Final Project: A Time Series Analysis of Weekly RSV Hospitalization Rate in Michigan

Abstract

We analyze weekly RSV hospitalization rates in Michigan from late 2018 to early 2026. The series has a clear annual cycle, a quiet 2020–2022 period that lines up with the COVID-19 pandemic, and a rebound peak in the 2022–2023 season. We fit two models in parallel. A SARIMA(4,0,1) $\times$ $(1,1,1)_{52}$ model with time and COVID-step regressors gives the lowest AIC and passes residual diagnostics. A seasonally forced SEIRS POMP model, fit by iterated filtering in pypomp, recovers latent and infectious period estimates of about 3.9 and 4.4 days, close to published RSV values. The SARIMA model fits the statistical structure more tightly on its own scale, but the SEIRS model gives parameters we can interpret in terms of transmission, which is why we included a POMP analysis in the first place.

1 Introduction

RSV (Respiratory Syncytial Virus) remains a significant public health concern, with high mortality rates in vulnerable populations such as older adults (over 65) and infants under 6 months (World Health Organization, n.d.). Understanding when and how RSV spreads is critical for optimizing the timing and targeting of preventive measures, including vaccination campaigns and monoclonal antibody distribution. The goal of this project is to analyze the structural and temporal dependencies of weekly RSV activity in Michigan. We aim to determine whether a simple mechanistic transmission model can reproduce the observed seasonal patterns.

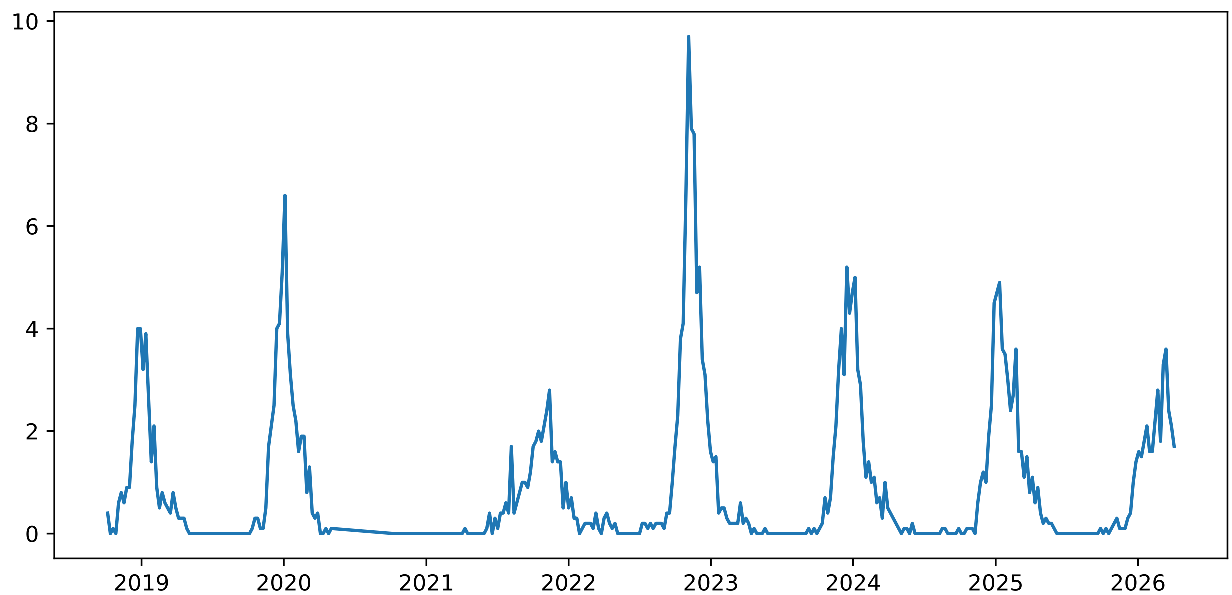
We use data from CDC RSV-NET (Centers for Disease Control and Prevention 2026), a surveillance system for laboratory-confirmed RSV hospitalizations. Our analysis focuses on Michigan, using a weekly time series of hospitalization rates from late 2018 through early 2026. We evaluate the data using two complementary approaches: a SARIMA model for statistical seasonality and a seasonally forced SEIRS POMP model to recover interpretable transmission parameters. A similar combination of SARMA and a seasonally forced SEIRS POMP has been applied to Michigan weekly flu cases in a previous 531 project (Anonymous 2025a); we follow the same general structure and use the same cosine seasonal forcing form, but target RSV hospitalizations and use pypomp in place of R pomp.

2 Data Preprocessing

The raw RSV-NET file provides weekly hospitalization rates stratified by location, sex, and age group. We kept the rows where Age Category is “All” and averaged within each week ending date, which gives one rate per week for the Michigan population as a whole. A few epiweek-52 entries appeared twice in some years; these were collapsed by taking the mean. All recorded rates are non-negative, so for the seasonal decomposition and the regression step we used $\log(\text{rate} + 0.01)$, which keeps the zero-rate weeks finite without introducing arbitrary jumps.

