

基于SARIMAX-AttLSTM和BEAST模型对重庆市秀山县流感特征的预测分析

张浩琳¹, 敖玉琪², 曾浩航¹, 杨淑燕², 杨川^{2*}, 曾庆^{1*}

¹重庆医科大学公共卫生学院, 重庆

²秀山土家族苗族自治县疾病预防控制中心传染病防制科, 重庆

收稿日期: 2025年10月22日; 录用日期: 2025年11月13日; 发布日期: 2025年11月24日

摘要

背景: 流感是一种急性呼吸道传染病, 对公共卫生系统构成严重威胁。明确区域流感流行特征并准确预测发病趋势, 对制定防控策略具有重要意义。目的: 分析重庆市秀山土家族苗族自治县(秀山县)流感流行特征及发病率变化趋势, 为区域防控提供依据。方法: 利用2018年第1周至2024年第26周的流感报告卡数据, 结合同期气象数据分析流行特征。采用BEAST算法识别趋势与季节变化点, 并构建SARIMAX-AttLSTM混合模型预测发病趋势。结果: 共报告流感病例35,057例, 男性52.7%、女性47.3% ($\chi^2 = 53.8$, $P < 0.001$)。主要发病年龄为1~9岁 (52.9%, $\chi^2 = 2957.0$, $P < 0.001$), 学生和幼托儿童为主要受影响人群。发病呈明显季节性, 冬春为高发期。BEAST算法识别出4个季节性与5个趋势性变化点。SARIMAX-AttLSTM模型表现最佳(MAE = 8.064, RMSE = 11.724, $R^2 = 0.922$), 引入温度和降雨量后预测精度进一步提升。模型预测2024年第27周后发病率将上升。结论: 秀山县流感呈冬春高发、儿童和学生易感的特征。SARIMAX-AttLSTM模型可有效捕捉流感季节与短期波动, 具有较高的预警与决策支持应用价值。

关键词

流感发病率, SARIMAX-AttLSTM模型, BEAST算法, 秀山县, 预测

Analysis and Prediction of Influenza Characteristics in Xiushan County, Chongqing Based on the SARIMAX-AttLSTM Model and BEAST Algorithm

Haolin Zhang¹, Yuqi Ao², Haohang Zeng¹, Shuyan Yang², Chuan Yang^{2*}, Qing Zeng^{1*}

¹School of Public Health, Chongqing Medical University, Chongqing

*共同通讯作者。其中, 曾庆为第一通讯作者。

文章引用: 张浩琳, 敖玉琪, 曾浩航, 杨淑燕, 杨川, 曾庆. 基于 SARIMAX-AttLSTM 和 BEAST 模型对重庆市秀山县流感特征的预测分析[J]. 统计学与应用, 2025, 14(11): 312-325. DOI: 10.12677/sa.2025.1411332

²Department of Communicable Disease Prevention and Control, Center for Disease Control and Prevention, Xiushan Tujia and Miao Autonomous County, Chongqing

Received: October 22, 2025; accepted: November 13, 2025; published: November 24, 2025

Abstract

Background: Influenza is an acute respiratory infectious disease that poses a serious threat to public health systems. Understanding regional influenza transmission patterns and accurately predicting incidence trends are essential for developing effective prevention and control strategies. **Objective:** To analyze the epidemiological characteristics and temporal trends of influenza incidence in Xiushan Tujia and Miao Autonomous County (Xiushan County), Chongqing, and to provide evidence-based support for regional influenza control. **Methods:** Weekly influenza case report data from the 1st week of 2018 to the 26th week of 2024 were analyzed in conjunction with meteorological data to characterize influenza patterns. The BEAST algorithm was applied to identify trend and seasonal change points, and a hybrid SARIMAX-AttLSTM model was developed to predict future influenza incidence. **Results:** A total of 35,057 influenza cases were reported, with males and females accounting for 52.7% and 47.3%, respectively ($\chi^2 = 53.8, P < 0.001$). The 1~9-year age group had the highest incidence (52.9%, $\chi^2 = 2957.0, P < 0.001$), with students and preschool children being the most affected populations. Influenza incidence showed clear seasonality, peaking in winter and spring. The BEAST algorithm identified four seasonal and five trend change points. The SARIMAX-AttLSTM hybrid model achieved the best performance (MAE = 8.064, RMSE = 11.724, $R^2 = 0.922$), and its accuracy improved further after incorporating temperature and rainfall as exogenous variables. The model predicted an upward trend in influenza incidence after the 27th week of 2024. **Conclusions:** Influenza in Xiushan County exhibits strong winter-spring seasonality, with children and students as the main susceptible groups. The SARIMAX-AttLSTM hybrid model effectively captures both seasonal cycles and short-term fluctuations, providing a valuable tool for early warning and public health decision-making.

Keywords

Influenza Incidence, SARIMAX-AttLSTM Model, BEAST Algorithm, Xiushan County, Prediction

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

流行性感冒(简称流感)是一种由流感病毒引起的急性呼吸道传染病,主要通过飞沫传播,该疾病具有传播速度快、季节性强等特点[1]。据世界卫生组织统计,全球每年约有 10 亿例季节性流感病例,其中 300 万至 500 万例为重症,每年因流感所导致的死亡人数大约 30~60 万例[2]。

流感也是我国重要的公共卫生问题。一项基于 82 个内陆城市五年监测数据的分析显示,流感病例占同期呼吸道疾病总病例的 72.7% [3]。全国每年因流感导致的超额死亡约 9.6 万~24 万例[4]。重庆市作为中国西南地区的重要直辖市,其流感负担长期处于较高水平。2023 年,全市流感发病率高达 1560.38/10 万[5]。国家法定传染病报告系统通过网络直报提高了传染病报告的及时性与完整性,增强了疫情早期处置能力[6]。然而,该系统主要反映疾病的趋势性变化,而社区和哨点医院调查数据更贴近实际情况,基

于调查的发病率通常高于法定报告数据[4] [7]。传染病报告系统的发病率可能会导致对流感疫情的低估，从而影响医疗资源配置和公共卫生应对。

本研究基于重庆市秀山土家族苗族自治县(简称秀山县)流感报告数据，分析并探究该区域流感的分布特征，并构建了基于重庆市秀山县的 SARIMAX-AttLSTM 混合模型，该模型可结合流感发病率与气象数据，捕捉流感季节性规律及短期波动，实现流感的早期预警与趋势预测。此外，通过 BEAST 算法模型对流感发病率变化点进行分析，可辅助疾控部门识别流行特征转折点、明确高风险时段，为优化监测体系、制定防控策略提供科学依据，从而提升区域流感防控能力。

