https://doi.org/10.46344/JBINO.2025.v14i01.05

MODELING THE MONKEY POX OUTBREAK WITH THE REFINED SEIR MODEL INCLUDING VITAL DYNAMICS FOR THE US

Bryne Tan Jin-Yon Bryne & Fabiano de Abreu Agrela Rodrigues

Mestre emCiências- Northwestern University
https://0009-0000-0308-5183

Pós-PhdemNeurociências - Califórnia University FCE
Aveiro – Portugal
deabreu.fabiano@hotmail.com
https://orcid.org/0000-0002-5487-5852

ABSTRACT

The Monkeypox (Mpox) outbreak has renewed interest in using infectious disease modeling to predict and manage its spread. In this study, we apply an enhanced Susceptible-Exposed-Infectious-Recovered (SEIR) model to the Mpox epidemic in the United States, incorporating vital dynamics such as births and deaths to account for long-term population changes. By extending the traditional SEIR framework, our model captures more realistic infection dynamics, offering deeper insights into transmission patterns, peak infection periods, and the eventual decline of the outbreak. This approach aims to provide a clearer understanding of Mpox transmission dynamics and to assist health authorities in developing strategies to mitigate the spread. Our findings highlight the importance of early interventions and illustrate the potential for long-term epidemic control through strategic planning The SEIR model provides a basic framework for understanding the evolution of an epidemic [4] that divides the population into four segments: Susceptible (S), Exposed (E), Infectious (I), and Recovered (R). To model infectious diseases like Mpox more realistically, especially over an extended period, it is important to include vital dynamics, such as birth and death rates, in the SEIR model.

SEIR Model with Vital Dynamics

In the SEIR model with vital dynamics, the population is not fixed but changes over time due to births and deaths. This is particularly relevant for long-term modeling of diseases like Monkeypox.

Mathematically, the SEIR model is expressed as a system of coupled ordinary differential equations given by:

$$\frac{dS(t)}{dt} = \mu N - \beta \frac{S(t)I(t)}{N} - \mu S(t)$$
$$\frac{dE(t)}{dt} = \beta \frac{S(t)I(t)}{N} - \sigma E(t) - \mu E(t)$$
$$\frac{dI(t)}{dt} = \sigma E(t) - \gamma I(t) - \mu I(t)$$
$$\frac{dR(t)}{dt} = \gamma I(t) - \mu R(t)$$

where:

- $S(t)$, $E(t)$, $I(t)$, $R(t)$ represent the populations of susceptible, exposed, infectious, and recovered individuals at time t, respectively.
- β is the transmission rate of the disease, reflecting how effectively the disease spreads.
- σ is the rate at which exposed individuals become infectious.
- γ is the recovery rate, indicating how quickly infectious individuals recover and move into the recovered class.
- μ represents the natural death rate, assumed to be equal across all compartments.
- $N = S(t) + E(t) + I(t) + R(t)$ is the total population at time t, assumed constant in this model.

This set of coupled equations models the evolution of the epidemic by tracking the transitions between the different states ($S(t)$, $E(t)$, $I(t)$, and $R(t)$). Over time, it is expected that the disease dynamics will reach an equilibrium state.

Modeling Monkeypox with SEIR and Vital Dynamics

Mpox can be modeled via the SEIR framework with vital dynamics, particularly if the disease is expected to persist in the population over a long period. This approach allows the model to account for changes in population due to births and deaths, which are crucial for understanding the long-term behavior of the disease. Specifically, the population is not fixed but changes over time due to these vital dynamics [1, 2].

Steps for Modeling

1. **Determine Parameters:**
Transmission Rate (β): Reflects how easily Mpox spreads within the population.
Incubation Rate (σ): The rate at which exposed individuals become infectious.
Recovery Rate (γ): The rate at which infected individuals recover and become immune or are removed from the susceptible pool.
Birth and Death Rates: Include natural population birth and death rates, along with disease-related mortality if applicable.

2. **Implement the SEIR Model with Vital Dynamics:**
Use the differential equations provided above.
Implement the model numerically, using numerical solvers in Python.

3. **Analyze Results:**
Simulate the model over time to observe how monkey pox spreads, stabilizes, or declines in the population.
Analyze the impact of different scenarios, such as vaccination, isolation, or changes in transmission rate, on the outbreak's dynamics.

4. **Interpret the Results:**
Understand the long-term impact of monkeypox on the population, considering vital dynamics.
Explore public health strategies to mitigate the outbreak's effects, including vaccination strategies, quarantine measures, and public health awareness.

