Forecasting Respiratory Syncytial Virus Activity Using Seasonal ARIMA Models: Evidence from CDC Weekly Surveillance Data

Yiwen Yang *

School of Economics, University of Edinburgh, Edinburgh, United Kingdom

* Corresponding Author Email: sakura_yang2023@163.com

Abstract. Respiratory syncytial virus (RSV) imposes seasonal operational pressure on paediatric care and hospital healthcare systems. This study evaluates whether weekly RSV test percent positivity data from the Centers for Disease Control and Prevention (CDC) National Respiratory and Enteric Virus Surveillance System (NREVSS) can support short-term forecasting of RSV transmission activity. The national weekly surveillance dataset spanning 3 January 2015 to 27 June 2020 was analyzed with four forecasting approaches: seasonal autoreggressive integrated moving average (SARIMA), non-seasonal autoregressive integrated moving average (ARIMA), exponential smoothing (ETS), and a seasonal naive benchmark model. Model predictive performance was validated via a 52-week chronological holdout sample and rolling-origin forecasts for 1–4 week ahead horizons. Multiple evaluation metrics were adopted, including root mean square error (RMSE), mean absolute error (MAE), symmetric mean absolute percentage error (sMAPE), directional accuracy, and peak-timing error. Exploratory time-series analyses verified prominent annual seasonality and short-term serial dependence embedded in RSV percent positivity time series. Among all candidate models, SARIMA yielded the smallest RMSE and MAE on the holdout dataset, and exhibited relatively robust performance as the forecast window extended. The overall findings demonstrate that surveillance-driven SARIMA forecasting can facilitate early epidemic warning and advance healthcare resource planning. Nevertheless, further external validation using post-pandemic and regional RSV surveillance datasets is required, as the sample only covers the pre-pandemic and early coronavirus disease 2019 (COVID-19) period.

Keywords: Respiratory syncytial virus, SARIMA, time-series forecasting, public health surveillance, percent positivity.

1. Introduction

1.1. Research Background

Respiratory syncytial virus (RSV) is a major cause of acute lower respiratory infection in young children and can cause severe disease among clinically vulnerable groups [1, 2]. Seasonal outbreaks increase demand for paediatric consultations, emergency care, beds, testing, and staff. Because activity can rise rapidly, forecasts issued before the seasonal peak may give healthcare providers time to prepare.

RSV activity in the United States usually follows an annual cycle, but peak timing and magnitude vary between seasons [3]. This study uses weekly reports from the Centers for Disease Control and Prevention (CDC) National Respiratory and Enteric Virus Surveillance System (NREVSS) laboratory reports [4]. Percent positivity is treated as a relative indicator of viral activity rather than population prevalence, and its strong seasonal structure makes it suitable for seasonal time-series modelling.

1.2. Research Problem and Research Questions

Historical surveillance describes RSV activity after it occurs, whereas operational planning requires prospective information. The research problem is whether routinely reported weekly percent positivity can provide sufficiently accurate short-term forecasts despite year-to-year variation and changes in testing or reporting.

Using national CDC NREVSS data from 2015 to 2020, the study compares seasonal autoregressive integrated moving average (SARIMA) with seasonal naive, exponential smoothing, and non-seasonal autoregressive integrated moving average (ARIMA) models under a common training and evaluation design.

The study addresses three questions:

(1) Does weekly RSV percent positivity exhibit a clear and identifiable seasonal structure?

(2) Does SARIMA provide more accurate short-term forecasts than the benchmark models?

(3) To what extent do the forecasting results indicate potential value for RSV early warning and healthcare resource planning?

The hypothesis is that modelling annual seasonality will allow SARIMA to outperform non-seasonal ARIMA and match or modestly improve on simpler seasonal benchmarks. The study therefore tests the added forecasting value of explicit seasonal structure for public-health surveillance.

2. Literature Review

2.1. RSV Seasonality and Laboratory-based Surveillance

Zhang et al. used ARIMA, generalized additive, and least absolute shrinkage and selection operator (LASSO)-based models to study paediatric RSV infections and identified marked seasonal peaks, with meteorological factors contributing to prediction [5]. Thongpan et al. similarly found recurrent annual RSV positivity in Thailand and successfully applied seasonal ARIMA models [6].

Together, these studies show that RSV contains strong temporal dependence, although epidemic timing differs across geographical and climatic settings. Their emphasis was mainly on seasonal description and associated predictors rather than standardized out-of-sample comparison.

Seasonality is also not fixed. US evidence from 2021-2023 and paediatric studies during and after COVID-19 documented disrupted timing and burden across regions, highlighting the need to test forecasts on genuinely unseen observations [3, 7, 8].

2.2. Forecasting Clinical Burden and Seasonal Peaks

Cutrera et al. modelled laboratory-confirmed infections and RSV-related hospitalizations in a large Italian paediatric hospital and developed a SARIMA forecast for the next seasonal peak [9]. This demonstrated the potential operational value of seasonal models for preparing paediatric services.

However, hospitalizations and laboratory positivity represent different outcomes. Admissions directly reflect severe burden but are affected by clinical practice and may lag viral circulation. Percent positivity is not prevalence, yet it may provide an earlier surveillance signal before hospital demand rises.

