Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Data-Driven Change-Point SEIR Modeling of the 2025 COVID-19 Epidemic Wave in China: Model Comparison, Robustness Analysis, and Residual Diagnostics

Lilin Xu

2021109613@qq.com

Santa Monica College

Research Article

Keywords: SEIR model, COVID-19, change-point analysis, time-varying transmission, sensitivity analysis, residual diagnostics, China

Posted Date: August 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10804129/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Abstract

Background: National COVID-19 surveillance in China showed a pronounced epidemic wave during spring 2025, followed by a sustained decline through the summer. This study examined whether that trajectory could be adequately represented by a standard SEIR model with a constant transmission rate or whether the data supported a structural change in transmission intensity.

Methods: We analyzed 153 consecutive days of nationally reported COVID-19 cases from 1 May through 30 September 2025. Daily observations were digitized from official China CDC figures and calibrated to the corresponding monthly totals. A standard susceptible-exposed-infectious-removed (SEIR) model was fitted under two specifications: a constant-beta baseline and a piecewise-beta model with a change point selected by profiling candidate dates and minimizing the Akaike Information Criterion (AIC). The main analysis fixed the latent period at 3 days, the infectious period at 5 days, and used an observation scaling factor q to connect model incidence with reported cases. Model performance was assessed using R-squared, RMSE, AIC, and BIC. Robustness was evaluated through change-point, observation-scale, and biological-parameter sensitivity analyses, followed by residual diagnostics for day-of-week structure and temporal autocorrelation.

Results: The AIC profile selected 24 May 2025 as the optimal change point, two days before the observed case peak on 26 May. The piecewise model substantially outperformed the constant-beta model (R-squared 0.8690 vs. 0.6689; RMSE 1,861.07 vs. 2,959.03; AIC 2,313.85 vs. 2,451.74; BIC 2,329.00 vs. 2,460.83). The model-implied reproduction index declined from 1.4006 before the change point to 0.8725 afterward. This directional transition from expansion to contraction was preserved across all tested change-point dates, observation-scale values, and nine latent/infectious-period combinations, although goodness of fit deteriorated under some extreme biological assumptions. Residuals retained a strong weekly reporting pattern and positive autocorrelation.

Conclusions: A data-driven piecewise transmission rate provided a markedly better description of the 2025 epidemic wave than a constant-rate SEIR model. The strongest inference is a robust structural transition in transmission dynamics in late May rather than a precise estimate of any single biological parameter or a causal attribution to a specific intervention.

1. Introduction

Compartmental models are useful because they force a complicated epidemic into a small number of explicit assumptions. The classical SIR framework of Kermack and McKendrick provided the basic idea: people move between epidemiological states, and those transitions can be written mathematically [1, 2]. For COVID-19, I considered the SEIR form a better starting point than the simpler SIR model because it includes an exposed stage. That extra compartment represents the delay between becoming infected and becoming infectious. It is still a simple model, but it captures one feature of respiratory infection that matters for the timing of an epidemic wave.

The difficulty is that a simple SEIR model can also be too simple. A common assumption is that the transmission rate stays constant over the entire study period. Real epidemics rarely behave that way. Viral evolution, immunity, vaccination, seasonality, changes in contact behavior, mobility, testing, and reporting can all alter the effective transmission environment. If one constant value of beta is forced to describe both a rising and a falling phase, the model may end up fitting neither phase especially well. This is one reason time-varying, piecewise, and non-autonomous transmission models have become common in COVID-19 research [3, 10].

This issue becomes more important in later pandemic waves. Omicron-lineage viruses differ from earlier SARS-CoV-2 variants, and their epidemiological time scales are not identical. Published studies report shorter serial intervals and incubation periods, with pooled Omicron incubation estimates close to three days in several analyses [5, 9]. I therefore did not want to treat parameter values taken from 2020 or 2021 as fixed truths for a 2025 epidemic. Instead, I used literature-supported values as a baseline and then tested how much the conclusions changed when those assumptions were varied.

The literature also showed that I was not dealing with a new modeling idea. Chinese COVID-19 studies have already used meta-SEIRS systems, delayed SIR models, and other time-dependent formulations to describe reinfection and changing Omicron-era transmission [3, 10]. What interested me was a narrower problem. If a transmission change is introduced, how should its timing be chosen? Simply looking at the epidemic curve, selecting the observed peak as the change point, and then claiming that the model reproduces that peak would make the argument partly circular.

My interest in the 2025 wave came from the shape of the China CDC surveillance curve itself. Reported cases rose quickly in spring 2025, reached a high level in late May, and then declined for several months [6]. The change was visually obvious. That led to a simple question: could one constant transmission rate really explain the entire May-to-September trajectory? If not, the next question was whether the timing of a change could be estimated from the data instead of being chosen by eye.

For that reason, the change point was treated as part of model selection. I did not fix it at the observed peak. Candidate dates were fitted one by one, and the date with the lowest AIC was used as the preferred change point [12]. The observed peak was kept only as a reference for interpretation. This separation matters. A peak comparison based on the same series used for fitting can be an internal consistency check, but it is not independent validation. A true external validation would require another surveillance series that was not used to fit the model.

A second issue appeared when I tried to connect the SEIR equations to reported case counts. The model describes an underlying infection process, but reported cases are filtered through testing, health-care seeking, surveillance definitions, and reporting delays. Many SARS-CoV-2 infections are also never documented [15]. With only one national case series, a scale factor linking model incidence to reported cases is difficult to separate from the assumed initial numbers of exposed and infectious individuals. I therefore use q only as an observation scaling factor. I do not interpret it as an estimated biological reporting probability, and I later test the model across a wide range of q values.

The study therefore asks three main questions. First, does allowing one change in transmission improve the SEIR fit enough to justify the added complexity? Second, when the change date is selected from the data rather than fixed at the observed peak, where does it fall? Third, does the conclusion of expansion before the change and contraction afterward survive reasonable changes in model assumptions? I also examine the residuals rather than stopping at a good-looking fitted curve. The aim is not to claim that a simple SEIR model explains every feature of national COVID-19 surveillance. It is to identify which conclusions are stable, and which ones depend on choices made during modeling.

2. Data Sources and Preprocessing

2.1 Data source and study period

The analysis used publicly available national COVID-19 surveillance reports released by China CDC under the series “National COVID-19 Epidemic Situation.” The study window extended from 1 May through 30 September 2025, yielding 153 consecutive daily observations [6]. This interval captured the high-incidence phase around the spring epidemic peak as well as the subsequent multi-month decline, and therefore contained information from the late growth phase, the turning region, and the decay phase of a single national epidemic wave.

China CDC monthly reports present daily newly reported confirmed cases graphically and provide monthly totals in the accompanying text. Because a machine-readable daily CSV series was not published on the corresponding pages, the official figures were used as the source of within-month daily variation, while the official monthly totals served as calibration constraints. This design preserves the temporal shape visible in the official surveillance figures while anchoring the overall number of reported cases to the totals stated by the public-health authority [6].

2.2 Digitization and monthly calibration

Daily values were digitized from the official bar charts using WebPlotDigitizer [7]. Let x_d denote the digitized value for day d within month m, let X_m be the sum of all digitized daily values in that month, and let T_m be the official monthly total. Each digitized daily value was rescaled according to $x^*_d = x_d(T_m/X_m)$. This multiplicative calibration preserves the relative day-to-day pattern within the month while aligning the monthly scale with the official total. The preprocessing pipeline then standardized dates, checked for missing and non-numeric records, sorted observations chronologically, and generated the final 153-day calibrated daily case series.