Methodology

The primary data on Mpox cases was sourced from the Global Health website, specifically through their publicly accessible repository. The dataset, provided in CSV format, included daily reports on confirmed cases, recoveries, and other critical epidemiological metrics across various regions. This dataset was instrumental in serving as the foundation for the model's input parameters. For the purposes of this study, we focused primarily on the US data due to its extensive and comprehensive coverage. By analyzing the trends and dynamics within the US, we aimed to generate insights that could be applied to other large, densely populated regions, offering predictions about potential outbreaks in similar environments.

Data Cleaning and Pre-processing

A Python script was utilized to refine the raw dataset, prioritizing variables crucial for the SEIR model. The refinement process was guided by the need to focus analysis on specific regions with significant data availability and clarity. The following steps were implemented:

• **Handling Missing Data**: Missing values in critical fields such as daily cases and deaths were managed using interpolation methods to maintain data integrity and continuity for time-series analysis. This approach ensures that temporal trends in the data are preserved, which is essential for accurate disease modeling.

• **Removing Duplicates**: Duplicate records were identified and removed based on unique identifiers including date and country. This step was crucial to ensure the reliability of the dataset, avoiding skewed data interpretation.

• **Date Formatting**: All dates were standardized to the ISO format (YYYY-MM-DD) to facilitate chronological analysis and modeling.

• **Feature Selection and Exclusion**: We streamlined the dataset by retaining only the most relevant features for epidemic modeling, such as Date, Location, New Cases, and Total Cases. Broad regional aggregates like continents were excluded to hone in on more granular, country-specific data, enhancing the model's specificity and applicability to targeted public health responses.

• **Data Verification**: Post-refinement, data was verified for consistency and accuracy by visualizing trends in daily new cases across selected high-impact countries. This step not only validated the refinement process but also provided preliminary insights into the disease's transmission dynamics.

The refined dataset was then exported for use in the SEIR model.

SEIR Model Setup

The SEIR model consists of four compartments:

• **S (Susceptible)**: The number of individuals susceptible to infection.
• **E (Exposed)**: The number of individuals

exposed but not yet infectious.
• **I (Infectious)**: The number of individuals actively spreading the disease.
• **R (Recovered)**: The number of individuals that have recovered and are no longer infectious.

The transitions between these compartments are governed by parameters that represent the rate of infection, latency period, and recovery rate:
• β (Infection Rate): The rate at which susceptible individuals become infected after contact with infectious individuals.
• σ (Latency Rate): The rate at which exposed individuals progress to the infectious stage.
• γ (Recovery Rate): The rate at which infectious individuals recover and move to the recovered category.

The cleaned data was fed into the model, with population-specific parameters estimated for different regions based on the US epidemiological data.

Parameter Estimation and Fitting
The SEIR model parameters (β, σ, γ) were estimated using the following techniques:
• **Nonlinear Least Squares Fitting**: This was applied to minimize the difference between the model output and the observed data.

• **Sensitivity Analysis**: Performed to understand how changes in parameters affect the model's accuracy.

• The SEIR model parameters (β, σ, γ) were estimated using techniques such as nonlinear least squares fitting [3].

Implementation
The cleaned data and estimated parameters were used in the Python package SciPy for numerical integration of the differential equations governing the SEIR model. The model was simulated over the time period corresponding to the dataset.

Step I:Optimisation with a Gaussian filter
We applied the following approach:
• Data Smoothing: A Gaussian filter was applied to the raw data to smooth out irregularities caused by reporting delays and anomalies. This pre-processing step helps in reducing noise in the observed data and produces a more accurate representation of the epidemic curve.
• Parameter Optimization: The transmission rate (β), recovery rate (γ), and exposure rate (σ) were adjusted iteratively to minimize the least squares difference between the modeled infections and the smoothed actual data. This process refines the model to capture the infection dynamics accurately.

Results
Two values of N were used $N = 230000000$ and $N = 330000000$ (to see if changing N had any effect on the optimization. For both cases, the value of maximum β was 0.25 from the sensitive analysis graph shown below.