2.3. Machine-learning Forecasting and the Scope of the Present Study

Recent studies have applied automated machine learning and neural networks to respiratory-disease hospitalizations. Albrecht et al. reported lower errors for some machine-learning models than for seasonal naive methods, although the contribution of laboratory-virus predictors varied [10]. Morbey et al. found that low average error did not necessarily imply accurate peak timing and that atypical seasons challenged models dependent on historical patterns [11].

The present study addresses a narrower question using one routinely available univariate series: whether explicit annual seasonality improves forecasting under equal information conditions. Machine-learning methods require different predictors and tuning procedures, so the direct comparison is limited to interpretable seasonal naive, ETS, ARIMA, and SARIMA models.

2.4. Research Gaps and Contribution of the Present Study

Four gaps motivate the analysis. First, in-sample fit does not demonstrate forecasting ability, so the study combines a chronological holdout with rolling-origin evaluation for one- to four-week horizons.

Second, previous studies use non-equivalent outcomes. This study focuses on weekly percent positivity and interprets it only as an early laboratory-surveillance indicator.

Third, seasonal and non-seasonal statistical models are compared on the same data, horizon, and evaluation procedure, allowing the value of SARIMA's seasonal structure to be isolated.

Finally, root mean square error (RMSE) and mean absolute error (MAE) are supplemented with directional accuracy and peak-timing error. These measures assess whether forecasts identify weekly movement and the seasonal maximum, linking statistical accuracy to early warning.

The contribution is therefore a transparent test of whether SARIMA adds practical forecasting value over simpler interpretable benchmarks, rather than a claim of universal model superiority.

3. Methodology

3.1. Outcome Variable

Weekly RSV percent positivity was calculated by summing positive detections across the included regions and diagnostic categories, dividing by the corresponding summed number of tests reported to CDC NREVSS, and multiplying by 100. It partially adjusts for testing volume and provides a readily interpretable indicator of reported viral activity.

Because positivity remains affected by laboratory participation, testing behavior, and reporting, it was not interpreted as population prevalence. The national weekly series used an annual seasonal frequency of approximately 52 weeks.

3.2. Data Preprocessing

Data were checked for duplicate or inconsistent weeks, missing observations, invalid values, and positive counts exceeding total tests. The final series contained no missing weeks, so interpolation was unnecessary.

Extreme observations corresponding to plausible seasonal epidemic peaks were retained. Observations would have been removed or corrected only if they represented duplicate weeks, impossible values, or internally inconsistent test counts.

The original percentage scale was retained to preserve public-health interpretability; no transformation was applied because the series did not show a clear increase in variance with its level. Early-2020 observations were interpreted cautiously because testing and healthcare-seeking behavior may have changed.

3.3. Forecasting Models

Four models were estimated under the same framework: seasonal naive, exponential smoothing (ETS), non-seasonal ARIMA, and SARIMA. This design separates the contribution of annual seasonality from short-term autocorrelation while preserving interpretability.

Seasonal naive used the value from the same week one year earlier. ETS modelled level, trend, and seasonality; ARIMA captured non-seasonal dependence; and $SARIMA(p,d,q)(P,D,Q)[52]$ combined short-term dynamics with a 52-week cycle.

Trigonometric seasonality, Box-Cox transformation, autoregressive and moving-average errors, trend and seasonal components (TBATS), Prophet, and machine-learning models were excluded because the dataset contained one relatively short univariate series with one dominant seasonal cycle. Adding substantially more complex models would weaken comparability and increase overfitting risk.

3.4. Model Identification and Estimation

Time plots, seasonal decomposition, the autocorrelation function (ACF), and the partial autocorrelation function (PACF) were examined for trend, variance, seasonality, and unusual observations. Conventional stationarity was assessed with Augmented Dickey-Fuller (ADF) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests.

ADF and KPSS do not directly test whether seasonal differencing is required. Despite non-seasonal stationarity, the seasonal plot, seasonal and trend decomposition using logically estimated scatterplot smoothing (STL), and ACF showed strong annual recurrence. The forecast package's seasonal-strength procedure selected one lag-52 seasonal difference, while ordinary differencing remained $d = 0$. This avoided unnecessary differencing and retained interpretable low-frequency information.

Candidate ARIMA and SARIMA orders were identified from ACF/PACF patterns and compared using the corrected Akaike information criterion (AICc), with automated selection and nearby simpler specifications examined. Parameters were estimated by maximum likelihood. Residual plots, residual ACF through the annual seasonal lag, and a lag-24 Ljung-Box test assessed adequacy [12, 13]. The Ljung-Box test targeted short- to medium-term residual dependence, while the extended residual ACF was used to inspect remaining annual autocorrelation. Analysis used R/RStudio with forecast, tseries, tidyverse, lubridate, and zoo.

3.5. Experimental Setup and Evaluation

The first 235 observations formed the training sample and the final 52 weeks formed a chronological holdout, preventing information leakage and testing performance across one unseen seasonal cycle. This fixed holdout assessed performance over a complete cycle rather than operational short-horizon accuracy.