2. 资料与方法

2.1. 数据来源与处理

本研究所用数据主要来源于中国流感监测信息系统，包括重庆市秀山县 2018 年第 1 周至 2024 年第 26 周的流感报告卡数据，共计 35,057 例，该数据经秀山县疾病预防控制中心审核整理，确保数据准确性与完整性。

$$\text{流感发病率} = \frac{\text{一定时期内报告的流感病例数}}{\text{同期地区常住人口数}} \times 10^5$$

常住人口数据由秀山县统计局提供，涵盖 2018 年至 2024 年按年龄分组的人口分布情况。标准人口结构使用重庆市第七次全国人口普查数据。为了计算分年龄组的流感发病率，本研究将各年龄组每周的发病人数除以相应周的常住人口数，并进一步采用直接标化法对年龄结构差异进行调整，最终得到每周的标化流感发病率。具体公式如下：

$$\text{标准化流感发病率} = \frac{\sum \left(\frac{a_i}{n_i} \times P_i^{std} \right)}{\sum P_i^{std}} \times 10^5$$

其中， a_i 为秀山县第 i 年龄别周流感病例数， n_i 为该组别常住人口数， P_i^{std} 为第 i 年龄别标准人口数。

本研究所使用的气象数据由重庆市气象局提供，包含 2018 年 1 月 1 日至 2024 年 6 月 30 日的秀山县国家气象观测站的日平均气温、日平均风速、日平均相对湿度及日累计降水量数据。气象站点监测范围覆盖全县区域，能较反映研究地区的整体气候状况。为保持与流感监测数据的时间粒度一致，将日气象数据按周聚合，采用算术平均法计算各变量的周均值，构建与流感数据时间分辨率匹配的周气象数据集(包括周均温度(WAT)、周均风速(WAW)、周累积降雨量(WCR)、周均相对湿度(WAH))。

2.2. 研究方法

本研究基于重庆市秀山县 2018 年第 1 周至 2024 年第 26 周流感报告卡数据进行基本特征描述，并使用 BEAST 算法模型对流感发病率时间序列进行变化点分析，识别季节性变化点与趋势性变化点。最后，结合流感的同期气象数据，在 SARIMAX-LSTM 混合模型的基础上引入注意力机制，构建 SARIMAX-AttLSTM 混合模型以提升预测模型的精度和泛化能力。本研究以 9:1 比例划分训练集与验证集，构建并比较了 SARIMAX、AttentionLSTM、SARIMAX-AttLSTM 混合模型。最后，采用平均绝对误差(Mean absolute error, MAE)、均方根误差(Root mean square error, RMSE)和决定系数(R^2)作为评估模型预测性能指标。

本研究在 Python (3.9.13) 和 R (4.3.2) 环境下完成数据处理、建模与分析，主要的工具包如下：statsmodels (0.14.0)、pandas (1.5.3)、numpy (1.25.0)、scikit-learn (1.2.2)、tensorflow (1.13.0)、keras (2.13.0)、matplotlib (3.7.1)。其中，用于流感发病率变化点分析工具包为 Rbeast (0.5.3)。

3. 结果

3.1. 2018~2024 年秀山县流感发病的基本特征

3.1.1. 人群特征

2018 年第 1 周~2024 年第 26 周, 秀山县共报告流感病例 35,057 例。其中, 性别分布显示, 男性 18,473 例, 女性 16,584 例, 男性和女性的构成比分别为 52.7%和 47.3%, 比例为 1.1:1。除 2024 年上半年外, 其余年度的男性病例数均多于女性病例数, 性别构成差异有统计学意义($\chi^2 = 53.8, P < 0.001$) (详见表 S1)。年龄分布显示, 1~9 岁年龄段的人群流感发病例数最高, 占比 52.9%, 其次是 10~19 岁年龄组(占 24.3%)。其中, 患者最小年龄为 4 天, 最大年龄为 96 岁。病例年龄组构成差异有统计学意义($\chi^2 = 2957.0, P < 0.001$) (见表 S1)。职业分布显示, 居前三位的职业分别为学生(15319 例, 占比 43.7%)、幼托儿童(7631 例, 占比 21.8%)、散居儿童(5278 例, 占比 15.1%)。除 2020 年和 2022 年散居儿童人群病例构成比居于首位外, 其余年份职业为学生的病例占比位于首位, 十年间的职业构成差异有统计学意义($\chi^2 = 4987.3, P < 0.001$) (见表 S1)。

3.1.2. 时间特征

2018 年第 1 周~2024 年第 26 周秀山县流感报告发病率总体呈上升趋势。其中, 流感发病率在 2018 年至 2019 年逐渐上升, 2020~2021 年有所回落, 2022 年再次上升, 2023 年达到最高峰值(19,326 例, 发病率 = 3104.09/10 万), 且在 2023 年第 11 周达近 6 年流感峰值(696.96/10 万)。从图 1 可以发现秀山县流感发病率高峰可能出现在每年冬季(第 47 周至次年第 8 周)至次年春季(第 9 周至第 20 周)。该地区的流感疫情具有明显季节性, 除 2022、2023 年外, 其余年份发病呈现典型冬季(第 47 周至次年第 8 周)高峰模式。2023 年呈现“双峰”特征: 流感流行全年呈现春季(第 9 周至第 20 周)主高峰, 冬季(第 47 周至次年第 8 周)次高峰的特点; 2022 年则流感流行表现为夏秋季(第 21 周至第 26 周)高峰, 流感流行高峰为 6 月, 其余月份疫情趋于平稳, 详见图 1。

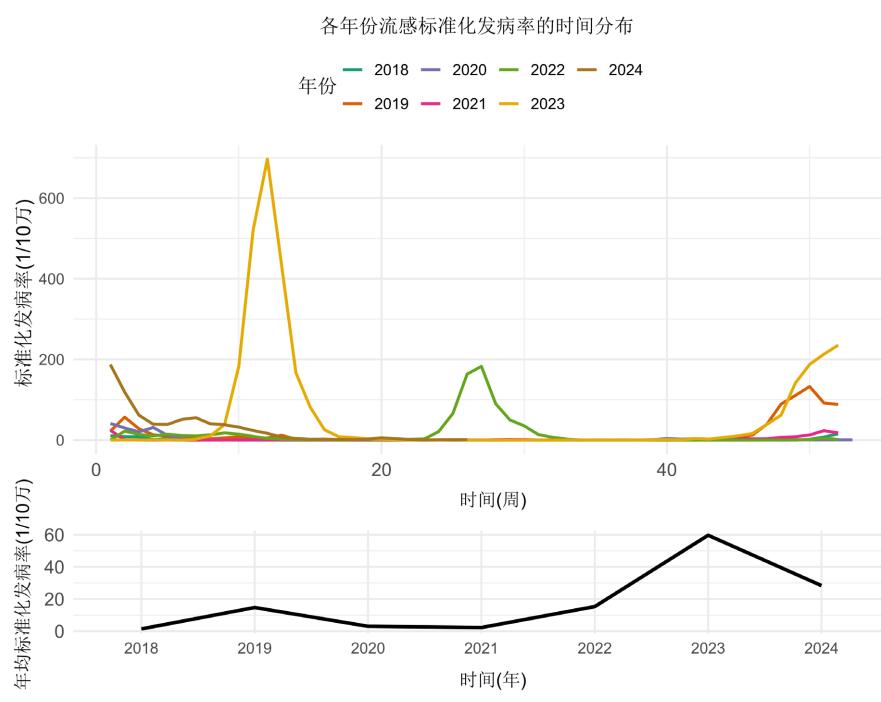


Figure 1. Temporal trend of influenza incidence in Xiushan County from 2018 to 2024
图 1. 2018~2024 年秀山县流感发病率时间趋势图