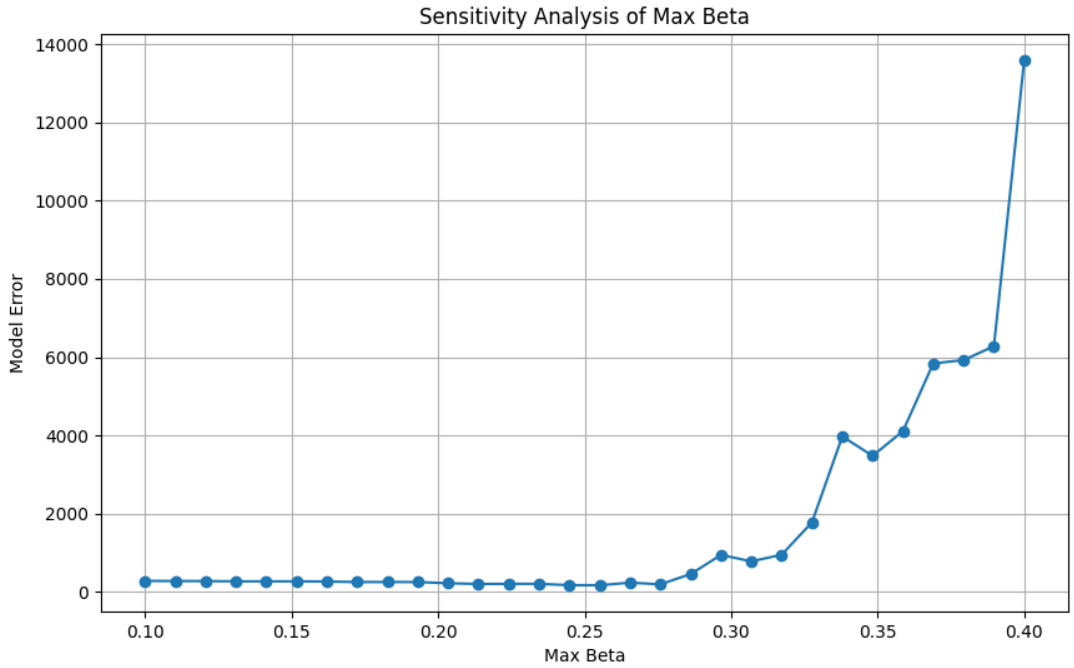

Figure 1: Sensitivity Analysis graphs shows that the optimized value of β is 0.25

Comparison of optimized SEIR model Gaussian Fit with data

The optimized SEIR model closely approximated the smoothed Mpox data from the United States. Key observations include:

- The early exponential growth of the epidemic was initially overestimated, likely due to limitations in early data.
- The model effectively captured the plateau phase, reflecting a slowing of infections attributed to public health measures and changes in population behavior.
- The eventual decline in case numbers was modeled with reasonable accuracy, offering valuable insights into the trajectory of the epidemic and the timing of control measures.

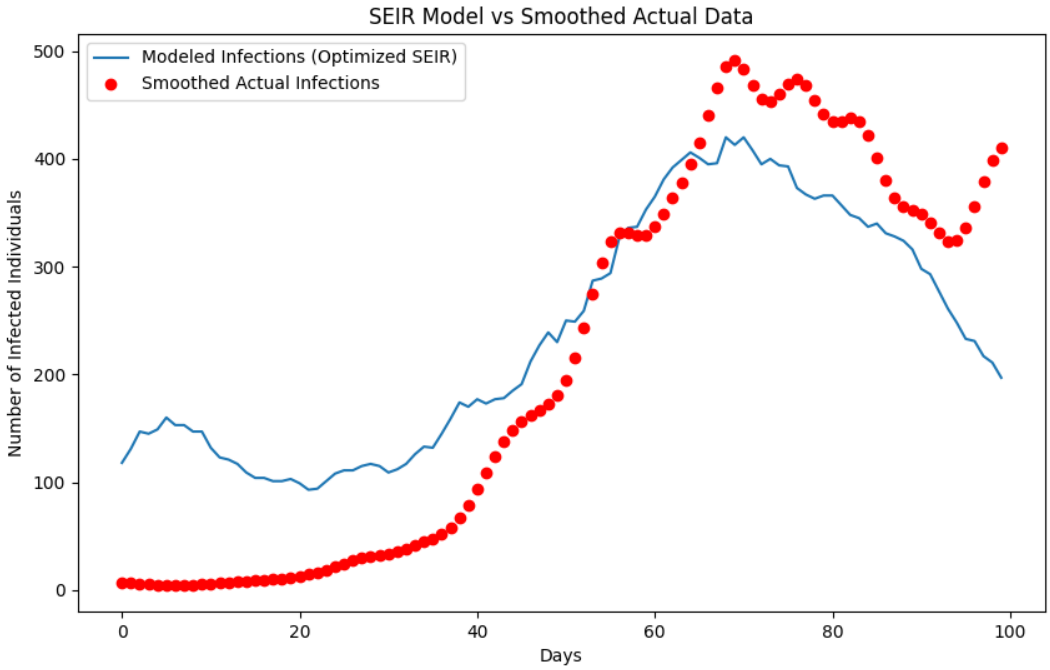

Figure 2: Comparison of Modeled Infections (Optimized SEIR) for $N = 213000000$ versus Smoothed Actual Data. The figure highlights the close alignment between the model's predictions and the reported cases, showing the epidemic's key phases: growth, plateau, and decline.