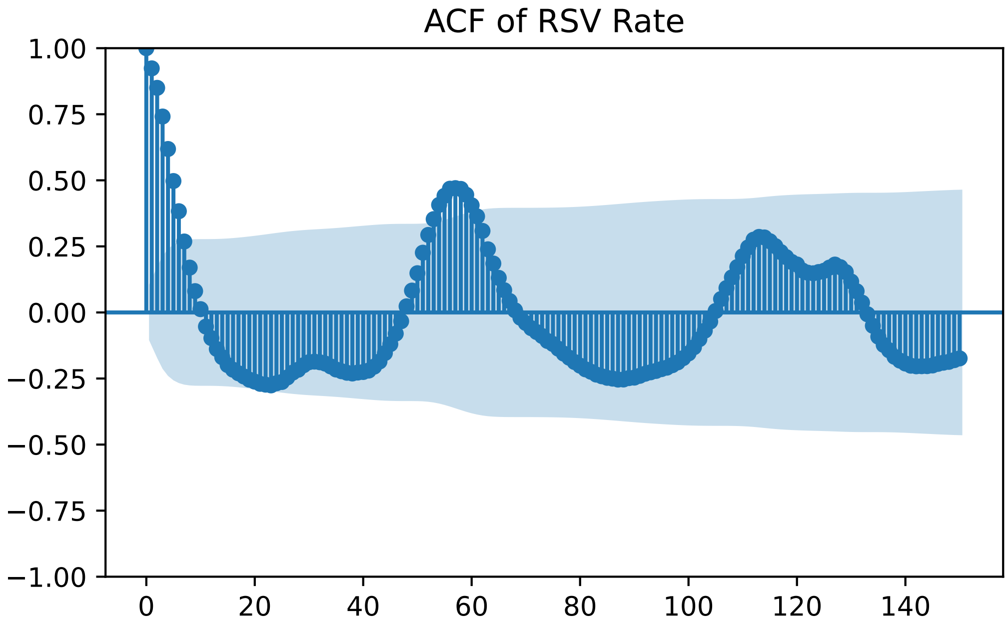
3 Exploratory Data Analysis

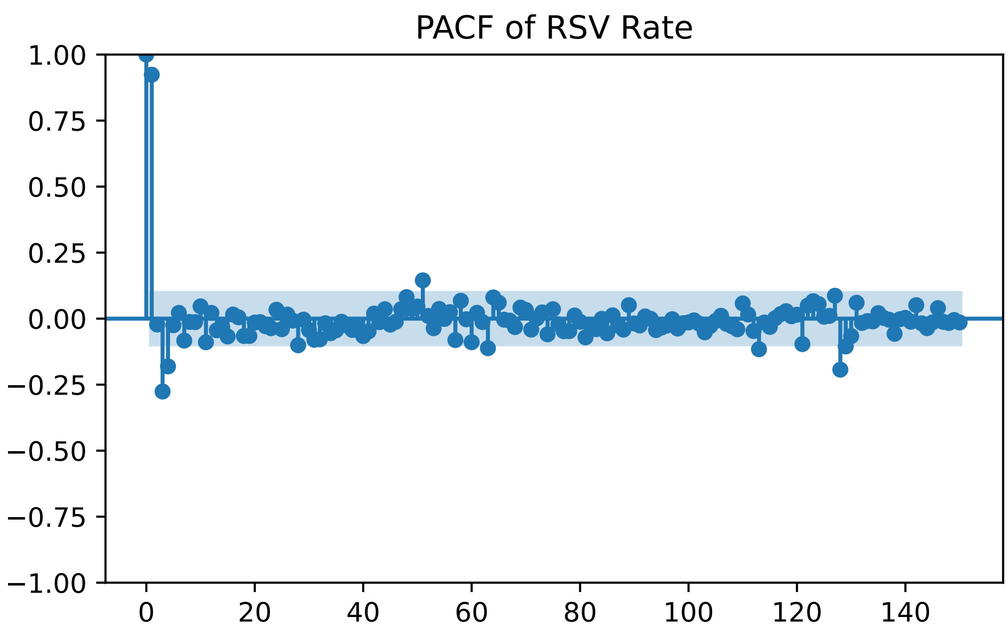
3.1 Time plot



The time plot shows a clear annual cycle: each year's peak sits somewhere around the end of the calendar year. The 2022–2023 peak is unusually tall relative to other seasons, and the period from early 2020 through 2021 looks structurally different, with the usual winter peak almost absent. Both observations are consistent with the COVID-19 pandemic disrupting RSV circulation during 2020 and producing a rebound in 2022. A seasonal plot broken out by year appears in the Supplementary Material and makes this pattern easier to read.

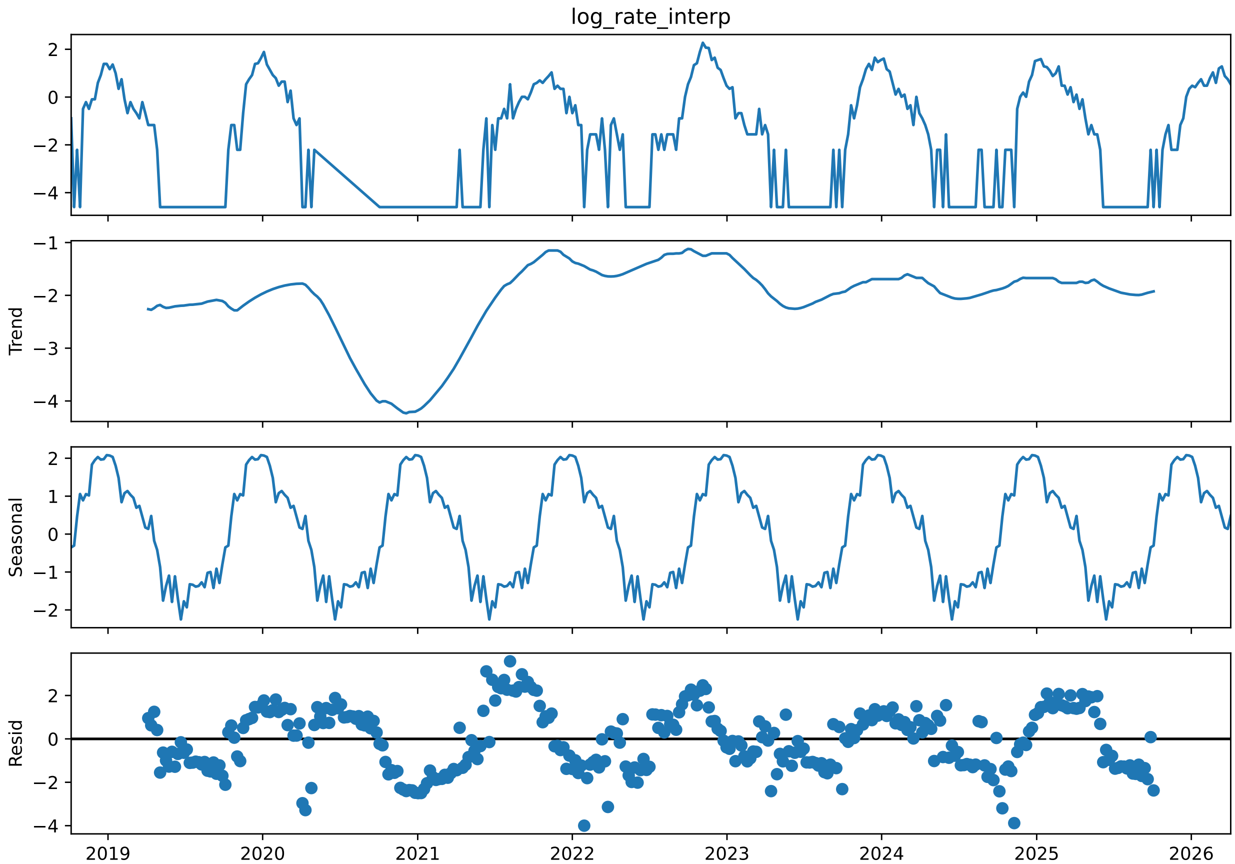
3.2 ACF and PACF





The ACF shows strong short-lag autocorrelation and an oscillating pattern with bumps near lag 52 and 104. This is evidence of an annual seasonality in the RSV series. The PACF has a dominant spike at lag 1, suggesting that an AR component is important even before we add the seasonal structure. Together these plots motivate a SARIMA specification with both nonseasonal and seasonal terms. The spectral density in the Supplementary Material confirms a 52-week period, which we use as the seasonal period s in the SARIMA model.

3.3 Decomposition



An additive decomposition of the log rate gives a seasonal component that repeats cleanly on a 52-week cycle, a trend component that dips during 2020–2022 and spikes around 2023, and a residual component that still has noticeable structure. The trend shape supports the COVID-disruption reading of the time plot. The residual being this large also tells us a standard decomposition is not enough on its own, which is part of why we move on to SARIMA and then to a mechanistic model.