Operational accuracy was evaluated separately through expanding-window rolling origins for forecasts one to four weeks ahead. Origins advanced one week at a time across the holdout wherever the full forecast horizon was observable. At each origin, the previously selected model specifications were re-estimated using only information available at that date.

Accuracy measures included mean absolute error (MAE), root mean square error (RMSE), and symmetric mean absolute percentage error (sMAPE). Directional accuracy was the proportion of forecast weeks for which the signs of predicted and observed week-to-week changes were identical. Peak-timing error was the absolute difference between the weeks of maximum predicted and observed positivity in the holdout.

Robustness was judged from model rankings across the fixed holdout, repeated origins, and increasing horizons. A reliable model should maintain its relative performance as the horizon lengthens.

4. Dataset

The study used public weekly RSV laboratory-surveillance data from CDC NREVSS, which receives reports from participating US clinical, public-health, and commercial laboratories. Participation and testing practices may vary over time. Table 1 summarizes the dataset characteristics and the experimental design.

The national series aggregated RSV detections and total tests across US Department of Health and Human Services (HHS) regions and diagnostic test types. National weekly percent positivity was computed from the ratio of aggregated positive detections to aggregated tests. The series covered 3 January 2015 to 27 June 2020, comprising 287 weekly observations; the main outcome was percent positivity.

Table 1. Summary of the dataset and experimental design

Dataset component	Specification
Data source	CDC NREVSS RSV laboratory-surveillance data
Frequency	Weekly
Geographical level	United States national level
Study period	3 January 2015–27 June 2020
Number of observations	287 weeks
Primary variable	Weekly RSV percent positivity
Supporting variables	Total tests and positive detections
Seasonal period	Approximately 52 weeks
Test set	Final 52 weeks
Forecast horizons	1–4 weeks ahead
Software	R/RStudio; forecast, tseries, tidyverse, lubridate, and zoo

NREVSS reflects reported laboratory activity rather than population prevalence, hospitalization, severity, or demographic risk. Forecasts were therefore interpreted as predictions of reported RSV activity. Changes in laboratory participation, testing behavior, and the composition of diagnostic test types may affect comparability over time. The data were public, aggregated, and de-identified, so individual consent and institutional ethical approval were not required.

5. Experiments and Results

5.1. Exploratory Data Analysis

The national series comprised 287 weekly observations, with the last 52 reserved for testing. Fig. 1 shows repeated peaks and the chronological holdout.

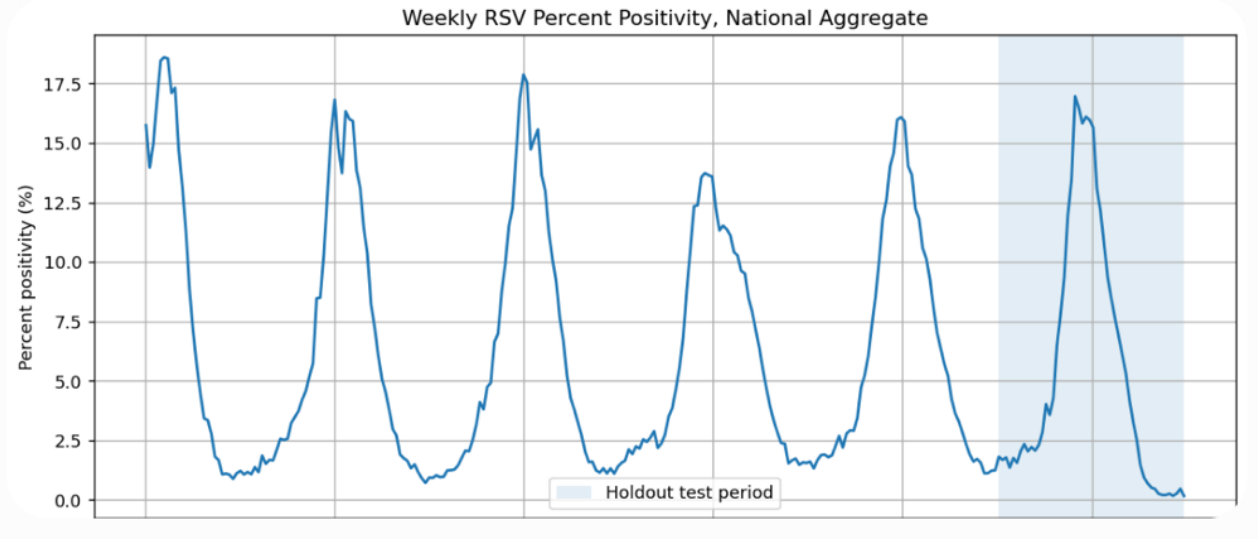


Fig. 1 Weekly RSV percent positivity, national aggregate, 2015–2020 (Picture credit: Original)

RSV positivity rose sharply during winter and declined after each peak. Although the annual pattern was clear, peak magnitude and timing varied, supporting a model that combines seasonal recurrence with short-term updating. Fig. 2 displays the seasonal profiles by week and the average seasonal pattern.