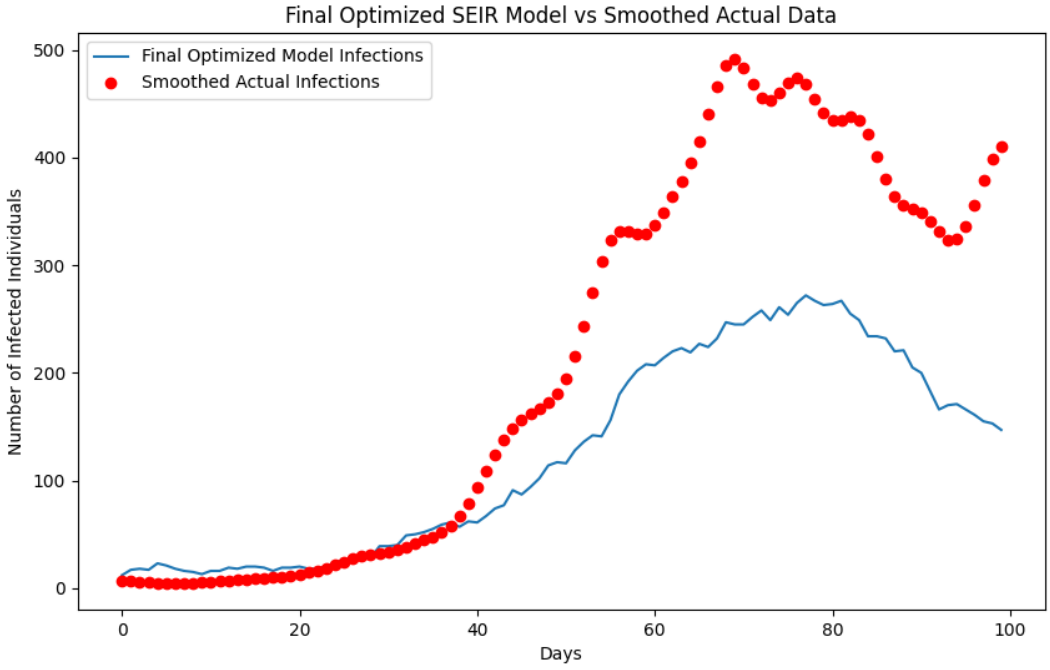

Figure 3: Comparison of Modeled Infections (Optimized SEIR) for $N = 330000000$ versus Smoothed Actual Data.

Comparison between the two graphs

• **Initial Phase (Up to Day ~40):**

The $N = 213$ million graph models the initial phase of the data well.
The $N = 330$ million graph overestimates the initial phase of the data.

• **Peak Phase (Day ~40 to Day ~70):**
For $N = 213$ million: The model slightly underestimates the peak compared to the actual data but follows a similar shape.
For $N = 330$ million: The model significantly underestimates the peak and diverges more from the actual data than the model with $N = 213$ million.

• **Decline and Subsequent Waves (Day ~70 Onwards):**
For $N = 213$ million: Post-peak, the model shows a decline that is smoother and more gradual compared to the actual data, which shows more fluctuation.
For $N = 330$ million: The model's decline is sharper than the actual data, and it fails to capture the secondary rise observed in the actual infection data.

Discussion
The results underscore the effectiveness of a hybrid approach for optimizing the SEIR model to fit real-world epidemic data. By integrating data smoothing and parameter refinement, the model achieved a closer fit to the Mpox data. The refined model provided:
• A clearer delineation of the epidemic phases, particularly illustrating the impact of public health interventions.
• Enhanced predictions for future case trends, supporting improved planning for healthcare resource allocation and epidemic management.

Future improvements could include incorporating additional data sources such as vaccination rates or the geographic distribution of cases. Furthermore, exploring alternative smoothing techniques could help to better address issues such as missing data or irregular reporting patterns.

Conclusion
The optimized SEIR model, enhanced with Gaussian filtering, provides a robust framework for analyzing the Mpox epidemic in the United States. By aligning the model more closely with actual case data, it offers valuable insights for public health strategies and intervention planning. Continued refinement and validation using emerging data will be essential for maintaining the model's accuracy and predictive power.

Step 2: Optimization of the Beta Parameter
Based on the earlier analysis, it was observed that the β parameter, which represents the transmission rate of the infection, is relatively independent of the population size N. This independence suggests that optimizing β does not require adjusting for different population sizes, allowing us to focus on its impact under a fixed population assumption.

