases were of the Victoria lineage (86.39%) and 200 cases were of the Yamagata lineage (13.61%). Notably, following the onset of the COVID-19 pandemic in the spring of 2020, there was a significant reduction in the incidence of influenza. The annual incidence rates for influenza in 2020 and 2021 were 4.54 per 100,000 and 5.59 per 100,000, respectively, markedly lower than the pre-pandemic rates observed in 2018 (16.01 per 100,000) and 2019 (18.74 per 100,000), suggesting that public health measures implemented during the COVID-19 pandemic, such as mask-wearing and social distancing, were highly effective in decreasing the transmission of respiratory infections, including influenza.

In terms of seasonal distribution, surveillance data (***Figure 2A***) revealed a notable dynamic interplay between influenza A and B viruses over the study period. In 2018, influenza A H1N1 was the dominant strain. However, by 2019, cases of the influenza B Victoria lineage gradually increased, surpassing H1N1 during the summer to become the predominant strain. In 2020, there was a sharp decline in influenza cases, with influenza A H3N2 emerging as the leading strain. The dominance of influenza B Victoria returned in 2021, only to be overtaken by a resurgence of influenza A H3N2 in the summer of 2022. By the winter of 2023, both influenza A H3N2 and influenza B Victoria exhibited concurrent surges, highlighting a dual epidemic. Interestingly, contrary to the conventional understanding that influenza A typically drives larger epidemics, our findings show that influenza B accounted for 52.18%, 100%, and 64.76% of total influenza cases in 2019, 2021, and 2022, respectively.

In the demographic analysis (***Figure 2B***), among the 3155 laboratory-confirmed influenza cases, males (1811 cases; 57.40%) outnumbered females (1344 cases; 42.60%), with a male-to-female ratio of 1.35:1. Significant gender differences were observed across age groups ($\chi^2=37.36$, $p<0.001$), with males constituting 60.68% of cases in individuals under 18 years, whereas females were more prevalent among adults (51.53%). A statistically significant difference was noted in the gender distribution between influenza A and B ($\chi^2=6.64$, $p=0.01$), with males comprising a higher proportion (59.86%) of influenza B cases. No statistically significant gender differences were found in the distribution of influenza subtypes ($\chi^2=2.43$, $p=0.49$). Additionally, children and adolescents under 18 years of age were identified as the most at-risk population, comprising 73.12% of all cases. Influenza B was more prevalent in this age group (50.33% vs 49.67%), while influenza A predominated in the adult population (63.56% vs 36.44%). These findings suggest that susceptibility to different influenza virus types may vary across age groups.

Overall, our findings suggest that in subtropical regions, the epidemic intensity of Influenza B may rival or even exceed that of Influenza A. Furthermore, the widespread implementation of non-pharmaceutical interventions (e.g. mask-wearing) during the COVID-19 pandemic significantly curtailed the transmission of respiratory viruses, including influenza. Continuous monitoring of

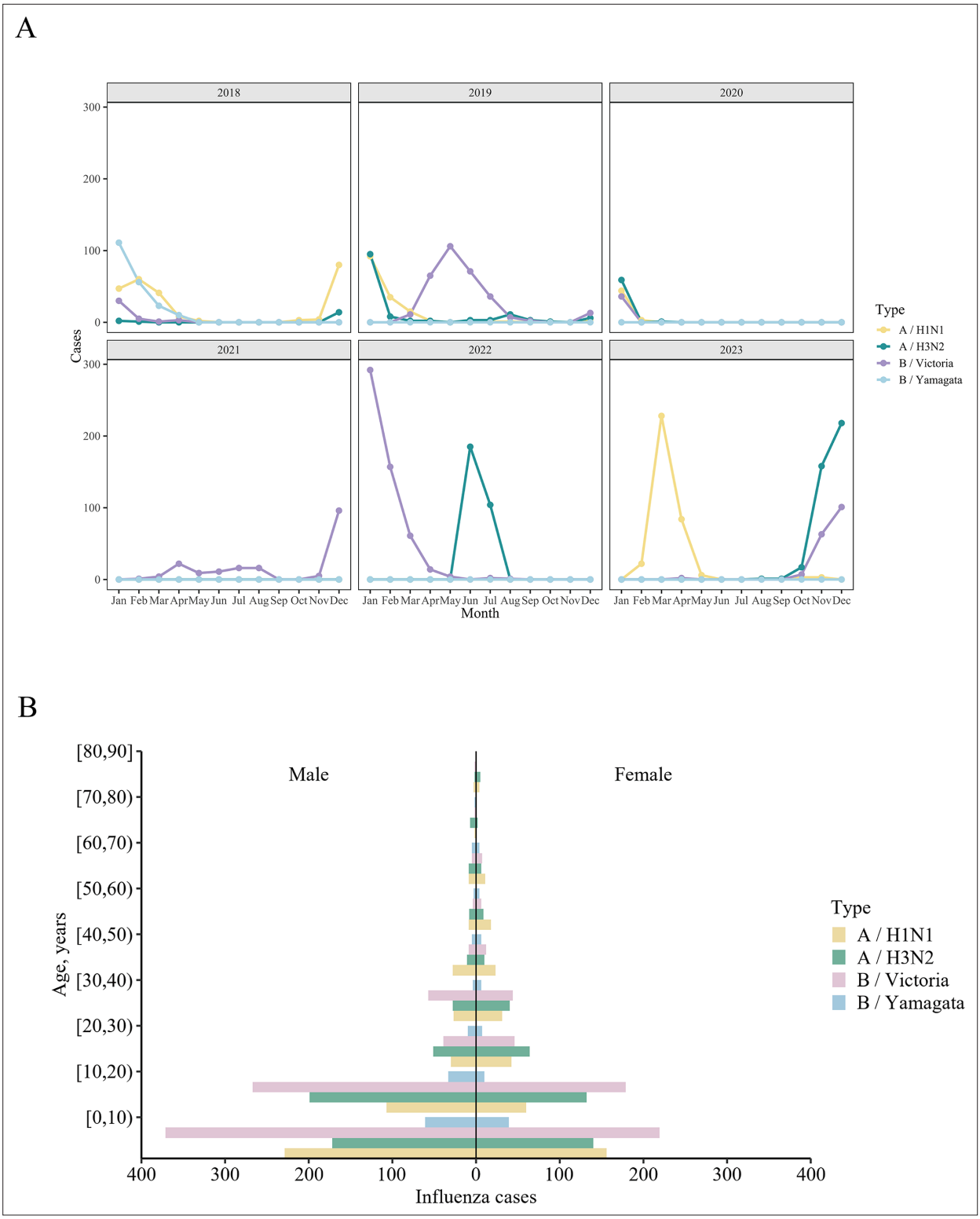


Figure 2. Temporal trends and demographic distribution of confirmed influenza cases in Putian city (2018–2023). This figure presents the temporal and demographic characteristics of influenza cases in Putian City over the period from 2018–2023. (**A**) Line plots illustrating the monthly case counts of various influenza subtypes (Influenza A: H1N1, H3N2; Influenza B: Victoria, Yamagata) across the six-year study period. The transition from historical heatmaps to precise line plots enhances the visualization of distinct epidemiological surges, notably capturing the profound interruption of typical seasonal transmission during the peak COVID-19 pandemic restrictions (2020–2021) and the subsequent viral rebound. (**B**) A population pyramid detailing the age and sex distribution of the 3155 laboratory-confirmed influenza cases. The visualization bifurcates cases by gender (males on the

Figure 2 continued on next page

Figure 2 continued

left, females on the right) across age intervals, underscoring the differential susceptibility of specific demographic cohorts, particularly highlighting the heightened vulnerability of children (0–10 years) and the elderly (60 years and above).

Influenza A (H1N1, H3N2) and Influenza B Victoria strains is essential, alongside heightened surveillance and preventive measures targeting children and adolescents, particularly in densely populated settings such as schools. These insights are crucial for informing vaccine strain selection, vaccination strategies, and targeted public health interventions.

Temporal trends in meteorological factors and influenza incidence

Between 2018 and 2023, influenza incidence in Putian city exhibited distinct seasonal patterns. The majority of cases occurred between December and April of the following year, accounting for over 60% of the annual cases, with a secondary peak observed from May to September. Concurrently, meteorological factors demonstrated clear cyclical variations (***Figure 3***), with detailed descriptive statistics provided in ***Appendix 1—table 1***. During the summer months (July–August), the average daily temperature peaked above 32 °C, while the winter months (December to February) experienced the lowest temperatures, dropping below 10 °C (***Figure 3A***). This period was also marked by a significant increase in diurnal temperature variation, with the maximum daily range exceeding 13 °C (***Figure 3E***). The annual average relative humidity was approximately 76.10%, with the highest levels observed during the monsoon season from April to June, reaching or exceeding 97% (***Figure 3B***). Precipitation showed a pronounced seasonal distribution, peaking from June to September, while the period from November to February was characterized by markedly reduced rainfall (***Figure 3C***). Solar radiation (***Figure 3D***) and UV index (***Figure 3F***) both reached their annual maxima between July and September. High wind speeds were frequently observed in Putian city from July to September during the study period, which may correlate with the increased frequency of typhoons during this period (***Figure 3G***). Throughout the year, the easterly wind direction was predominant, with infections in every month (***Figure 3H***).

A comparative analysis of the temporal patterns of influenza cases and meteorological factors revealed several noteworthy correlations. Firstly, peaks in influenza A and B cases were predominantly observed during January to March, characterized by lower average temperatures and larger diurnal temperature ranges (***Figure 3A and E***). Interestingly, influenza B occasionally exhibited summer outbreaks, indicating a potentially greater adaptability to higher temperatures. Secondly, extreme rainfall events and periods of high humidity were often temporally aligned with short-term surges in influenza A and B cases (***Figure 3B and C***). For instance, the peak in influenza B cases in June 2022 coincided with an episode of extreme rainfall. Moreover, the majority of influenza outbreaks occurred during periods of reduced solar radiation and lower UV indices (***Figure 3D and F***). Variations in wind speed were not significantly correlated with influenza incidence in our analysis (***Figure 3G***). During periods of easterly winds, there was a noticeable decrease in the incidence of influenza A and B cases in Putian city (***Figure 3H***). Our findings indicate that the incidence of influenza in the subtropical city of Putian was closely associated with various meteorological conditions.

DLNM analysis for the cumulative and lagged effects of multiple meteorological factors on influenza incidence risks

Prior to analytical modeling, we assessed the pairwise associations and potential collinearity among all meteorological variables using a comprehensive correlation heatmap and scatterplot matrix (***Figure 4—figure supplement 1***). Notably, certain variables, such as solar radiation and the UV index, exhibited high positive correlations. To avoid coefficient instability typically caused by multicollinearity in regression models, all DLNM analyses adopted a univariate approach for meteorological factors, sequentially evaluating the non-linear and lagged effects of individual meteorological predictors while adjusting for time trends and including other fixed covariates. Conversely, all variables were retained during the LSTM forecasting phase, as the non-linear gating architecture of recurrent neural networks inherently exhibits robust regularization against potential collinearity among input features.

To elucidate the non-linear associations between meteorological factors and influenza transmission dynamics, DLNM models were constructed using influenza positivity rates as the outcome variable. For each meteorological variable, the median exposure level was used as the reference. For each

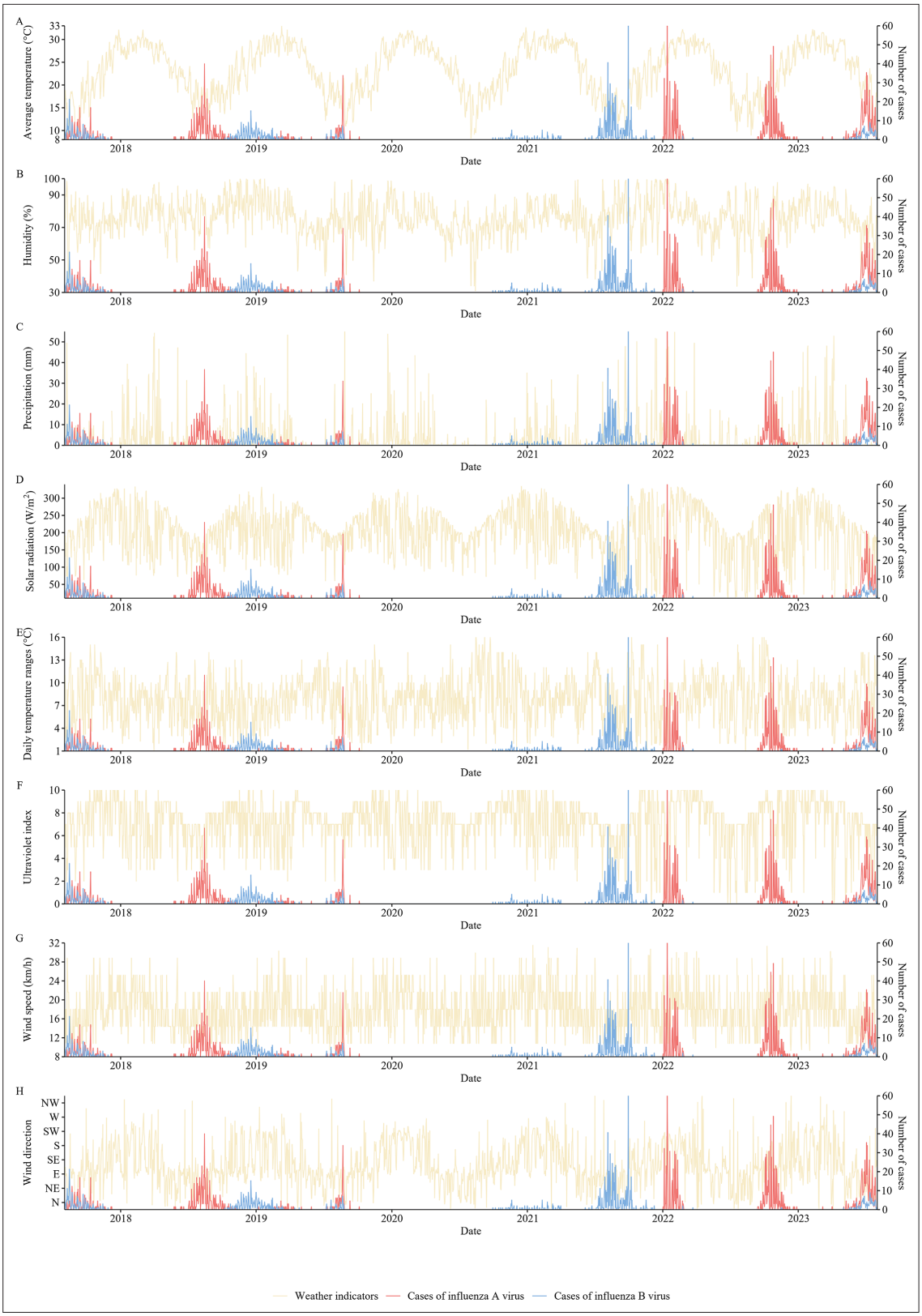


Figure 3. Time-series dynamics of meteorological factors and absolute influenza incidence in Putian city from 2018 to 2023. This figure vertically aligns temporal fluctuations of influenza-related and key meteorological variables, January 2018-December 2023. From top to bottom: (**A**) average temperature, (**B**) humidity, (**C**) precipitation, (**D**) solar radiation, (**E**) daily temperature ranges, (**F**) ultraviolet index, (**G**) wind speed, and (**H**) wind direction, plotted continuously over the 5-year surveillance period. In each panel, the left y-axis corresponds to the observed values of the respective

Figure 3 continued on next page

Figure 3 continued

meteorological factor (yellow lines), while the right y-axis delineates the absolute number of confirmed cases for Influenza A (red lines) and Influenza B (blue lines). These descriptive plots visually contextualize the raw seasonal correlations and climatic extremes preceding major influenza outbreaks. Note the substantial suppression of influenza activity during 2020–2022 (pandemic intervention period) and the pronounced post-restriction rebound during 2023, particularly for influenza A, which exceeded pre-pandemic peak levels.

influenza-specific DLNM model, the combination of df_x for the exposure-response spline and df_{time} for the long-term trend spline that yielded the minimum QAIC was selected as the optimal configuration and applied to the final model. Detailed QAIC values for all tested combinations are provided in ***Appendix 1—table 2***. Overall, the associations between the eight evaluated meteorological factors and the cumulative risk of influenza positivity exhibited pronounced non-linear and subtype-specific characteristics over the 15-day lag period (***Figure 4A and B***). Influenza A was primarily associated with lower temperature, reduced diurnal temperature range, and intermediate UV levels, whereas influenza B showed broader meteorological sensitivity, with increased risks observed under lower temperature, higher humidity, lower solar radiation, lower UV index, and specific wind directions.