3.1.3. 地区特征

2018 年第 1 周~2024 年第 26 周, 秀山县各乡镇街道流感平均发病率在 282.86/10 万至 2000.95/10 万之间, 呈现显著地区差异($P<0.001$), 反映县域内流感流行强度不均衡。发病率居前三位的乡镇分别为中和街道、清溪场街道和乌杨街道。详见表 S2。

3.2. 使用 BEAST 算法模型的流感变化点检测

利用 BEAST 算法对秀山县流感发病率时间序列的变化点分析结果如图 2 所示。该方法将序列分解为季节性、趋势和误差分量, 识别各成分变化点位置并量化分布概率大小。图 2(a)为 2018 年第 1 周至 2024 年第 26 周秀山县周流感发病率时间序列图。该部分显示过去 6.5 年间, 流感发病率实际波动情况。可以看出该地区 2022、2023、2024 年均有较为明显的流感发病率波动, 其中 2023 年波动幅度最大。图 2(b)、图 2(c)分别显示了季节性成分及季节性成分中变化点的位置和概率。结果表明, 季节成分有 4 个明显变化点, 分别出现在 2020 年第 4 周、2022 年第 25 周、2023 年第 14 周、2024 年第 1 周, 对应的概率分别为 0.298、0.513、0.527、0.809 (见表 1)。这 4 个变化点将秀山县流感发病率序列的季节性因子划分为 5 个部分, 即 2018 年第 1 周至 2020 年第 4 周, 该地区有轻微流感季节波动。2020 年第 5 周至 2022 年第 24 周, 无明显没季节波动。而 2022 第 25 周后以及 2023 年和 2024 年的季节性波动较为明显。图 2(d)、2e 表示趋势分量及趋势分量中变化点的位置和概率。该地区在 2018 年第 1 周至 2024 年第 26 周期间共出现了 5 个变化点(2018 年第 46 周、2020 年第 1 周、2022 年第 26 周、2023 年第 11 周和 51 周), 每个趋势变化点的概率值分别为 0.398、0.998、0.635、0.557、0.523 (见表 1)。这 5 个趋势将序列划分为 6 个部分, 其中, 在 2018~2020 年期间, 流感发病率呈轻微上升趋势。而在 2020 年第 1 周至 2022 年第 25 周期间, 未观察到明显流感发病率趋势变化。从 2022 年第 26 周开始流感发病率有明显上升趋势, 其中在 2023 年第 11 周趋势上升幅度最大, 之后到 2023 年第 52 周趋势逐渐放缓。最后, 图 2(f)表示序列分解后的误差分量。

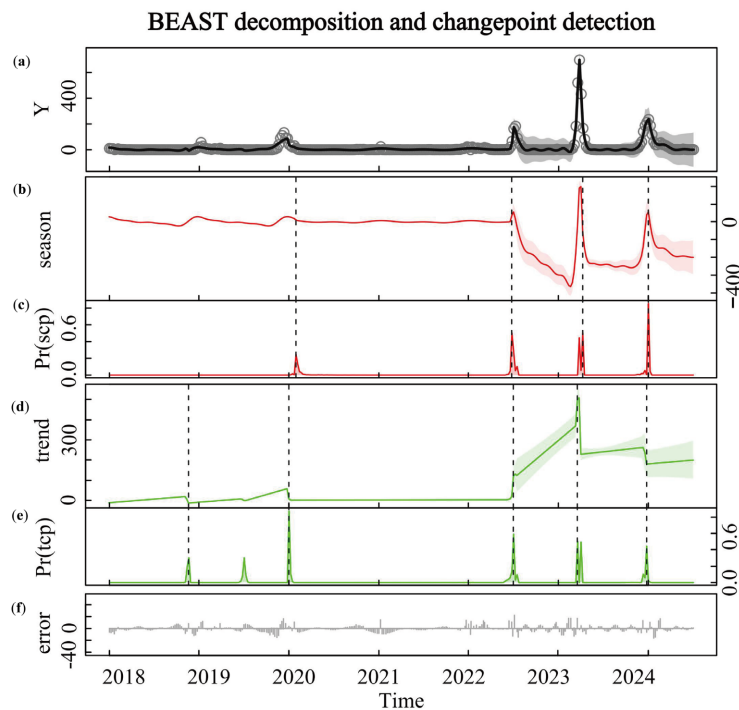


Figure 2. BEAST results of the influenza incidence time series in Xiushan County from 2018 to 2024
图 2. 2018~2024 年秀山县流感发病率时间序列 BEAST 结果

Table 1. Locations and probabilities of seasonal and trend change points
表 1. 季节与趋势变化点位置及概率

Scp ¹	Time_Scp ³	Pr_Scp ⁴	Tcp ²	Time_Tcp ³	Pr_Tcp ⁴
1	2020 (4 week)	0.298	1	2018 (46 week)	0.398
2	2022 (25 week)	0.513	2	2020 (1 week)	0.998
3	2023 (14 week)	0.527	3	2022 (26 week)	0.635
4	2024 (1 week)	0.809	4	2023 (11 week)	0.557
\	\	\	5	2023 (51 week)	0.523

¹ Scp 是季节性成分的变化点数量；² Tcp 是趋势成分的变化点数量；³ Time_表示各成分变化点位置；⁴ Pr_表示各个成分变化点的位置概率。

3.3. SARIMAX-AttLSTM 混合模型的流感预测

3.3.1. SARIMAX-AttLSTM 混合模型构建与比较

在对时间序列进行建模前，首先对流感发病率序列进行了平稳性检验，根据单位根结果显示，该序列整体上具有良好的平稳性($ADF = -5.899, P < 0.05$; $KPSS = 0.233, P = 0.10$)，满足建模要求。随后，结合自相关函数(ACF)与偏自相关函数(PACF)图(见图 3)，并采用网格搜索法以最小赤池信息准则(AIC)为模型筛选标准(模型参数筛选结果见表 S3)，最终选定 SARIMAX (2, 0, 2) (0, 1, 1) 52 为最优模型($AIC = 399.794$)。