Justification for the Independence of β

from N

The transmission rate β is a measure of the probability that an infectious individual will transmit the disease to a susceptible person. This rate is the a-priori contact probability and is independent of the total population size N.

Moreover, the maximum value of β remained stable across different population sizes ($N = 213{,}000{,}000$ and $N = 330{,}000{,}000$). This can be attributed to the fact that β is influenced primarily by the disease's inherent infectiousness and social dynamics, rather than by population size itself.

The second step in refining the SEIR model involved a detailed sensitivity analysis of the β parameter. Optimizing β is crucial because it directly influences the model's effectiveness in capturing the dynamics of the infectious spread accurately. By focusing on this parameter, we aim to enhance the model's predictive accuracy regarding how rapidly and extensively the infection spreads through the population.

For this phase of the analysis, we utilized a fixed population size of $N = 213{,}000{,}000$. This choice was based on preliminary findings that suggested a consistent model behavior across different population values, thus simplifying the optimization process. The sensitivity analysis aimed to identify the optimal β value that minimizes error between the model predictions and actual observed data, thereby ensuring that the model can reliably simulate the epidemic under study.

Methodology

The sensitivity analysis was conducted by systematically varying the beta parameter within a predetermined range and observing the resultant changes in the model's output. The objective was to minimize the discrepancy between the modeled infections and the smoothed actual data. The range for beta was set from 0.10 to 0.40, with increments of 0.01.

Results

The analysis revealed a nonlinear relationship between the beta parameter and the model error. Initially, the model error decreased significantly as beta increased from 0.10 to 0.25, indicating an improved fit to the actual data. However, beyond a beta of 0.25, the error began to increase sharply, suggesting over-fitting at higher rates of transmission.

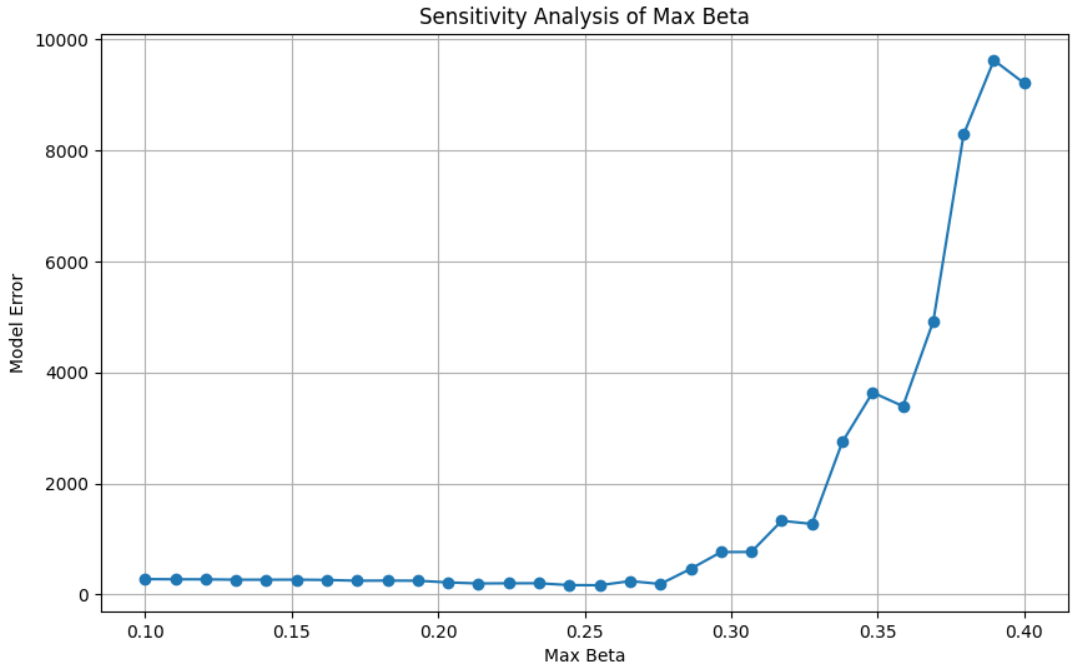

Figure 4: Sensitivity Analysis of the Beta Parameter

The optimal beta value was identified at approximately 0.25, where the model error was minimized, aligning closely with the trend observed in the smoothed actual infection data. This value was then used to adjust the SEIR model parameters for subsequent simulations.

Discussion

This sensitivity analysis underscores the importance of precise parameter estimation in epidemiological models. By optimizing beta, the SEIR model not only aligns more closely with actual data but also enhances its utility in predicting future outbreaks under similar conditions. The optimal beta reflects a balance between model sensitivity to new infections and realistic long-term predictions.