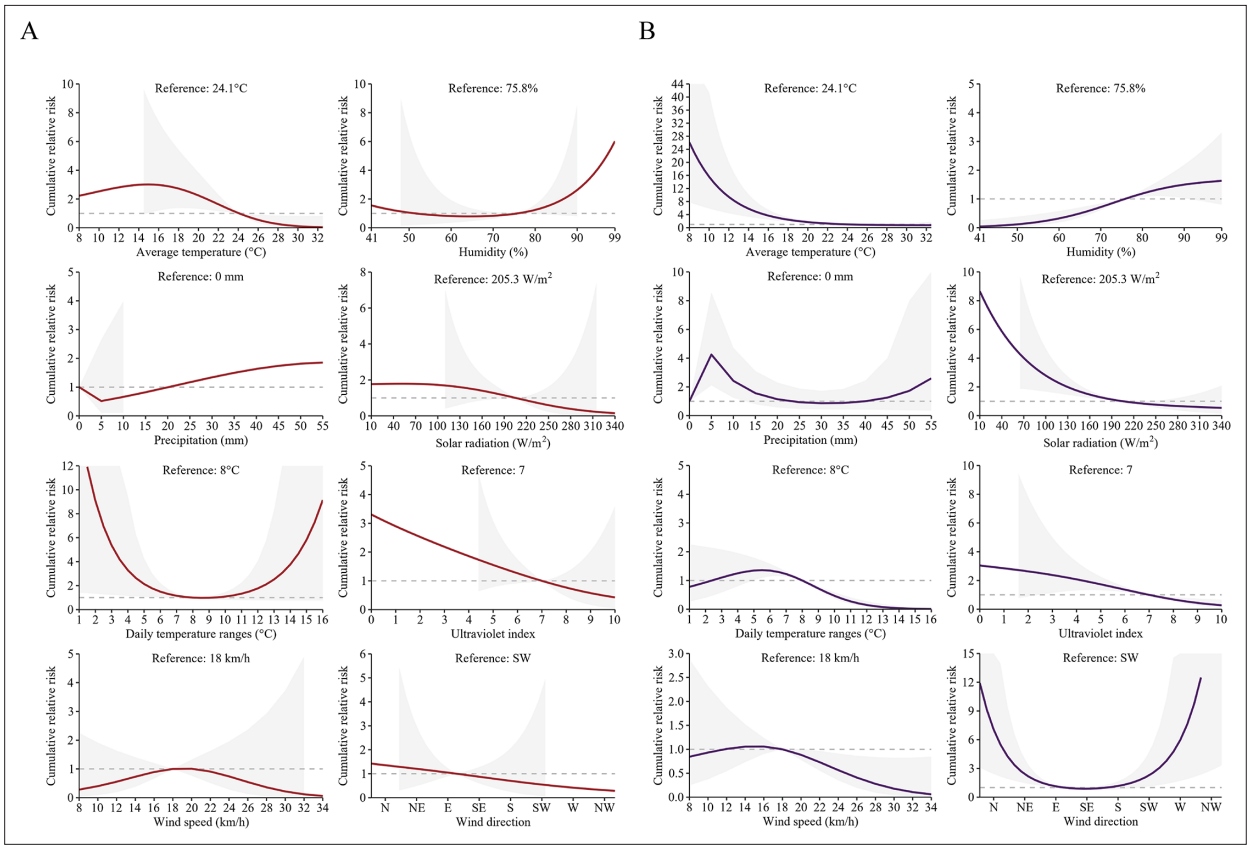


Figure 4. Cumulative risk of meteorological factors on influenza A and B positivity rates over a 15-day lag period. This figure illustrates the non-linear, cumulative exposure-response relationships between eight individual meteorological drivers and the 15-day cumulative relative risk of influenza A and B positivity, estimated via distributed lag non-linear models (DLNM). Models were adjusted for long-term trends, weekly detection volumes, and other fixed covariates. Panels in (**A**) depict the non-linear cumulative risks for Influenza A, while panels in (**B**) show the corresponding risks for Influenza B. In each subplot, the x-axis represents the observed continuous range of a specific meteorological factor, and the y-axis indicates the cumulative relative risk compared to a predefined median reference value (e.g. 24.1 °C for average temperature, 75.8% for humidity). The solid line reflects the estimated non-linear trend of risk variation, and the shaded regions denote the 95% confidence intervals (CIs).

The online version of this article includes the following figure supplement(s) for figure 4:

Figure supplement 1. Correlation heatmap and scatterplot matrix illustrating pairwise associations and potential collinearity among influenza indicators and meteorological variables.

Figure supplement 2. Sensitivity analysis of the cumulative relative risk of meteorological factors on influenza A and B positivity rates upon removal of the weekly detection covariate.

Regarding temperature effects, using 24.1 °C as the reference, the cumulative risk for influenza A showed a non-linear inverse-U pattern, peaking at 15 °C with a relative risk (RR) of 3.01 (95% CI: 1.03–8.78) and gradually decreasing across the 15–24°C range. In contrast, the cumulative risk for influenza B peaked at a much lower temperature of 8 °C (RR = 26.12, 95% CI: 7.79–87.61), with the risk steadily declining between 8°C and 24°C.

Humidity demonstrated divergent effects between the two subtypes. While relative humidity showed no significant association with influenza A risk (using 75.8% as a reference), the cumulative risk for influenza B increased gradually within the high-humidity range of 76–91%, reaching its peak at 91% (RR = 1.52, 95% CI: 1.00–2.29). Similarly, precipitation showed no significant impact on influenza A. However, for influenza B, precipitation within the 0–10 mm range exhibited a rise-then-fall risk pattern, peaking at 5 mm (RR = 4.25, 95% CI: 2.11–8.58). These findings suggest that influenza A is predominantly driven by specific temperature bands, whereas influenza B transmission is more closely linked to cool-wet meteorological regimes.

Solar radiation and the UV index also manifested distinct subtype affinities. For influenza A, neither low (75 W/m²) nor high (305 W/m²) solar radiation yielded significant cumulative risks. In contrast, low solar radiation (75 W/m²) posed a significant risk for influenza B, peaking at 9.6 days (RR = 1.16, 95% CI: 1.08–1.23). For the UV index (reference = 7), influenza A only showed significant risk exactly at an index of 7 (RR = 1.00, 95% CI: 1.00–1.00), whereas influenza B risk progressively decreased within the UV index range of 2.6–7, peaking at 2.6 (RR = 2.49, 95% CI: 1.04–5.96).

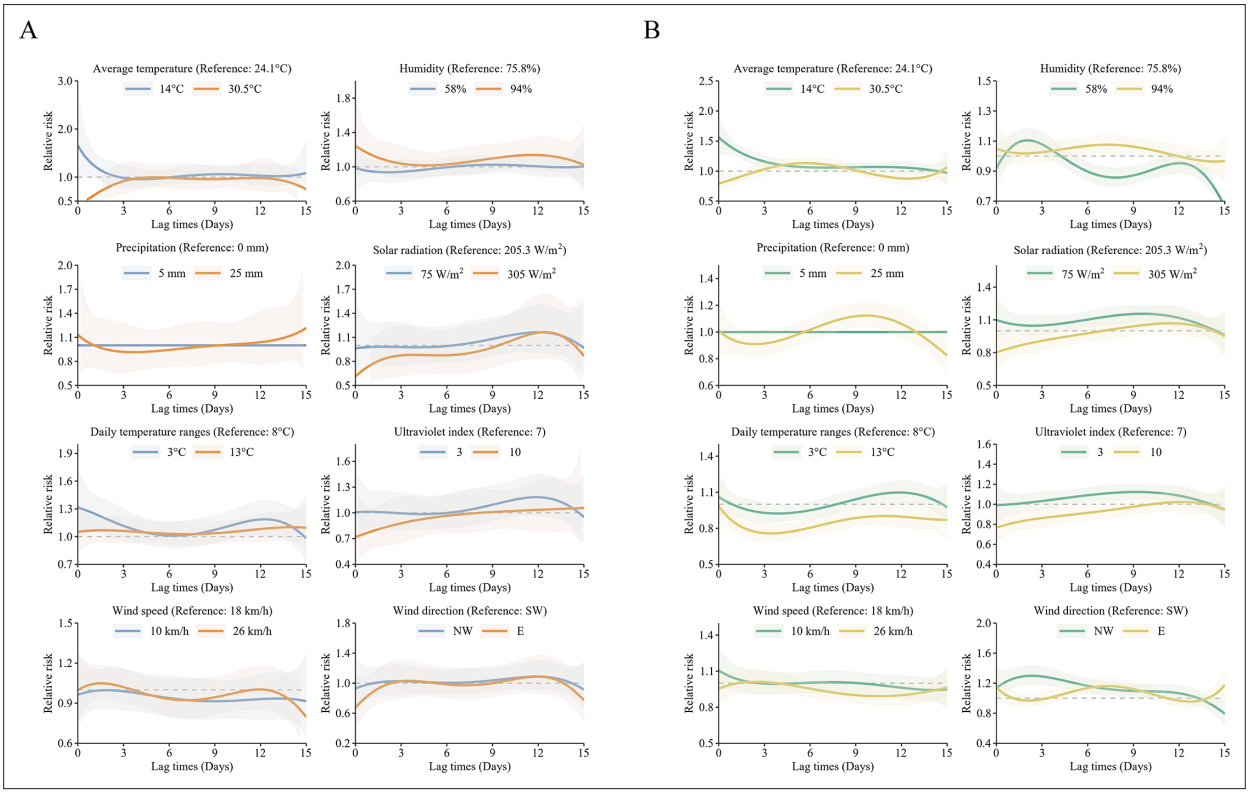


Figure 5. Lagged temporal effects of extreme meteorological conditions on influenza A and B positivity rates. This figure visualizes the distribution of relative risks over a 15-day lag period following exposure to extreme meteorological conditions, derived from distributed lag non-linear model (DLNM). The analysis compares the temporal lag effects at the extreme high (95th percentile) and extreme low (5th percentile) values of selected weather variables for influenza A (**A**) and influenza B (**B**). In each subplot, the x-axis represents the specific lag days (from days 0 to 15), and the y-axis shows the relative risk compared to the median reference condition. Solid orange lines (for influenza A) and yellow lines (for influenza B) represent extreme high-value conditions (e.g. extreme heat), whereas solid blue lines (for influenza A) and green lines (for influenza B) represent extreme low-value conditions (e.g. extreme cold). The shaded areas surrounding the curves represent the 95% confidence intervals. These lag-response curves elucidate the dynamic delay between specific extreme weather events and the subsequent peaks in influenza test positivity.

The online version of this article includes the following figure supplement(s) for figure 5:

Figure supplement 1. Sensitivity analysis of the lagged temporal effects of extreme meteorological conditions on influenza A and B positivity rates upon removal of the weekly detection covariate.

Diurnal temperature range (DTR) was associated with both influenza subtypes, although the estimates for influenza A were imprecise. Relative to 8 °C, the cumulative risk of influenza A was highest at 1 °C (RR = 15.69, 95% CI: 1.42–173.00), while influenza B risk peaked at 5.5 °C (RR = 1.36, 95% CI: 1.09–1.69). Wind speed was not significantly associated with either influenza subtype. Wind direction was not associated with influenza A, but northerly, easterly, and westerly winds were associated with elevated influenza B risk, with the highest estimate observed for northerly winds (RR = 27.33, 95% CI: 3.33–224.41).

Lagged effects of extreme meteorological conditions on influenza positivity risks

We further investigated the dynamic temporal variations in influenza positivity risks under extreme meteorological conditions (evaluating the 5th and 95th percentiles of each variable) to characterize type-specific vulnerability windows (***Figure 5A and B***). Low temperature was associated with short-term increases in influenza A risk, with significant effects between lag 0.2 and 1 day and a peak at lag 0.2 days (RR = 1.55, 95% CI: 1.01–2.39). High temperature was not associated with influenza A. For influenza B, low temperature showed both immediate and delayed associations, with elevated risks at lag 0–5.6 days and 9.2–10.2 days. The strongest effect occurred at lag 0 days (RR = 1.56, 95% CI: 1.26–1.93). High temperature was also associated with influenza B risk at lag 4.2–7.2 days, peaking at lag 5.8 days (RR = 1.13, 95% CI: 1.04–1.23).

Under extreme humidity, neither very low (58%) nor very high (94%) humidity generated significant lagged risks for influenza A. However, for influenza B, low humidity posed a rapid risk (1.2–3 days, peak RR = 1.11 at day 2), while high humidity resulted in a delayed risk (5–9.2 days, peak RR = 1.08 at day 7.2). For precipitation, heavy rainfall (25 mm) uniquely impacted influenza B, inducing a lagged risk between 7.8 and 11 days (peak RR = 1.12 at day 9.6), with no significant effects observed for influenza A.

Low solar radiation and low UV index were consistently associated with delayed increases in influenza B risk. Low solar radiation increased influenza B risk from lag 4.6 to 12.6 days, peaking at lag 9.6 days (RR = 1.16, 95% CI: 1.08–1.23). Low UV index was associated with elevated influenza B risk from lag 5 to 11.8 days, peaking at lag 9.2 days (RR = 1.12, 95% CI: 1.05–1.20). No significant lagged associations were observed between these radiation-related variables and influenza A.

Extreme DTR showed limited but subtype-specific lagged effects. Low DTR was associated with influenza A risk at lag 0.2–2 days and 11–12.8 days, with peak estimates at lag 0.2 days (RR = 1.31, 95% CI: 1.02–1.66) and lag 12.2 days (RR = 1.19, 95% CI: 1.01–1.40). For influenza B, low DTR was associated with increased risk at lag 10.2–12.2 days, peaking at lag 11.8 days (RR = 1.10, 95% CI:

Table 1. Optimal LSTM hyperparameters for predicting influenza A and B positivity rates.