The calibrated series reached its highest observed daily value on 26 May 2025. That date is epidemiologically informative, but it was deliberately not used as the primary model change point. The separation between “observed peak” and “model-selected change point” is central to the analysis. If the same peak date were used both to define the model and to judge the model's temporal performance, the resulting evidence would be partly self-referential. Here, the peak date is reported only after the change-point profile has been evaluated.

2.3 Data quality and interpretation limits

The dataset has two important measurement limitations. First, the daily series was reconstructed from graphical information. Monthly calibration controls the aggregate scale, but small day-specific digitization errors may remain because of image resolution, axis reading, or point placement. Second, reported confirmed cases are an observation of the surveillance system rather than a direct census of infection events. Daily counts can be influenced by testing access, health-care seeking, laboratory turnaround, reporting delays, administrative schedules, and day-of-week effects [14, 15]. Accordingly, the analysis emphasizes relative temporal structure and changes in transmission state rather than attempting to reconstruct the true total number of infections in China.

These limitations motivate the distinction between a mechanistic transmission model and an observation model. The SEIR equations represent a latent disease-transmission process, whereas the surveillance data represent the filtered, delayed, and incompletely observed output of that process. The present study uses a deliberately simple observation scaling relation and then examines the residuals for evidence of reporting structure that this simplification fails to capture. This approach does not eliminate measurement uncertainty, but it makes the limitations explicit and testable.

3. Mathematical Model and Statistical Methods

3.1 SEIR transmission model

I used a standard SEIR structure with four compartments: susceptible S(t), exposed E(t), infectious I(t), and removed R(t). The model assumes homogeneous mixing at the national scale. That assumption is clearly a simplification for a country as large as China, but it gives a transparent baseline for asking whether a change in transmission is needed. The equations are

$$dS/dt = -\beta(t)S(t)I(t)/N$$

$$dE/dt = \beta(t)S(t)I(t)/N - \sigma E(t)$$

$$dI/dt = \sigma E(t) - \gamma I(t)$$

$$dR/dt = \gamma I(t)$$

Here, N is the total population. The parameter beta(t) controls transmission, sigma moves individuals from the exposed to the infectious compartment, and gamma removes individuals from the infectious compartment [1, 2]. The quantities 1/sigma and 1/gamma can be read as the model's mean latent and infectious periods. For the main analysis I used 3 days and 5 days, respectively, so sigma = 1/3 day^-1 and gamma = 1/5 day^-1. I chose these as working biological assumptions, not as values estimated precisely from this case series. They are consistent with Omicron-era literature and are later varied in sensitivity analysis [5, 9, 10].

The differential equations were solved numerically with SciPy's solve_ivp using LSODA [8]. I tracked the cumulative flow from E to I and then took daily differences to obtain model incidence. This detail matters because the number currently in I is a prevalence-like quantity, whereas the surveillance series records newly reported cases. Using daily flow rather than the size of the infectious compartment keeps those two ideas separate.

3.2 Observation model and scaling factor q

The SEIR equations do not directly produce reported confirmed cases, so I added a simple observation layer. A scaling factor q maps model-generated daily incidence onto the scale of the reported series. In compact form, expected reported cases are proportional to q*sigma*E(t), or to the corresponding daily increment in the cumulative E-to-I flow. This is only a bridge between the mechanistic model and the surveillance data. It is not meant to reproduce the full testing and reporting system [15].

I intentionally avoid calling q a reporting rate. With one aggregate case series, q is strongly confounded with the starting sizes of E and I; several combinations can generate very similar fitted curves. The main analysis therefore fixes q = 0.30 as a reference scale. I then repeat the fit for q values from 0.10 to 0.99. If the main conclusion hardly changes, then the result is being driven mainly by the timing and shape of the epidemic curve rather than by one arbitrary assumption about absolute scale.

3.3 Constant-transmission baseline

I started with the simplest transmission assumption: beta(t)=beta for all 153 days. This constant-beta model is the baseline, not the model I expected to win. Its purpose is to ask whether one transmission environment can describe the late growth phase, the turning region, and the long decline at the same time. If it can, a change point is unnecessary. If it cannot, then allowing beta to change becomes a testable extension rather than complexity added for its own sake.

3.4 Piecewise transmission model and data-driven change-point search

The extended model uses two transmission rates, beta1 before a change point t_c and beta2 afterward. I did not set t_c to 26 May, even though that is the observed peak date. Instead, I searched all candidate dates that left at least 14 observations on both sides. For each date, the piecewise model was refitted, the residual sum of squares was recalculated, and AIC was recorded. The date with the lowest AIC was selected as the preferred change point [12].

This search is central to the study. It lets the data answer a question that could otherwise be decided subjectively. At the same time, the selected date should not be treated as a literal claim that one real-world event happened on that exact day. In a two-rate model, an abrupt break is a compact approximation to changes that may actually be gradual and caused by several processes at once.

3.5 Parameter estimation and model-comparison criteria

Model parameters were fitted by minimizing squared differences between observed and predicted daily cases. I used several fit measures because each answers a slightly different question. RMSE keeps the error on the scale of daily cases. R-squared summarizes how much of the observed variation is reproduced. AIC and BIC ask whether improved fit is large enough to justify extra model flexibility, with BIC penalizing complexity more strongly [12]. No single number was treated as sufficient on its own.

I did not use a standard one-degree-of-freedom likelihood-ratio test as the main comparison between the constant-beta and free-change-point models. Under the no-change model, the location of t_c does not exist as an identifiable parameter, so the usual regularity conditions behind that test are not fully satisfied. For that reason, the main comparison relies on AIC, BIC, RMSE, and R-squared. A fixed-date piecewise model can still be used as a secondary reference comparison, but I do not treat that result as the main evidence for a freely estimated change point [12, 13].

3.6 Model-implied reproduction index

Because gamma is fixed in the main analysis, beta/gamma provides a convenient within-model summary of transmission strength. I report R_before=beta1/gamma and R_after=beta2/gamma. Values above one indicate an expanding regime in this model, while values below one indicate contraction. I deliberately call these model-implied reproduction indices rather than classical R0 values. The Chinese population in 2025 was not fully susceptible, and the model does not represent all immunity, age, contact, or spatial structure [4]. Their main use here is comparison before and after the estimated change.

3.7 Sensitivity analyses

I used three sensitivity checks. First, I moved the change point to 20 May, 23 May, 26 May, 29 May, and 1 June and refitted the model each time. Second, I repeated the analysis with q = 0.10, 0.30, 0.50, 0.70, and 0.99. Third, I crossed latent periods of 2, 3, and 4 days with infectious periods of 4, 5, and 7 days, giving nine biological scenarios. For every run I recorded both the direction of R_before and R_after and the overall fit. I kept those two ideas separate on purpose: a conclusion can remain directionally stable even when a particular parameter combination fits the data poorly.

3.8 Parameter-bound and identifiability diagnostic

I also performed one diagnostic fit with wider biological bounds: 1–7 days for the latent period and 3–10 days for the infectious period. This was not an attempt to estimate clinical durations from national case counts. The narrower question was whether the optimizer was simply being pushed against the limits I had imposed. If the solution remained away from those boundaries, that would reduce concern about a purely numerical boundary effect, while still leaving the underlying identifiability problem in place.