3.4 Regression Analysis on COVID time step

To check whether the 2020–2022 window behaved differently from the rest of the record, we fit a simple OLS regression of the log rate on a linear and quadratic time trend, a COVID step dummy for 2020–2022, and interaction terms with time. The COVID step is data-driven, motivated by the visibly quiet 2020–2022 period in the time plot, and is consistent with reports that RSV circulation patterns shifted during that season for reasons beyond the virus itself (Garg I 2022). All five covariates are significant at the 5% level, which supports the idea that RSV activity really did change during the pandemic period. However, the overall R^2 is only 0.095, so linear and quadratic trend terms alone explain very little of the variation. Most of the structure sits in the seasonality and in higher-frequency fluctuations that a pure trend model cannot capture.

4 Methodology

(Ionides 2026) We use two models in parallel. The first is a seasonal ARIMA model with exogenous regressors,

$$\text{SARIMA}(p, d, q) \times (P, D, Q)_s,$$

fit by maximum likelihood and selected by AIC, where

$$\text{AIC} = 2k - 2\hat{\ell}$$

with $\hat{\ell}$ the maximized log-likelihood and k the number of estimated parameters. We search over $p, q \in \{0, 1, 2, 3, 4\}$, $P, Q \in \{0, 1\}$, and $d, D \in \{0, 1\}$, with the seasonal period fixed at $s = 52$. The full AIC tables are given in the Appendix and the detailed SARIMA model equations appear in the Supplementary Material.

The second model is a partially observed Markov process (POMP) of SEIRS form with seasonal forcing. SEIRS is a natural fit for RSV: it has a latent stage (E), an infectious stage (I), and allows waning immunity so that recovered individuals can return to the susceptible pool, which is consistent with the annual recurrence of RSV. The stochastic transitions and state updates are defined as follows:

$$\begin{aligned}
\Delta N_{SE} &\sim \text{Binomial}(S, 1 - e^{-\beta I/N \Delta t}) \\
\Delta N_{EI} &\sim \text{Binomial}(E, 1 - e^{-\mu_{EI} \Delta t}) \\
\Delta N_{IR} &\sim \text{Binomial}(I, 1 - e^{-\mu_{IR} \Delta t}) \\
\Delta N_{RS} &\sim \text{Binomial}(R, 1 - e^{-\mu_{RS} \Delta t})
\end{aligned}$$

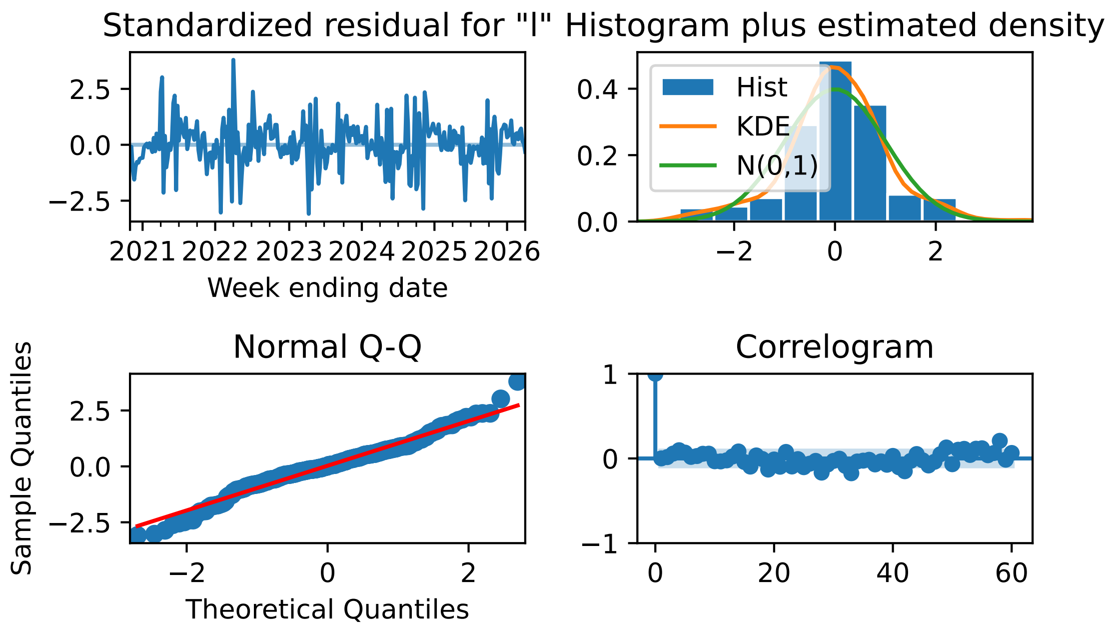
$$\begin{aligned}
S_{t+\Delta t} &= S_t - \Delta N_{SE} + \Delta N_{RS} \\
E_{t+\Delta t} &= E_t + \Delta N_{SE} - \Delta N_{EI} \\
I_{t+\Delta t} &= I_t + \Delta N_{EI} - \Delta N_{IR} \\
R_{t+\Delta t} &= R_t + \Delta N_{IR} - \Delta N_{RS} \\
H_{t+\Delta t} &= H_t + \Delta N_{IR}
\end{aligned}$$

We fit the POMP model with iterated filtering (IF2) in pypomp (pypomp developer team, n.d.), starting from both a local neighborhood of a reasonable initial guess and from a wider random search.

5 SARIMA Model AIC Table

After searching different parameters and printing the AIC table, there are several candidate sets of parameters that have similar small AIC values. Then, in the residual diagnostic plots of these sets of parameters, applying differences to seasonal components significantly enhances the residual distribution for the SARIMA models. Considering minimizing AIC values while normalizing residuals presentation in Q-Q plot, $SARIMA(4, 0, 1) \times (1, 1, 1)_52$ is selected. We further examine this model using the diagnostic plots below.

5.1 Residual Diagnostics



The diagnostic panel suggests that the selected SARIMA model captures most of the temporal dependence in the log RSV series. The standardized residuals fluctuate around zero with no obvious drift, and the residual correlogram shows little remaining autocorrelation at the displayed lags up to 60. After seasonal differencing, the Q-Q plot and the histogram align reasonably well with the normal reference lines, suggesting that the residuals are approximately normal.

6 POMP (SEIRS) MODEL

6.1 Model specification

The latent process follows a seasonally forced SEIRS compartmental structure with state variables S_t, E_t, I_t, R_t, and an accumulator H_t representing the accumulated infection-related observation process. Individuals move according to the cycle:

$$S \to E \to I \to R \to S$$

(Ionides 2026)

The transmission rate $\beta(t)$ is modeled as a periodic function to reflect the annual seasonality of RSV:

$$\beta(t) = \beta_0 \left[1 + a \cos(2\pi(t + \phi)/52)\right]$$

(Anonymous 2025a)

where β_0 is the baseline rate, a is the seasonal amplitude, and ϕ is the phase shift.