Hyperparameter category	The number of hyperparameter for predicting influenza A positivity rate	The number of hyperparameter for predicting influenza B positivity rate
Layers	1	2
Neurons in layer 1	256	512
Neurons in layer 2	-	256
Neurons in layer 3	-	-
Activation function	sigmoid	softplus
Batch size	8	8
Learning rate	0.0001	0.0001
Epochs	587	931

LSTM: Long short-term memory. Hyperparameters were optimized through grid search. This table presents the optimal LSTM network structure (number of layers and neurons), activation function, batch size, learning rate, and number of epochs for predicting both influenza A and B infections. The number of neurons refers to the count of LSTM cells in each layer; detailed LSTM architecture is provided in Appendix 1.

1.00–1.20). High DTR showed no significant association with either subtype. Wind-related effects were observed primarily for influenza B. Southwesterly winds were associated with increased influenza B risk from lag 0.4–10.4 days, peaking at lag 2.4 days (RR = 1.30, 95% CI: 1.17–1.44), while easterly winds were associated with risk from lag 5–9.8 days, peaking at lag 7.4 days (RR = 1.16, 95% CI: 1.07–1.26). Wind speed was not significantly associated with either influenza A or B.

Crucially, sensitivity analyses indicated that after removing the 'weekly detection volumes' covariate, the reconstructed DLNM models remained structurally highly consistent with the original fully adjusted models. This robust concordance confirms that the identified meteorological drivers represent intrinsic environmental triggers of viral transmission, independent of fluctuating surveillance intensity. Detailed methodological descriptions and results of the sensitivity analysis are presented in ***Figure 4—figure supplement 2***, ***Figure 5—figure supplement 1***, and Appendix 1.

Influenza prediction using LSTM network

Building upon the previous analysis of the impact of meteorological factors on influenza dynamics, this study further aimed to construct a multifactorial influenza prediction network utilizing LSTM. To optimize the LSTM network, we employed Bayesian optimization through the Hyperopt framework, which allowed us to fine-tune key hyperparameters. This optimization process led to the identification of the optimal configurations for influenza A and B prediction algorithm (***Table 1***).

For influenza A, the optimal LSTM network consisted of one hidden layer with 256 neurons, a learning rate of 0.0001, and a batch size of 8. In contrast, the influenza B LSTM network had two hidden layers with 512 and 256 neurons, a learning rate of 0.0001, and a batch size of 8. Leveraging these optimal hyperparameters, alongside a combination of meteorological data, weekly detection volumes, DOW, non-local population proportion, and mask-wearing stringency indices, we trained the LSTM network using historical data from 2018–2022. A 10% of the data were randomly selected as the validation set, while data from 2023 were subsequently used to test the network's performance. The loss curves plotted during training revealed the convergence behavior of the LSTM. As shown in ***Figure 6C and D***, the training loss and validation loss for both the influenza A and B LSTM networks decreased progressively with increasing epochs, initially displaying a steep decline followed by stabilization. This pattern indicates that the algorithms effectively learned the characteristics of the data and converged to a stable solution. Furthermore, the close alignment between training and validation loss curves suggests that both LSTM networks for influenza A and B performed well and exhibited minimal overfitting, demonstrating strong learning capabilities and generalizability.

Following adequate training, the LSTM networks were employed to predict influenza A and B positivity rates in Putian city for the year 2023. The results revealed a strong match between the predicted and observed values (***Figure 6A and B***). The LSTM networks accurately captured both the timing and magnitude of these non-linear epidemic surges, including the two outbreak peaks of influenza A during February–March and November–December of 2023, as well as the peak of influenza B in November–December, demonstrating good predictive performance. The predictive performance of the LSTM networks was quantitatively assessed using metrics such as MAE, RMSE, MAPE, and SMAPE. For influenza A, the MAE was 0.009, RMSE was 0.035, MAPE was 0.158, and SMAPE was 0.521; for influenza B, the MAE was 0.002, RMSE was 0.011, MAPE was 0.17, and SMAPE was 0.484 (***Table 2***). These findings underscore the precision of the LSTM in characterizing and predicting influenza trends.

To quantify the contributions of epidemiological and behavioral covariates to the network's handling of pandemic non-stationarity, we conducted a covariate-ablation sensitivity analysis (***Table 2***). The LSTM networks were retrained sequentially by removing mask-wearing stringency indices and weekly detection volumes, while maintaining identical Bayesian-optimized hyperparameters. For influenza A, the exclusion of mask-wearing stringency indices resulted in a pronounced degradation of predictive performance compared to the full model: MAE, MAPE, and SMAPE increased by 33.3, 56.3, and 62.4%, respectively. The removal of weekly detection volumes similarly degraded performance, increasing MAE by 22.2% and SMAPE by 51.8%. This systemic shift in absolute errors (MAE) and the amplification of relative errors (SMAPE) clearly indicate that LSTM actively leverages mask mandates and testing intensities to capture the relative trends of influenza A. For influenza B, the ablation of covariates revealed an even higher dependency on the surveillance context. Removing weekly detection volumes proved highly detrimental, triggering a 100% surge in MAE, an 18.2% increase in RMSE, and a drastic 199.6% spike in SMAPE. This confirms that testing volumes act as a critical constraint

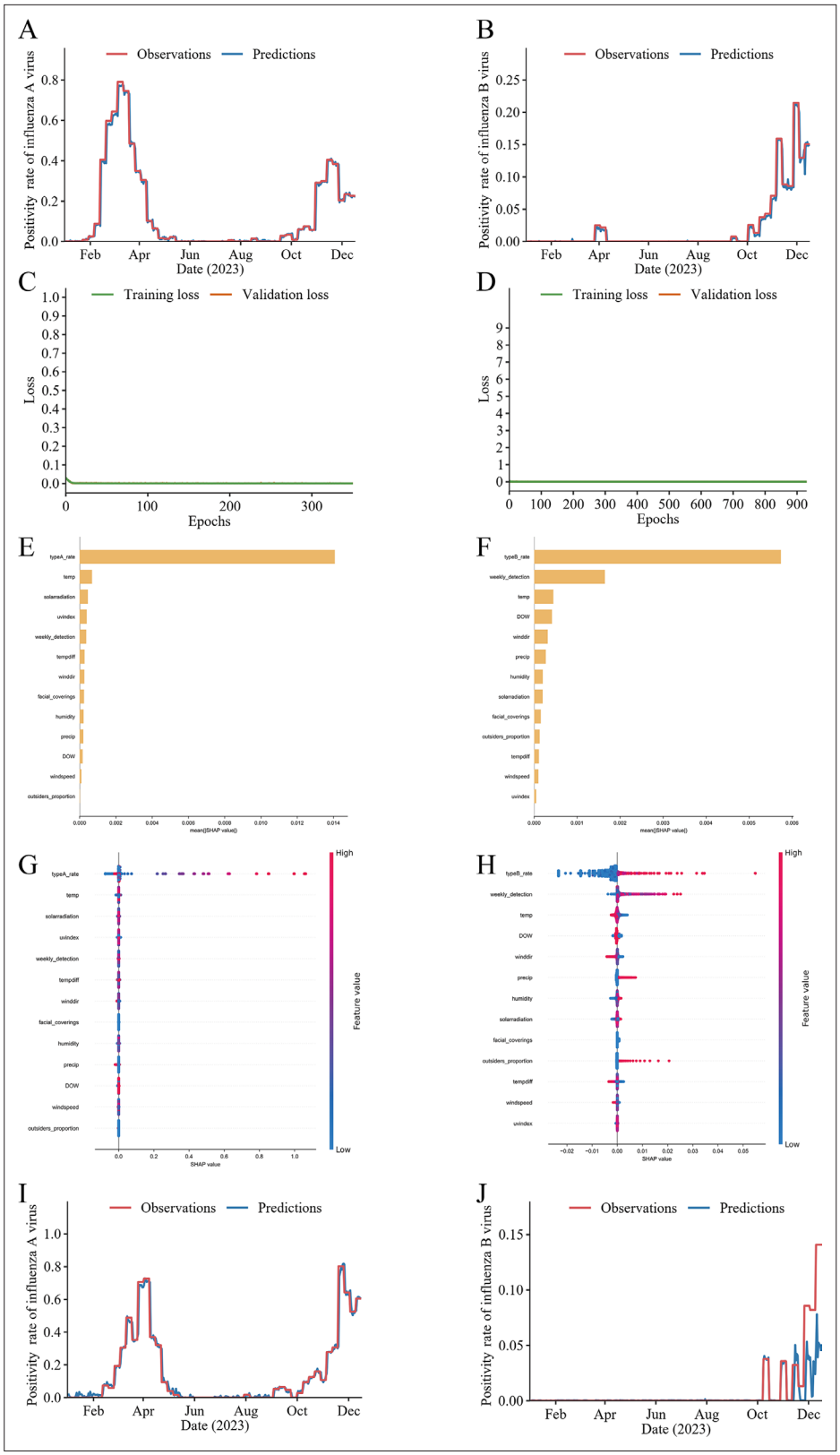


Figure 6. High-fidelity forecasting of the influenza positive rate using the LSTM network and SHAP interpretability. This comprehensive panel visualizes the predictive performance and feature importance of the LSTM model configured to forecast the influenza positivity rate. (**A, B**) Internal Out-of-Time validation (2023) comparing the LSTM-predicted positive rate (blue lines) against the observed positive rate (red lines) for Influenza A and B. The

Figure 6 continued on next page

Figure 6 continued

x-axis represents the date, and the y-axis reflects the positivity rates. This comparison visually demonstrates the network adeptly captures the non-linear viral rebound dynamics. (**C, D**) Training loss curves: These panels show the variation in loss values during the training of the LSTM network. The x-axis represents the number of training epochs, while the y-axis indicates the loss values. The smooth decline and stabilization of the loss curves indicate effective learning by the network, with no significant signs of overfitting. This is a crucial measure of the network's stability and effectiveness throughout the training process. (**E, F**) SHAP (SHapley Additive exPlanations) feature importance rankings. The bar charts quantify the mean absolute SHAP values, identifying historical subtype positivity rates (typeA_rate / typeB_rate), testing volume (weekly_detection), average temperature (temp), and mask-wearing stringency indices (facial_coverings) as paramount predictive drivers. The x-axis denotes the mean SHAP value, while the y-axis lists the variables. (**G, H**) SHAP summary plots illustrating the distribution of impacts for each feature on the model's output, where color indicates feature value (red = high, blue = low). This graph provides a visual understanding of how and to what extent different features influence the influenza predictions, offering critical insights into the network's decision-making process. (**I, J**) External spatial validation utilizing 2023 surveillance data from the neighboring city of Sanming, confirming the LSTM's robust generalizability in forecasting independent epidemic curves. The date is plotted on the x-axis, and the positivity rates are plotted on the y-axis.

The online version of this article includes the following figure supplement(s) for figure 6:

Figure supplement 1. Observed versus multivariate ARIMA-predicted weekly positivity rates for influenza A and B in Putain City across the 2023 out-of-time validation period.

Figure supplement 2. Observed versus XGBoost-predicted weekly positivity rates for influenza A and B in Putian City across the 2023 out-of-time validation period.

variable for forecasting influenza B. Omitting mask-wearing indices similarly increased MAE by 50% and SMAPE by 77.1%.

Notably, a divergence was observed in the MAPE metric for influenza B, which paradoxically decreased (by 20.1% and 29.4%) upon the removal of these two covariates. This phenomenon is a well-documented mathematical artifact of the MAPE calculation, which utilizes the true observed value as the denominator. Given the extremely low baseline positivity rate of influenza B during certain periods, infinitesimal absolute deviations are asymmetrically magnified into massive percentage errors, causing MAPE to become highly unstable and misleading. In contrast, SMAPE, which utilizes the mean of the predicted and actual values in the denominator, effectively circumvents this asymmetric 'low base-rate effect'. The profound deterioration of SMAPE across all ablated models corroborates the positive, stabilizing contribution of NPI and testing-volume covariates. These findings validate that our unified LSTM architecture does not merely rely on meteorological inputs, but effectively contextualizes epidemiological regime shifts to optimize forecasting precision.

To rigorously evaluate the spatial generalizability of the LSTM framework, we conducted an external validation utilizing concurrent 2023 surveillance and meteorological data from Sanming city. The network demonstrated robust transferability, achieving an MAE of 0.015, RMSE of 0.044, MAPE

Table 2. Sensitivity analysis of the LSTM forecasting network evaluating the impact of epidemiological covariate ablation on predictive performance.

Performance evaluation metric	Predicting influenza A positivity rate			Predicting influenza B positivity rate		
	Full model (All variables included)	Mask-wearing index removed	Weekly detection volumes removed	Full model (All variables included)	Mask-wearing index removed	Weekly detection volumes removed
MAE	0.009	0.012	0.011	0.002	0.003	0.004
RMSE	0.035	0.035	0.034	0.011	0.012	0.013
MAPE	0.158	0.247	0.203	0.170	0.135	0.120
SMAPE	0.521	0.846	0.791	0.484	0.857	1.450

Note: MAE (Mean Absolute Error) and RMSE (Root Mean Square Error) represent absolute predictive deviations. MAPE (Mean Absolute Percentage Error) and SMAPE (Symmetric Mean Absolute Percentage Error) assess relative accuracy. The full model incorporates meteorological variables alongside epidemiological covariates (mask-wearing stringency indices, weekly detection volumes, DOW, and migrant population proportion). Ablated models demonstrate the degradation in predictive accuracy upon the targeted removal of specific covariates, highlighting their necessity in adjusting for epidemiological non-stationarity.

of 0.176, and SMAPE of 0.674 for influenza A; and an MAE of 0.005, RMSE of 0.018, MAPE of 0.487, and SMAPE of 1.792 for influenza B. As illustrated in ***Figure 6I and J***, the predicted trajectories successfully mirrored the actual observed incidence patterns in the new region, confirming the framework's operational stability beyond its internal validation domain.

To benchmark the predictive value added by the LSTM architecture, we constructed two baseline models using identical training (2018–2022) and validation (2023) positivity-rate datasets, inclusive of all contextual covariates: the multivariate ARIMA model and the XGBoost gradient-boosting model. The LSTM model demonstrated decisive superiority over both baselines across all evaluated metrics (***Appendix 1—table 3***). For influenza A, the LSTM achieved an MAE of 0.009 and SMAPE of 0.521, compared with the ARIMA's MAE of 0.136 (SMAPE 1.212) and XGBoost's MAE of 0.138 (SMAPE 1.081), representing approximately 15-fold error reductions relative to both benchmarks. For influenza B, the LSTM yielded an MAE of 0.002 (SMAPE 0.484), compared with 0.049 (SMAPE 0.810) for ARIMA and 0.070 (SMAPE 0.737) for XGBoost, approximately 25- to 35-fold reductions. Notably, ARIMA and XGBoost produced errors of comparable magnitude despite their disparate assumptions regarding linearity, suggesting that the LSTM's advantage derives not from non-linear modeling capacity per se, but from architectural properties specific to recurrent sequential processing.

Beyond global error metrics, visual comparison of forecasting trajectories (***Figure 6***; ***Figure 6—figure supplements 1 and 2***) reveals critical behavioral disparities. For influenza A, both the linear ARIMA model and the non-linear XGBoost model generated conservatively smoothed forecasts that entirely failed to capture the sudden, explosive peaks of the 2023 post-restriction viral rebound. For influenza B, ARIMA produced continuous spurious fluctuations during non-epidemic periods (when true positivity was near zero), likely overreacting to covariate variations, while XGBoost generated largely flat trajectories that missed the mid-year outbreak peak. In stark contrast, the LSTM's recurrent gating mechanisms successfully filtered out covariate noise during low-transmission periods while accurately tracking extreme epidemiological phase transitions, demonstrating a qualitative advantage in handling non-stationary regime shifts.