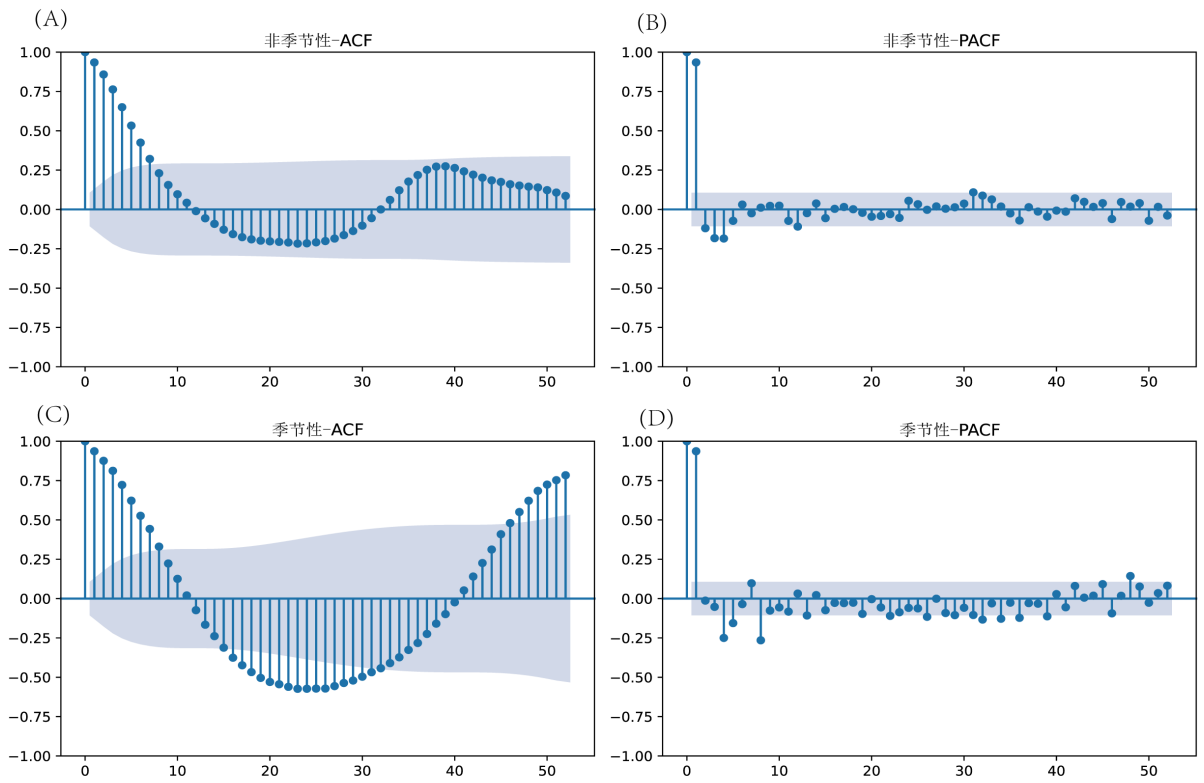


Figure 3. ACF and PACF results of the SARIMAX model: non-seasonal autocorrelation function (A); non-seasonal partial autocorrelation function (B); seasonal autocorrelation function (C); seasonal partial autocorrelation function (D)
图 3. SARIMAX 模型 ACF 与 PACF 结果：非季节性自相关函数图(A)；非季节性偏自相关函数图(B)；季节性自相关函数图(C)；季节性偏自相关函数图(D)

在构建嵌入注意力机制的 LSTM 模型及 SARIMAX-AttLSTM 混合模型中的神经网络部分, 本文统一采用双层隐藏层结构。每层包含 128 个神经元, 学习率设定为 0.001, 批量大小为 8, 最大训练轮数为 300。损失函数采用均方误差, 并引入 Dropout 机制以增强模型的泛化能力, 防止过拟合, 丢弃率设置为 0.2。

本研究将 SARIMAX 模型的预测结果与 Attention-LSTM 模型对残差的拟合结果融合, 构建 SARIMAX-AttLSTM 混合模型, 其整体拟合效果见图 4。对比三种模型的完整时间序列拟合趋势(图 4(A)、图 4(C)、图 4(E))及验证集 95%置信区间拟合趋势(图 4(B)、图 4(D)、图 4(F)), 发现混合模型在训练集拟合精度更优, 验证集泛化能力更强, 且能较准确捕捉发病率变化趋势(图 4(F))。相较而言, SARIMAX 模型与 Attention-LSTM 模型在部分波动区间存在偏差; 而混合模型通过整合 SARIMAX 模型处理季节性与短期依赖特征的能力, 及 Attention-LSTM 模型中注意力机制对不同时间滞后信息重要性的动态分配优势, 能更全面捕捉流感发病率的时间变化关键特征, 进而提升模型整体预测效果。