Per-step transition probabilities are $1 - \exp(-\text{rate} \cdot \Delta t)$, where the rates for each compartment are defined as:

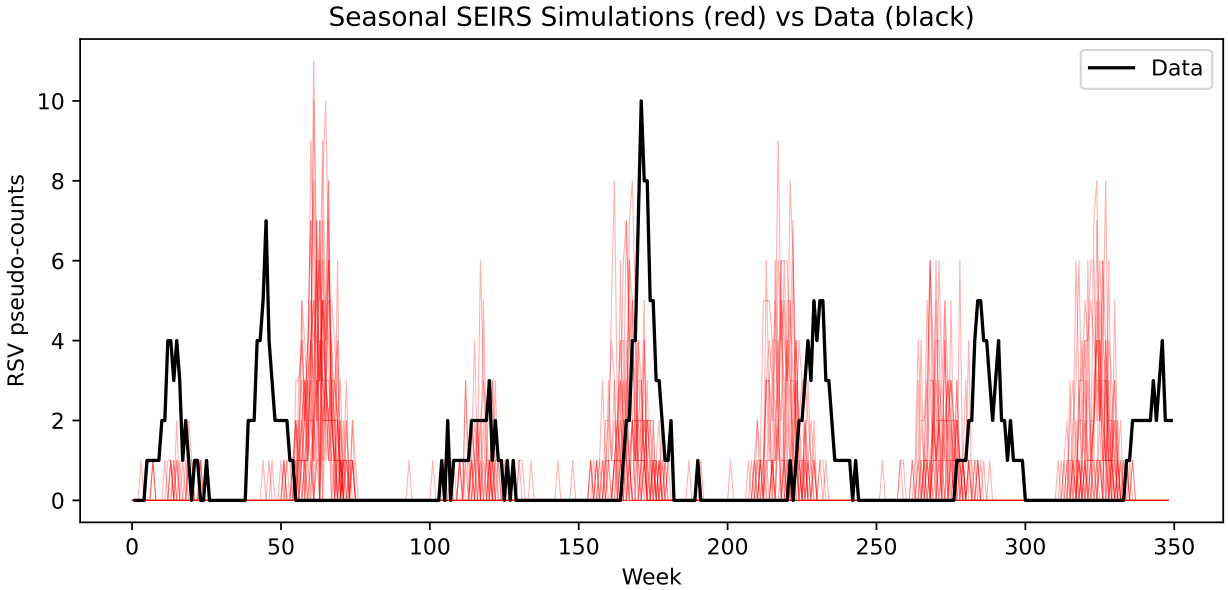
- $S \to E$: $\beta(t)(I + \iota)/N$
- $E \to I$: μ_{EI}
- $I \to R$: μ_{IR}
- $R \to S$: μ_{RS}

The observation model follows a negative binomial distribution to account for overdispersion in the surveillance data:

$$\text{cases}_t \sim \text{NegBin}(\text{mean} = \rho H_t,\ \text{size} = k)$$

The parameter ι handles occasional imported infections to prevent the system from becoming trapped at $I = 0$, while μ_{RS} is fixed at $1/52$ to reflect an average immunity duration of one year. Weekly rates are treated as pseudo-counts with a fixed population $N = 100{,}000$; this simplification is discussed further in the limitations section.

6.2 Simulation at the initial guess



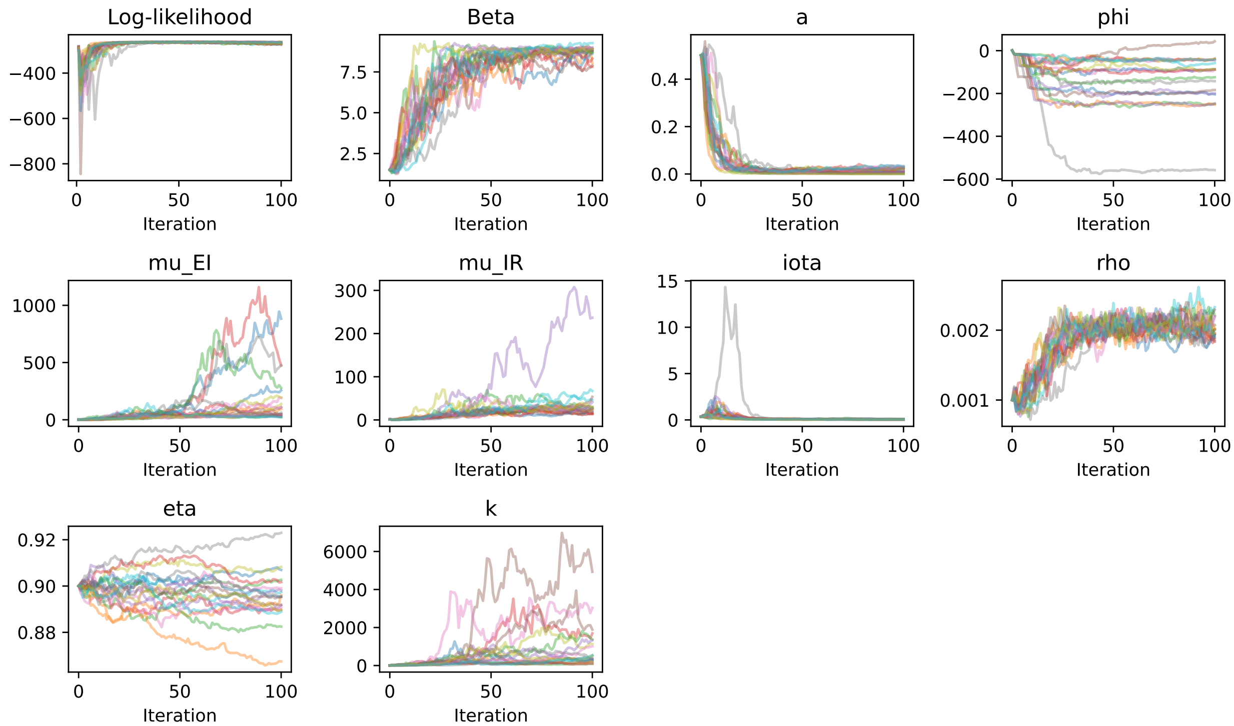
Twenty simulations from the initial guess produce seasonal peaks of roughly the right height and frequency, but the phase is clearly off and the simulated peaks are noisier than the observed series. This is expected from an unfitted guess; the point is just to check that the model is capable of producing the right shape.

6.3 Particle filter at the initial guess

```
Mean Log-Likelihood: -866.3
```

Using the initial parameter guess ($\beta_0 = 1.5, a = 0.5, \phi = 1, \mu_{EI} = 7/3, \mu_{IR} = 7/5, \mu_{RS} = 1/52, \iota = 0.35, \eta = 0.9, \rho = 0.001, k = 10$), we ran 10 replicate particle filters with $J = 5000$ particles. We set μ_{EI}, μ_{IR}, and μ_{RS} based on the inverse of the incubation period (3-5 days), infectious period (3-8 days), and re-susceptible period (52 weeks) (Eiland 2009). The resulting mean log-likelihood was -866.3. This is a weak starting point but gives a baseline for the iterated filtering searches below.

