Research Article

Mouhamadou Djima Baranon*, Patrick Guge Oloo Weke, Judicaël Alladatin, Boni Maxime Ale

Markov modeling on dynamic state space for genetic disorders and infectious diseases with mutations: Probabilistic framework, parameter estimation, and applications

https://doi.org/10.1515/cmb-2024-0005
received February 28, 2024; accepted April 18, 2024

Abstract: The emergence and dynamic prevalence of genetic disorders and infectious diseases with mutations pose significant challenges for public health interventions. This study investigated the parameter estimation approach and the application of the dynamic state-space Markov modeling of these conditions. Using extensive simulations, the model demonstrated robust parameter estimation performance, with biases and mean-squared errors decreasing as sample size increased. Applying the model to COVID-19 data revealed distinct temporal patterns for each variant, highlighting their unique emergence, peak dominance, and decline or persistence trajectories. Despite the absence of clear trends in the data, the model exhibited a remarkable accuracy in predicting future prevalence trends for most variants, showcasing its potential for real-time monitoring and analysis. While some discrepancies were observed for specific variants, these findings suggest the model's promise as a valuable tool for informing public health strategies. Further validation with larger datasets and exploration of incorporating additional factors hold the potential for enhancing the model's generalizability and applicability to other evolving diseases.

Keywords: Markov process, dynamic state space, genetic disorders, infectious disease, L-BFGS algorithm

MSC 2020: 62F30, 62K05, 62P10

1 Introduction

Markov processes are stochastic processes with the Markov property, where all the information needed to predict the future is fully contained in the current state without depending on previous states (i.e., the system does not have "memory") [32]. They are named after their creator, Andrey Markov (1856–1922), who presented

*** Corresponding author: Mouhamadou Djima Baranon,** Department of Mathematics and Statistics, Pan African University Institute for Basic Sciences, Technology, and Innovation (PAUSTI), Nairobi, Kenya; Ecole Nationale de Statistique, de Planification et de Démographie (ENSPD), Université de Parakou, Parakou, Bénin, e-mail: djima.mouhamadou@students.jkuat.ac.ke
Patrick Guge Oloo Weke: School of Mathematics, University of Nairobi, Nairobi, Kenya, e-mail: pweke@uonbi.ac.ke
Judicaël Alladatin: Faculté des Sciences de l'Education, Université de Montréal, Montréal, Canada; Consortium Siabanni pour la Formation, la Recherche et le Développement (Consortium SFR-D), Abomey-Calavi, Bénin, e-mail: judicael.alladatin.1@ulaval.ca
Boni Maxime Ale: Institute of Tropical and Infectious Diseases, University of Nairobi, Nairobi, Kenya; Strathmore University Business School, Strathmore University, Nairobi, Kenya; Research Department, Holo Global Health Research Institute, Nairobi, Kenya; Research Department, Health Data Acumen, Nairobi, Kenya; Research Department, MOI University, Nairobi, Kenya, e-mail: bonim@uonbi.ac.ke
ORCID: Mouhamadou Djima Baranon 0009-0005-7783-653X; Patrick Guge Oloo Weke 0000-0002-6283-4567; Judicaël Alladatin 0000-0001-7230-9953; Boni Maxime Ale 0000-0002-8449-3310

the first results about Markov chains with a finite state space in 1906. An extension to a countably infinite state space was formulated later [10,29]. These processes are associated with Brownian motion and the ergodic hypothesis [26,33], two crucial concepts in statistical physics that significantly contributed to that field in the early twentieth century [42].

In addition, Markov processes have evolved thanks to multiple scientific works, leading to several types: discrete-time Markov process, continuous-time Markov process, hidden Markov process, semi-Markovian Markov process [22], Markov process Markov decision [15], finite memory Markov process, etc. Then they find their applications in various fields such as biology, voice recognition, finance, insurance, engineering, population dynamics, health.

The main reason for their adoption in healthcare is the "memoryless property," since complete information often does not exist in patients suffering from a given medical condition [7,9]. Considering a probability space $(\Omega, F, \mathbb{P})$ with a filtration $(\mathcal{F}_t, t \in I)$ and $(S, \mathcal{S})$ a measurable space, a stochastic process $Z = \{Z_t : \Omega \to S\}_{t\in I}$ adapted to $\mathcal{F}_t$ is said to have the Markov property if

$$P_r(Z_t \in B|F_s) = P_r(Z_t \in B|Z_s), \tag{1}$$

for each $B \in \mathbb{S}$, $s, t \in I$, with $s < t$.