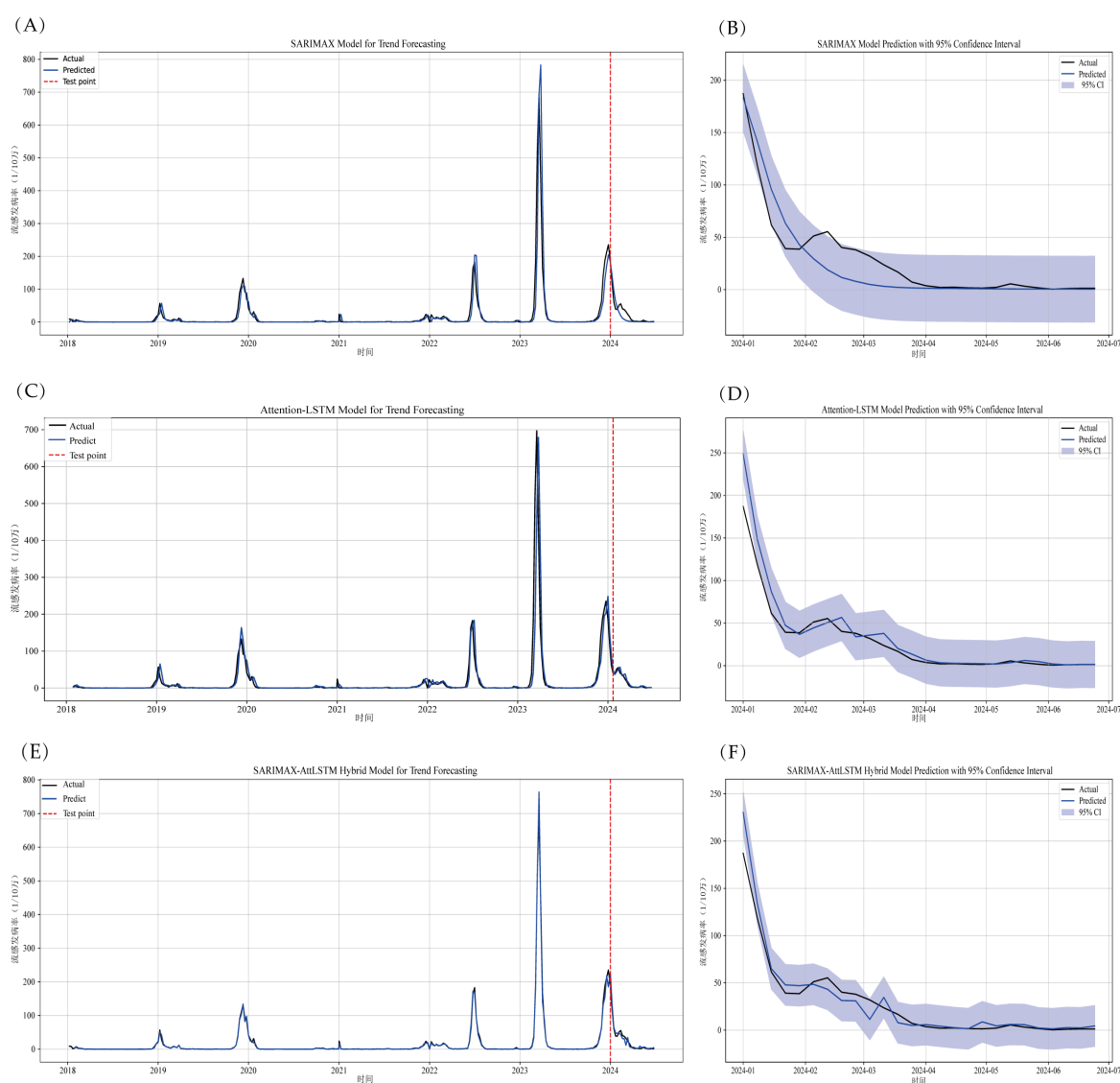


Figure 4. Comparison of the performance of SARIMAX, Attention-LSTM, and SARIMAX-AttLSTM models in fitting influenza incidence

图 4. SARIMAX、Attention-LSTM 与 SARIMAX-AttLSTM 模型在流感发病率拟合中的性能比较

3.3.2. 模型性能评估

通过对三种模型的预测性能进行量化评估(表 2)，发现 SARIMAX-AttLSTM 混合模型在预测流感发病率方面表现最优(MAE = 8.064, RMSE = 11.724, $R^2 = 0.922$)，该模型在误差控制与拟合精度方面具有显著优势。

此外，为评估气象因素对预测准确性的影响，本研究对比了各模型中不同气象因素组合的预测结果(见附表 S4)。结果显示：温度单独作为外生变量时，SARIMAX 模型与 Attention-LSTM 模型拟合效果更优；而温度与降雨量共同作为外生变量纳入时，混合模型拥有最佳预测性能。

Table 2. Comparison of evaluation metrics for the three models
表 2. 三个模型的评估指标对比

模型	MAE	RMSE	R^2
SARIMAX	7.149	28.453	0.846
AttLSTM	9.032	16.909	0.838
SARIMAX-AttLSTM	8.063	11.723	0.922

3.3.3. SARIMAX-AttLSTM 混合模型预测趋势

基于 2018 年第 1 周~2024 年第 26 周的流感发病率及气象数据，本研究利用训练完成的 SARIMAX-AttLSTM 混合模型，对 2024 年第 27 周~2025 年第 52 周(未来 26 周)的流感发病趋势进行预测。因该时段缺乏实测气象数据，研究采用季节性分解法模拟填补气温与降雨量，生成的预测数据作为外生变量输入模型。预测趋势见图 5，由图中红线可知，未来 26 周流感发病率呈明显上升趋势，提示可能出现流感疫情峰值。虽该模型无法精准预测各时点具体发病率，但在识别中短期流行趋势上仍具有重要参考价值。

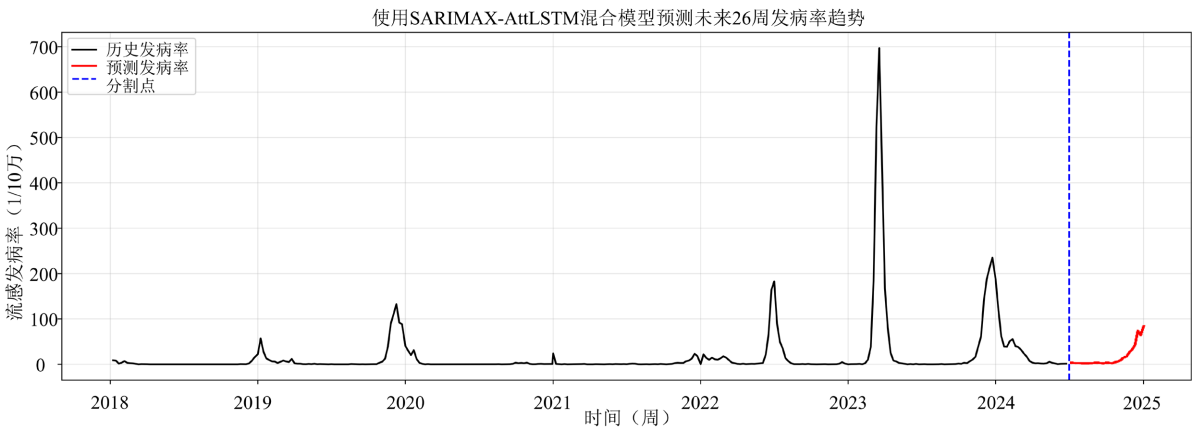


Figure 5. Forecast of influenza incidence for the next 26 weeks using the SARIMAX-AttLSTM hybrid model
图 5. 使用 SARIMAX-AttLSTM 混合模型对未来 26 周流感发病率的预测结果

3.4. 预测模型在流感监测与预警中的应用

为实现模型成果的公共卫生转化，本研究探讨了如何将 SARIMAX-AttLSTM 混合预测模型嵌入到当地疾病监测与预警工作流程中。首先，可将模型纳入秀山县疾控中心的常规流感监测系统中，利用模型对每周流感发病率进行实时预测，并与监测系统中报告的实际发病数据进行动态对比。其次，可根据模型预测值设定分级预警阈值：当预测值接近或超过历史同期发病率均值的一个标准差时，可判定为“轻度预警”；超过两个标准差时为“中度预警”；超过三个标准差则为“重度预警”。该分级预警机制可帮

助疾控部门在流感流行前期提前识别风险信号,及时启动针对性防控措施,如加强重点人群健康监测、储备防控物资及发布公众健康提示。此外,模型输出的趋势变化信息还可为上级卫生行政部门提供科学决策支持,从而构建“数据驱动-智能预测-主动干预”的流感监测预警体系。

4. 讨论

本研究基于 2018 年第 1 周至 2024 年第 26 周重庆市秀山县流感报告卡数据及同期气象数据,系统分析了流感的流行特征、时序变化点位置及概率,并基于注意力机制构建了流感发病率预测模型。结果显示,研究期间的流感发病率总体呈上升趋势,季节性特征明显,且存在显著的人群与地区差异。BEAST 模型识别出多处变化点,与疫情防控策略调整及气象条件波动密切相关。构建的 SARIMAX-AttLSTM 混合模型在短期预测中表现优异,拟合精度显著高于单一模型。