6.4 Local search



We ran IF2 from 20 starting points around the initial guess, with $M = 100$ iterations, $J = 5000$ particles, a cooling factor of 0.5, and random-walk standard deviations of 0.02 on all estimated parameters (0.5 on ϕ). The parameters μ_{RS} and N were held fixed. The search took about 165 seconds on a Colab GPU and reached a best endpoint log-likelihood of roughly -275, a clear jump from the starting value near -866.

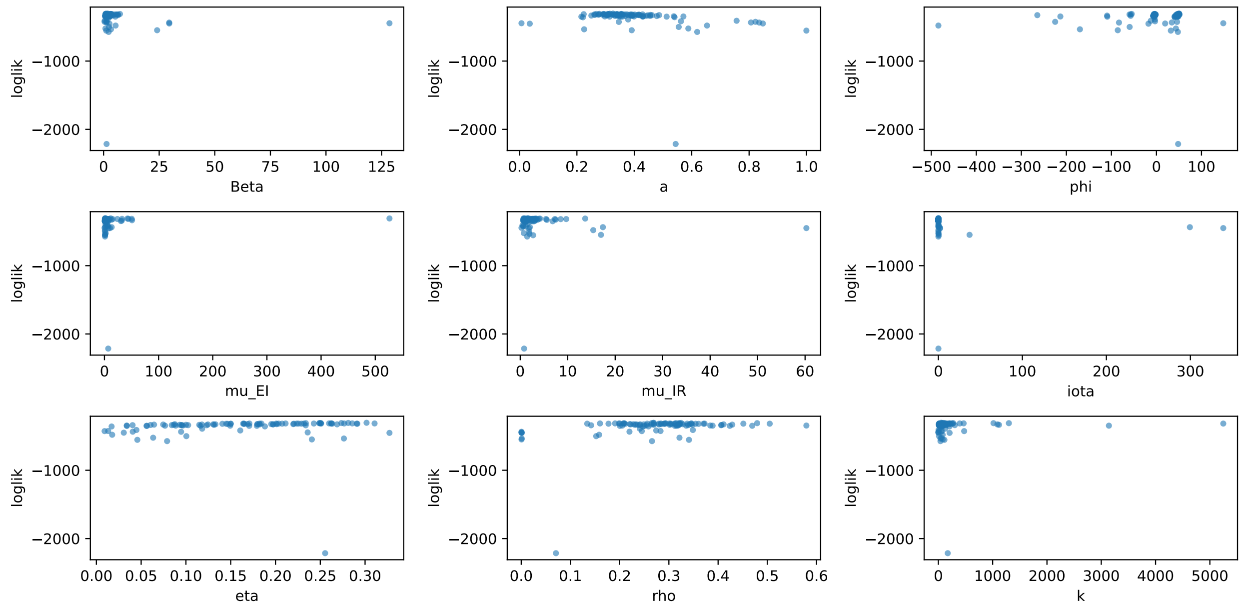
The trace plots are informative in a different way. Across many replicates, μ_{EI} drifts above 1000 and μ_{IR} above 300, which is not biologically reasonable: these values imply individuals pass through the Exposed and Infectious compartments almost instantly. Known RSV biology says the latent period is roughly 4 days and the infectious period around 5 days. This is a classic case of weak identifiability. Weekly incidence data alone do not carry enough information to separate the latent and infectious period durations, so MIF inflates both rates to gain likelihood without recovering meaningful values. The phase ϕ also drifts to large negative values across runs, but that is harmless: ϕ enters the model only inside $\cos(2\pi(t + \phi)/52)$, so it is identifiable only modulo 52, and different replicates simply converge to equivalent phases.

6.5 Global search

```
Best global log-likelihood: -304.8
```

For the global search we ran IF2 from 100 random starting points, drawn from the uniform ranges $\beta_0 \in [1, 50]$, $a \in [0.05, 0.9]$, $\phi \in [0, 52]$, $\mu_{EI} \in [7/5, 7/3]$, $\mu_{IR} \in [7/8, 7/3]$, $\iota \in [0.05, 1]$, $\eta \in [0.01, 0.3]$,

$\rho \in [0.05, 0.6]$, and $k \in [1, 30]$. The IF2 schedule was kept the same as in the local search, and final log-likelihoods were evaluated with a particle filter at $J = 10{,}000$.




```
Best parameters from global search:
  Beta = 1.7761
  a = 0.3259
  phi = 50.2452
  mu_EI = 1.7986
  mu_IR = 1.5951
  iota = 0.0474
  eta = 0.3017
  rho = 0.2688
  k = 51.3568
  loglik = -304.8
```

The best global result reached a log-likelihood of about -305. The corresponding parameters are $\beta_0 = 1.78$, $a = 0.33$, $\phi = 50.25$, $\mu_{EI} = 1.80$, $\mu_{IR} = 1.60$, $\iota = 0.047$, $\eta = 0.30$, $\rho = 0.27$, and $k = 51.4$. These are biologically sensible: $\mu_{EI} = 1.80$ per week corresponds to a latent period of about 3.9 days, and $\mu_{IR} = 1.60$ per week gives an infectious period of about 4.4 days. Both match what is reported in the RSV literature. The reporting fraction $\rho \approx 0.27$ is also in a plausible range for a surveillance rate relative to the $N = 100{,}000$ pseudo-population. The scatter plots show that high-likelihood points cluster in a reasonably narrow region of β_0, a, ρ, and η, while several other parameters remain weakly identified, which is consistent with what we saw in the local search.

6.6 Model Comparison

The SARIMA and SEIRS models are fit on different scales, so their fit metrics are not directly comparable. SARIMA is fit on the interpolated log rate and reaches an AIC of about 806. The SEIRS model is fit on pseudo case counts and reaches a log-likelihood of about −305. A direct comparison of the two numbers would not be meaningful.

What we can compare is what each model gives us. The SARIMA fit is tight in a statistical sense: the residuals behave like white noise at the displayed lags, the COVID-step regressors come out significant, and the model is simple to extend with more covariates. But it does not say anything about how RSV actually spreads in the population. The SEIRS model is looser in fit, but returns quantities we can interpret: a latent period of about 4 days, an infectious period of about 4.4 days, a seasonal amplitude, and a reporting fraction. These values are consistent with what is reported in the RSV literature. Which model is more useful depends on the question: SARIMA for description and short-term forecasting, SEIRS for mechanism.

7 Conclusions

Michigan RSV hospitalizations from 2018 to 2026 follow a clear annual cycle, disrupted in the 2020–2022 period and with a rebound in 2022–2023. A SARIMA(4,0,1) $\times$ $(1,1,1)_{52}$ model with COVID-step regressors captures this structure well at a statistical level. A seasonally forced SEIRS POMP model, fit with iterated filtering, recovers parameter values close to known RSV biology: a latent period around 4 days, an infectious period around 4.4 days, and a reporting fraction of about 0.27. Both models point to the same main features of the data, namely the annual seasonality and the pandemic disruption. The SEIRS fit shows that a simple transmission model is at least consistent with the observed pattern, although the separation of the latent and infectious rates is only weakly identified from weekly incidence alone.