To open the computational 'black box' and elucidate the predictive logic driving the LSTM networks, we conducted a SHAP analysis. The global feature importance rankings (***Figure 6E and F***) revealed that the endogenous autoregressive component, historical influenza positivity rates, was the overwhelmingly dominant predictor for both subtypes. Beyond historical rates, the networks exhibited distinct, subtype-specific dependencies on meteorological and surveillance covariates. For influenza A (***Figure 6E***), ambient temperature emerged as the second most impactful predictor, followed by solar radiation, the UV index, and weekly detection volumes. Conversely, for influenza B (***Figure 6F***), weekly detection volumes emerged as the second most critical feature overall, followed by temperature and DOW. Crucially, these feature hierarchies confirm that the deep learning algorithm did not merely memorize spurious, weather-driven testing surges. Rather, it autonomously recognized the profound confounding impact of surveillance intensity bias. By assigning high predictive weight to weekly detection volumes and behavioral covariates, the LSTM network dynamically calibrated its epidemiological forecasts, effectively contextualizing intrinsic meteorological triggers within the rapidly shifting pandemic-era surveillance landscape.

Discussion

This study characterized the population-level relationship between meteorological conditions and influenza activity in Putian, a representative non-megacity subtropical coastal city in southeastern China, using time-series (2018–2023) of integrated surveillance and meteorological data. As an ecological investigation, it was designed to describe environment–influenza associations and to develop an interpretable short-horizon forecasting framework, rather than to infer individual-level causality. By combining DLNM-based association analysis with Bayesian-optimized, subtype-specific LSTM forecasting, and by explicitly incorporating covariates related to testing intensity and NPI stringency, this work addresses three persistent gaps in subtropical influenza research: the isolated application of inferential and predictive tools, the limited use of principled hyperparameter optimization in LSTM-based influenza models, and the incomplete accommodation of pandemic-era non-stationarity in climate-sensitive forecasting. The findings deliver subtype-specific insights into meteorology–influenza coupling and establish a methodologically rigorous foundation for future operational surveillance systems in subtropical settings.

Unlike temperate regions, subtropical areas exhibit year-round influenza circulation characterized by considerable epidemiological heterogeneity. A Hong Kong-based study reported that local influenza A typically peaks between January and April, whereas influenza B peaks from May to July (***Kim et al., 2020***). A study from Okinawa, Japan, identified analogous trends, with influenza A peaking from December to March and influenza B from April to June (***Iha et al., 2016***). Our 6-year dataset revealed that both viruses circulated year-round in Putian, with peaks concentrated in winter-spring and additional summer activity. Influenza A predominantly peaked between December and April, whereas influenza B outbreaks frequently occurred outside this conventional window, suggesting that extreme weather fluctuations, rather than mere seasonality, act as primary outbreak triggers (***Dave and Lee, 2019***). The infection burden was concentrated among individuals under 18 years (73.12%), with male predominance in childhood and female predominance in adulthood, consistent with heightened exposure risk in school-aged children through close-contact group settings (***Jackson et al., 2013***). Notably, the significant suppression of influenza positivity observed in 2020 likely reflects the impact of strict NPIs and altered population contact networks implemented during the early COVID-19 period. Collectively, these patterns motivated the subsequent quantitative characterization of subtype-specific meteorological drivers.

The meteorology-influenza relationship remains a critical question in environmental epidemiology, with ambient temperature, relative humidity, precipitation, DTR, and wind parameters known to modulate virion survival, host respiratory susceptibility, and population contact behaviors (***Polozov et al., 2008***; ***Sooryanarain and Elankumaran, 2015***). Consistent with the intrinsic non-linearity of these relationships, Pearson correlation coefficients between influenza positivity rates and meteorological variables in our cohort (***Figure 4—figure supplement 1***) were uniformly weak (all $r<0.3$), reinforcing the inadequacy of linear approaches. DLNM cross-basis analysis, in contrast, revealed clear non-linear exposure–response and lag–response structures that differed markedly between subtypes.

Cumulative exposure–response patterns diverged notably between subtypes. Influenza A positivity was associated with warmer conditions between 15°C and 24°C, with the cumulative relative risk peaking at 15 °C (RR = 3.01, 95% CI: 1.03–8.78). Conversely, the cumulative risk for influenza B progressively declined from 8°C to 24°C, peaking sharply at the colder threshold of 8 °C (RR = 26.12, 95% CI: 7.79–87.61). This distinct thermal preference contrasts with findings from Shanghai, where both subtypes showed maximal cumulative risk near 1.4 °C (***Zhang et al., 2020***). Enhanced transmission at lower temperatures is broadly consistent with greater envelope stability in cold air (***Lowen and Steel, 2014***), impaired mucociliary clearance (***Eccles, 2002***), and increased indoor confinement facilitating aerosol exchange (***Liao et al., 2005***). Conversely, elevated temperatures can also inadvertently increase transmission through reduced outdoor activity and greater reliance on closed-cycle air conditioning (***Lofgren et al., 2007***). The subtype-specific divergence may further reflect fundamental virological and immunological differences, including the relative stability of influenza A glycoproteins under warmer, humid conditions and a structural preference of influenza B for cooler, wetter environments (***Marr et al., 2019***). Immunologically, high ambient temperatures can impair adaptive immune responses, potentially facilitating influenza A infection (***Moriyama and Ichinohe, 2019***), whereas cooler settings preserve innate responses, aiding the control of influenza B (***Lofgren et al., 2007***). Complex host-environment interactions may also suppress specific immune components, such as immunoglobulin M, inadvertently creating a permissive environment for influenza B transmission (***Xu et al., 2017***).

High relative humidity has been widely documented to enhance influenza risk by promoting viral transmission, as corroborated by multi-city studies in the Sichuan Basin (***Zhou et al., 2022***). Cellular models indicate that humidity influences respiratory droplets evaporation, altering solute concentration and influenza A virus viability (***Yang et al., 2012***), while findings from Arizona indicate increased influenza risk during the rainy season (***Soebiyanto et al., 2010***). Consistent with these observations, our results showed that the cumulative risk for influenza B gradually escalated within a high-humidity range of 76–91%, peaking at 91% (RR = 1.52, 95% CI: 1.00–2.29). The cumulative risk for influenza B also progressively increased with decreasing solar radiation, peaking at 10 W/m² (RR = 8.66, 95% CI: 2.14–35.00), suggesting that insufficient sunlight exposure may specifically promote influenza B transmission, a pattern absent for influenza A. Extensive literature supports that adequate solar radiation and subsequent UV exposure reduce airborne influenza viability and support host vitamin D synthesis and immune modulation (***Cannell et al., 2006***; ***Sagripanti and Lytle, 2007***). However, contradictory

effects of UV and solar radiation on subtypes A and B reported in pediatric Vitamin D supplementation trials (***Urashima et al., 2010***) necessitate further mechanistic investigation into the interplay between radiant energy and specific viral lineages. Physical mechanical factors such as DTR and wind metrics also demonstrated nuanced impacts. Prior research indicates that DTR serves as a critical risk factor, with severe fluctuations (11–14°C) increasing infection risk by compromising respiratory epithelial barriers and modulating inflammatory mucus production (***Park et al., 2020***). Our analysis partially corroborated this, showing that narrower DTRs (1–8°C) heightened cumulative risks for both subtypes, potentially reflecting sustained susceptibility under chronically stable yet suboptimal thermal conditions (***Cheng et al., 2014***). Wind direction influenced only influenza B, while wind speed exerted no significant cumulative effect on either subtype, consistent with literature deeming wind speed a noncritical epidemiological variable (***Ianevski et al., 2019***).

Lag structures further distinguished the subtypes within the 15-day post-exposure window. Low average temperatures of 14 °C were associated with a rapid 1 day lagged effect for influenza A, but induced a prolonged, bimodal lagged risk for influenza B (0–5 days and 9–10 days). Conversely, high averages of 30.5 °C conferred a delayed risk for influenza B only at 4.2–7.2 days, a timeline closely mirroring Wang et al.'s reported 7.3 day lag effect for influenza B at 28 °C (***Wang et al., 2023***). Extreme high precipitation (25 mm) induced a delayed risk for influenza B at 7.8–11 days, consistent with meta-analyses confirming that extreme rainfall events in the subtropics trigger lagged infectious disease surges (***Aune et al., 2021***). Notably, our data revealed a profound paradox regarding diurnal temperature variation. While a high DTR (13 °C) exhibited significant cumulative relative risk, it produced no distinct lagged risk peaks. In stark contrast, a low DTR (3 °C) triggered pronounced lagged risks for both influenza A (0–2 and 11–12.8 days) and influenza B (10.2–12.2 days). The intrinsic physiological mechanisms underlying this divergence, where broad temperature swings drive overall cumulative burden, but narrow temperature bands dictate specific temporal outbreak windows, warrant deeper investigation. Overall, these results identify the 1-to-2-week window following specific meteorological anomalies as an operationally meaningful period for anticipatory action, while emphasizing that thresholds and lag structures require local calibration. The fixed 15-day maximum lag adopted here, although consistent with prior literature, may not capture all longer-range environmental effects; systematic sensitivity analyses varying the lag window would be a valuable extension.

A central methodological feature of our framework is the deliberate division of labor between the DLNM and LSTM components. Multivariate DLNMs efficiently expose transparent, lag-resolved main-effect surfaces for each meteorological variable, but cannot accommodate high-dimensional interactions among meteorological, autoregressive, and socio-behavioral factors without parameter inflation and severe multicollinearity. The LSTM stage was therefore not intended to replace DLNM inference, but to complement it by learning joint non-linear structure across concurrent covariates. Furthermore, extensive environmental inputs inevitably introduce severe collinearity, such as the strongly correlated solar radiation and UV index. While traditional multivariate models are highly vulnerable to such overlapping variances, the recurrent, weighted representation learned by the LSTM is comparatively tolerant of such redundancy, allowing broader covariate integration than in previous efforts (***Zhu et al., 2022***). Crucially, this LSTM stage explicitly incorporates weekly detection volumes, mask-wearing stringency indices, non-local population proportion, and DOW effects alongside meteorological inputs, and uses influenza positivity rates rather than absolute case counts as the modeling endpoint. Together, these design choices are intended to mitigate, rather than fully eliminate, the surveillance-related biases that can distort count-based forecasting during periods of fluctuating testing intensity.

To maximize this architectural advantage and ensure mathematical robustness, we conducted refined hyperparameter optimizations using the Hyperopt framework (***Li et al., 2023***). By employing a Bayesian strategy based on Tree-based Parzen Estimators (TPE; ***Hanifi et al., 2022***), our approach efficiently navigated the complex, high-dimensional hyperparameter space. Unlike traditional heuristic or manual grid search methods, this Bayesian approach systematically optimized the learning rate, network depth, and activation functions to strike an optimal balance between model fitting capacity and the prevention of overfitting (***Hanifi et al., 2024***). To our knowledge, this represents a novel application of Hyperopt-based tuning for an epidemiological LSTM forecasting network, ensuring that our algorithm's superior performance is grounded in mathematically rigorous architectural configuration rather than arbitrary parameter selection.

The optimized LSTM algorithm demonstrated strong predictive performance, with loss curves for both subtype-specific networks exhibiting favorable convergence, consistent with prior work (***Du et al., 2023***). To rigorously isolate the predictive contribution attributable to the LSTM's recurrent architecture, we benchmarked it against two covariate-matched baselines representing distinct methodological paradigms: the multivariate ARIMA model and the XGBoost gradient-boosting ensemble. This design controls simultaneously for linearity (ARIMA→LSTM contrast) and for non-linearity without recurrent memory (XGBoost→LSTM contrast), allowing us to attribute observed performance gains to specific architectural inductive biases rather than to informational asymmetry or model non-linearity in general. The LSTM substantially outperformed both baselines on the 2023 validation window. For influenza A, it achieved an MAE of 0.009, compared with 0.136 for ARIMA and 0.138 for XGBoost, approximately 15-fold reductions. For influenza B, the LSTM yielded an MAE of 0.002, vs 0.049 for ARIMA and 0.070 for XGBoost, 25- to 35-fold reductions. Critically, ARIMA and XGBoost produced errors of comparable magnitude despite their disparate assumptions regarding linearity, and both systematically under-predicted the explosive 2023 post-NPI rebound for influenza A while generating flat or spurious trajectories for influenza B (***Figure 6—figure supplements 1 and 2***). That parallel failure indicates the LSTM's advantage under pandemic-era non-stationarity derives not from non-linearity per se, but from four architectural properties intrinsic to recurrent gated networks yet absent in tree ensembles: (i) threshold-based, sequence-insensitive splits fail to encode present–past dynamics; (ii) XGBoost lacks forget–input–output gates to propagate and re-weight historical states across time lags; (iii) gradient-boosted trees assume near-independence, contradicting the pronounced temporal autocorrelation in epidemiological time series; (iv) ARIMA and LSTM are sequential models with different functional forms, whereas XGBoost is non-sequential. The joint failure of ARIMA and XGBoost, despite differing non-linear treatment, isolates recurrent sequential memory as the critical architectural feature for forecasting under non-stationarity. We emphasize that this interpretation applies specifically to the present non-stationary influenza forecasting task and does not constitute a general dismissal of gradient-boosted ensembles, which retain state-of-the-art performance across many structured prediction domains. Rather, it highlights that for surveillance time series exhibiting pronounced temporal dependencies and abrupt regime shifts, such as the 2023 post-NPI rebound, explicit sequential memory becomes functionally essential. This mechanistic reading, together with the LSTM's comparative edge over previously reported ARIMA-based (***Li et al., 2024***) and LSTM-based influenza prediction models (***Zhu et al., 2022***), positions our framework as a substantive methodological advance in predictive modeling for climate-sensitive diseases.

External evaluation using independent Sanming CDC data yielded MAE values of 0.0152 for influenza A and 0.0054 for influenza B, providing preliminary evidence of model portability within the southeastern Chinese subtropical context, although this rests on a single external city and should not be interpreted as broad spatial generalizability. SHAP analysis (***Figure 6E and F***) further enhanced LSTM interpretability, identifying historical influenza positivity rates, weekly detection volumes, temperature, and solar radiation as the most highly weighted predictive features, a pattern consistent with temperature- and radiation-mediated influenza virus inactivation mechanisms (***Sagripanti and Lytle, 2007***), and corroborating the network's autonomous capacity to calibrate predictions against fluctuating surveillance intensities. While SHAP's underlying assumption of feature independence may not fully capture temporal interactions within the LSTM network, treating each meteorological variable's lagged sequence as a unified feature partially preserves temporal dependencies. Future work could explore temporal SHAP extensions or alternative interpretability methods designed for sequential data.

Beyond its forecasting accuracy, the DLNM–LSTM framework offers several potential contributions for subtropical public health practice. By delineating subtype-specific meteorological thresholds, characterizing their lagged temporal effects, and dynamically integrating behavioral and surveillance covariates, it enables the anticipation of high-risk transmission windows 1–2 weeks in advance of clinical signal emergence. Such foresight can inform pre-emptive vaccination campaigns for vulnerable populations, optimize antiviral stockpiling and healthcare resource allocation, and intensify surveillance during predicted risk windows to accelerate case detection and containment. The underlying dual-stage architecture is also methodologically extensible to other climate-sensitive infectious diseases, including dengue and hand-foot-and-mouth disease, in which non-linear environmental coupling and lagged effects similarly dominate transmission dynamics. In a warming climate, where influenza

seasonality may shift in response to evolving weather regimes, the framework's modular design permits periodic recalibration to maintain predictive validity. Importantly, however, this framework should be regarded as a methodological foundation for future operational developments rather than as a deployable early-warning system: routine use in public health practice would require prospective recalibration, integration with operational surveillance infrastructure, and additional validation beyond the scope of the present study.