在人群特征方面,本研究通过对不同性别的流感阳性例数对比,发现男女性别比为 1.1:1,男性总阳性例数略高于女性,性别差异具有统计学意义,这与已有研究结果一致[8]。有研究表明,女性的免疫系统在某些方面更为活跃,考虑可能与雌激素对免疫细胞的调节作用相关,雄激素可能对免疫系统的调节作用较弱[9]。但该结果与上海市研究结果不一致[10],可能由流感毒株的差异导致,一般而言,甲型流感病毒在男性群体中比女性群体感染风险高,而乙型流感病毒在性别中无明显差异,而本研究主要患者感染甲型流感病毒为主[11]。与湖北省研究结果一致[12],本研究中 1~9 岁儿童为主要发病人群,其流感总阳性率高远高于成人,儿童免疫系统发育不全,对流感病毒抵抗力较弱可能是导致这一现象的主要原因之一[13]。同时学校、托幼机构的高聚集性也增加了传播风险[14]。提示校园和托幼机构是流感防控的重点场所,应加强疫苗接种与健康教育[15]。从时间分布来看,流感发病率在 2020~2021 年显著下降,可能与新冠疫情期间的口罩佩戴、社交距离及学校停课等综合防控措施有关[16];2022 年后流感发病率迅速回升,并于 2023 年达到峰值,提示防控措施放松后人群易感性增加及病毒重新传播[17]。空间分布上,各个乡镇/街道流感发病率差异具有统计学意义,其中发病率前三的为中和街道、石耶镇以及清溪场街道,分析原因,可能与地区人口密度相关,中心地区人口聚集风险大,增加病毒传播风险[18],该结果为流感分区分层防控工作提供了依据。

通过 BEAST 算法模型分析发现,在 2018 年第 1 周至 2024 年第 26 周期间,秀山县流感发病率呈现出五个明显的趋势性突变以及四个季节性结构突变。具体来看,流感发病率在 2020 年第 1 周趋势表现为下降。本研究推测可能与新冠疫情防控措施的全面实施有关。为遏制疫情传播,各地积极采取减少人员流动的措施,同时大力推广戴口罩、勤洗手等健康防护举措,这些措施有效降低了病毒传播风险[16]。在 2022 年第 26 周,流感发病率呈现上升趋势。考虑是由于新冠疫情进入后期阶段,防控力度有所减弱,人群防护意识降低,导致病毒传播机会增加[17]。此外,在 2022 年第 25 周,秀山县出现夏季流感高发期。流感哨点医院监测数据显示,该时段内流感病毒主要优势毒株为甲型 H3N2 型。由于该毒株在近几年流行强度较低,人群免疫水平普遍偏低,因此导致秀山县在该时期出现了夏季流感流行季。

模型比较结果表明,SARIMAX-AttLSTM 混合模型在拟合精度与预测能力方面均优于 SARIMAX 和 Attention-LSTM 单一模型,尤其在捕捉季节性波动方面表现更佳。混合模型通过引入时间维度的注意力机制,有效识别了不同滞后期对流感预测的相对重要性,其中 $t-1$ 时刻的发病率对当前预测影响最大,提示流感传播具有明显的时间依赖特征。模型分析还表明,秀山县流感具有明显的季节性规律,高发期集中于冬春季,与已有研究结果一致[19]。气象因素分析进一步发现,低温和少雨环境下流感发病风险显著升高,这可能与流感病毒在低温低湿环境中的存活延长及人群室内聚集增加有关[20],提示应基于气象条件优化流感预警与干预机制。最后,基于混合模型的未来 26 周预测结果拟合良好,显示出较高的实用价值。

然而,本研究仍存在一定局限。首先,受数据缺失的影响,本研究构建的模型未纳入空气质量、人群流动及疫苗接种率等关键影响因素,上述因素在流感传播中可能发挥着重要作用。例如,有研究表明,臭氧可能为流感的保护因素,而 PM2.5 与 PM10 则与流感传播风险的上升存在关联[21]。在节假日期间,集会、返乡、出游等人群流动可能会影响人群接触流感病毒的机会[22],本研究忽略了该因素可能会导致模型低估传播范围。另外,疫苗接种作为预防流感传播的重要公共卫生干预手段[23],在流感疫苗接种率较高的年份或地区,未纳入该因素可能会使模型高估流感发病率。上述这些情况都可能在一定程度上限制模型的解释力与外推性。其次,使用 SARIMAX-AttLSTM 混合模型对未来 26 周流感发病趋势进行预测时,因部分气象数据缺失而采用模拟值填补,可能会降低模型短期预测的精准性。最后,Attention-LSTM 模型在超参数设定上仍依赖专业经验,模型稳定性与泛化能力有待进一步提升。未来研究可融合搜索指数和空气质量、人群流动、疫苗接种等多源数据,并引入集成学习与自适应调参策略,以进一步提高预测的精度与稳定性。

5. 结论

本研究系统揭示了重庆市秀山县流感的时空特征与流行规律,并验证了 SARIMAX-Att LSTM 混合模型在短期预测中的精度与稳定性。研究结果为地方流感防控策略提供了数据支持,也为多源数据驱动的传染病预测模型研究提供了方法学参考。未来应进一步融合多维动态数据,构建实时更新的智能化流感监测与预警体系,以实现更精准的公共卫生决策支持。

致 谢

本研究承蒙重庆市秀山土家族苗族自治县疾病预防控制中心的大力支持,特此向该中心工作人员在数据收集过程中所给予的宝贵帮助致以诚挚谢意。同时,感谢重庆市气象局提供的气象数据,为本研究顺利开展提供了重要保障。