3.9 Residual diagnostics

Finally, I checked what was left in the errors. A smooth epidemic model can fit the broad curve and still miss a strong reporting pattern. COVID-19 surveillance often differs between weekdays and weekends

because testing, clinic use, laboratory processing, and reporting schedules are not constant across the week [14]. I defined residuals as observed minus predicted cases, compared mean weekend and weekday residuals with Welch's t test, and calculated the lag-7 residual correlation and Durbin-Watson statistic. If a weekly pattern remained, I interpreted it as evidence that the transmission model was capturing the broad epidemic shape while the observation process was still incomplete.

4. Results

4.1 Data-driven change point and AIC profile

Profiling candidate change-point dates produced a pronounced low-AIC region in late May 2025. The minimum occurred on 24 May, which was therefore selected as the preferred transmission change point. This date was obtained from fitting the complete case series and was not fixed from the observed daily maximum. The observed peak occurred on 26 May, approximately two days later. The ordering is compatible with the general idea that a decline in transmission intensity can precede the visible maximum of reported incidence because infections already generated before the change continue to progress through latent and reporting processes. Nevertheless, the two-day difference should not be overinterpreted as evidence for a specific causal lag.

The full AIC profile also demonstrates why the change-point analysis is more informative than simply assigning the observed peak as the split date. Candidate dates outside the late-May region generally produced substantially higher AIC values. This pattern indicates that the data favor a structural transition in a restricted temporal window rather than responding equally well to any arbitrary split. The resulting profile is therefore part of the primary model-selection evidence rather than a decorative visualization.

4.2 Comparison of constant-beta and free-change-point models

The constant-beta SEIR model captured part of the overall decline but was unable to reproduce the rapid increase during May, the high-incidence turning region, and the subsequent sustained decrease. Allowing the transmission rate to change at the data-selected date improved every reported fit measure. R-squared increased from 0.6689 in the constant-beta model to 0.8690 in the free-change-point model. RMSE decreased from 2,959.03 to 1,861.07 daily cases, a reduction of about 37.1%. AIC decreased from 2,451.74 to 2,313.85, while BIC decreased from 2,460.83 to 2,329.00.

The information-criterion differences are large relative to the penalty imposed for the additional model flexibility. The AIC improvement is 137.89 points and the BIC improvement 131.83 points. Because both criteria penalize additional parameters, the preferred piecewise model is not simply benefiting from having more adjustable quantities. Instead, the improvement indicates that the temporal structure added by the change in β explains features of the epidemic curve that a constant transmission rate cannot represent efficiently.

Visual comparison of the fitted trajectories reinforces this interpretation. The constant-beta curve tends to compromise between the May rise and the later decline, whereas the piecewise model follows the turning region and subsequent decay more closely. The piecewise model does not pass through every daily fluctuation, nor should it be expected to do so. The observed series contains clear high-frequency reporting variation, while the mechanistic SEIR structure is intended to represent the smoother underlying epidemic process. The fit is therefore evaluated in terms of broad trajectory, information criteria, and residual structure rather than pointwise interpolation.

Metric	Constant-beta model	Free-change-point model
R^2	0.6689	0.8690
RMSE	2,959.03	1,861.07
AIC	2,451.74	2,313.85
BIC	2,460.83	2,329.00

4.3 Transmission-rate estimates and stage-specific transmission state

Under the reference biological assumptions, the free-change-point model estimated β1 = 0.280118 before 24 May and β2 = 0.174493 afterward. With γ fixed at 0.2 day^-1, these values correspond to R_before = 1.4006 and R_after = 0.8725. The fitted transmission rate therefore fell by about 37.7% between the two regimes. Within the model, the pre-change value above one is consistent with an expanding transmission state, whereas the post-change value below one is consistent with the sustained decline observed over the following months.

These ratios should be interpreted as model-implied stage-specific transmission indices rather than as universal estimates of the basic reproduction number of the circulating SARS-CoV-2 lineage. Classical R0 is defined under conditions close to complete susceptibility, whereas the Chinese population in 2025 had accumulated immunity through vaccination and multiple previous waves of infection. In addition, the national model averages over geographic, demographic, and contact heterogeneity. The strongest use of R_before and R_after is therefore internal comparison: under the same model and observation assumptions, the system changes from a regime above the epidemic threshold to one below it.

Parameter	Estimate/assumption	Interpretation
t_c	24 May 2025	AIC-selected change point
β1	0.280118	Pre-change transmission rate
β2	0.174493	Post-change transmission rate
Latent period	3 days	Fixed in primary analysis
Infectious period	5 days	Fixed in primary analysis
q	0.30	Reference observation scale
R_before	1.4006	Expanding regime
R_after	0.8725	Contracting regime

4.4 Internal consistency of peak timing

The reconstructed model incidence reached its maximum around 23 May, whereas the observed reported-case peak occurred on 26 May. The two dates are close, but their proximity should be described as an internal consistency check rather than an external validation. Both the fitted model and the peak comparison use the same underlying surveillance series. Importantly, the 26 May observed peak was not used to set the primary change point, so the agreement is not mechanically built into the model-selection procedure. Even so, an independent validation claim would require a separate data source that did not contribute to parameter fitting.

4.5 Sensitivity to prespecified change-point dates

When the change point was fixed at 20 May, 23 May, 26 May, 29 May, or 1 June and the model was refitted, every scenario retained the same directional pattern: R_before remained above one and R_after remained below one. Thus, the central inference of a transition from expansion to contraction does not depend on placing the structural change exactly on 24 May or exactly on the observed peak date.

At the same time, the models were not equally well supported. Moving the change point too far from the favored late-May region reduced goodness of fit. In particular, a 1 June split produced a noticeably poorer representation of the observed trajectory. This distinction is important. Robustness of the sign or direction of an inferred transition does not imply that all alternative change points are equally plausible. The profile and sensitivity analyses jointly indicate both a robust qualitative transition and a preferred temporal location for that transition.

4.6 Sensitivity to the observation scaling factor

Varying q from 0.10 to 0.99 had little effect on the fitted transmission-state indices. Across this wide observation-scale range, R_before remained close to 1.40, and R_after remained near 0.87–0.88. The stability is consistent with the fact that the temporal shape of the reported epidemic curve, rather than

its absolute latent infection scale, carries most of the information used to distinguish the pre- and post-change transmission regimes.

This result also reinforces the decision not to interpret q as an epidemiological reporting-rate estimate. The analysis does not claim that 30%, 50%, or 99% of infections were reported in China during the study period. Instead, q serves as a nuisance scale linking latent model incidence to the observed series. The fact that the qualitative transmission conclusion persists across nearly an order of magnitude in q suggests that the principal result is not an artifact of choosing a particular scale.

4.7 Sensitivity to latent and infectious period assumptions

The nine combinations formed from latent periods of 2, 3, and 4 days and infectious periods of 4, 5, and 7 days all preserved $R_{before} > 1$ and $R_{after} < 1$. The direction of the inferred transition is therefore robust to a broad set of plausible biological assumptions. This is an important result because aggregate national case data generally do not contain enough independent information to estimate latent and infectious periods with high precision while simultaneously estimating transmission and observation parameters.

However, goodness of fit was not uniformly strong across all nine scenarios. Across the biological-parameter sensitivity runs, R-squared ranged from − 0.295 to 0.869, and the poorest scenario therefore produced a negative R-squared. A negative value indicates that, under that biological assumption, the fitted model can perform worse than a simple benchmark based on the mean of the observations. I did not remove or average away these weak fits. They are useful because they show that directional robustness and quantitative model adequacy are different properties.

The most defensible interpretation is therefore that the case curve strongly supports a transition from an expanding to a contracting regime, while providing weaker support for extreme combinations of latent and infectious duration. The reference 3-day latent and 5-day infectious assumptions provide a useful baseline, but the sensitivity results argue against presenting those values as if they had been precisely learned from the national surveillance series. The model is more informative about structural transmission change than about fine-scale biological timing.