Further analysis could explore the interaction between beta and other model parameters, such as the incubation rate and recovery rate, to holistically improve the model's predictive capability.

Validation Against Empirical Data

Post-optimization, the SEIR model was validated against smoothed actual infection data. This step was crucial to confirm the model's efficacy in capturing the dynamics of the Monkeypox outbreak. The figure below illustrates the comparison between the modeled infections and the actual data, highlighting the model's ability to replicate the observed epidemic curve accurately.

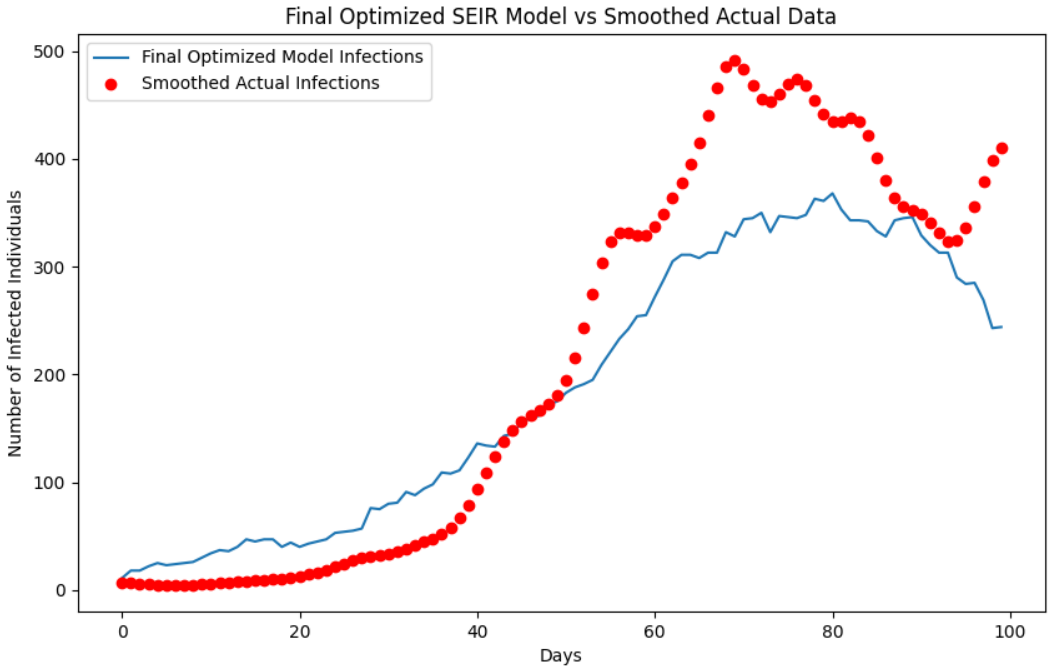

Figure 5: Comparison between the improved optimized SEIR model and smoothed actual infection data, demonstrating how the transmission rate affects the data fitting

These results indicate that the SEIR model, with an optimized β parameter, exhibits a closer alignment with the initial transmission phase and shows reasonable agreement during the plateau phase.

Step 3: Refined Approach to SEIR Modeling

The refined approach to modeling the SEIR model involved synthesizing data smoothing techniques and dynamic parameter optimization to better fit the observed epidemic data. The model was designed to capture the realistic dynamics of the epidemic by incorporating variations in the transmission rate over time, reflecting real-world phenomena such as public health interventions and behavioral changes.

Data Smoothing

The Gaussian smoothing technique allowed us to focus on the average behavior of transmission, rather than being influenced by daily fluctuations. This pre-processing step, implemented using a Gaussian filter, was essential in reducing noise and stabilizing the input data, enabling the model to capture the underlying trends without being skewed by daily reporting anomalies. The following code demonstrates the implementation:

```
smoothed_data = gaussian_filter1d(original_data, sigma=2)
```

Refined SEIR Model

The SEIR model equations were refined to include a time-dependent transmission rate ($\beta(t)$), which changes in response to interventions and other factors affecting the spread of the disease. This dynamic $\beta(t)$ was modeled as an exponential decay function starting from an initial value, adjusting over time to simulate the effects of Interventions or public intervention measures. The exponential

decay reflect the decrease in the transmission rate after the initialisation of the Intervention.

Python Implementation:

```
def dynamic_beta(t, intervention_start, intervention_strength): initial_beta = 0.3 if t <intervention_start: return initial_beta else: return initial_beta * exp(-intervention_strength * (t - intervention_start))
```

Parameter Optimization

Parameter optimization was performed using the *minimize* function from SciPy's optimization module. The objective function was designed to minimize the root mean square error (RMSE) between the model's predicted infections and the smoothed actual data. This method ensures that the model parameters are tuned to best fit the actual epidemic curve.