This study has several limitations that should be considered when interpreting the findings. Firstly, our primary analysis relies on influenza surveillance data collected from seven sentinel hospitals in Putian, which inherently captures only a fraction of all influenza cases occurring in the broader community. Although employing positivity rates as the primary outcome substantially mitigates the surveillance bias inherent in count-based analyses, residual selection effects may persist if testing patterns shift differentially across demographic subgroups with systematically divergent positivity profiles. Our inclusion of weekly detection volumes as an explicit LSTM covariate, complemented by parallel DLNM sensitivity analyses validating the independence of meteorological effects from testing volumes, was designed to characterize and partially account for this residual bias, but cannot fully eliminate it. Consequently, positivity rates likely underestimate the true burden of community influenza infection. Although the surveillance infrastructure in Putian remained uniform and uninterrupted throughout the 2018–2023 timeline, mitigating measurement bias and temporal confounding risks, the transferability of our findings to other global subtropical regions with different socioeconomic structures, healthcare systems, or population behaviors requires cautious, region-specific calibration. Accordingly, our framework should be strictly interpreted as a high-fidelity tool designed to forecast the observable public health surveillance signal, which is the most operationally relevant target for public health agencies, rather than for estimating unobserved, absolute community disease burden. Future research should aim to integrate broader data sources, such as community-level surveillance or syndromic data streams where available, and pursue multi-center studies across diverse global subtropical regions to disentangle core meteorological determinants that transcend geographical boundaries from those requiring local adaptation.

Secondly, the interpretation of our framework's predictive performance during the 2023 validation period requires careful epidemiological and methodological contextualization. The year of 2023 represented an anomalous, post-restriction 'rebound' period characterized by rapid NPI relaxation and the release of accumulated population-level immunity debt, resulting in an atypical influenza surge that exceeded pre-pandemic peaks. The framework's high accuracy across this period should therefore be interpreted as evidence of algorithmic agility and adaptive capacity during a highly volatile transitional phase, rather than as definitive proof of long-term predictive validity under a stabilized post-2024 epidemiological regime. Methodologically, while our strict chronological OOT data partitioning prevented temporal information leakage, a critical requirement for LSTM integrity, the reliance on a single, fixed chronological split point (December 31, 2022) intrinsically limits our evaluation to one specific structural break. This fixed-split approach may not exhaustively probe the DLNM–LSTM framework's resilience against all forms of future epidemiological non-stationarity. Consequently, naive extrapolation of the reported 2023 performance metrics to future surveillance years should be avoided absent prospective recalibration. Future studies should consider employing expanding-window or rolling-origin cross-validation frameworks to provide a more continuous characterization of algorithmic robustness. Continuous integration of accumulating 2024 and 2025 data, combined with adaptive learning architectures capable of detecting regime shifts in real time, will be essential before any operational deployment of this, or similar forecasting frameworks, for routine public health surveillance.

Thirdly, beyond the surveillance coverage limitation noted above, the aggregate-level nature of the available data further constrained our ability to adjust for individual-level confounders, including personal vaccination status, detailed comorbidities, healthcare-seeking behavior, and socioeconomic status. This limitation was particularly exacerbated by the profound epidemiological disruptions during the COVID-19 pandemic, where raw numbers of confirmed cases became heavily confounded by surveillance intensity (i.e. fluctuating testing volumes) rather than solely reflecting underlying viral transmission. Our adoption of influenza positivity rates as the primary modeling endpoint and incorporation of weekly detection volumes, face-covering stringency, and non-local population proportion as dynamic covariates rigorously mitigated these aggregate-level biases and linked our methodological

design directly to the forecasting outcomes. As unequivocally demonstrated by our covariate-ablation sensitivity analyses (***Table 2***), failing to account for mask mandates and testing volumes leads to severe, mathematically predictable deviations in absolute forecasting accuracy. Furthermore, while daily case counts in a single city can occasionally be small, sporadic, and driven by external importations, our incorporation of non-local population proportion effectively adjusted for these localized importation risks. Consequently, although our findings characterize population-level associations between meteorological factors and influenza activity, they should not be interpreted as evidence of micro-level causal mechanisms at the individual patient level. Ultimately, the DLNM–LSTM framework's robust performance across the non-stationary transition out of NPI policies highlights the absolute necessity of integrating behavioral and virological baseline metrics into future climate-driven predictive surveillance systems.

Fourthly, our benchmarking strategy was deliberately structured as a paradigmatically layered three-way comparison (linear-sequential ARIMA, non-linear non-sequential XGBoost, non-linear sequential LSTM) rather than an exhaustive algorithmic survey. This design prioritized methodological clarity, isolating the contribution of recurrent sequential inductive bias over horizontal coverage. Nonetheless, our evaluation does not extend to Transformer-based attention architectures, nor to hybrid ensemble strategies (e.g. LSTM–XGBoost stacking or multi-model Bayesian model averaging) that may offer complementary strengths. Systematic benchmarking against these emerging architectures, alongside prospective recalibration on additional subtropical surveillance streams and operational stress-testing under real-time data latency, will be essential next steps as the framework evolves toward operational deployment.

Finally, it is imperative to explicitly delineate the generalizability of our forecasting framework within geographic and epidemiological contexts. The successful external validation in Sanming provides direct empirical evidence that our LSTM architecture and the identified meteorological thresholds are robustly transferable across the subtropical southeastern Chinese context, which shares comparable monsoon climates, year-round influenza circulation, and standardized public health intervention frameworks. However, we explicitly define the boundaries of this transferability. The specific meteorological coefficients, lag structure, and predictive parameters derived in this study should not be indiscriminately extrapolated to other regions, even within subtropical latitudes, where climatic regimes (e.g. non-monsoon subtropics) or public health contexts (e.g. higher baseline influenza vaccination coverage or distinct NPI strategies) differ substantially. In such disparate settings, directly applying our pretrained model weights may yield substantial systematic biases. While the underlying dual-modeling architecture remains methodologically transferable, its parameterization requires rigorous local recalibration using region-specific surveillance data. Acknowledging these boundaries ensures that the framework can be deployed safely and appropriately to realize targeted, climate-sensitive infectious disease forecasts.

In conclusion, this study elucidates the distinct, non-linear meteorological drivers of influenza A and B transmission in a subtropical Chinese urban setting through an integrated dual-stage DLNM–LSTM framework. By adopting influenza positivity rates as the primary outcome and integrating socio-behavioral covariates, including mask-wearing stringency indices, weekly detection volumes, and non-local population proportion indicators, our approach mitigates surveillance-related biases and accommodates pandemic-era non-stationarity. It shows lower forecast errors than covariate-matched baselines, both linear autoregressive (ARIMA) and non-linear gradient-boosting (XGBoost), and provides preliminary evidence of portability within southeastern subtropical China, combining the interpretability of distributed lag modeling with the flexibility and sequential inductive bias of deep learning. The framework offers an interpretable, climate-informed methodological foundation for future operational surveillance developments in subtropical settings.

Acknowledgements

We acknowledge the invaluable contribution of the Putian CDC and its affiliated sentinel hospitals to this study. We express our deep respect and gratitude to the dedication of all participating healthcare workers, disease prevention personnel, clinical medical staff, and laboratory technicians involved in case screening, data recording, sample collection, and laboratory testing for their significant efforts and contributions to this study. The study was financially supported by a grant from Guangdong Provincial Science and Technology Program (grant number: 2022B1111020003) awarded to Xi-Wen

Jiang. Our study was carried out with complete autonomy, with the assurance that the funding bodies had no influence at any stage of the research process.

Additional information

Competing interests

Jin-Lin Tan, Zhe-Qiang Xue, Jun-Ju Huang, Ze-Fan Ruan, Xi-Wen Jiang: Was affiliated with DAAN Gene Co, Ltd; no other competing interests to declare. Zhi-Shen Wu: Affiliated with DAAN Gene Co, Ltd; no other competing interests to declare. The other authors declare that no competing interests exist.

Funding

Funder	Grant reference number	Author
Guangdong Provincial Department of Science and Technology	2022B1111020003	Xi-Wen Jiang

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Long Xie, Conceptualization, Formal analysis, Validation, Visualization, Methodology, Writing – original draft, Writing – review and editing; Meng-Jie Zhang, Data curation, Formal analysis, Supervision, Validation, Investigation, Methodology; Jin-Lin Tan, Data curation, Software, Formal analysis, Validation, Visualization, Methodology, Writing – original draft; Yi-Xin Ling, Data curation, Formal analysis, Supervision, Validation, Investigation; Zhe-Qiang Xue, Formal analysis, Validation; Zhi-Shen Wu, Data curation, Software, Formal analysis, Validation, Visualization, Methodology; Jun-Ju Huang, Jian-Ling Chen, Data curation, Investigation, Visualization; Ze-Fan Ruan, Software, Investigation, Visualization; Jing Qian, Data curation, Software, Investigation, Visualization; Hai-Yong Pan, Data curation, Software, Investigation; Xiao Han, Conceptualization, Resources, Formal analysis, Validation, Methodology, Project administration; Sheng Xiong, Conceptualization, Resources, Formal analysis, Methodology, Writing – review and editing; Long-Mei Ling, Conceptualization, Resources, Supervision, Funding acquisition, Project administration, Writing – review and editing; Xi-Wen Jiang, Conceptualization, Resources, Supervision, Funding acquisition, Methodology, Project administration, Writing – review and editing

Author ORCIDs

Long Xie https://orcid.org/0000-0002-2138-888X
Xi-Wen Jiang https://orcid.org/0009-0002-7977-4510

Ethics

This study was institutionally approved by the Research Ethics Committee of the Putian Centre for Disease Control and Prevention (Putian CDC, approval number: Ethics Review 2020-003), Putian, and the Medicine and Biological Engineering Technology Research Centre of the Ministry of Health (MBETRC), Guangzhou, China. All procedures complied with the later amendments of the 1964 Declaration of Helsinki and relevant ethical guidelines for biomedical research. A major methodological strength of this study lies in its robust, uninterrupted longitudinal data collection framework spanning January 1, 2018, to December 31, 2023. While the analytical timeline encompasses a "retrospective" phase (January 1, 2018 - October 13, 2020) prior to formal ethical approval, and a "prospective" phase thereafter, we emphasize that this distinction represents a purely administrative demarcation regarding the timing of ethical approval. It does not reflect any shift in demographic cohorts, sentinel hospital locations, or data collection methodologies. Importantly, the historical data (2018-2020) were not subjected to the recall biases or misclassification risks typical of traditional retrospective chart reviews. Rather, they were systematically extracted from a continuously operating, highly standardized public health sentinel surveillance network. From the inception of data collection through the end of

2023, the local CDC maintained absolute uniformity in clinical influenza-like illness (ILI) definitions, nasopharyngeal swabbing procedures, and real-time reverse transcription polymerase chain reaction (RT-PCR) diagnostic assays. Consequently, the pre-2020 data possess the high-fidelity characteristics of a strict prospective cohort, ensuring unparalleled longitudinal consistency and mitigating temporal measurement bias across the entire pre-pandemic, pandemic, and post-restriction timeline. Rigorous procedures were employed to ensure patient confidentiality and data privacy. Influenza surveillance data were fully anonymized prior to being accessed by the research team. The dataset included only essential epidemiological information (such as onset date, gender, age, and residential district) necessary for the analysis, excluding any personal identifiable information. Additionally, individual-level data on vaccination status or detailed socioeconomic profiles were not routinely collected. Any potential linkage between coded data and original patient identifiers maintained by Putian CDC was strictly inaccessible to the researchers. Consequently, the Research Ethics Committee waived the requirement for individual informed consent. Throughout the data collection process, the diagnosis of all ILI cases, information and specimen collection, and influenza virus nucleic acid testing strictly followed prevailing national standards, including the Diagnostic Criteria for Influenza (source: National Health Commission of the People's Republic of China) and the National Influenza Surveillance Technical Specifications (source: China CDC). All participating medical professionals, CDC personnel, and laboratory technicians underwent professional training to ensure research quality. Meteorological and COVID-19 pandemic-related covariate data were sourced from publicly available databases, ensuring no privacy concerns. The anonymized data were securely managed, used solely for the stated research purposes, and were not disclosed to any third party.

Peer review material

Version of Record:
Reviewer #2 (Public review): https://doi.org/10.7554/eLife.107767.4.sa1
Author response https://doi.org/10.7554/eLife.107767.4.sa2

Reviewed preprint v3
Reviewer #2 (Public review): https://doi.org/10.7554/eLife.107767.3.sa1
Author response: https://doi.org/10.7554/eLife.107767.3.sa0
eLife Assessment https://doi.org/10.7554/eLife.107767.3.sa2

Reviewed preprint v2
Reviewer #2 (Public review): https://doi.org/10.7554/eLife.107767.2.sa1
Author response: https://doi.org/10.7554/eLife.107767.2.sa0
eLife Assessment https://doi.org/10.7554/eLife.107767.2.sa2

Reviewed preprint v1
Reviewer #1 (Public review): https://doi.org/10.7554/eLife.107767.1.sa2
Reviewer #2 (Public review): https://doi.org/10.7554/eLife.107767.1.sa1
Author response: https://doi.org/10.7554/eLife.107767.1.sa0
eLife Assessment https://doi.org/10.7554/eLife.107767.1.sa3

Additional files

Supplementary files

MDAR checklist

Data availability

The influenza surveillance data analysed in this study were collected and are held by the Putian Centre for Disease Control and Prevention (Putian CDC) under its statutory infectious disease surveillance mandate. Because these data contain pathogen test results linked to individual cases, they are classified as sensitive personal information under the Personal Information Protection Law of the People's Republic of China and are subject to restrictions on public sharing under Chinese scientific data management regulations. Public deposition in open-access repositories is therefore not permitted under these regulations, and this restriction was a condition of the ethical approval for this study. A

fully de-identified version of the dataset cannot be shared openly, because the data retain variables including age, sex, residential district, and date of consultation, the combination of which presents a residual re-identification risk at the scale of a single city. De-identified data are, however, available to qualified researchers for the purpose of verifying or reproducing the findings reported in this article. Requests should be submitted in writing to the corresponding author Long-Mei Lin at Putian CDC (ptcdc001@163.com) and should include the applicant's name and institutional affiliation, a description of the intended analysis, and evidence of ethical approval or exemption from the applicant's institution where applicable. Each request will be reviewed and must receive approval from the Research Ethics Committee of Putian CDC. Approved applicants will receive the data through secure, private channels under a data use agreement, rather than via public repositories. There are no other restrictions on institutional or geographic eligibility. Meteorological data are publicly available from Visual Crossing (https://www.visualcrossing.com/). Mask-wearing stringency data are available from Our World in Data (https://ourworldindata.org/coronavirus). All computational codes (R and Python scripts) used for data preprocessing, DLNM construction, LSTM hyperparameter optimization, the ARIMA and XGBoost baseline algorithm implementation, and performance evaluation have been deposited in a publicly accessible repository (https://github.com/Long-Xie/DLNM-LSTM-Influenza-Forecasting-Putian, copy archived at ***Xie, 2026***) under an open-source MIT license.