4.8 Parameter-bound and identifiability diagnostic

When the allowable latent and infectious period ranges were widened for diagnostic purposes, the fitted solution did not appear to be driven to the imposed boundaries. This finding reduces concern that the reference solution is merely a numerical consequence of an artificially narrow search space. It does not, however, establish that the national case series identifies the biological durations precisely. Given the known confounding among transmission, initial-state, observation-scale, and progression parameters, the more conservative strategy remains to anchor the primary biological durations to external evidence and use sensitivity analysis to determine whether substantive conclusions depend on them.

4.9 Residual structure and day-of-week effects

Despite the improved overall fit, residuals from the free-change-point model were strongly structured in time. Mean weekend residuals were − 1,793.8 cases, whereas mean weekday residuals were + 690.0 cases. The difference was strongly significant in the Welch comparison ($p = 6.694 \times 10^{-15}$). The lag-7 residual correlation was 0.701 and the Durbin-Watson statistic was 0.998, both indicating substantial serial structure rather than independent random errors.

The sign pattern is consistent with systematic overprediction on weekends and underprediction on at least some weekdays. Such a weekly cycle is more plausibly explained by testing, health-care utilization, laboratory processing, and data-reporting schedules than by true biological transmission changing sharply on fixed days of every week. Similar day-of-week effects have been documented in COVID-19 surveillance data [14]. The residual analysis therefore clarifies the scope of the mechanistic model: it captures the broad multi-week epidemic trajectory, but it does not fully model the observation process that converts infections into daily reported case counts.

This residual structure also has statistical implications. Ordinary least-squares fitting implicitly treats errors much more simply than the observed residual series warrants. Positive autocorrelation can affect uncertainty estimates, and count data often exhibit mean-variance relationships and overdispersion. The residual diagnostics therefore motivate future work using explicit day-of-week reporting multipliers, delay distributions, or Poisson/negative-binomial observation models rather than implying that the present SEIR fit is a complete description of the surveillance-generating mechanism.

5. Discussion

5.1 Principal findings

The principal finding of this study is structural rather than numerical. Three results support the same interpretation. First, the change-point profile identifies late May as a preferred transition region without using the observed peak date as a fixed input. Second, the piecewise-beta model provides a markedly better description of the observed national trajectory than the constant-beta model, even after information criteria penalize additional complexity. Third, the inferred transition from $R_{before} > 1$ to $R_{after} < 1$ persists across alternative change points, observation scales, and biological parameter combinations. Taken together, these results support a change in the effective transmission environment during late May 2025.

This conclusion is stronger than a statement based solely on a high R-squared. A model can achieve high apparent fit for many reasons, including overparameterization. Here, the improvement is accompanied by large reductions in AIC and BIC, a data-driven profile for the structural date, multiple sensitivity analyses, and explicit residual diagnostics. These additional steps do not remove all uncertainty, but they distinguish the conclusions that are robustly supported from quantities that depend more heavily on modeling assumptions.

The analysis also illustrates why a deliberately modest mechanistic model can still be useful. More complex epidemic models can include age structure, spatial mobility, vaccination states, multiple variants, reinfection, waning immunity, and reporting delays. Such models may be more realistic when the necessary data are available. The purpose of the present study is different: to ask whether the simplest interpretable SEIR framework requires a structural change in transmission to describe one national epidemic wave. For that question, a low-dimensional piecewise model provides a transparent comparison that can be scrutinized and reproduced.

5.2 Relationship to previous SEIR and time-varying transmission studies

The finding that a time-varying transmission parameter better describes an epidemic wave is consistent with a large body of COVID-19 modeling work. Studies in China and elsewhere have used time-dependent or non-autonomous parameters to represent changes in transmission conditions, including reinfection and Omicron-era dynamics [3, 10]. The present analysis contributes to that general literature not by proposing a new compartmental system, but by combining a data-driven structural break, explicit model comparison, sensitivity testing, and residual diagnosis on a 2025 national surveillance series.

The estimated pre-change reproduction index of approximately 1.40 should not be directly compared with published early-Omicron basic reproduction numbers as if the quantities were identical. Liu and Rocklöv reported substantially larger R0 estimates for early Omicron, but R0 is defined under conditions approximating a fully susceptible population [4]. By 2025, the Chinese population had experienced vaccination and multiple infection waves, including large Omicron-era outbreaks, so susceptibility and contact structure were fundamentally different [11]. The present β/γ ratios are better viewed as stage-specific indices within a partially immune, highly aggregated population.

This distinction matters because apparent discrepancies in reproduction numbers can otherwise be misinterpreted as evidence that one study is “wrong.” In reality, reproduction measures depend on population immunity, contact patterns, model structure, observation period, and the specific definition being used. The value of R_before and R_after in the present study lies primarily in their within-wave contrast under common assumptions. Their crossing of the threshold at one is more informative for the study's research question than their absolute numerical comparison with estimates from earlier pandemic phases.

5.3 Why the decline in beta should not be interpreted as a specific intervention effect

The estimated decrease in β during late May indicates a reduction in the composite transmission intensity represented by the model, but it does not identify the cause of that reduction. The model contains no direct covariates for public-health policy, mobility, contact behavior, meteorological conditions, variant composition, testing intensity, or changing immunity. Any of these mechanisms,

acting alone or together, could contribute to a shift in the effective transmission environment. The analysis therefore does not support a claim that a particular intervention “caused” the decline.

For the same reason, a counterfactual simulation that holds the pre-change β1 constant throughout the study period should not be labeled a “no-intervention” scenario. A more accurate description is “a scenario in which the estimated transmission-rate reduction does not occur.” This wording retains the mathematical meaning of the counterfactual while avoiding causal language that the data and model cannot justify. Establishing a specific causal driver would require additional external information, such as intervention timing, mobility data, contact surveys, viral lineage frequencies, or other covariates, together with an identification strategy designed for causal inference.

This caution is not merely semantic. In epidemic modeling, the transmission parameter often absorbs multiple processes that are not explicitly represented. A decrease in β may summarize fewer contacts, lower per-contact transmission probability, changes in population susceptibility, changes in who is being tested, spatial reweighting of regional outbreaks, or several of these effects simultaneously. Treating β as an interpretable composite parameter is appropriate; assigning it to one real-world cause without supporting data is not.

5.4 What the sensitivity analyses contribute to model credibility

The sensitivity analyses reveal an important distinction between robust qualitative inference and robust quantitative fit. Moving the change point within the late-May window leaves the expansion-to-contraction pattern intact. Varying the observation scale over a very wide range also leaves this pattern essentially unchanged. Biological sensitivity analysis likewise preserves the direction of the transition in all nine scenarios. These results indicate that the central structural inference is not produced by one narrowly tuned parameter choice.

At the same time, the biological scenarios demonstrate that not every assumption is equally compatible with the data. Some combinations substantially degrade R-squared, including at least one scenario with a negative value. This is scientifically useful rather than embarrassing. It shows that the model has a well-defined domain of adequacy and that some biological assumptions are difficult to reconcile with the observed trajectory. A transparent manuscript should report this heterogeneity rather than present sensitivity analysis as a ritual in which every alternative produces an equally attractive result.

The q analysis has a different interpretation. The near-invariance of R_before and R_after across q values suggests a scale non-identifiability: changing the observation scale can be compensated by changes in latent compartment sizes without substantially altering the temporal transmission pattern. This is exactly why q should remain a nuisance observation parameter rather than be promoted to a biological estimate. Sensitivity analysis is therefore doing two jobs simultaneously: demonstrating robustness of the main conclusion and revealing which parameters the data are not well suited to identify.