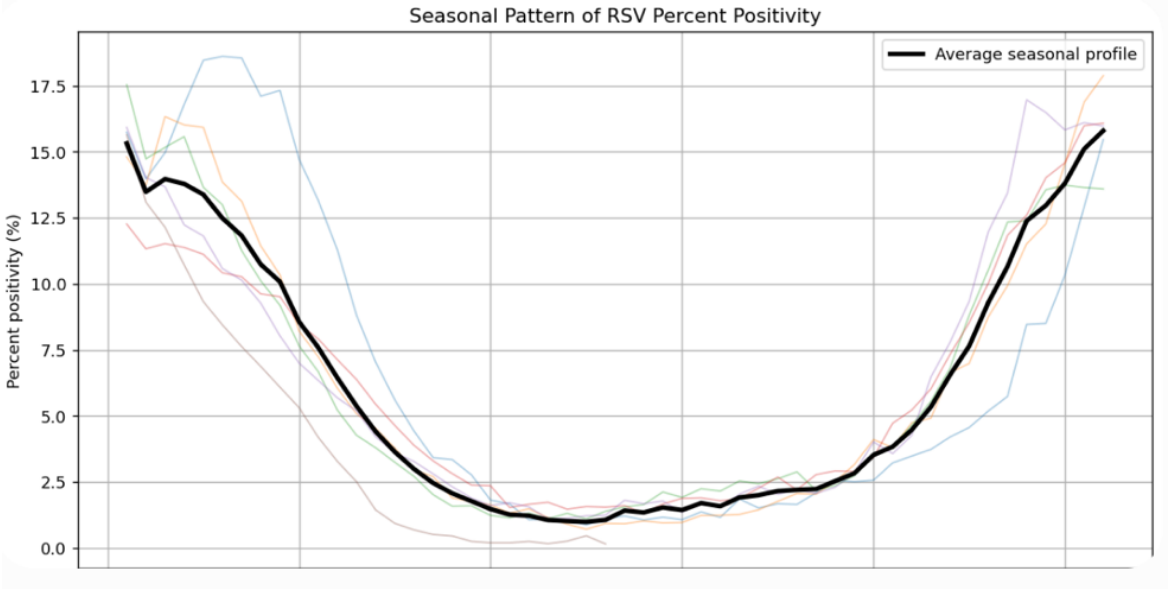


Fig. 2 Seasonal pattern of weekly RSV percent positivity (Picture credit: Original)

Table 2 reports calendar-year maxima for description only. Because an RSV season spans two calendar years, adjacent rows may represent the same epidemic season. Maxima occurred from late November to early February.

Table 2. Calendar-year maxima of RSV percent positivity

Calendar year	Peak week ending date	Calendar week number	Peak percent positivity (%)
2015	2015-02-07	6	18.61
2016	2016-12-31	52	17.88
2017	2017-01-07	1	17.54
2018	2018-12-29	52	16.08
2019	2019-11-30	48	16.96
2020	2020-01-04	1	15.64