In other words, the Markov property essentially asserts that the conditional probability of the process being in a certain state at a time t, given all the information available up to the time s, is equal to the conditional probability of the process being in that state at time t, given only the current state at the time s.

If the set S is discrete, then

$$P_r(Z_n = z_n|Z_{n-1} = z_{n-1}, Z_{n-2} = z_{n-2}, \ldots, Z_0 = z_0) = P_r(Z_n = z_n|Z_{n-1} = z_{n-1}). \tag{2}$$

Furthermore, the world faces many diseases in which cells or viruses mutate over time, leading to new variants. Genetic disorders and infectious diseases are notorious for such behaviors. In such a context, a treatment plan for a disease could be ineffective when, in response, the cells change their background (mutation) [3,19,36]. Therefore, to control such diseases, it is important to conduct studies highlighting their progression. Several mathematical models have been developed for this purpose [13,25,48]. They are mostly based on differential equations. Often limited by the availability of complete information on the patients' past, Markov processes are adapted to face that challenge. Many works have then used the Markov approach in modeling diseases such as cancer [8,31,47], hepatitis [45], diabetes [6], malaria [34], HIV [27], and cardiovascular diseases [18,40].

However, most of those Markov models assume that state spaces and propagation rates are constant. Such hypotheses are not plausible when it comes to genetic disorders and infectious diseases with mutations. Indeed, mutations lead to the appearance of new types of cells whose inclusion changes the state space [17]. The need to develop models that, in addition to being able to only take into account information from the present to predict the future, can also adjust to changes in state spaces is imperative [3]. These limitations pave way to the need of Markov processes in genetics disorders, and infectious diseases modeling. Markov processes in one dynamic state spaces would meet this need. Nevertheless, modeling the progression of these diseases (genetic disorders and infectious diseases with mutations) by considering dynamic state spaces involves increasing the number of parameters over time.

Furthermore, parameter estimation, a critical aspect of modeling, plays an important role in guaranteeing the accuracy and predictive capabilities of models [24,37]. The precision with which model parameters are estimated directly influences the fidelity of predictions, making it an essential focal point in research endeavors [20,30] aimed at comprehensively understanding and mitigating the impact of genetic disorders and infectious diseases. In the context of genetic disorders and infectious diseases, the incorporation of mutations introduces an additional layer of complexity. Maximum-likelihood estimation (MLE) methods are, moreover, renowned for non-linear optimizations, in particular with their properties of consistency, asymptotic efficiency, asymptotically normal distribution, and maximum informativeness [2,41]. They use likelihood functions, logarithms, as well as derivatives. In particular, for Markov models in dynamic spaces, the complexity of the likelihood functions makes the search for an analytical solution very complicated. Hence, there is need to search for a numerical solution. This study aims to develop a dynamic state-space Markov model to enhance

the understanding and prediction of the progression of genetic and infectious diseases, with a particular focus on mutations associated with COVID-19.

The outcomes of this study hold the potential to not only enhance fundamental understanding of genetic disorders and infectious diseases but also inform the development of targeted interventions and therapeutic strategies. In investigating dynamic state-space modeling in the context of genetic disorders and infectious diseases, this research provides a robust foundation for advancing precision in medicine and public health initiatives.

This article is structured into five main sections. The first one is about the definition of the key concepts used. The second section is related to the materials and methods. The third one describes the parameters estimation approach. In the fourth section, the results of the simulation study are presented. The last section is about the application with real data severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2).

2 Definitions

Definition 1. A **Markov process** is a type of stochastic process in which the future probabilities are only determined by the process's current state, independent of any previous states. Formally, it is a sequence of random variables $Z_1, Z_2, \ldots$ possessing the Markov property, i.e.,

$$P_r(Z_n = z | Z_{n-1} = z_{n-1}, Z_{n-2} = z_{n-2}, \ldots, Z_0 = z_0) = P(Z_n = z | Z_{n-1} = z_{n-1}), \tag{3}$$

with $P_r(B|A)$ being the conditional probability of B given A. The set of all the possible values that Z can take is called the **state space**. The transition probabilities are the probabilities of moving from a given state to another one. When they are constant over time, the Markov chain is said to be time-homogeneous.