References

Absar N, Uddin N, Khandaker MU, Ullah H. 2022. The efficacy of deep learning based LSTM model in forecasting the outbreak of contagious diseases. *Infectious Disease Modelling* **7**:170–183. DOI: https://doi.org/10.1016/j.idm.2021.12.005, PMID: 34977438

Aune KT, Davis MF, Smith GS. 2021. Extreme precipitation events and infectious disease risk: A scoping review and framework for infectious respiratory viruses. *International Journal of Environmental Research and Public Health* **19**:165. DOI: https://doi.org/10.3390/ijerph19010165, PMID: 35010425

Bengio Y. 2012. Practical recommendations for gradient-based training of deep architectures. In: Montavon G, Orr GB, Müller KR (Eds). *Neural Networks: Tricks of the Trade*. Springer. p. 437–478. DOI: https://doi.org/10.1007/978-3-642-35289-8_26

Bloom-Feshbach K, Alonso WJ, Charu V, Tamerius J, Simonsen L, Miller MA, Viboud C. 2013. Latitudinal variations in seasonal activity of influenza and respiratory syncytial virus (RSV): a global comparative review. *PLOS ONE* **8**:e54445. DOI: https://doi.org/10.1371/journal.pone.0054445, PMID: 23457451

Cannell JJ, Vieth R, Umhau JC, Holick MF, Grant WB, Madronich S, Garland CF, Giovannucci E. 2006. Epidemic influenza and vitamin D. *Epidemiology and Infection* **134**:1129–1140. DOI: https://doi.org/10.1017/S0950268806007175, PMID: 16959053

Chen T, Guestrin C. 2016. XGBoost: a scalable tree boosting system. Association for Computing Machinery San Francisco, California, USA. DOI: https://doi.org/10.1145/2939672.2939785

Cheng J, Xu Z, Zhu R, Wang X, Jin L, Song J, Su H. 2014. Impact of diurnal temperature range on human health: a systematic review. *International Journal of Biometeorology* **58**:2011–2024. DOI: https://doi.org/10.1007/s00484-014-0797-5, PMID: 24535132

Chicco D, Warrens MJ, Jurman G. 2021. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science* **7**:e623. DOI: https://doi.org/10.7717/peerj-cs.623, PMID: 34307865

Dave K, Lee PC. 2019. Global geographical and temporal patterns of seasonal influenza and associated climatic factors. *Epidemiologic Reviews* **41**:51–68. DOI: https://doi.org/10.1093/epirev/mxz008, PMID: 31565734

Deyle ER, Maher MC, Hernandez RD, Basu S, Sugihara G. 2016. Global environmental drivers of influenza. *PNAS* **113**:13081–13086. DOI: https://doi.org/10.1073/pnas.1607747113, PMID: 27799563

Du X, Si L, Li P, Yun Z. 2023. A method for detecting the quality of cotton seeds based on an improved ResNet50 model. *PLOS ONE* **18**:e0273057. DOI: https://doi.org/10.1371/journal.pone.0273057, PMID: 36791128

Eccles R. 2002. An explanation for the seasonality of acute upper respiratory tract viral infections. *Acta Oto-Laryngologica* **122**:183–191. DOI: https://doi.org/10.1080/00016480252814207, PMID: 11936911

Gasparrini A, Armstrong B, Kenward MG. 2010. Distributed lag non-linear models. *Statistics in Medicine* **29**:2224–2234. DOI: https://doi.org/10.1002/sim.3940, PMID: 20812303

Gasparrini A. 2011. Distributed lag linear and non-linear models in R: the package dlnm. *Journal of Statistical Software* **43**:1–20 PMID: 22003319.

Groves HE, Piché-Renaud PP, Peci A, Farrar DS, Buckrell S, Bancej C, Sevenhuysen C, Campigotto A, Gubbay JB, Morris SK. 2021. The Impact of the COVID-19 pandemic on influenza, respiratory syncytial virus, and other seasonal respiratory virus circulation in Canada. *Infectious Diseases* **1**:100015. DOI: https://doi.org/10.1101/2021.04.15.21255591

Guo Q, Dong Z, Zeng W, Ma W, Zhao D, Sun X, Gong S, Xiao J, Li T, Hu W. 2019. The effects of meteorological factors on influenza among children in Guangzhou, China. *Influenza and Other Respiratory Viruses* **13**:166–175. DOI: https://doi.org/10.1111/irv.12617, PMID: 30407738

Hanifi S, Lotfian S, Zare-Behtash H, Cammarano A. 2022. Offshore wind power forecasting—a new hyperparameter optimisation algorithm for deep learning models. *Energies* **15**:6919. DOI: https://doi.org/10.3390/en15196919

Hanifi S, Cammarono A, Zare-Behtash H. 2024. Advanced hyperparameter optimization of deep learning models for wind power prediction. *Renewable Energy* **221**:119700. DOI: https://doi.org/10.1016/j.renene.2023.119700

Ianevski A, Zusinaite E, Shtaida N, Kallio-Kokko H, Valkonen M, Kantele A, Telling K, Lutsar I, Letjuka P, Metelitsa N, Oksenych V, Dumpis U, Vitkauskiene A, Stašaitis K, Öhrmalm C, Bondeson K, Bergqvist A, Cox RJ, Tenson T, Merits A, et al. 2019. Low temperature and low UV indexes correlated with peaks of influenza virus activity in northern Europe during 2010. *Viruses* **11**:207. DOI: https://doi.org/10.3390/v11030207, PMID: 30832226

Iha Y, Kinjo T, Parrott G, Higa F, Mori H, Fujita J. 2016. Comparative epidemiology of influenza A and B viral infection in a subtropical region: a 7-year surveillance in Okinawa, Japan. *BMC Infectious Diseases* **16**:650. DOI: https://doi.org/10.1186/s12879-016-1978-0, PMID: 27821090

Iuliano AD, Roguski KM, Chang HH, Muscatello DJ, Palekar R, Tempia S, Cohen C, Gran JM, Schanzer D, Cowling BJ, Wu P, Kyncl J, Ang LW, Park M, Redlberger-Fritz M, Yu H, Espenhain L, Krishnan A, Emukule G, van Asten L, et al. 2018. Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. *Lancet* **391**:1285–1300. DOI: https://doi.org/10.1016/S0140-6736(17)33293-2, PMID: 29248255

Jackson C, Vynnycky E, Hawker J, Olowokure B, Mangtani P. 2013. School closures and influenza: systematic review of epidemiological studies. *BMJ Open* **3**:e002149. DOI: https://doi.org/10.1136/bmjopen-2012-002149, PMID: 23447463

Jagtap AD, Karniadakis GE. 2022. How important are activation functions in regression and classification? a survey, performance comparison, and future directions. *arXiv*. https://arxiv.org/abs/2209.02681

Javanian M, Barary M, Ghebrehewet S, Koppolu V, Vasigala V, Ebrahimpour S. 2021. A brief review of influenza virus infection. *Journal of Medical Virology* **93**:4638–4646. DOI: https://doi.org/10.1002/jmv.26990, PMID: 33792930

Kim BI, Park O, Lee S. 2020. Comparison of influenza surveillance data from the Republic of Korea, selected northern hemisphere countries and Hong Kong Special Administrative Region SAR (China) from 2012 to 2017. *Western Pacific Surveillance and Response Journal* **11**:1–9. DOI: https://doi.org/10.5365/wpsar.2019.10.2.015, PMID: 33936854

Li Y, Reeves RM, Wang X, Bassat Q, Brooks WA, Cohen C, Moore DP, Nunes M, Rath B, Campbell H, Nair H, RSV Global Epidemiology Network, RESCEU investigators. 2019. Global patterns in monthly activity of influenza virus, respiratory syncytial virus, parainfluenza virus, and metapneumovirus: a systematic analysis. *The Lancet. Global Health* **7**:e1031–e1045. DOI: https://doi.org/10.1016/S2214-109X(19)30264-5, PMID: 31303294

Li Q, Zhang Z, Ma Z. 2023. Raman spectral pattern recognition of breast cancer: A machine learning strategy based on feature fusion and adaptive hyperparameter optimization. *Heliyon* **9**:e18148. DOI: https://doi.org/10.1016/j.heliyon.2023.e18148

Li G, Li Y, Han G, Jiang C, Geng M, Guo N, Wu W, Liu S, Xing Z, Han X, Li Q. 2024. Forecasting and analyzing influenza activity in Hebei Province, China, using a CNN-LSTM hybrid model. *BMC Public Health* **24**:2171. DOI: https://doi.org/10.1186/s12889-024-19590-8, PMID: 39135162

Liao CM, Chang CF, Liang HM. 2005. A probabilistic transmission dynamic model to assess indoor airborne infection risks. *Risk Analysis* **25**:1097–1107. DOI: https://doi.org/10.1111/j.1539-6924.2005.00663.x, PMID: 16297217

Liu J, Chen E, Zhang Q, Shi P, Gao Y, Chen Y, Liu W, Qin Y, Shen Y, Shi C. 2020. The correlation between atmospheric visibility and influenza in Wuxi city, China. *Medicine* **99**:e21469. DOI: https://doi.org/10.1097/MD.0000000000021469, PMID: 32769879

Liu X, Liu C, Huang R, Zhu H, Liu Q, Mitra S, Wang Y. 2021. Long short-term memory recurrent neural network for pharmacokinetic-pharmacodynamic modeling. *Int. Journal of Clinical Pharmacology and Therapeutics* **59**:138–146. DOI: https://doi.org/10.5414/CP203800, PMID: 33210994

Lofgren E, Fefferman NH, Naumov YN, Gorski J, Naumova EN. 2007. Influenza seasonality: underlying causes and modeling theories. *Journal of Virology* **81**:5429–5436. DOI: https://doi.org/10.1128/JVI.01680-06, PMID: 17182688

Lowen AC, Steel J. 2014. Roles of humidity and temperature in shaping influenza seasonality. *Journal of Virology* **88**:7692–7695. DOI: https://doi.org/10.1128/JVI.03544-13, PMID: 24789791

Marr LC, Tang JW, Van Mullekom J, Lakdawala SS. 2019. Mechanistic insights into the effect of humidity on airborne influenza virus survival, transmission and incidence. *Journal of the Royal Society, Interface* **16**:20180298. DOI: https://doi.org/10.1098/rsif.2018.0298, PMID: 30958176

Moriyama M, Ichinohe T. 2019. High ambient temperature dampens adaptive immune responses to influenza A virus infection. *PNAS* **116**:3118–3125. DOI: https://doi.org/10.1073/pnas.1815029116, PMID: 30718396

Moriyama M, Hugentobler WJ, Iwasaki A. 2020. Seasonality of respiratory viral infections. *Annual Review of Virology* **7**:83–101. DOI: https://doi.org/10.1146/annurev-virology-012420-022445, PMID: 32196426

Ng H, Li Y, Zhang T, Lu Y, Wong C, Ni J, Zhao Q. 2022. Association between multiple meteorological variables and seasonal influenza A and B virus transmission in Macau. *Heliyon* **8**:e11820. DOI: https://doi.org/10.1016/j.heliyon.2022.e11820, PMID: 36468127

Panchal G, Ganatra A, Kosta YP, Panchal D. 2011. Behaviour analysis of multilayer perceptronswith multiple hidden neurons and hidden layers. *International Journal of Computer Theory and Engineering* **3**:332–337. DOI: https://doi.org/10.7763/IJCTE.2011.V3.328

Park JE, Son WS, Ryu Y, Choi SB, Kwon O, Ahn I. 2020. Effects of temperature, humidity, and diurnal temperature range on influenza incidence in a temperate region. *Influenza and Other Respiratory Viruses* **14**:11–18. DOI: https://doi.org/10.1111/irv.12682, PMID: 31631558

Peci A, Winter AL, Li Y, Gnaneshan S, Liu J, Mubareka S, Gubbay JB. 2019. Effects of absolute humidity, relative humidity, temperature, and wind speed on influenza activity in Toronto, Ontario, Canada. *Applied and Environmental Microbiology* **85**:e02426-18. DOI: https://doi.org/10.1128/AEM.02426-18, PMID: 30610079

Polozov IV, Bezrukov L, Gawrisch K, Zimmerberg J. 2008. Progressive ordering with decreasing temperature of the phospholipids of influenza virus. *Nature Chemical Biology* **4**:248–255. DOI: https://doi.org/10.1038/nchembio.77, PMID: 18311130

Ryu S, Cowling BJ. 2021. Human influenza epidemiology. *Cold Spring Harbor Perspectives in Medicine* **11**:a038356. DOI: https://doi.org/10.1101/cshperspect.a038356, PMID: 32988982

Sagripanti JL, Lytle CD. 2007. Inactivation of influenza virus by solar radiation. *Photochemistry and Photobiology* **83**:1278–1282. DOI: https://doi.org/10.1111/j.1751-1097.2007.00177.x, PMID: 17880524

Shoji M, Katayama K, Sano K. 2011. Absolute humidity as a deterministic factor affecting seasonal influenza epidemics in Japan. *The Tohoku Journal of Experimental Medicine* **224**:251–256. DOI: https://doi.org/10.1620/tjem.224.251, PMID: 21705850

Si X, Mengersen K, Ye C, Hu W. 2024. Interactive effect of air pollutant and meteorological factors on seasonal influenza transmission, Shanghai, China. *Atmospheric Environment* **318**:120208. DOI: https://doi.org/10.1016/j.atmosenv.2023.120208

Soebiyanto RP, Adimi F, Kiang RK. 2010. Modeling and predicting seasonal influenza transmission in warm regions using climatological parameters. *PLOS ONE* **5**:e9450. DOI: https://doi.org/10.1371/journal.pone.0009450, PMID: 20209164

Soebiyanto RP, Gross D, Jorgensen P, Buda S, Bromberg M, Kaufman Z, Prosenc K, Socan M, Vega Alonso T, Widdowson M-A, Kiang RK. 2015. Associations between meteorological parameters and influenza activity in Berlin (Germany), Ljubljana (Slovenia), Castile and León (Spain) and Israeli districts. *PLOS ONE* **10**:e0134701. DOI: https://doi.org/10.1371/journal.pone.0134701, PMID: 26309214

Sooryanarain H, Elankumaran S. 2015. Environmental role in influenza virus outbreaks. *Annual Review of Animal Biosciences* **3**:347–373. DOI: https://doi.org/10.1146/annurev-animal-022114-111017, PMID: 25422855

Tamerius JD, Shaman J, Alonso WJ, Bloom-Feshbach K, Uejio CK, Comrie A, Viboud C. 2013. Environmental predictors of seasonal influenza epidemics across temperate and tropical climates. *PLOS Pathogens* **9**:e1003194. DOI: https://doi.org/10.1371/journal.ppat.1003194, PMID: 23505366

Urashima M, Segawa T, Okazaki M, Kurihara M, Wada Y, Ida H. 2010. Randomized trial of vitamin D supplementation to prevent seasonal influenza A in schoolchildren. *The American Journal of Clinical Nutrition* **91**:1255–1260. DOI: https://doi.org/10.3945/ajcn.2009.29094, PMID: 20219962

Uyeki TM, Hui DS, Zambon M, Wentworth DE, Monto AS. 2022. Influenza. *Lancet* **400**:693–706. DOI: https://doi.org/10.1016/S0140-6736(22)00982-5, PMID: 36030813

Venna SR, Tavanaei A, Gottumukkala RN, Raghavan VV, Maida A, Nichols S. 2018. A novel data-driven model for real-time influenza forecasting. *IEEE Access* **14**:5512. DOI: https://doi.org/10.1101/185512