Python Implementation:

```
result = minimize(objective_function, initial_params, bounds=bounds)
```

Comparison of SEIR model with different population sizes

This analysis compares two scenarios in the refined SEIR model: one with a population size of N = 213 million and the other with N = 330 million. The objective is to evaluate whether population size significantly impacts the model parameters and its fit to the observed Monkey pox epidemic data in the United State

Results

For the results, I tried the optimization with different values of the US population.

Comparison for the model fit to US data

The optimized SEIR model for both cases provided a much closer fit to the actual data, effectively capturing the initial phase, peak, and overall average trend of the epidemic. The results demonstrate the effectiveness of the refined hybrid approach in optimizing the SEIR model for real-world applications [1, 2]. The figure below illustrates the comparison between the optimized model's infections and the smoothed actual data, highlighting the model's accuracy in simulating the epidemic dynamics. However, there are a few key differences

For N=213000000:

The model slightly underestimated and anticipated the peak earlier than observed in the data. However, it demonstrated strong alignment with the data for t>200 days, accurately capturing the post-peak dynamics.

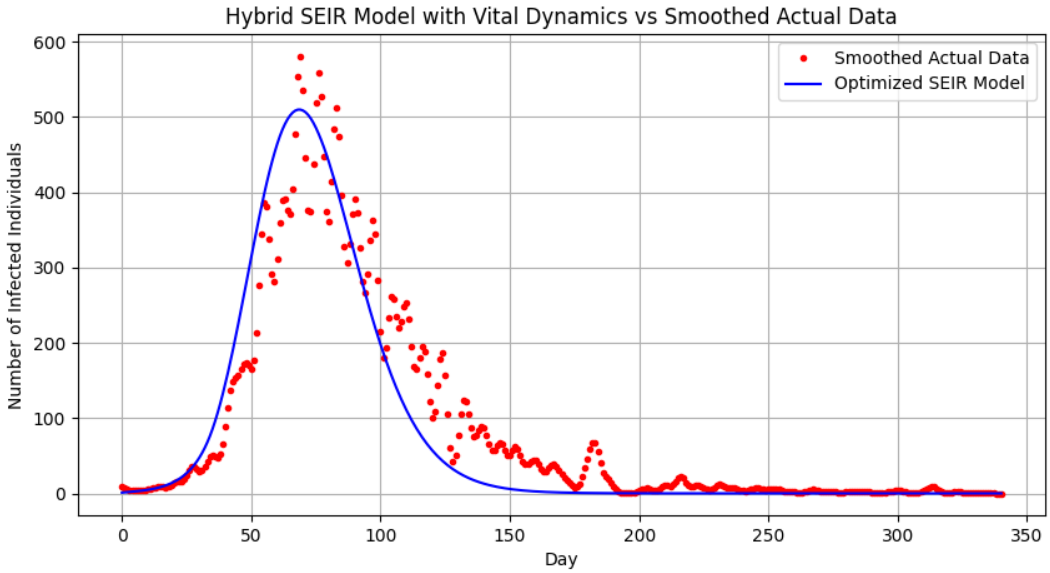

Figure 6: Comparison of the Optimized SEIR Model and Smoothed Actual Data

N=330000000

The model for the larger population better captured the peak and timing of the epidemic, aligning more closely with the US data. It also captured post-peak dynamics well.