参考文献

- [1] 流行性感冒诊疗方案(2025 年版)全科医学临床与教育[J]. 2025, 23(7): 581-583+589.
- [2] WHO (2025) Influenza (Seasonal). [https://www.who.int/zh/news-room/fact-sheets/detail/influenza-\(seasonal\)](https://www.who.int/zh/news-room/fact-sheets/detail/influenza-(seasonal))
- [3] Yu, L.J., Li, X.L., Wang, Y.H., et al. (2023) Short-term Exposure to Ambient Air Pollution and Influenza: A Multicity Study in China. *Environmental Health Perspectives*, **131**, Article 127010. <https://doi.org/10.1289/ehp.12146>
- [4] 李文娟, 王大燕. 我国流感疾病负担相关研究进展[J]. 中国人兽共患病学报, 2019, 35(10): 928-933.
- [5] 重庆市卫生健康委. 重庆市 2023 年度居民健康状况报告[R]. 2024. https://wsjkw.cq.gov.cn/zwgk_242/wsjklymsxx/jkfw_266458/gzxx_266460/202412/t20241211_13881316.html
- [6] 张洪龙, 曾令佳, 赖圣杰, 等. 2016 年国家传染病自动预警信息系统运行情况分析[J]. 疾病监测, 2018, 33(2): 159-167.
- [7] 陈长华, 谢金水. 福建省沙县医院法定传染病网络直报数据与实际卡片数据差异调查报告[J]. 疾病监测, 2006(11): 613+616.
- [8] 李静怡. 2018-2024 年河源地区流感流行病学分析[J]. 实验室检测, 2025, 3(14): 248-250.
- [9] Creisher, P.S., Seddu, K., Mueller, A.L. and Klein, S.L. (2023) Biological Sex and Pregnancy Affect Influenza Pathogenesis and Vaccination. In: *Current Topics in Microbiology and Immunology*, Springer, 111-137. https://doi.org/10.1007/978-3-031-35139-6_5
- [10] 陈婉琦, 马飞飞, 王广鹤, 等. 2013-2023 年上海市嘉定区流行性感流行病学分析[J]. 上海预防医学, 2024, 36(12): 1130-1136.
- [11] Karolyi, M., Pawelka, E., Kelani, H., Funk, G.C., Lindner, B., Porpaczy, C., et al. (2021) Gender Differences and Influenza-Associated Mortality in Hospitalized Influenza a Patients during the 2018/19 Season. *Infection*, **49**, 103-110. <https://doi.org/10.1007/s15010-020-01537-x>

- [12] 葛胜辉, 石玉琴, 张玲, 等. 湖北省 2020-2024 年流行性感冒流行病学特征分析[J]. 现代预防医学, 2025, 52(13): 2349-2353.
- [13] Zheng, K., Zeng, Z., Tian, Q., Huang, J., Zhong, Q. and Huo, X. (2023) Epidemiological Evidence for the Effect of Environmental Heavy Metal Exposure on the Immune System in Children. *Science of the Total Environment*, **868**, Article 161691. <https://doi.org/10.1016/j.scitotenv.2023.161691>
- [14] Kajjume, T., Mukai, S., Toyota, N., Kanazawa, I., Kato, A., Akimoto, E., *et al.* (2022) Effectiveness of Seasonal Influenza Vaccine in Elementary and Middle Schools: A 10-Year Follow-Up Investigation. *BMC Infectious Diseases*, **22**, Article No. 909. <https://doi.org/10.1186/s12879-022-07898-y>
- [15] 李小燕, 解春宝. 四川某医院 2019-2024 年儿童流感的流行病学特征[J]. 中华医院感染学杂志, 2025(22): 3431-3435.
- [16] Cai, W., Deng, Y., Xu, P. and Xue, J. (2024) Molecular Epidemiological Analysis of Influenza Viruses in Influenza-Like Illness Cases: A Retrospective Study in Chongqing Hi-Tech Zone, China (2021-2024). *Virology Journal*, **21**, Article No. 340. <https://doi.org/10.1186/s12985-024-02594-6>
- [17] Li, Z., Xiong, Y., Long, J., Li, T., Fu, X., Yang, S., *et al.* (2023) Resurgence of Influenza during COVID-19 in Chongqing, China: A Retrospective Analysis. *Journal of Medical Virology*, **95**, e29249. <https://doi.org/10.1002/jmv.29249>
- [18] Yang, J., Chen, Q., Zheng, X., Sun, A., Sun, M., Zhou, Q., *et al.* (2025) Comparative Analysis of Epidemiological and Spatiotemporal Patterns in Seasonal Influenza and COVID-19 Outbreaks. *Scientific Reports*, **15**, Article No. 8602. <https://doi.org/10.1038/s41598-025-93372-z>
- [19] Zhu, A.Q., Li, Z.J. and Zhang, H.J. (2023) Spatial Timing of Circulating Seasonal Influenza a and B Viruses in China from 2014 to 2018. *Scientific Reports*, **13**, Article No. 7149. <https://doi.org/10.1038/s41598-023-33726-7>
- [20] Ruan, W., Liang, Y., Sun, Z. and An, X. (2025) Climate Warming and Influenza Dynamics: The Modulating Effects of Seasonal Temperature Increases on Epidemic Patterns. *npj Climate and Atmospheric Science*, **8**, Article No. 72. <https://doi.org/10.1038/s41612-025-00968-3>
- [21] Tao, Y.Y., Li, J.X., Hu, Y.M., *et al.* (2021) Quadrivalent Influenza Vaccine (Sinovac Biotech) for Seasonal Influenza Prophylaxis. *Expert Review of Vaccines*, **20**, 1-11. <https://doi.org/10.1080/14760584.2021.1875823>
- [22] Ishola, D.A. and Phin, N. (2011) Could Influenza Transmission Be Reduced by Restricting Mass Gatherings? Towards an Evidence-Based Policy Framework. *Journal of Epidemiology and Global Health*, **1**, 33-60. <https://doi.org/10.1016/j.jegh.2011.06.004>
- [23] Tanner, A.R., Dorey, R.B., Brendish, N.J. and Clark, T.W. (2021) Influenza Vaccination: Protecting the Most Vulnerable. *European Respiratory Review*, **30**, Article 200258. <https://doi.org/10.1183/16000617.0258-2020>

附 录

Table S1. Annual distribution of influenza cases in Xiushan County, Chongqing (2018~2024) by year, sex, age group, and occupation