Wang C, Li Y, Feng W, Liu K, Zhang S, Hu F, Jiao S, Lao X, Ni H, Xu G. 2017. Epidemiological features and forecast model analysis for the morbidity of influenza in Ningbo, China, 2006-2014. *International Journal of Environmental Research and Public Health* **14**:559. DOI: https://doi.org/10.3390/ijerph14060559, PMID: 28587073

Wang D, Lei H, Wang D, Shu Y, Xiao S. 2023. Association between temperature and influenza activity across different regions of China during 2010-2017. *Viruses* **15**:594. DOI: https://doi.org/10.3390/v15030594, PMID: 36992303

Xiao J, Gao M, Huang M, Zhang W, Du Z, Liu T, Meng X, Ma W, Lin S. 2022. How do El Niño Southern Oscillation (ENSO) and local meteorological factors affect the incidence of seasonal influenza in New York state. *Hygiene and Environmental Health Advances* **4**:100040. DOI: https://doi.org/10.1016/j.heha.2022.100040, PMID: 36777308

Xie L. 2026. DLNM-LSTM-influenza-forecasting-putian. swh:1:rev:000b60da8174a679621479d94f83adf09e67f2d7. Software Heritage. https://archive.softwareheritage.org/swh:1:dir:1b525859fe6cb7fe6743831a4f083fc9862ed147;origin=https://github.com/Long-Xie/DLNM-LSTM-Influenza-Forecasting-Putian;visit=swh:1:snp:fcd02181e2444775de0940d217c81586d0a9e088;anchor=swh:1:rev:000b60da8174a679621479d94f83adf09e67f2d7

Xu DL, Hu XK, Tian YF. 2017. Effect of temperature and food restriction on immune function in striped hamsters (*Cricetulus barabensis*). *The Journal of Experimental Biology* **220**:2187–2195. DOI: https://doi.org/10.1242/jeb.153601, PMID: 28381582

Yang W, Elankumaran S, Marr LC. 2012. Relationship between humidity and influenza a viability in droplets and implications for influenza's seasonality. *PLOS ONE* **7**:e46789. DOI: https://doi.org/10.1371/journal.pone.0046789, PMID: 23056454

Zhang Y, Ye C, Yu J, Zhu W, Wang Y, Li Z, Xu Z, Cheng J, Wang N, Hao L, Hu W. 2020. The complex associations of climate variability with seasonal influenza A and B virus transmission in subtropical Shanghai, China. *Science of The Total Environment* **701**:134607. DOI: https://doi.org/10.1016/j.scitotenv.2019.134607, PMID: 31710904

Zhang R, Peng Z, Meng Y, Song H, Wang S, Bi P, Li D, Zhao X, Yao X, Li Y. 2022. Temperature and influenza transmission: Risk assessment and attributable burden estimation among 30 cities in China. *Environmental Research* **215**:114343. DOI: https://doi.org/10.1016/j.envres.2022.114343, PMID: 36115415

Zhao Z, Zhai M, Li G, Gao X, Song W, Wang X, Ren H, Cui Y, Qiao Y, Ren J, Chen L, Qiu L. 2023. Study on the prediction effect of a combined model of SARIMA and LSTM based on SSA for influenza in Shanxi Province, China. *BMC Infectious Diseases* **23**:71. DOI: https://doi.org/10.1186/s12879-023-08025-1, PMID: 36747126

Zhou L, Yang H, Pan W, Xu J, Feng Y, Zhang W, Shao Z, Li T, Li S, Huang T, Wang C, Li W, Li M, He S, Zhan Y, Pan M. 2022. Association between meteorological factors and the epidemics of influenza (sub)types in a subtropical basin of Southwest China. *Epidemics* **41**:100650. DOI: https://doi.org/10.1016/j.epidem.2022.100650, PMID: 36375312

Zhu H, Chen S, Lu W, Chen K, Feng Y, Xie Z, Zhang Z, Li L, Ou J, Chen G. 2022. Study on the influence of meteorological factors on influenza in different regions and predictions based on an LSTM algorithm. *BMC Public Health* **22**:2335. DOI: https://doi.org/10.1186/s12889-022-14299-y, PMID: 36514013

Appendix 1

Supplementary methods, results, and tables

For: "Integrating distributed lag non-linear models and deep learning reveals meteorological drivers of influenza A and B positivity in a 6-year subtropical surveillance study". Figures cited in this appendix are provided as figure supplements to the indicated main-text figures.

This appendix contains six sections. Section 'Supplementary methods' outlines the methodology for nucleic acid extraction and real-time RT-PCR, the principles and formulas of the LSTM cell, and the hyperparameter optimization strategy. Section 'Supplementary results tables' provides supplementary results tables (***Appendix 1—table 1***, descriptive statistics of meteorological factors; ***Appendix 1—table 2***, QAIC values for DLNM degrees-of-freedom combinations). Section 'Linear association analysis' reports the linear association analysis. Section 'Sensitivity analysis of the DLNM model' details the sensitivity analysis of the DLNM. Section 'Construction and analysis of the autoregressive integrated moving average (ARIMA) model' describes the construction and analysis of the multivariate ARIMA baseline model. Section 'Comparative analysis with an eXtreme gradient boosting (XGBoost) machine-learning baseline' describes the XGBoost machine-learning baseline and its comparative analysis against the LSTM, with ***Appendix 1—table 3*** detailing comparative predictive performance of LSTM, multivariate ARIMA, and XGBoost on the 2023 validation set.

Supplementary methods

Nucleic acid extraction

A commercial magnetic bead-based kit (Da'an Gene, Guangzhou, China) was used, following the manufacturer's protocol. The extracted nucleic acid was dissolved in 50 μL of elution buffer, with concentration and purity measured using a Nanodrop 2000 spectrophotometer (Thermo-Fisher Scientific, Waltham, MA). Acceptable nucleic acid purity ratios were defined as an optical density (OD)260/280 ratio between 1.7 and 2.5, and an OD260/230 ratio between 0.5 and 2.5. The integrity of the extracted nucleic acids was confirmed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA), with RNA Integrity Number (RIN) values ≥8 and 28 S/18 S ratios ≥1.5 considered indicative of high-quality samples.

Real-time reverse transcription polymerase chain reaction (RT-PCR)

Qualified nucleic acids underwent one-step real-time fluorescence RT-PCR detection within two hours, utilizing commercially available detection kits for influenza A (H1N1) virus (2009) RNA, seasonal influenza virus H3N2 subtype, and influenza B (Victoria and Yamagata lineages), all purchased from Da'an Gene (Guangzhou, China). The detection procedure was performed on a QuantStudio real-time fluorescence quantitative PCR instrument (Thermo-Fisher Scientific, Waltham, MA), following the manufacturer's instructions. Each detection batch contained negative, positive, and internal control replicates to ensure the reliability of the results. The PCR results were interpreted adhering to the assay manufacturer's protocols, with amplification curves and cycle threshold (Ct) values used to determine positivity. Positive results were determined if typical amplification curves were observed in both the internal reference fluorescence channel and the target gene fluorescence channel, with corresponding Ct values within the reference range. Specific positive criteria were as follows: for influenza A H1N1, FAM channel amplification with Ct value ≤38; for H3N2, FAM channel amplification with Ct value ≤38; for influenza B Victoria, FAM channel amplification with Ct value ≤38; for Yamagata, VIC channel amplification with Ct value ≤38.

Specific principles and formulas for the cell structure of the LSTM network

The core component of the long short-term memory (LSTM) network is the LSTM cell. The LSTM network incorporates multiple LSTM cells per layer. Each LSTM cell encompasses three gating mechanisms: the forget gate, the input gate, and the output gate (***Figure 1—figure supplement 1***), which collaboratively control the flow and updating of information between different time steps, enabling the LSTM to selectively retain or discard information, thereby effectively processing long sequence data.

The forget gate controls the amount of information retained from the previous time step and discards irrelevant data. Mathematically, it is represented as:

$$f_t = \sigma\left(W_f \cdot \left[h_{t-1}, input_t\right] + b_f\right)$$

The input gate determines which information from the current time step is added to the cell state, where $C_t^{'}$ represents the candidate state that could be incorporated. This gate is governed by:

$i_t = \sigma\left(W_i \cdot \left[h_{t-1}, input_t\right] + b_i\right)$ and $C_t' = function\left(W_C \cdot \left[h_{t-1}, input_t\right] + b_C\right)$, and the cell state is updated as: $C_t = f_t \times C_{t-1}^{'} + i_t \times C_t^{'}$

The output gate modulates the final cell state and controls the transmission of information to the next time step, expressed as:

$$O_t = \sigma(W_o \cdot [h_{t-1}, input_t] + b_o) \text{ and } h_t = O_t \cdot tan\left(C_t\right)$$

where $input_t$ denotes the input sequence at time t, the symbol $\cdot$ represents element-wise multiplication, h_t and h_{t-1} are the hidden states at time t and t-1 respectively, C_t is the cell state at time t, W represents weight matrices, and b represents bias vectors. $function$ represents the activation function for the candidate state, which is selected based on hyperparameter tuning.

This structure allows the model to capture long-term dependencies within time-series data while efficiently filtering out irrelevant information through the three gates.

Hyperparameter optimization strategy and search space

To ensure optimal convergence and predictive stability of the LSTM networks, we conducted rigorous hyperparameter tuning using the Hyperopt framework. Hyperopt employs a Bayesian optimization strategy based on Tree-based Parzen Estimators (TPE), which significantly outperforms traditional grid search and random search methods by systematically evaluating past performance to propose the next optimal set of hyperparameters (***Hanifi et al., 2022***). Previous comparative analyses have demonstrated Hyperopt's superior performance in LSTM applications compared to other optimization libraries, such as Scikit-opt and Optuna (***Hanifi et al., 2024***). In our framework, key hyperparameters, including the learning rate, number of hidden layers, number of neurons, and activation functions, were meticulously tuned within established theoretical ranges using Mean Squared Error (MSE) as the reference metric.

Learning rate

Following Bengio, who identified the learning rate as a critical hyperparameter and recommended a search range of 10^{-6} to 1 (***Bengio, 2012***), and corroborated by recent findings demonstrating that rates between 0.0001 and 0.01 yield superior convergence in LSTM networks (***Du et al., 2023***), we established a discrete learning rate search space of (0.000001, 0.00001, 0.0001, 0.001, 0.01, 0.1). This specific range balances the trade-off between computational cost and fine-grained optimization.

Network architecture

Theoretically, increasing the number of hidden layers and neurons enhances the algorithm's fitting capacity; however, excessive complexity risks overfitting and impedes convergence. Based on Panchal's findings, which suggest that 1–2 hidden layers suffice for most applications and 3 layers are considered optimal for maximizing accuracy without severe overfitting (***Panchal et al., 2011***), we restricted our architectural exploration to configurations with 1–3 hidden layers. Within each hidden layer, the candidate set for the number of neurons was selected as powers of 2 (i.e. [8, 16, 32, 64, 128, 256, 512]) to optimize computational memory efficiency. Furthermore, the batch size for each training session was also tuned covering the range (8, 16, 32, 64, 128, 256, 512), with the number of epochs flexibly controlled within a natural number interval from 300 to 1000.

Activation functions

Recognizing that no single activation function universally optimizes network performance across all sequential tasks (***Jagtap and Karniadakis, 2022***), we evaluated a comprehensive suite of activation functions. The search space included traditional functions (Sigmoid, Rectified Linear Unit [ReLU], and

Hyperbolic Tangent [TanH]) alongside newer, advanced variants (Softplus, Exponential Linear Unit [ELU], and Swish) for automatic selection during the Bayesian optimization process.

Supplementary results tables

Appendix 1—table 1. Descriptive statistics of meteorological factors in Putian City (2018–2023).

Meteorological factor	Minimal	5th percentile	Median	Mean	SD	95th percentile	Maximum
Average temperature (°C)	7.6	13.8	24.1	23.3	5.6	30.7	32.9
Humidity (%)	31	57.95	75.8	76.1	10.8	94.15	99.8
Precipitation (mm)	0	0	0	3.9	10.9	23.394	144.41
Wind speed (km/h)	6.8	10.8	18	17.9	5.4	25.2	116.8
Wind direction *	NW	-	SW	-	-	-	E
Solar radiation (W/m^2)	5.8	73.1	205.3	202.0	69.1	305.9	341.3
Ultraviolet index	0	3	7	7.1	2.1	10	10
Daily temperature ranges (°C)	1	3	8	7.8	3.0	13	16

*Note: Wind direction is described based on the frequency of occurrence of the eight cardinal directions. NW: north-westerly; SW: south-westerly; E: easterly.

Appendix 1—table 2. QAIC values for DLNM models across degrees of freedom combinations.

Influenza type	Meteorological factor	df_x	df_{time} 6	7	8	9	10
		2	-	1.04E+36	7921.38	4113.21	3170.75
		3	12562353	3317.23	10240.29	5143.05	3153.05
		4	26540151	9.74E+17	42517.57	81578405	3096.05
	Humidity	5	3.11E+08	3253.86	494344.9	3442.48	3068.09
		2	-	-	-	-	-
		3	51466.64	3531.96	4224.61	3933.90	3103.71
		4	4.19E+134	3535.98	4834.92	4409.85	3103.31
	Precipitation	5	210749.33	3572.73	5381.03	4498.39	3104.24
		2	213743.68	3255.26	4605.74	3623.64	3181.35
		3	-	2.1E+120	5183.57	3792.24	3181.52
		4	-	3246.15	4806.17	3753.06	3173.27
	Solar radiation	5	239783	3243.05	6172.39	3773.43	3167.59
		2	-	3340.64	4410.01	3497.04	3169.77
		3	293680.27	3384.87	4808.02	3349.16	3175.01
		4	-	1.07E+173	4833.78	3325.88	3160.23
	Daily temperature range	5	-	3384.08	5610.18	3297.89	3123.95
		2	40091169	3313.98	4954.56	1.58E+40	3188.91
		3	303643.15	3218.40	4898.15	4195.35	3096.06
		4	68152.83	3240.96	4327.51	3769.00	3104.44
	Temperature	5	-	3252.09	4135.43	1.59E+256	3089.88
		2	211637.62	3207.10	4757.41	32337.63	3156.51
		3	152137.84	3194.74	4386.31	3540.17	3154.23
		4	1.7E+09	3197.19	5059.73	4097.16	3136.33
	Ultraviolet index	5	61743113	3194.97	3217.16	3225.33	3122.32
		2	215264.74	5.2E+118	4243.30	3730.69	3135.29
		3	-	7.81E+185	4232.54	7.96E+08	3105.84
		4	4.16E+16	3249.18	3817.01	7.07E+91	3097.46
	Wind direction	5	2.1E+109	3211.63	3842.97	3324.891	3093.36
		2	370034.51	3344.84	4504.67	2.33E+20	3177.68
		3	409397.77	3447.91	4404.46	4491.16	3134.36
		4	4.75E+16	3494.94	4729.04	5548.65	3129.81
Influenza A	Wind speed	5	539380.65	539306.70	4193.76	4.17E+18	3122.94

Appendix 1—table 2 Continued on next page

Appendix 1—table 2 Continued

Influenza type	Meteorological factor	df_x	df_{time} 6	7	8	9	10
		2	3171.40	-	3175.68	3173.96	-
		3	3175.22	-	3176.62	3173.47	3232.91
		4	3115.23	-	3120.42	3117.40	3095.73
	Humidity	5	3077.17	-	2.91E+184	3081.48	3060.79
		2	-	-	-	-	-
		3	3119.93	-	3102.16	3097.36	3072.71
		4	3059.24	-	3012.14	3003.45	2983.99
	Precipitation	5	3054.20	-	2993.22	2992.69	2963.95
		2	3062.49	-	3058.97	3049.41	3028.98
		3	3048.52	-	3040.77	3023.35	1.43E+54
		4	3026.96	-	3028.70	3001.70	-
	Solar radiation	5	3029.44	-	1.91E+80	3005.49	3013.18
		2	3112.52	-	3116.68	3121.76	1.21E+138
		3	3100.92	4.06E+12	3101.86	3101.91	7.24E+09
		4	3018.37	-	3.12E+13	3013.11	3008.44
	Daily temperature range	5	2985.16	6.89E+262	9.54E+221	2977.85	2973.01
		2	3104.50	-	-	3049.23	3038.55
		3	3048.00	-	-	3022.72	6.91E+84
		4	3040.63	-	3047.21	3010.52	2997.21
	Temperature	5	3039.46	5.26E+152	3045.00	3008.85	2.5E+108
		2	3084.96	-	3082.84	3066.21	3033.36
		3	3101.43	-	3080.68	3060.74	1.58E+152
		4	3185.54	-	3038.69	2992.05	-
	Ultraviolet index	5	3077.97	-	2983.66	2980.61	1.7E+113
		2	3159.66	8.1E+27	3155.93	3099.48	5375.50
		3	3160.55	4.07E+121	-	3071.69	3046.36
		4	3213.52	-	3101.80	3047.93	3016.08
	Wind direction	5	3161.81	5.50E+229	3091.10	3042.39	2996.10
		2	3199.67	21479.97	3204.81	3190.44	3.37E+130
		3	3168.48	-	-	3167.41	4.2E+99
		4	3152.88	3139.90	3173.00	3159.70	3.51E+171
Influenza B	Wind speed	5	3144.57	-	3165.02	3155.37	2.5E+108

Note: DLNM (distributed lag non-linear models), QAIC (quasi-Akaike information criterion), degrees of freedom (df). '-' indicates poor model fit or non-convergence at the specified df combination, precluding QAIC calculation.