Limitations and Future Prospects

The simpler SARIMA model is not directly comparable to the mechanistic SEIRS model because the two are fit on different scales (log rate vs. pseudo case counts), which limits how much the comparison can say about relative fit.

The rebound in 2023 after a quiet 2020–2022 is also worth noting. It suggests that COVID may not have systematically changed RSV transmission, but rather that cases were underreported during the pandemic period. A SEIUR-type model with an unreported infected compartment, as introduced in (Chen 2022), would probably handle this better and bring more explanatory power to the data.

Acknowledgments

Data were sourced from the CDC RSV-NET surveillance system (Centers for Disease Control and Prevention 2026). The statistical methodology and core code patterns follow the STATS 531 course notes (Ionides 2026).

AI tools (Gemini, ChatGPT) were utilized for debugging code, clarifying modeling concepts, and refining the report's prose. All final modeling decisions, parameterizations, and interpretations were made by the authors.

Bibliography

Anonymous. 2025a. "A Time Series Analysis of Flu Cases in Michigan." https://ionides.github.io/531w25/final_project/project03/blinded.html.

———. 2025b. "Analysis of Malaria Cases in Florida." https://ionides.github.io/531w25/final_project/project05/blinded.html.

Centers for Disease Control and Prevention. 2026. "RSV-NET: Respiratory Syncytial Virus Hospitalization Surveillance Network." https://www.cdc.gov/rsv/php/surveillance/rsv-net.html.

Chen, Feng, Z. 2022. "SEIR Model with Unreported Infected Population and Dynamic Parameters for the Spread of COVID-19." https://doi.org/10.1016/j.matcom.2022.02.025.

Eiland, L. S. 2009. "Respiratory Syncytial Virus: Diagnosis, Treatment and Prevention." https://doi.org/10.5863/1551-6776-14.2.75.

Garg I, Sheikh AB, Shekhar R. 2022. "Impact of COVID-19 on the Changing Patterns of Respiratory Syncytial Virus Infections." https://doi.org/10.3390/idr14040059.

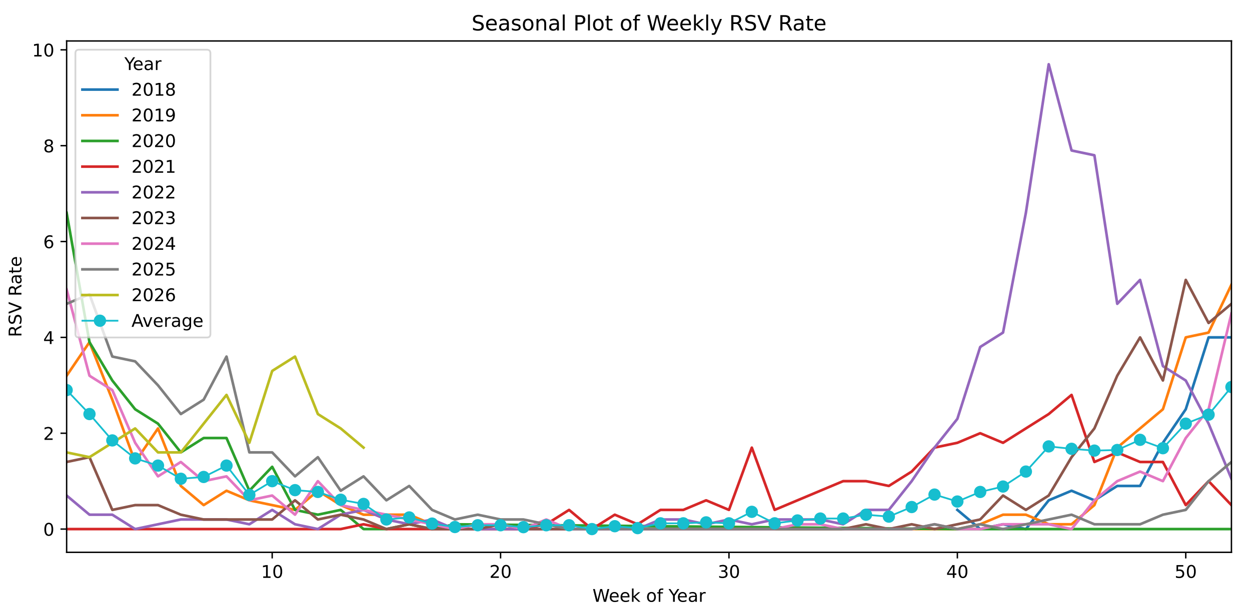
Ionides, Edward. 2026. "Modeling and Analysis of Time Series Data (STATS 531) Course Notes." https://ionides.github.io/531w26/.

pypomp developer team. n.d. *Modeling and Inference Using Partially Observed Markov Process Models in Python*. https://github.com/pypomp/pypomp.

World Health Organization. n.d. "Respiratory Syncytial Virus (RSV)." https://www.who.int/news-room/fact-sheets/detail/respiratory-syncytial-virus-(rsv)#:~:text=Respiratory%20syncytial%20virus%20(RSV)%20is,access%20to%20supportive%20medical%20care.

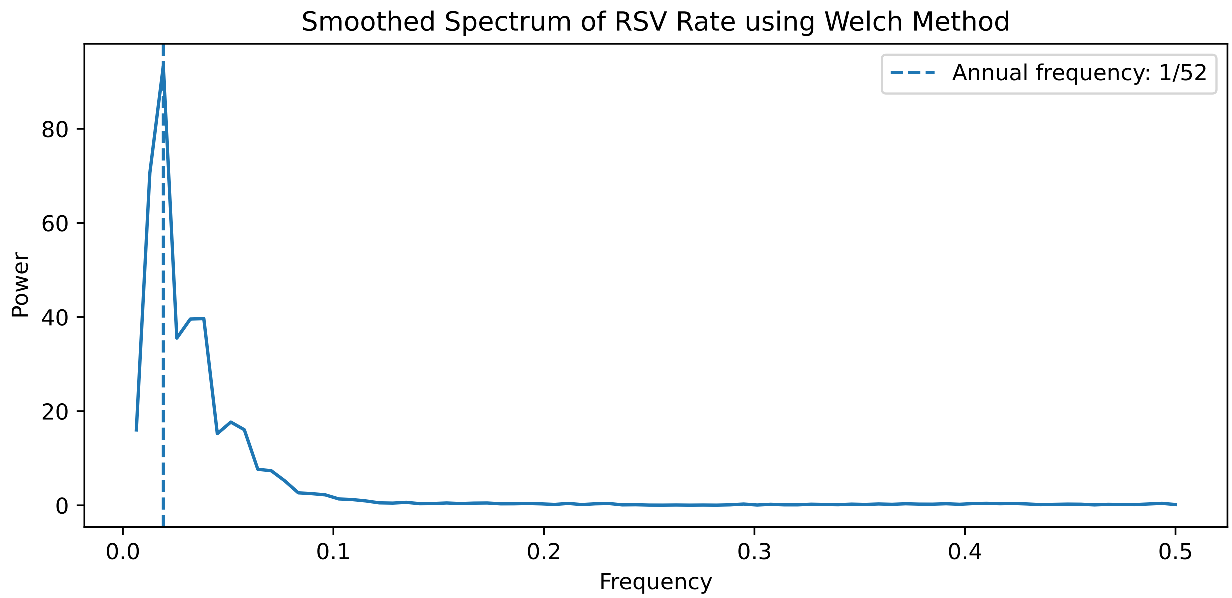
8 Supplementary material

8.1 Seasonal Plot



The seasonal plot of the weekly RSV rate aligns with the seasonal pattern observed in the time plot. The peaks mostly happen around the ends and the beginnings of the years.