5.5 Residual diagnostics and the observation process

The residual analysis is one of the most informative components of the study because it shows what the mechanistic model fails to explain. If only the fitted curve were considered, the piecewise model might appear to account for most of the epidemic dynamics. The strong weekly pattern in the residuals demonstrates that the surveillance series contains another layer of structure. The disease-transmission process and the case-reporting process are not the same stochastic mechanism.

Day-of-week effects can arise from reduced weekend testing, laboratory staffing, clinic attendance, administrative reporting, and catch-up reporting on subsequent weekdays. Simpson and colleagues documented substantial weekday variation in COVID-19 surveillance data, illustrating that these patterns can be large enough to affect time-series interpretation [14]. The present residual pattern is consistent with this broader literature. It also explains why a smooth SEIR curve should not be expected to pass through every daily observation: some of the high-frequency variation likely reflects reporting practice rather than equally rapid changes in biological transmission.

A natural extension is to separate the latent transmission model from a probabilistic observation model. One option would include day-of-week reporting multipliers. Another would convolve latent incidence with a reporting-delay distribution. For count data, Poisson or negative-binomial likelihoods could replace squared-error fitting, with the negative binomial particularly useful when the observed variance exceeds the mean. Such extensions would permit uncertainty intervals that better reflect the data-generating process and could alter formal model-comparison statistics, although the current structural findings provide a useful baseline for that work.

5.6 Broader value and generalizability of the modeling framework

From an applied perspective, the study demonstrates a relatively lightweight but transparent workflow for public surveillance data. Official graphical reports can be digitized and calibrated to stated totals; a standard mechanistic model can provide an interpretable dynamic backbone; information-criterion profiling can replace subjective selection of a structural model; sensitivity analyses can test dependence on key assumptions; and residual diagnostics can identify patterns that the model leaves unexplained. The workflow is therefore useful not because the exact parameter estimates should be transported to other epidemics, but because the sequence of analytical checks is reproducible.

The same strategy could be applied to other single-wave or stage-structured infectious-disease time series, provided that disease-specific biological assumptions and surveillance systems are reconsidered. Latent and infectious periods, reporting delays, age structure, spatial heterogeneity, and observation practices differ across pathogens and settings. The transferable lesson is methodological: begin with the simplest interpretable model, test whether additional structural complexity is supported by the data, and then constrain the claims through sensitivity and residual analysis.

6. Limitations

First, the daily case series was digitized from official graphical reports rather than obtained as a raw machine-readable daily dataset. Monthly calibration preserves official totals and reduces scale error, but it cannot eliminate small day-specific extraction errors [6, 7]. Future work would benefit from an official daily data release or from independent replication of the digitization procedure.

Second, reported confirmed cases are not equivalent to incident infections. The observation scaling factor q cannot be identified from the national case series as a true reporting probability because it is confounded with initial latent-state sizes and other model quantities [15]. Consequently, the model should not be used to infer the true cumulative number of infections in China from this fit alone.

Third, the model assumes homogeneous mixing at a national scale. China is geographically and demographically heterogeneous, and provincial epidemic waves need not be synchronized. A single national β averages over differences in age distribution, population density, urban-rural contact patterns, mobility, local immunity, and regional timing. The fitted structural change is therefore an aggregate national description rather than evidence that every region changed transmission simultaneously.

Fourth, β(t) is represented by a single abrupt change between two constants. Real transmission intensity may vary continuously or undergo several changes. The piecewise model is a low-dimensional approximation chosen for interpretability. A smoother time-varying β(t), spline formulation, state-space model, or multiple-change-point model could potentially describe the series more flexibly, but would require stronger regularization and a more careful treatment of overfitting.

Fifth, the latent and infectious periods are primarily literature-informed assumptions rather than parameters precisely identified from the national incidence curve. Although the sensitivity analysis supports the direction of the transmission-state change, poor fit under some biological scenarios shows that the observed series has limited ability to distinguish all plausible combinations [5, 9, 10].

Sixth, residuals exhibit clear weekly periodicity and positive serial dependence. These patterns violate the simplest independent-error interpretation of least-squares fitting and indicate that the observation model is incomplete [14]. Formal uncertainty intervals from a more realistic count likelihood with day-of-week and reporting-delay components would therefore be preferable in a future inferential analysis.

Seventh, the similarity between the model peak and observed peak is not an independent validation. The same case series contributes to fitting and to the comparison of peak timing. A stronger validation would compare model-inferred timing with an external surveillance measure such as sentinel-hospital positivity that was not used to fit the model.

Eighth, the study analyzes one national wave from May through September 2025. The estimated β values and reproduction indices should not be generalized automatically to other periods, viral lineages, or countries. Replication across additional epidemic waves would be necessary to determine which findings are persistent and which are specific to this episode.

7. Future Work

A priority is to replace the simple squared-error observation model with a probabilistic count model. A negative-binomial likelihood would allow for overdispersion, while explicit day-of-week multipliers and a reporting-delay distribution could separate transmission variation from administrative reporting patterns. Re-estimating the constant and piecewise models under such an observation process would test whether the model-selection advantage remains under more realistic error assumptions.

Second, uncertainty should be quantified for β1, β2, R_before, R_after, and the change point itself. Parametric bootstrap, residual bootstrap adapted to autocorrelation, profile intervals, or Bayesian posterior inference could provide interval estimates rather than single values. Change-point uncertainty is especially important because the AIC minimum lies within a broader late-May region rather than representing a physically exact discontinuity.

Third, an independent surveillance series should be used for external validation. Sentinel-hospital SARS-CoV-2 positivity, if available at appropriate temporal resolution, could be compared with the timing of the model-inferred structural transition. Agreement across distinct surveillance systems would provide stronger evidence than internal consistency within the same case-count series.

Fourth, provincial data would permit a multi-region or meta-population extension. Such a model could test whether the national curve is generated by a synchronized countrywide transition or by the aggregation of regional waves with different timings. Spatial structure would also allow mobility and regional population size to enter the transmission mechanism more explicitly.

Finally, if data on viral lineage composition, mobility, policy, weather, or contact behavior can be assembled, the model could be extended to investigate potential drivers of β(t). Such analyses should maintain a clear distinction between correlation and causation. Identifying a temporal association between a covariate and the estimated transmission change is not equivalent to establishing that the covariate caused the change.

8. Conclusion

Using 153 days of nationally reported COVID-19 surveillance data from China, this study compared a standard constant-transmission SEIR model with a piecewise-transmission model whose change point was selected directly from the data. The AIC profile identified 24 May 2025 as the preferred structural transition, approximately two days before the observed case peak. Allowing β to change at that point substantially improved fit, increasing R-squared from 0.669 to 0.869 and reducing RMSE, AIC, and BIC.

Under the reference biological assumptions, the model-implied reproduction index declined from 1.4006 before the change point to 0.8725 afterward, indicating a transition from an expanding to a contracting transmission regime. This directional conclusion remained stable across alternative change-point dates, observation-scale values, and nine latent/infectious-period scenarios, although some extreme biological

assumptions produced poor overall fit. Residual diagnostics additionally revealed a strong weekly reporting pattern and positive temporal dependence, demonstrating that the SEIR model captures the broad epidemic trajectory more successfully than the short-term reporting process.