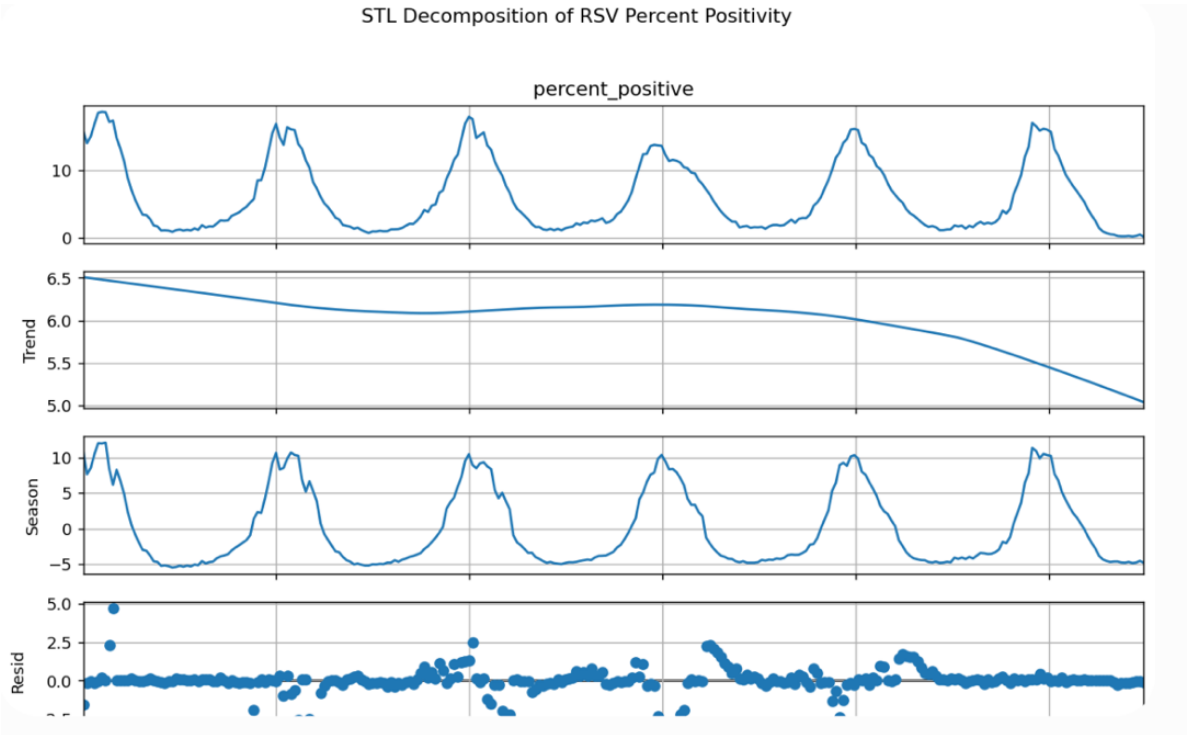


Fig. 3 STL decomposition of weekly RSV percent positivity (Picture credit: Original)

The seasonal plot and STL decomposition confirm a persistent annual component. The ACF shows an oscillating decay and strong correlation near lag 52, while the PACF has large early-lag spikes, indicating both seasonality and short-term dependence. Fig. 3 presents the STL decomposition, and Fig. 4 shows the corresponding ACF and PACF patterns.

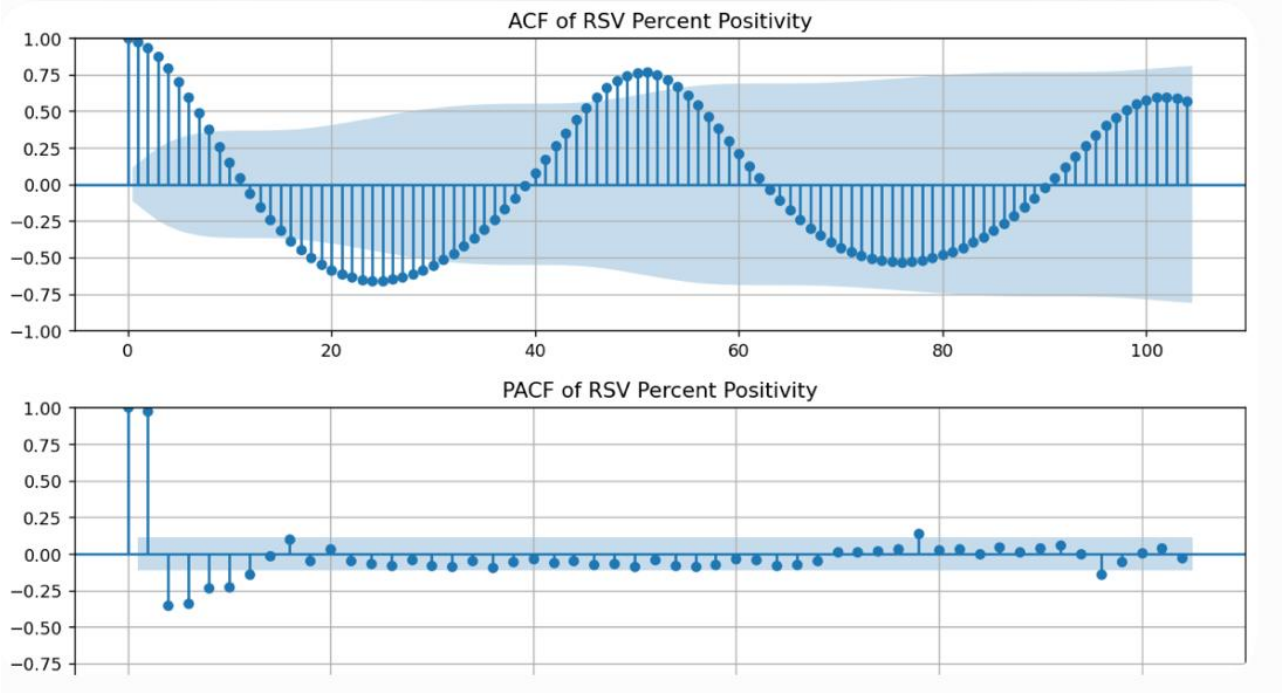


Fig. 4 ACF and PACF plots of weekly RSV percent positivity (Picture credit: Original)

ADF and KPSS tests were applied to the original and lag-52 differenced series. ADF tests a unit-root null, whereas KPSS tests a stationarity null; evidence favors stationarity when ADF rejects and KPSS does not.

As Table 3 shows, the original series rejected the ADF unit-root null ($p < 0.001$) and did not reject KPSS stationarity ($p > 0.10$). The seasonally differenced series gave the same conclusion. $D = 1$ was therefore selected from seasonal diagnostics and model selection, not from treating ADF or KPSS as seasonal tests.

Table 3. Stationarity test results

Series	ADF statistic	ADF p-value	KPSS statistic	KPSS p-value	Conclusion
Original series	-5.391	< 0.001	0.054	> 0.10	Stationary by non-seasonal tests, but strong seasonality remains
Seasonally differenced series	-5.510	< 0.001	0.125	> 0.10	Stationary after seasonal differencing

5.2. Model Selection Results

The comparison included ARIMA, SARIMA, ETS, and seasonal naive models. ARIMA represented short-term dependence, whereas SARIMA and the two seasonal benchmarks incorporated annual recurrence.