Linear association analysis

Method

Pearson correlation analysis was conducted to assess the linear relationships between meteorological factors and influenza incidence, with correlation coefficient matrices plotted to reveal underlying association patterns.

Result

Pearson correlation coefficients were calculated between various meteorological factors and influenza cases, and scatter plots, density plots, and correlation coefficient matrices were generated. The analysis revealed that influenza A cases were statistically significantly negatively correlated with average temperature (r=–0.163), wind speed (r=–0.093), solar radiation (r=–0.100), UV index (r=–0.072), and wind direction (r=–0.064). Similarly, influenza B cases also exhibited comparable negative correlations with these meteorological factors: average temperature (r=–0.197), wind speed (r=–0.091), solar radiation (r=–0.063), UV index (r=–0.071), and wind direction (r=–0.078). These results revealed that the linear correlations between the meteorological factors and the number of influenza A cases, influenza B cases, and influenza-like illness cases were weak (***Figure 4—figure supplement 1***).

Sensitivity analysis of the DLNM model

Rationale and methodological framework

In the principal DLNM analysis, weekly testing volumes were incorporated as a covariate to account for variations in surveillance intensity that could potentially confound the estimated associations between meteorological factors and influenza positivity rates. To rigorously assess the robustness of the DLNM findings against the influence of this covariate and to determine whether the principal conclusions are sensitive to its inclusion, we conducted a comprehensive sensitivity analysis by reconstructing the DLNM with the weekly detection variable removed while preserving all other model specifications.

Sensitivity analysis methodology

The sensitivity analysis was implemented in two complementary dimensions. First, we evaluated the impact of weekly detection volumes variable removal on the 15-day cumulative relative risk estimates across all meteorological factors for both influenza A and B subtypes. Second, we examined the lag-specific relative risk patterns under extreme meteorological conditions (defined as the 95th and 5th percentiles of each variable) following variable removal. The reconstructed DLNMs retained identical degrees of freedom for the cross-basis splines, lag structure (15-day maximum lag), reference values, and remaining covariates as those specified in the original model. Comparisons between the reconstructed and original models focused on the directionality, magnitude, and statistical significance of the estimated associations.

Results of cumulative risk sensitivity analysis

The sensitivity analysis for cumulative relative risk estimates demonstrated substantial concordance between the reconstructed DLNM (with the weekly detection volumes variable removed) and the original model. The non-linear response patterns of temperature, humidity, precipitation, wind speed, solar radiation, and daily temperature range on the 15-day cumulative risks of both influenza A and B positivity rates remained largely consistent in shape, magnitude, and statistical significance (***Figure 4—figure supplement 2***). Two notable deviations were observed: (1) for influenza A, the ultraviolet index exhibited a non-monotonic trajectory in the reconstructed model, initially increasing before declining, whereas the original model showed a continuous decreasing trend; however, the increasing segment in the reconstructed model did not reach statistical significance; and (2) for influenza B, the protective effect of wind speeds exceeding 28 km/hr on cumulative risk lost statistical significance in the reconstructed model, whereas this protective association was statistically significant in the original model. These minor deviations notwithstanding, the principal findings regarding the dominant meteorological drivers of influenza A (warm-humid conditions) and influenza B (cool-wet conditions) remained robust to the removal of the weekly detection volumes covariate.

Results of extreme condition lag-effect sensitivity analysis

The sensitivity analysis for lag-specific relative risk patterns under extreme meteorological conditions revealed that the reconstructed DLNM produced results substantially consistent with those of the original model (***Figure 5—figure supplement 1***). The temporal heterogeneity of meteorological

effects on influenza A and B positivity rates, including the delayed peak effects of extreme high temperature on influenza A and the sustained risk elevations under extreme low temperature for influenza B, was preserved across both modeling configurations. The lag-response patterns for extreme precipitation, humidity, wind speed, solar radiation, ultraviolet index, and daily temperature range exposures similarly demonstrated qualitative and quantitative agreement between the original and reconstructed models.

Interpretation and implications

Collectively, the sensitivity analysis findings substantiate the robustness of the principal DLNM conclusions regarding type-specific meteorological drivers of influenza transmission. The minor deviations observed in select parameters (ultraviolet index for influenza A; wind speed for influenza B) suggest that while weekly testing volumes may exert subtle modifying effects on certain associations, the overall meteorological response architecture identified in the original analysis is not critically dependent on the inclusion of this covariate. These results enhance confidence in the validity of the meteorological driver characterizations presented in the main analysis and support the interpretive framework adopted in the principal manuscript.

Construction and analysis of the autoregressive integrated moving average (ARIMA) model

Method

To establish a mathematically fair baseline against the deep learning framework, the conventional Autoregressive Integrated Moving Average (ARIMA) model was extended to a multivariate formulation to incorporate external regressors. The model was fitted using the auto.arima() function from the R package forecast. Crucially, the identical multidimensional covariate matrix used in the LSTM, comprising meteorological factors, weekly detection volumes, mask-wearing stringency indices, day of the week (DOW), and non-local population proportions, was explicitly fed into the model via the xreg argument. The function autonomously optimized the autoregressive, differencing, and moving average parameters (p, d, q) based on the Akaike Information Criterion (AIC/QAIC). Data partitioning was strictly consistent with the LSTM pipeline: 2018–2022 data were utilized for training, while 2023 data served as the validation set for predicting Influenza A and B positivity rates. Performance was assessed using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Symmetric Mean Absolute Percentage Error (SMAPE), and compared with that from the Bayesian-optimized LSTM neural network.

The four metrics, MAE, RMSE, MAPE, and SMAPE, are calculated as:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i|$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2}$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n} \left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100$$

$$\mathrm{SMAPE} = \frac{1}{n}\sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{(y_i + \hat{y}_i)/2} \times 100$$

with adjustments for zero denominators.

Results

The MAE for the covariate-adjusted ARIMA model of influenza A was 0.136, RMSE was 0.238, MAPE was 1.115, and SMAPE was 1.212. For influenza B, the MAE was 0.049, RMSE was 0.057, MAPE was 1.426, and SMAPE was 0.810. Overall, the predictive performance of the ARIMA model was substantially inferior to that of the Bayesian-optimized LSTM (***Appendix 1—table 3***). As illustrated in ***Figure 6—figure supplement 1***, the linear model struggled significantly with the non-stationary dynamics of the 2023 viral rebound. It either failed entirely to capture the extreme, explosive non-linear peaks (as seen in Influenza A) or generated continuous spurious predictions during zero-case periods due to mechanical linear reactions to covariate inputs (as seen in Influenza B).

Comparative analysis with an eXtreme gradient boosting (XGBoost) machine-learning baseline

Rationale

While the multivariate ARIMA model detailed in Additional File 5 constitutes the canonical linear statistical benchmark for epidemiological time-series forecasting, gradient-boosted decision-tree ensembles, most notably eXtreme Gradient Boosting (XGBoost), have emerged as one of the leading non-deep-learning approaches for structured predictive tasks (***Chen and Guestrin, 2016***). To subject our Bayesian-optimized LSTM network to a rigorous machine-learning benchmark and to isolate the incremental predictive value contributed by the recurrent architecture itself, we constructed an XGBoost regression model using the identical multidimensional covariate matrix employed in the deep-learning pipeline. Importantly, to guarantee methodological comparability, the XGBoost model was supplied with the identical set of exogenous covariates as the Bayesian-optimized LSTM, meteorological variables, weekly testing volumes, mask-wearing stringency indices, migrant population proportions, and day-of-the-week indicators, without any additional lag features, autoregressive terms, or engineered temporal representations, so as to strictly isolate the predictive contribution attributable to sequential inductive bias.

Methods

Model Implementation

XGBoost provides a scalable gradient-boosting framework built upon sparse classification and regression trees (CART), representing the state-of-the-art non-deep-learning approach for structured predictive tasks. The model was fitted in R (xgboost package, version 3.2.1.1), with hyperparameter tuning conducted via grid search over tree depth, learning rate, and L2 regularization parameters, as detailed in the Methods (Statistical analysis).

Input feature space

The predictor set was strictly matched to the LSTM covariate architecture and comprised: (i) eight meteorological variables (average temperature, humidity, precipitation, wind speed, wind direction, solar radiation, ultraviolet index, and daily temperature range); (ii) weekly influenza testing volumes; (iii) the mask-wearing stringency index; (iv) a day-of-week indicator; and (v) the proportion of non-local (migrant) population. Critically, no additional lag features, sliding-window statistics, or autoregressive terms were introduced, thereby ensuring that the information set available to XGBoost was identical to the contemporaneous covariate input consumed by the LSTM at each time step. This design deliberately isolates the predictive contribution of *architectural inductive bias*, specifically, the LSTM's intrinsic sequential memory versus XGBoost's memory-free tree-partitioning geometry, from any confounding due to hand-engineered temporal feature construction.

Data partitioning

The training-validation split strictly followed the LSTM and ARIMA protocols: data from 2018 to 2022 constituted the training set, while 2023 served as the held-out validation set for out-of-sample forecasting of weekly Influenza A and B positivity rates.

Performance evaluation

Predictive accuracy was quantified using mean absolute error (MAE), root mean squared error (RMSE), mean absolute percentage error (MAPE), and symmetric mean absolute percentage error (SMAPE), calculated identically to the metrics reported for the ARIMA baseline in Appendix 1—section 5.

Results

The XGBoost model produced substantially higher forecasting errors than the LSTM across both influenza subtypes (***Appendix 1—table 3***; ***Figure 6—figure supplement 2***). For *influenza A*, XGBoost yielded MAE = 0.1379, RMSE = 0.2161, MAPE = 1.8195, and SMAPE = 1.0805, errors of a magnitude broadly comparable to the multivariate ARIMA baseline (MAE = 0.136) and approximately 15-fold

higher than the LSTM (MAE = 0.009). For *influenza B*, XGBoost achieved MAE = 0.0704, RMSE = 0.0739, MAPE = 1.9172, and SMAPE = 0.7371, exceeding both the ARIMA benchmark (MAE = 0.049) and the LSTM (MAE = 0.002) by approximately 35-fold relative to the LSTM.

Appendix 1—table 3. Comparative predictive performance of LSTM, multivariate ARIMA, and XGBoost on the 2023 validation set.

Model	Subtype	MAE	RMSE	MAPE	SMAPE
Bayesian-optimized LSTM	Influenza A	0.009	0.035	0.158	0.521
Multivariate ARIMA	Influenza A	0.136	0.238	1.115	1.212
XGBoost	Influenza A	0.138	0.216	1.820	1.081
Bayesian-optimized LSTM	Influenza B	0.002	0.011	0.170	0.484
Multivariate ARIMA	Influenza B	0.049	0.057	1.426	0.810
XGBoost	Influenza B	0.070	0.074	1.917	0.737

Note: All three models received the identical multidimensional covariate matrix. Best-performing values are in bold.

Interpretation

The systematically inferior performance of XGBoost, despite receiving the identical covariate matrix as the LSTM, provides architecturally informative insight into the architectural properties required for post-pandemic influenza forecasting. Four fundamental limitations of the gradient-boosted tree framework, articulated in the Discussion section of the main manuscript, jointly account for this performance gap:

1. *Absence of intrinsic temporal ordering*. Decision trees construct predictions through threshold-based partitions along feature dimensions; their splitting logic is insensitive to the sequential ordering of observations and cannot structurally express the temporal dependency 'the current time step depends on prior history' in the manner naturally accommodated by recurrent neural architectures.
2. *Inability to model long-range temporal dependencies*. XGBoost lacks any mechanism analogous to the LSTM's gated recurrence, through which historical states are propagated, filtered, and dynamically re-weighted according to their relevance to the current prediction. Consequently, XGBoost cannot represent the non-linear, long-horizon temporal dependencies that characterize influenza dynamics.
3. *Violation of the independence assumption*. Tree-based ensembles implicitly assume that samples are independent and identically distributed, an assumption fundamentally at odds with the strong serial autocorrelation intrinsic to epidemiological time series. XGBoost typically neglects the temporal structure of residuals and cannot effectively model serial dependence.
4. *Non-comparable time-modeling paradigms*. The central methodological question of our forecasting analysis was whether a sequential deep-learning model offers substantive gains over traditional approaches. Because XGBoost belongs to a fundamentally non-sequential modeling paradigm, its comparison with the LSTM primarily illuminates the limitations of memory-free machine learning under non-stationary epidemiological conditions, rather than an equivalent 'horse race' between competing sequential frameworks. The ARIMA benchmark, which like the LSTM, is a sequential model that builds temporal dependence directly into its structure, therefore provides the more paradigmatically comparable reference against which the deep-learning architecture's advantages can be most directly discerned.

These results reinforce our central methodological conclusion: the LSTM's decisive predictive advantage during the volatile 2023 post-NPI transition arises not from privileged access to information, but from its architectural capacity to represent evolving, non-stationary, autocorrelated dynamics, capacities absent in both linear autoregressive (ARIMA) and non-linear memory-free machine-learning (XGBoost) baselines. Our finding has methodological implications beyond the present dataset: it suggests that recurrent gated architectures constitute a principled design choice

for public-health forecasting under regimes of pronounced non-stationarity, such as those induced by the pandemic-era disruption of respiratory-virus circulation.