8.2 Spectral Density Plot




```
Maximum smoothed spectral density: 93.39779725921727
Frequency at maximum: 0.019230769230769232 cycles per week
Period at maximum: 52.0 weeks
```

The smoothed spectrum estimated using Welch's method shows its maximum spectral density at frequency 0.01923 cycles per week, corresponding to a period of approximately 52 weeks. This confirms that the dominant periodic component in the RSV series is annual. We therefore used a seasonal period of 52 weeks in the SARIMA modeling.

8.3 Methodology: SARIMA

Following (Centers for Disease Control and Prevention 2026), the weekly case rate is defined as as:

$$\text{Rate} = \frac{\text{cases}_{\text{weekly}}}{100,000}$$

Model Selection SARIMA models are selected based on the Akaike Information Criterion (AIC):

$$\text{AIC} = 2k - 2\hat{\ell}$$

where $\hat{\ell}$ is the maximized log-likelihood estimate and k is the number of estimated parameters.

SARIMA Model Specification The SARIMA$(p, d, q) \times (P, D, Q)_{52}$ model, incorporating a quadratic time trend and COVID-19 indicators, is expressed as:

$$\phi(B)\Phi(B^{52}) \left[(1-B)^d(1-B^{52})^D Y_n - (\beta_0 + \beta_1 t_n + \beta_2 t_n^2 + \beta_3 C_n + \beta_4 t_n C_n + \beta_5 t_n^2 C_n)\right] = \theta(B)\Theta(B^{52})\epsilon_n$$

The autoregressive and moving-average polynomials are defined as:

$$\phi(x) = 1 - \phi_1 x - \cdots - \phi_p x^p, \quad \theta(x) = 1 + \theta_1 x + \cdots + \theta_q x^q$$

$$\Phi(x) = 1 - \Phi_1 x - \cdots - \Phi_P x^P, \quad \Theta(x) = 1 + \Theta_1 x - \cdots - \Theta_Q x^Q$$

where β_n are regression coefficients, t_n is the time index, and C_n represents the COVID-19 time step. The parameter search space for the AIC table includes $p, q \in \{0, 1, 2, 3, 4\}$, $P, Q \in \{0, 1\}$, and $d, D \in \{0, 1\}$.

Autocorrelation Function The autocorrelation function (ACF) at lag h is defined as:

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}, \quad \gamma(h) = \mathrm{Cov}(Y_t, Y_{t+h})$$

where $\gamma(h)$ is the autocovariance.