The selected models were ARIMA (2,0,1), SARIMA (1,0,2) (0,1,0) [52], and additive-trend/additive-seasonality ETS. AICc guided selection within model classes, while final ranking relied on out-of-sample accuracy. Only SARIMA passed the lag-24 Ljung-Box test ($p = 0.528$); the other models had p-values below 0.05. The selected specifications and residual diagnostic results are summarized in Table 4.

Table 4. Model selection and residual diagnostics

Model	Selected specification	AICc	Ljung-Box p-value	Comment
ARIMA	ARIMA(2,0,1)	503.829	0.004	Captures short-term autocorrelation only
SARIMA	SARIMA(1,0,2)(0,1,0)[52]	375.590	0.528	Captures annual seasonality
ETS	Additive trend + additive seasonality	-151.97	$p < 0.001$	Baseline comparison
Seasonal naive	y[t−52]	Not applicable	$p < 0.001$	Simple seasonal benchmark

Because model classes use different likelihood structures, their AICc values were not treated as directly decisive. Holdout and rolling-origin performance provided the main cross-model comparison.

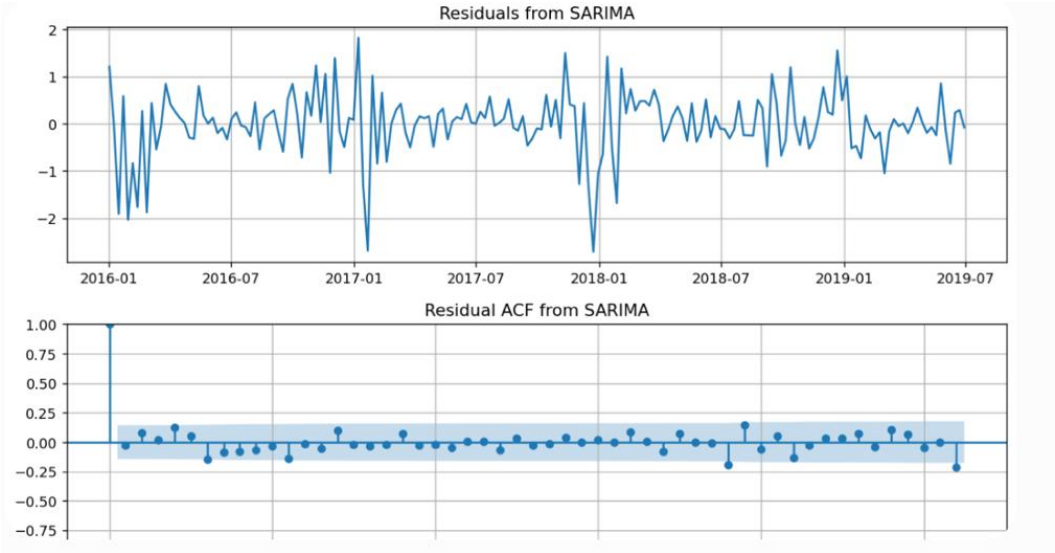


Fig. 5 Residual diagnostics for the selected SARIMA model (Picture credit: Original)

Fig. 5 shows SARIMA residuals centered near zero with little remaining autocorrelation. Together with the Ljung-Box result, this supports the adequacy of the selected temporal specification.

5.3. Forecast Accuracy

The final 52 weeks provided a fixed out-of-sample test across a complete seasonal cycle, whereas the rolling-origin analysis in Section 5.4 assessed operational one- to four-week forecasting performance. Fig. 6 compares observed values with forecasts from all four models.

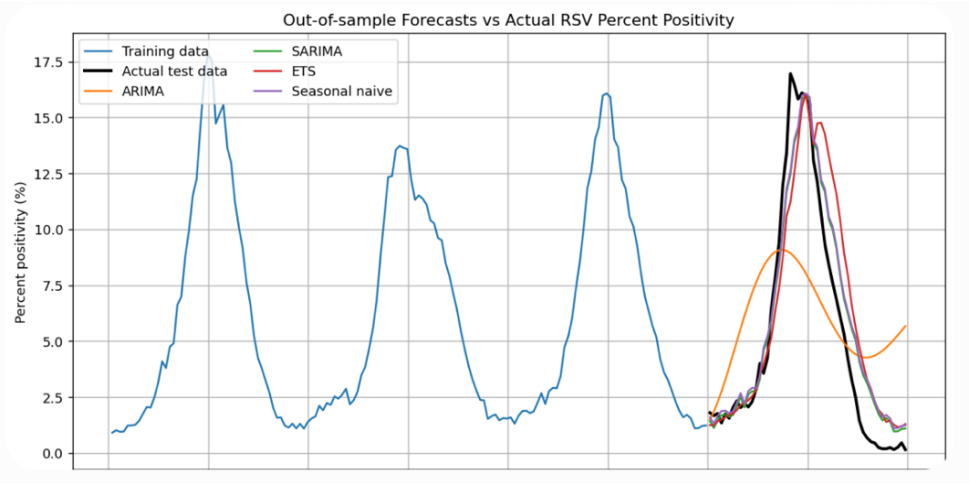


Fig. 6 Out-of-sample forecasts versus actual RSV percent positivity (Picture credit: Original)

RMSE and MAE measured numerical accuracy, while directional accuracy and peak-timing error assessed weekly movement and the timing of the maximum. No single measure captures all operational features.