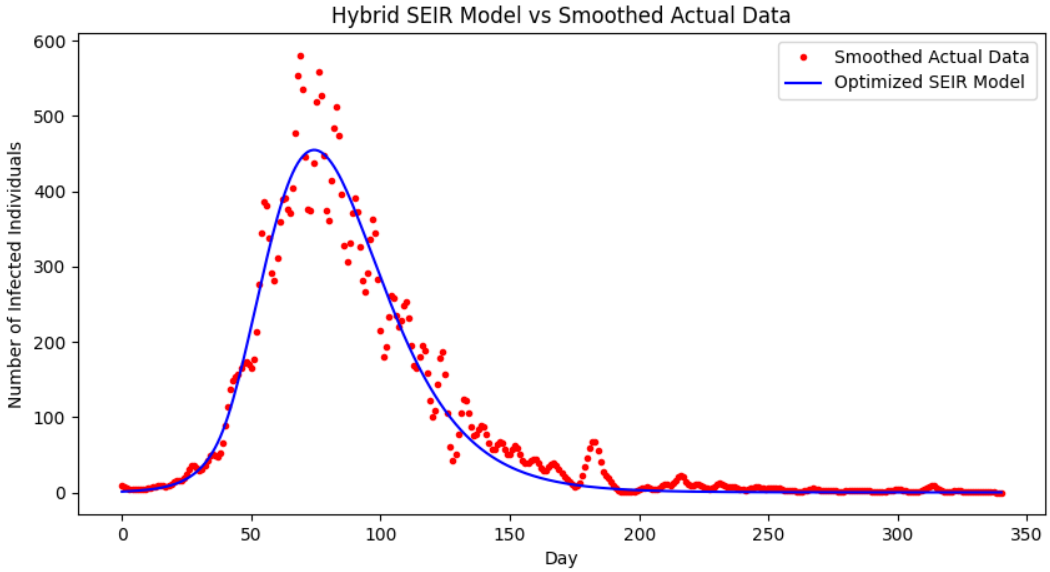

Figure 7: Comparison of the Optimized SEIR Model and Smoothed Actual Data

Comparison of the Optimized Parameters

The optimized parameters for both $N = 213000000$ and 313000000 for the SEIR model with dynamic transmission rates are as follows:

Parameter	N = 213 million	N = 330 million
Incubation Rate (σ)	0.120	0.239
Recovery Rate (γ)	0.05	0.0866

Intervention Start Day	38.8	34.6
Intervention Strength	0.0608	0.0403

Observations:

1. Incubation Rate (σ): For N = 330 million, the incubation rate (σ = 0.2389) was nearly double that for N = 213 million (σ = 0.1202). This could be a reflection of how the model optimizes the parameters to the data, rather than reflecting the true biological aspect of incubation.

2. Recovery Rate (γ): The recovery rate (γ) for N = 330 million was higher (γ = 0.0866) compared to N = 213 million (γ = 0.05).

2. Intervention Start Day: Interventions were modeled to begin earlier for N = 330 million (day ~34.6) than for N = 213 million (day ~38.8). This difference may reflect the need to expedite interventions in larger populations to curb faster epidemic progression and prevent healthcare systems from becoming overwhelmed.

4. Intervention Strength: The intervention strength was higher for N = 213 million (0.0608) compared to N = 330 million (0.0403), indicating stricter measures in the smaller population. Smaller populations may find it easier to implement stricter measures due to factors such as a lower absolute number of individuals to manage, more centralized resources, or enhanced capacity for enforcement and monitoring.

Equations of the Fit for N=330 million

The differential equations describing the SEIR model, with the dynamic transmission rate influenced by the Intervention parameters, are:

$$\frac{dS}{dt} = -\beta(t)\frac{SI}{N}$$
$$\frac{dE}{dt} = \beta(t)\frac{SI}{N} - \sigma E$$
$$\frac{dI}{dt} = \sigma E - \gamma I$$
$$\frac{dR}{dt} = \gamma I$$

where $\beta(t)$ is defined by the Intervention model:

$$\beta(t) = \begin{cases} \beta_{\text{initial}}, & t < t_{\text{intervention}} \\ \beta_{\text{initial}} \cdot e^{-\delta(t - t_{\text{intervention}})}, & t \geq t_{\text{intervention}} \end{cases}$$

with:

- $\beta_{initial} = 0.3$ (Initial transmission rate)
- $T = 34.6$ (Day intervention starts)
- $\delta = 0.0403$ (Intervention strength)

Conclusion

This study proves the effectiveness of the refined SEIR model with vital dynamics in capturing the progression of the Mpox outbreak in the United States. By incorporating a time-varying transmission rate,incubation period, and recovery rate, the model provides an accurate representation of epidemic dynamics, including onset, peak, and decline.

The comparison of population scenarios N=213 million and N=330 milliondemonstrates the sensitivity of model parameters to population size. The larger population scenario aligns more closely with observed data, emphasizing the importance of early and sustained interventions in larger populations.

The model highlights the critical role of early public health measures in reducing transmission rates and provides a robust tool for predicting epidemic trends and

informing intervention strategies.

References

[1] Transmission dynamics of Monkeypox virus: a mathematical modelling approach. *Modeling Earth Systems and Environment* (Springer), 2022.

[2] Epidemiologicalmodeling of monkeypox clades: a dual-strain SEIR approach with stability, bifurcation, and sensitivity analysis. *Modeling Earth Systems and Environment* (Springer), 2023.

[3] Mathematical assessment of monkeypox disease with the impact of vaccination using a fractional epidemiological modeling approach. *Nature*, 2023.

[4] Bjørnstad, O. N., Shea, K., Krzywinski, M., & Altman, N. The SEIRS model for infectious disease dynamics. *Nature Methods*, 17(6), 557–558, 2020. https://doi.org/10.1038/s41592-020-0856-2