8.4 SARIMA AIC Tables

The full SARIMA AIC tables below support the model choice in the SARIMA section. We searched $p, q \in \{0, 1, 2, 3, 4\}$, $P, Q \in \{0, 1\}$, and $d, D \in \{0, 1\}$(Anonymous 2025b), with $s = 52$. The lowest AIC (about 806) is reached at $d = 0$, $D = 1$, $(P, Q) = (1, 1)$, $(p, q) = (4, 1)$.

```
                            OLS Regression Results
================================================================================
Dep. Variable:     np.log(cases + 0.01)   R-squared:                       0.095
Model:                              OLS   Adj. R-squared:                  0.082
Method:                   Least Squares   F-statistic:                     7.184
Date:                  Tue, 21 Apr 2026   Prob (F-statistic):           2.07e-06
Time:                          22:15:53   Log-Likelihood:                -810.23
No. Observations:                   349   AIC:                             1632.
Df Residuals:                       343   BIC:                             1656.
Df Model:                             5
Covariance Type:              nonrobust
=================================================================================
                    coef    std err          t      P>|t|      [0.025      0.975]
---------------------------------------------------------------------------------
Intercept        -1.3163      0.455     -2.892      0.004      -2.211      -0.421
time             -0.0219      0.006     -3.589      0.000      -0.034      -0.010
time_sq         6.17e-05   1.77e-05      3.495      0.001     2.7e-05    9.64e-05
covid_step        4.5447      1.911      2.378      0.018       0.786       8.303
time_covid       -0.1146      0.036     -3.161      0.002      -0.186      -0.043
time_sq_covid     0.0006      0.000      3.769      0.000       0.000       0.001
==============================================================================
Omnibus:                      749.035   Durbin-Watson:                   0.337
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               42.043
Skew:                           0.520   Prob(JB):                     7.42e-10
```

```
Kurtosis:                       1.654   Cond. No.                     7.87e+05
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 7.87e+05. This might indicate that there are
strong multicollinearity or other numerical problems.


=== d=0, D=0, Seasonal (P,Q)=(0,0) ===
             MA0          MA1          MA2          MA3          MA4
AR0  1893.151170  1638.369200  1473.499275  1378.441906  1304.152847
AR1  1223.320159  1163.034516  1137.995299  1133.093144  1116.635630
AR2  1145.242047  1147.212998  1133.583554  1125.793770  1118.562917
AR3  1144.892241  1146.864693  1132.625625  1130.052598  1120.036496
AR4  1142.658020  1131.163879  1129.108157  1131.050161  1122.035691

=== d=0, D=0, Seasonal (P,Q)=(0,1) ===
             MA0          MA1          MA2          MA3          MA4
AR0  1481.351740  1340.600510  1194.381459  1121.643215  1072.347674
AR1  1016.772299   952.344177   902.315740   901.043004   893.149670
AR2   926.016122   924.069572   902.787852   899.196197   894.247460
AR3   925.406974   910.541895   900.728340   898.565016   895.477984
AR4   924.080797   911.780382   902.040203   899.390560   897.287100

=== d=0, D=0, Seasonal (P,Q)=(1,0) ===
             MA0          MA1          MA2          MA3          MA4
AR0  1430.985388  1306.241261  1185.243288  1126.130905  1083.125476
AR1  1016.755315   957.732479   926.964978   928.009581   922.798501
AR2   923.958365   924.021676   919.451004   924.125113   908.120038
AR3   904.714812   899.844847   900.000058   900.479664   900.446166
AR4   896.689546   897.092005   897.175042   899.013851   899.780324

=== d=0, D=0, Seasonal (P,Q)=(1,1) ===
             MA0          MA1          MA2          MA3          MA4
AR0  1426.642243  1293.186005  1174.028968  1112.820974  1066.011866
AR1  1018.726190   954.109436   903.758298   902.652157   894.623017
AR2   925.707148   926.002649   904.108559   900.826184   895.709521
AR3   906.336033   901.618363   901.834090   899.156487   896.794771
AR4   898.606112   898.362668   898.396079   900.250341   898.626631

=== d=0, D=1, Seasonal (P,Q)=(0,0) ===
             MA0          MA1          MA2          MA3          MA4
AR0  1460.005109  1350.103770  1253.894062  1209.090059  1165.212542
```

```
AR1  1222.110481  1167.878345  1125.586993  1122.082943  1110.606468
AR2  1141.590077  1143.587628  1126.471181  1123.338512  1112.423452
AR3  1134.886635  1136.674558  1112.784184  1122.204501  1102.219215
AR4  1127.855106  1108.940959  1112.076394  1114.189645  1102.979792

=== d=0, D=1, Seasonal (P,Q)=(0,1) ===
             MA0          MA1         MA2         MA3         MA4
AR0  1114.860068  1036.318771  943.079328  910.351363  882.447364
AR1   911.721194   859.682397  842.887396  832.220270  823.443681
AR2   845.351783   840.965273  842.850358  833.865549  824.578960
AR3   847.103855   842.965236  818.554249  816.523885  815.512683
AR4   853.749902   844.620485  820.554069  816.314187  817.243500

=== d=0, D=1, Seasonal (P,Q)=(1,0) ===
             MA0          MA1         MA2         MA3         MA4
AR0  1124.937526  1055.819301  967.067736  945.219352  925.264464
AR1   951.908239   905.019265  879.287699  880.672439  876.874151
AR2   865.630472   867.434367  868.424051  868.630131  868.124152
AR3   864.685374   848.956688  850.041761  852.031201  853.678306
AR4   857.499842   846.008168  847.942245  849.461301  850.682220

=== d=0, D=1, Seasonal (P,Q)=(1,1) ===
             MA0          MA1         MA2         MA3         MA4
AR0  1109.340817  1038.319204  951.003092  913.822421  889.484823
AR1   915.569652   862.883514  833.757075  825.607945  817.013427
AR2   836.998058   837.277984  832.895972  824.832874  806.716057
AR3   833.192575   810.161880  809.731369  807.942363  806.237800
AR4   824.018117   806.088616  808.037183  809.516459  807.729958

=== d=1, D=0, Seasonal (P,Q)=(0,0) ===
             MA0          MA1          MA2          MA3          MA4
AR0  1232.924854  1163.761870  1140.445635  1124.758371  1118.968376
AR1  1149.043889  1148.581641  1137.402921  1122.707598  1119.667925
AR2  1148.535889  1150.457844  1137.218068  1124.125726  1121.615430
AR3  1147.000483  1138.459874  1133.364876  1126.047625  1123.262622
AR4  1123.811389  1122.515315  1124.209375  1126.107082  1122.790528

=== d=1, D=0, Seasonal (P,Q)=(0,1) ===
             MA0         MA1         MA2         MA3         MA4
AR0  1021.828775  938.314504  902.764485  901.187437  894.015608
AR1   924.975339  907.043410  900.528377  899.612392  895.308383
AR2   924.945788  908.634203  901.367324  899.280780  897.202476
AR3   924.448015  904.357654  903.365089  900.396845  899.201977
AR4   925.450948  906.199766  904.296587  902.357343  900.787288
```

```
=== d=1, D=0, Seasonal (P,Q)=(1,0) ===
            MA0         MA1         MA2         MA3         MA4
AR0  1024.361858  958.803413  929.808939  930.386041  926.304263
AR1   925.094187  925.543726  925.425744  926.614269  923.748989
AR2   906.549081  909.001173  903.153148  904.565032  904.229900
AR3   899.635494  901.297112  902.851898  904.070213  903.294147
AR4   898.354034  899.409340  900.290933  902.271418  903.020159

=== d=1, D=0, Seasonal (P,Q)=(1,1) ===
            MA0         MA1         MA2         MA3         MA4
AR0  1023.814209  940.122463  904.244763  902.777848  895.587514
AR1   926.906671  908.995602  901.799807  901.200314  896.898413
AR2   908.536657  910.662996  902.669641  900.818315  898.815869
AR3   901.503534  903.140877  904.662980  901.888658  900.815778
AR4   900.141072  900.836915  902.025621  903.887324  902.394791

=== d=1, D=1, Seasonal (P,Q)=(0,0) ===
            MA0          MA1          MA2          MA3          MA4
AR0  1258.484894  1168.918804  1133.989155  1124.702007  1121.771733
AR1  1153.929996  1145.102097  1135.974191  1125.885159  1122.827551
AR2  1144.951697  1146.649658  1137.398106  1127.651564  1117.815832
AR3  1139.695238  1139.314100  1138.308076  1129.333098  1124.908370
AR4  1125.928121  1127.923362  1129.916497  1131.134823  1126.562998

=== d=1, D=1, Seasonal (P,Q)=(0,1) ===
           MA0         MA1         MA2         MA3         MA4
AR0  916.980134  859.915712  838.487224  836.772736  830.452971
AR1  843.257999  842.721646  840.416721  838.551198  832.452963
AR2  845.257938  844.406955  839.605900  839.811783  833.813678
AR3  847.257936  846.361227  833.084405  840.430442  835.052504
AR4  853.080065  848.972791  835.950507  835.266524  833.708483

=== d=1, D=1, Seasonal (P,Q)=(1,0) ===
           MA0         MA1         MA2         MA3         MA4
AR0  985.277372  912.193728  891.059229  891.126922  889.898624
AR1  872.983120  874.968801  876.803214  877.884153  879.880353
AR2  872.242673  873.065254  875.561850  876.323173  879.042480
AR3  866.345873  867.925740  867.418967  870.971598  872.936637
AR4  864.810765  866.687919  868.199717  868.099084  870.269027

=== d=1, D=1, Seasonal (P,Q)=(1,1) ===
           MA0         MA1         MA2         MA3         MA4
AR0  920.925065  863.636996  829.568994  827.757916  821.269350
```

AR1	840.638712	837.997557	829.929511	826.458423	822.978366
AR2	837.674016	840.068580	828.941959	826.663903	823.785029
AR3	829.697364	829.220473	830.152947	828.149548	826.866422
AR4	826.360031	827.129276	828.461955	828.445638	827.395603

8.5 Additional Limitations

Another issue is that using $N = 100{,}000$ and treating the rate as pseudo case counts may not be fully appropriate for a POMP model; actual case counts with a proper population denominator would be a more principled choice.

Finally, although the selected SARIMA model produces interpretable fits and reasonable residual diagnostics, the regression term uses only time and a COVID step. Other covariates, such as temperature, humidity, school-term indicators, or age-stratified hospitalization, could plausibly explain more of the variation and could be added in future work.