SARIMA achieved the lowest RMSE (1.547) and MAE (1.232), slightly improving on seasonal naive (1.603 and 1.279). ETS was less accurate, and non-seasonal ARIMA performed worst (RMSE 4.206; MAE 3.616), indicating that short-term dependence alone was insufficient across the full seasonal cycle. Table 5 reports the complete fixed-holdout accuracy results for all four models.

Table 5. Fixed-holdout forecast accuracy

Model	RMSE	MAE	sMAPE (%)	Peak timing error (weeks)	Directional accuracy (%)
ARIMA	4.206	3.616	79.948	2	68.6
SARIMA	1.547	1.232	46.656	4	76.5
ETS	2.454	1.843	52.819	4	74.5
Seasonal naive	1.603	1.279	47.582	4	76.5

SARIMA and seasonal naive also achieved the highest directional accuracy (76.5%), followed by ETS (74.5%) and ARIMA (68.6%). Seasonal information therefore improved identification of weekly changes.

sMAPE was retained as the percentage-based error measure because observed positivity approached zero during the off-season, making conventional mean absolute percentage error (MAPE) unstable. RMSE, MAE, sMAPE, and directional accuracy are therefore the more informative measures for this series.

Overall, the holdout results support the hypothesis that explicit annual seasonality improves forecasting. SARIMA produced the lowest average errors, but its advantage over seasonal naive was modest and no formal test of the difference in forecast loss was conducted. Its four-week peak-timing error also shows that low numerical error does not guarantee exact peak prediction. Conversely, ARIMA had poor average accuracy but a smaller two-week peak-timing error, confirming that peak timing and trajectory accuracy capture different dimensions of performance.

5.4. Robustness Check

Rolling-origin evaluation repeatedly re-estimated each model using only past data and produced forecasts one to four weeks ahead, approximating real-time surveillance updates (Table 6).

At one week, ARIMA and SARIMA had nearly identical RMSE (0.760 and 0.759), although ARIMA had slightly lower MAE. From two to four weeks, SARIMA was consistently best: its RMSE rose from 0.986 to 1.261, compared with 1.228 to 2.215 for ARIMA and 1.308 to 2.030 for ETS. Seasonal naive remained relatively stable across horizons and outperformed ARIMA and ETS at four weeks, although SARIMA retained the lowest RMSE and MAE.

Errors increased with horizon for the fitted statistical models, as expected, but SARIMA showed the slowest deterioration among them. Seasonal naive was nearly horizon-invariant because it repeated the corresponding value from the previous year. SARIMA's advantage became clearer at two- to four-week horizons, which are more relevant for advance planning.

The dataset ends in June 2020 and therefore excludes the strongly disrupted post-pandemic RSV seasons, although the final months include the early COVID-19 period. The results demonstrate robustness within a predominantly pre-pandemic series rather than universal performance under structural change; validation on later and regional data remains necessary.

Table 6. Rolling-origin forecast accuracy by horizon (RMSE / MAE)

Model	**1 week**	**2 weeks**	**3 weeks**	**4 weeks**
ARIMA	0.760 / 0.492	1.228 / 0.884	1.708 / 1.221	2.215 / 1.635
SARIMA	0.759 / 0.520	0.986 / 0.729	1.141 / 0.818	1.261 / 0.934
ETS	0.855 / 0.583	1.308 / 0.885	1.695 / 1.159	2.030 / 1.443
Seasonal naive	1.636 / 1.303	1.640 / 1.316	1.643 / 1.324	1.650 / 1.344

6. Public Health Implications and Discussion

6.1. Main Findings and Early-Warning Value

RSV positivity displayed a strong annual cycle, and seasonal modelling improved forecast accuracy. SARIMA had the lowest holdout RMSE and MAE, matched the best directional accuracy, passed residual diagnostics, and remained strongest at two- to four-week rolling horizons. However, its improvement over seasonal naive was modest and its peak-week estimate was not exact.

For surveillance, a predicted sustained rise may provide early warning before the seasonal maximum. Such forecasts should be interpreted as decision-support signals and combined with testing volume, hospital indicators, and local epidemiological judgement rather than used as stand-alone decisions.

6.2. Implications for Healthcare Planning

Short-term forecasts can support review of paediatric bed capacity, emergency staffing, diagnostic supplies, and communication plans before demand rises. The main value lies in preparedness rather than precise prediction of infections or admissions.

SARIMA is transparent, reproducible, and updateable from a routine weekly series. Its low data requirements make it practical where richer predictors are unavailable, although regional implementation would require local recalibration.