表 S1. 2018 年~2024 年重庆市秀山县流感病例分年份、性别、年龄组职业分布统计

变量	2018	2019	2020	2021	2022	2023	2024	合计	χ^2	P 值
性别									53.8	<0.001
男	252 (58.1)	2616 (55.0)	424 (50.0)	321 (57.3)	2553 (52.1)	10254 (53.1)	2053 (48.5)	18473 (52.7)		
女	182 (41.9)	2142 (45.0)	424 (50.0)	239 (42.7)	2345 (47.9)	9072 (46.9)	2180 (51.5)	16584 (47.3)		
年龄(岁)									2957.0	<0.001
<1	0 (0.0)	152 (3.2)	36 (4.2)	12 (2.1)	262 (5.3)	404 (2.1)	229 (5.4)	1095 (3.1)		
1~9	100 (23.0)	2588 (54.4)	423 (49.9)	297 (53.0)	2660 (54.3)	10549 (54.6)	1929 (45.6)	18546 (52.9)		
10~19	217 (50.0)	1500 (31.5)	111 (13.1)	165 (29.5)	527 (10.8)	5477 (28.3)	511 (12.1)	8508 (24.3)		
20~29	26 (6.0)	212 (4.5)	122 (14.4)	33 (5.9)	437 (8.9)	827 (4.3)	487 (11.5)	2144 (6.1)		
30~39	21 (4.8)	145 (3.0)	86 (10.1)	23 (4.1)	412 (8.4)	815 (4.2)	376 (8.9)	1878 (5.4)		
40~49	19 (4.4)	53 (1.1)	24 (2.8)	11 (2.0)	171 (3.5)	375 (1.9)	171 (4.0)	824 (2.4)		
50~59	25 (5.8)	47 (1.0)	14 (1.7)	5 (0.9)	179 (3.7)	417 (2.2)	208 (4.9)	895 (2.6)		
60~69	16 (3.7)	33 (0.7)	15 (1.8)	5 (0.9)	82 (1.7)	188 (1.0)	131 (3.1)	470 (1.3)		
70~79	8 (1.8)	20 (0.4)	11 (1.3)	8 (1.4)	116 (2.4)	207 (1.1)	137 (3.2)	507 (1.4)		
≥80	2 (0.5)	8 (0.2)	6 (0.7)	1 (0.2)	52 (1.1)	67 (0.3)	54 (1.3)	190 (0.5)		
职业										
学生	3 (0.7)	763 (16.0)	157 (18.5)	90 (16.1)	1141 (23.3)	4658 (24.1)	819 (19.3)	7631 (21.8)	4987.3	<0.001
散居儿童	8 (1.8)	632 (13.3)	256 (30.2)	71 (12.7)	1367 (27.9)	2060 (10.7)	884 (20.9)	5278 (15.1)		
幼托儿童	307 (70.7)	2854 (60.0)	151 (17.8)	316 (56.4)	948 (19.4)	9780 (50.6)	963 (22.7)	15319 (43.7)		
农民	79 (18.2)	301 (6.3)	176 (20.8)	50 (8.9)	821 (16.8)	1181 (6.1)	845 (20.0)	3453 (9.8)		
家务及待业	4 (0.9)	12 (0.3)	1 (0.1)	4 (0.7)	33 (0.7)	86 (0.4)	32 (0.8)	172 (0.5)		
离退休人员	7 (1.6)	71 (1.5)	11 (1.3)	6 (1.1)	74 (1.5)	360 (1.9)	71 (1.7)	600 (1.7)		
干部职工	6 (1.4)	21 (0.4)	11 (1.3)	0 (0.0)	59 (1.2)	104 (0.5)	28 (0.7)	229 (0.7)		
教师	15 (3.5)	31 (0.7)	19 (2.2)	4 (0.7)	114 (2.3)	267 (1.4)	97 (2.3)	547 (1.6)		
工人	1 (0.2)	46 (1.0)	56 (6.6)	9 (1.6)	224 (4.6)	538 (2.8)	345 (8.2)	1219 (3.5)		
医务人员	2 (0.5)	7 (0.1)	2 (0.2)	1 (0.2)	33 (0.7)	63 (0.3)	33 (0.8)	141 (0.4)		
其他	2 (0.5)	20 (0.4)	8 (0.9)	9 (1.6)	84 (1.7)	229 (1.2)	116 (2.7)	468 (1.3)		

Table S2. Influenza incidence in towns and subdistricts of Xiushan County, 2018~2024
表 S2. 2018 年~2024 年秀山县各乡镇街道流感发病情况

地区	2018	2019	2020	2021	2022	2023	2024	总例数	年均发病率 (1/10 万)
中和街道	88	1943	272	156	1751	7351	1270	12831	2000.95
乌杨街道	37	449	73	33	461	2112	405	3570	1138.94
平凯街道	41	395	85	32	409	1908	390	3260	1084.58
清溪场街道	50	312	73	143	667	2280	541	4066	1404.66
溶溪镇	16	97	14	16	219	818	151	1331	774.45
官庄街道	30	174	43	6	78	272	107	710	631.35
龙池镇	13	250	31	11	232	707	108	1352	822.55
石堤镇	6	58	19	1	46	133	32	295	402.24
洪安镇	11	69	19	4	58	237	88	486	405.21
石耶镇	7	65	14	70	167	383	99	805	1416.22
梅江镇	13	136	39	5	115	495	148	951	432.59
峨溶镇	2	48	16	1	44	134	59	304	282.86
雅江镇	2	49	18	7	33	133	31	273	360.53
隘口镇	6	81	14	4	66	254	127	552	603.53
兰桥镇	37	45	15	6	34	199	46	382	460.91
钟灵镇	6	90	19	3	54	295	138	605	578.01
膏田镇	3	42	6	1	38	164	83	337	646.82
溪口镇	6	95	15	1	72	279	82	550	818.34
妙泉镇	5	62	9	3	41	113	33	266	696.48
宋农镇	15	28	5	3	42	110	26	229	695.06
里仁镇	4	39	5	9	70	138	51	316	588.06
涌洞镇	2	46	10	1	33	149	49	290	823.58
中平乡	2	26	6	20	30	233	29	346	903.62
岑溪乡	4	55	11	0	29	127	27	253	751.61
海洋乡	19	17	7	13	17	36	19	128	369.35
龙凤坝镇	0	0	0	0	53	209	81	343	421.61
大溪乡	4	31	3	1	10	39	7	95	320.39

Table S3. Selected parameter results for the SARIMAX model
表 S3. SARIMAX 模型部分参数筛选结果

参数	MAE	RMSE	AIC
(1, 0, 0) (1, 0, 0) _{s2}	8.331	33.333	451.421
(1, 0, 1) (1, 0, 0) _{s2}	7.983	31.601	451.773
(1, 0, 1) (1, 0, 1) _{s2}	8.008	31.669	451.748
(2, 0, 0) (1, 0, 0) _{s2}	7.854	30.831	449.604

续表

(2, 0, 0) (1, 0, 1) ₅₂	7.832	30.764	450.868
(2, 0, 1) (1, 0, 0) ₅₂	7.684	30.176	441.858
(2, 0, 1) (1, 0, 1) ₅₂	7.499	29.981	443.291
(2, 0, 1) (0, 0, 1) ₅₂	7.546	30.059	441.191
(2, 0, 2) (0, 0, 1) ₅₂	7.429	32.210	432.861
(2, 0, 2) (0, 1, 1) ₅₂	7.149	28.453	399.794

Table S4. Comparison of meteorological factors in the SARIMAX, Attention-LSTM and SARIMAX-AttLSTM models
表 S4. SARIMAX、Attention-LSTM 和 SARIMAX-AttLSTM 模型气象因素比较

模型	MAE	RMSE
SARIMAX (WAT)	7.018	27.771
SARIMAX (WCR)	7.213	28.428
SARIMAX (WAT, WCR)	7.149	28.453
Attention-LSTM (WAT)	8.576	15.939
Attention-LSTM (WCR)	11.529	22.119
Attention-LSTM (WAT, WCR)	9.032	16.909
SARIMAX-AttLSTM (WAT)	11.205	16.547
SARIMAX-AttLSTM (WCR)	10.778	17.245
SARIMAX-AttLSTM (WAT, WCR)	8.063	11.723