The central contribution is therefore not a claim of precise biological parameter recovery or causal attribution to a particular policy. Rather, the analysis shows that a transparent data-driven structural change, evaluated with information criteria and challenged through sensitivity and residual diagnostics, provides a more defensible description of the 2025 epidemic wave than a single constant transmission rate. Future work using probabilistic count observation models, bootstrap or Bayesian uncertainty, regional data, and independent surveillance validation could strengthen the inferential framework while preserving the interpretability of the present approach.

Declarations

Data and Code Availability

The source surveillance information was obtained from publicly available monthly China CDC reports. We digitized daily values from official figures and calibrated them to the corresponding monthly totals. Data cleaning, numerical integration, change-point profiling, sensitivity analyses, residual diagnostics, and figure generation were implemented in Python using NumPy, Pandas, SciPy, and Matplotlib. For formal submission, the cleaned dataset, extraction/calibration script, model-fitting code, result tables, and figure-generation script should be archived together in a versioned public repository to support reproducibility [6–8].

Declaration of AI Use

During the preparation of this work, the author used generative AI tools (GPT-based assistants) to assist with Python code implementation and English-language translation and polishing. After using these tools, the author reviewed and edited all content as needed and takes full responsibility for this publication.

References

1. Kermack WO, McKendrick AG (1927) A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society A, 115*(772), 700–721. https://doi.org/10.1098/rspa.1927.0118
2. Hethcote HW (2000) The mathematics of infectious diseases. SIAM Rev 42(4):599–653. https://doi.org/10.1137/S0036144500371907
3. Zhu W et al (2024) Prediction of SARS-CoV-2 infection cases based on the meta-SEIRS model. Epidemiol Infect 152:e144. https://doi.org/10.1017/S0950268824001274

4. Liu Y, Rocklöv J (2022) The effective reproductive number of the Omicron variant of SARS-CoV-2 is several times relative to Delta. J Travel Med 29(3):taac037. https://doi.org/10.1093/jtm/taac037
5. Du Z, Liu C, Wang L, Bai Y, Lau EHY, Wu P, Cowling BJ (2022) Shorter serial intervals and incubation periods in SARS-CoV-2 variants than the SARS-CoV-2 ancestral strain. J Travel Med 29(6):taac052. https://doi.org/10.1093/jtm/taac052
6. Chinese Center for Disease Control and Prevention (2025) *National COVID-19 epidemic situation: Monthly reports, May–September 2025*.
https://www.chinacdc.cn/jksj/xgbdyq/202506/t20250605_307356.html
https://www.chinacdc.cn/jksj/xgbdyq/202507/t20250703_308176.html
https://www.chinacdc.cn/jksj/xgbdyq/202508/t20250806_309202.html
https://www.chinacdc.cn/jksj/xgbdyq/202509/t20250905_310229.html
https://www.chinacdc.cn/jksj/xgbdyq/202510/t20251011_312903.html
7. Rohatgi A *WebPlotDigitizer. WebPlotDigitizer - Copyright 2010–2024 Ankit Rohatgi*
8. Virtanen P et al (2020) SciPy 1.0: Fundamental algorithms for scientific computing in Python. Nat Methods 17:261–272. https://doi.org/10.1038/s41592-019-0686-2
9. Wu Y, Kang L, Guo Z et al (2022) Incubation period of COVID-19 caused by unique SARS-CoV-2 strains: A systematic review and meta-analysis. JAMA Netw Open 5(8):e2228008. https://doi.org/10.1001/jamanetworkopen.2022.28008
10. Li Z, Pei L, Duan G, Chen S (2024) A non-autonomous time-delayed SIR model for COVID-19 epidemics prediction in China during the transmission of Omicron variant. Electron Res Archive 32(3):2203–2228. https://doi.org/10.3934/era.2024100
11. Cai J, Zhang H, Zhu K et al (2024) Risk of reinfection and severity with the predominant BA.5 Omicron subvariant China, from December 2022 to January 2023. Emerg Microbes Infections 13(1):2292071. https://doi.org/10.1080/22221751.2023.2292071
12. Akaike H (1973) Information theory and an extension of the maximum likelihood principle. In *Proceedings of the 2nd International Symposium on Information Theory* (pp. 267–281)
13. Wilks SS (1938) The large-sample distribution of the likelihood ratio for testing composite hypotheses. Ann Math Stat 9(1):60–62. https://doi.org/10.1214/aoms/1177732360
14. Simpson RB, Lauren BN, Schipper KH, McCann JC, Tarnas MC, Naumova EN (2022) Critical periods, critical time points and day-of-the-week effects in COVID-19 surveillance data: An example in Middlesex County, Massachusetts, USA. Int J Environ Res Public Health 19(3):1321. https://doi.org/10.3390/ijerph19031321
15. Li R, Pei S, Chen B et al (2020) Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). Science 368(6490):489–493. https://doi.org/10.1126/science.abb3221

Figures

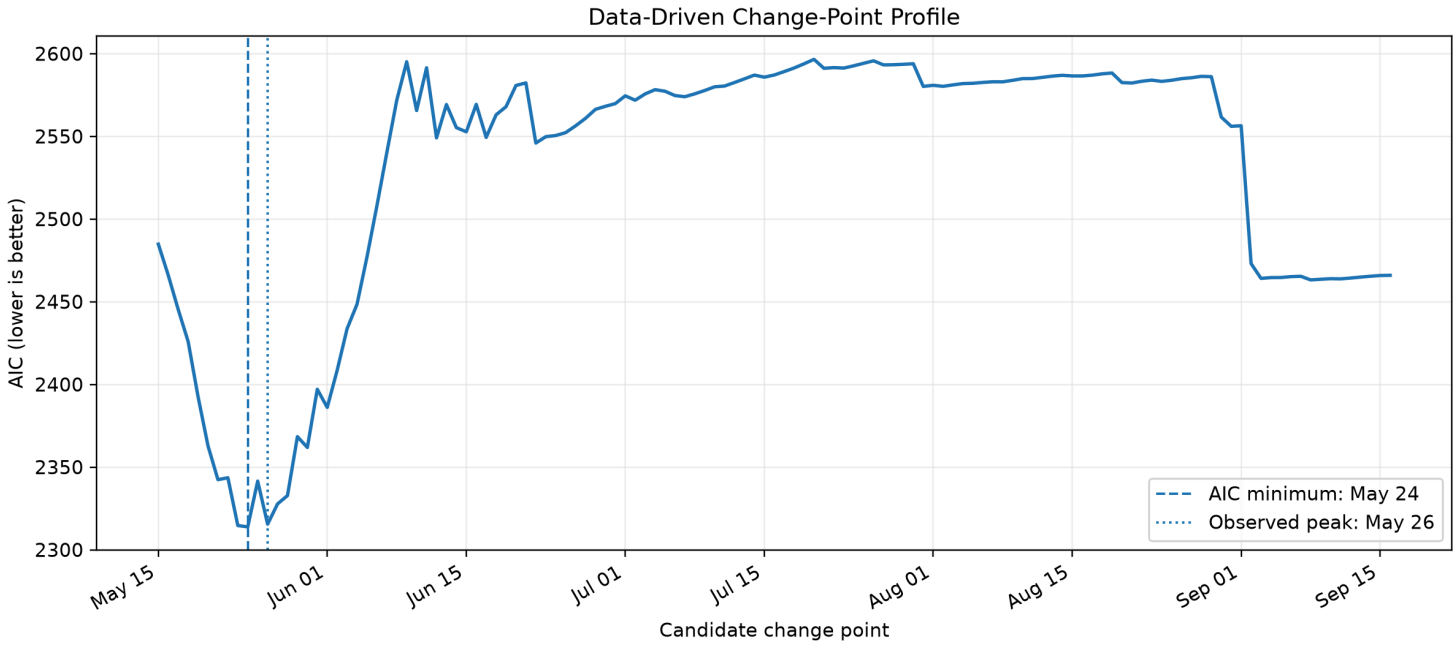


Figure 1

AIC values for candidate transmission change points. The minimum occurs on 24 May 2025; the observed case peak on 26 May is shown only as a contextual reference.