6.3. Comparison with Previous Studies and Benchmark Models

The winter recurrence agrees with earlier RSV studies across different settings, and the potential planning value is consistent with hospital-based peak forecasting [5, 6, 9]. Post-pandemic evidence nevertheless shows that historical seasonality can shift [3, 7, 8].

Benchmark results clarify the source of improvement. Non-seasonal ARIMA performed poorly across the full holdout and deteriorated rapidly with horizon. Seasonal naive was competitive but could not adapt to short-term deviations, while SARIMA combined annual recurrence with recent dynamics.

The mismatch between average error and peak timing supports Morbey et al.'s warning that the model with the lowest general error may not best predict epidemic peaks [11]. Operational evaluation should therefore retain both standard and event-based metrics.

6.4. Limitations and Future Research

The main limitation is the predominantly pre-pandemic national series, with early-pandemic observations through June 2020, which may not represent later structural changes or regional heterogeneity. Only about five and a half years of weekly data were included, providing a limited number of complete seasonal cycles for estimation and testing.

Percent positivity is affected by testing access, laboratory participation, reporting, the tested population, and changes in the composition of diagnostic test types. It does not directly measure prevalence, hospitalization, severity, or causal drivers, and no uncertainty from reporting revisions was modelled.

Future work should add post-pandemic observations, estimate HHS-region models, and compare SARIMA with methods using weather, mobility, school calendars, testing volume, or hospital data. Probabilistic calibration and real-time onset and peak alerts should also be evaluated.

7. Conclusion

This study compared ARIMA, SARIMA, ETS, and seasonal naive forecasts of weekly CDC NREVSS RSV percent positivity using a 52-week holdout and rolling-origin evaluation.

RSV activity showed strong annual seasonality. SARIMA achieved the lowest holdout RMSE and MAE and remained comparatively robust at two- to four-week horizons, although its average-error advantage over seasonal naive was modest and peak timing was less precise.

These forecasts may support early warning and healthcare preparation when combined with regional surveillance and clinical indicators. The findings are based on national data from 2015 to June 2020, including early-pandemic observations, and require validation on recent and region-specific series.

References

[1] Li Y., Wang X., Blau D. M., et al. Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in children younger than 5 years in 2019: a systematic analysis. The Lancet, 2022, 399(10340): 2047–2064.

[2] Curns A. T., Rha B., Lively J. Y., et al. Respiratory Syncytial Virus-Associated Hospitalizations Among Children <5 Years Old: 2016 to 2020. Pediatrics, 2024, 153(3): e2023062574.

[3] Hamid S., Winn A., Parikh R., Jones J. M., et al. Seasonality of Respiratory Syncytial Virus—United States, 2017–2023. MMWR Morbidity and Mortality Weekly Report, 2023, 72(14): 355–361. doi: 10.15585/mmwr.mm7214a1

[4] Centers for Disease Control and Prevention. Respiratory Syncytial Virus Laboratory Data (NREVSS). CDC Data, 2022. https://data.cdc.gov/Laboratory-Surveillance/Respiratory-Syncytial-Virus-Laboratory-Data-NREVSS/52kb-ccu2

[5] Zhang H., Wen S., Zheng J., Chen X., Lv F., Liu L. Meteorological factors affecting respiratory syncytial virus infection: A time-series analysis. Pediatric Pulmonology, 2020, 55(3): 713–718.

[6] Thongpan I., Vongpunsawad S., Poovorawan Y. Respiratory syncytial virus infection trend is associated with meteorological factors. Scientific Reports, 2020, 10: 10931.

[7] Daniels D., Wang D., Suryadevara M., et al. Epidemiology of RSV Bronchiolitis Among Young Children in Central New York Before and After the Onset of the COVID-19 Pandemic. Pediatric Infectious Disease Journal, 2023, 42(12): 1056–1062.

[8] Bourdeau M., Vadlamudi N. K., Bastien N., et al. Pediatric RSV-Associated Hospitalizations Before and During the COVID-19 Pandemic. JAMA Network Open, 2023, 6(10): e2336863.

[9] Cutrera R., Ciofi degli Atti M. L., Dotta A., et al. Epidemiology of respiratory syncytial virus in a large pediatric hospital in Central Italy and development of a forecasting model to predict the seasonal peak. Italian Journal of Pediatrics, 2024, 50: 65.

[10] Albrecht S., Broderick D., Dost K., et al. Forecasting severe respiratory disease hospitalizations using machine learning algorithms. BMC Medical Informatics and Decision Making, 2024, 24: 293.

[11] Morbey R. A., Todkill D., Watson C., Elliot A. J. Machine learning forecasts for seasonal epidemic peaks: Lessons learnt from an atypical respiratory syncytial virus season. PLOS ONE, 2023, 18(9): e0291932.

[12] Hyndman R. J., Khandakar Y. Automatic Time Series Forecasting: The forecast Package for R. Journal of Statistical Software, 2008, 27(3): 1–22. doi: 10.18637/jss.v027.i03

[13] Hyndman R. J., Athanasopoulos G. Forecasting: Principles and Practice. 2nd ed. Melbourne: OTexts, 2018. Online edition. Available at: https://otexts.com/fpp2/. Accessed 13 July 2026.