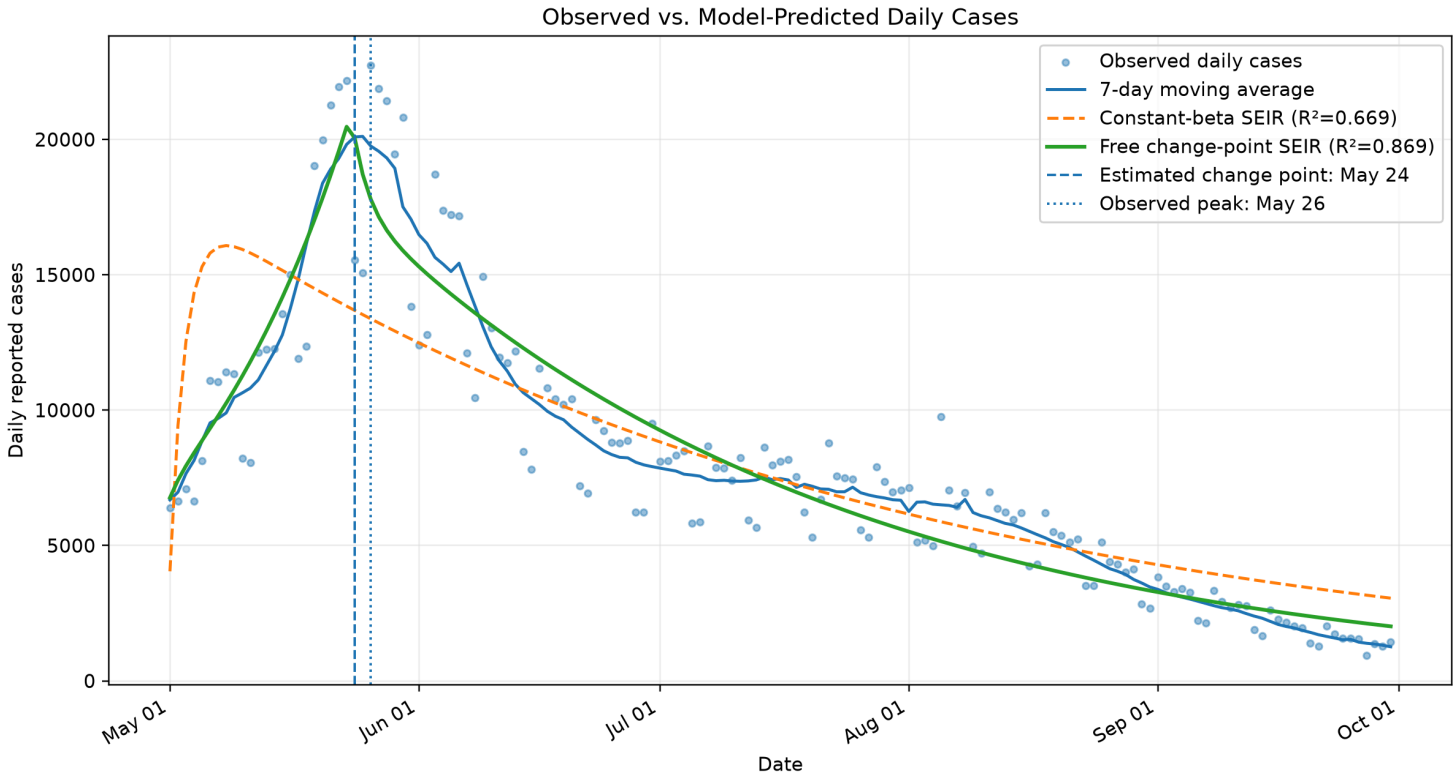


Figure 2

Observed daily cases, seven-day moving average, constant-beta SEIR fit, and free-change-point SEIR fit, with the estimated 24 May change point and the observed 26 May peak marked separately.

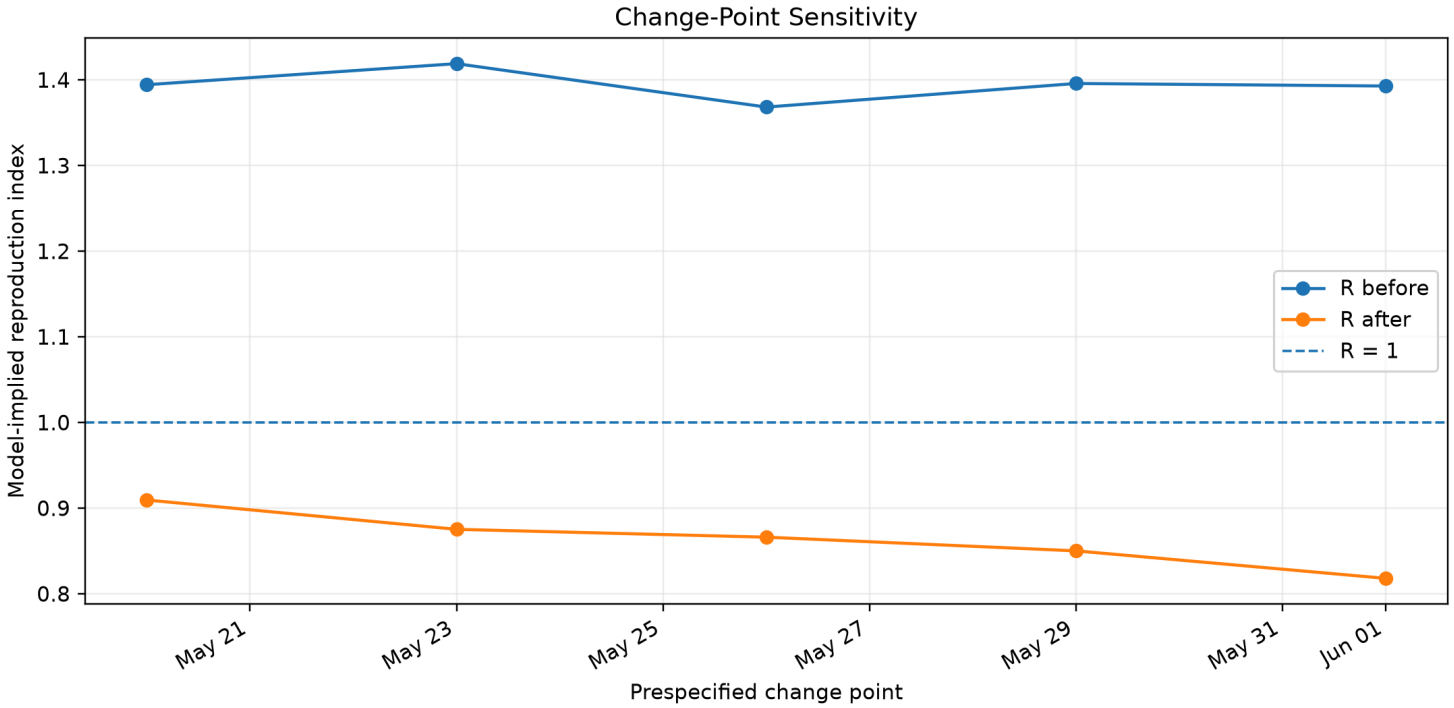


Figure 3

Sensitivity to prespecified change-point dates. Model-implied reproduction indices remain above one before and below one after each tested change point.

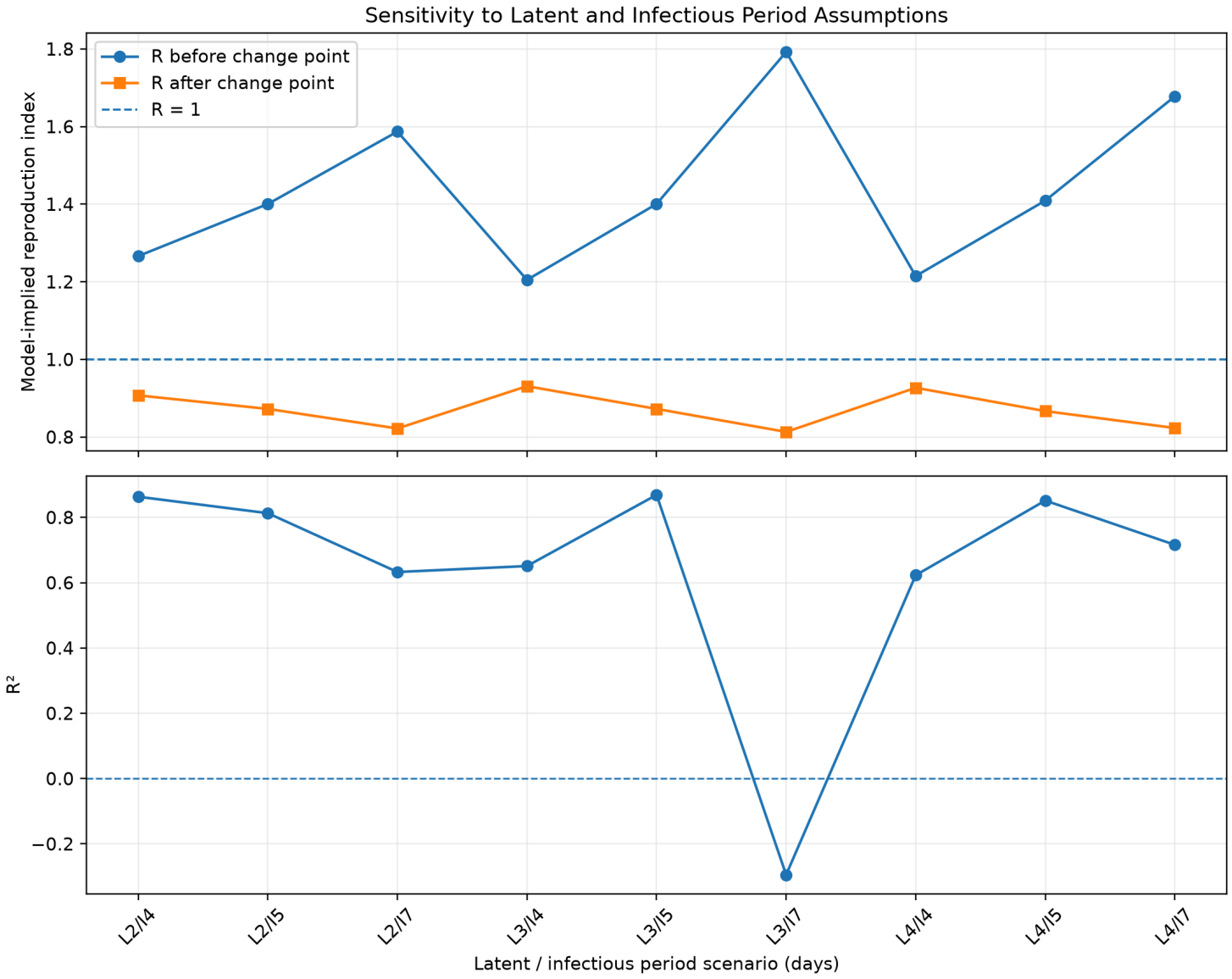


Figure 4

Top panel: R_before and R_after across nine biological scenarios. Bottom panel: corresponding R-squared values, illustrating that directional robustness can coexist with poor fit in some extreme scenarios.

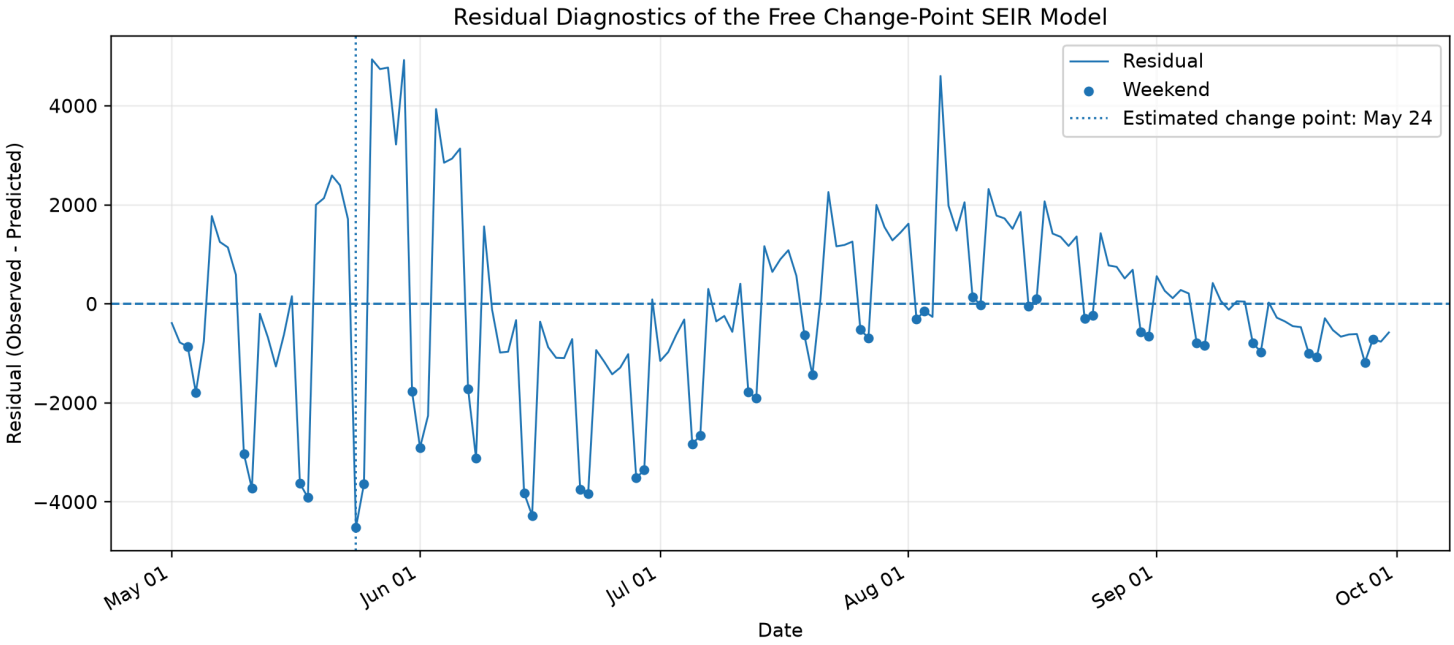


Figure 5

Daily residuals show systematic temporal structure and weekend effects, with the 24 May change point marked for reference.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FigS1qsensitivity.png
- FigS2finalSEIRcompartments.png
- FigS3counterfactual.png