Definition 2. Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm for MLE

Suppose that we want to estimate θ, a vector of the parameters, using the MLE method, which consists of maximizing the log-likelihood function, $\log L$:

$$\theta_{\text{MLE}} = \arg\max_{\theta} \{\log L(\theta)\}. \tag{4}$$

The BFGS algorithm applied to that optimization problem is as follows:

Initial step: Let $\varepsilon > 0$ be the stopping criteria. State initial values θ_1 and a symmetric positive definite matrix Q_1.

Step 1: If $\|\nabla \log L(\theta_k)\| < \varepsilon$, stop. If not, state $d_k = Q_k \nabla \log L(\theta_k)$, choose $\alpha_k > 0$ and compute $\theta_{k+1} = \theta_k - \alpha_k d_k$, with ∇, the gradient.

Step 2: Compute Q_{k+1}

$$Q_{k+1} = Q_k + \frac{a_k a_k^t}{a_k b_k^t} + \frac{b_k^t Q_k b_k a_k a_k^t}{(a_k b_k^t)^2} - \frac{a_k b_k^t Q_k + Q_k b_k a_k^t}{a_k^t b_k}, \tag{5}$$

with

$$a_k = \alpha_k d_k \equiv \theta_{k+1} - \theta_k, \tag{6}$$

$$b_k = \nabla \log L(\theta_{k+1}) - \nabla \log L(\theta_k). \tag{7}$$

We then replace k with $k + 1$ and return to **Step 1**.

Definition 3. Genetic disorders are medical conditions caused by changes in the DNA sequence of genes or chromosomes [11]. These changes can be inherited from one or both parents, or they can happen on their own. Genetic disorders are classified into three types: single-gene disorders, chromosomal disorders, and complex disorders. Mutations in a single gene cause single-gene disorders, whereas chromosomal disorders are caused by missing or altered chromosomes. Mutations in two or more genes, as well as environmental and lifestyle factors, cause complex disorders. Cystic fibrosis, sickle cell anemia, Down syndrome, and muscular dystrophy are examples of common genetic disorders [46].

Definition 4. Infectious diseases are illnesses or conditions caused by infectious agents (bacteria, viruses, fungi, or parasites) that enter the body and can cause an infection.

A **bacterium cell** is a type of prokaryotic cell with a simple internal structure and no nucleus. Bacteria can be found almost anywhere on Earth in various shapes (cocci: round, bacilli: capsule-shaped, spirilla: spiral-shaped) and sizes (0.2 to 2 µm). They are critical to the planet's ecosystems and human survival. Some of them can survive extreme temperatures and pressures. They are also used in biotechnology, food processing, and the production of antibiotics and other chemicals [1,4,23,44]. The harmful ones are called pathogenic bacteria. Furthermore, **viruses** are made up of nucleic acid fragments wrapped in a protein. They enter healthy cells and replicate using the cell's replication systems. The viruses cause cellular membranes to rearrange, resulting in the formation of specific intracellular compartments known as replication organelles (ROs), which are required for viral replication [28,39]. Viruses are much smaller than bacteria. Bacteria are on the microscopic scale, while viruses are on the nanoscopic scale (25–250 nm).

Fungi are found on decaying plant and animal matter and have two types of cells: yeast cells and mold cells. Yeast fungi (size between 3–5 µm) are found all over the world on plants, in soil, and in sugary mediums such as fruit. Mold fungi are composed of multicellular filaments known as hyphae [14,21]. Moreover, **parasites** are living organisms that gain an advantage by attaching to a host at the expense of the host's health. They can range in size from tiny parasites such as malaria, which is about 4 µm long, to much larger ones such as tapeworms, which can grow to be several meters long. Parasites can enter the body of a host and exploit its resources, often evading the host's immune system. Depending on the parasite's ability to adapt and suppress the host's immune response, this ability to evade the immune system can result in chronic or severe infections [38,43].

3 Materials and methods

3.1 Genetic disorders and infectious diseases with mutation progression model

We consider a cellular population composed of mutant cells and non-mutant cells. The following parameters and variables are used (Table 1).

For d types of mutant cells ($d \geq 1$), the non-mutant cell number increasing and decreasing rates are, respectively, $\lambda(1 - \sum_{s=1}^{d}\gamma_s)(l - \sum_{s=1}^{d}m_s)$ and $\mu(l - \sum_{s=1}^{d}m_s)$. For the mutant cells, they are, respectively. $\lambda(\sum_{s=1}^{k}\gamma_s)(l - \sum_{s=1}^{d}m_s) + \sum_{s=1}^{d}\alpha_s m_s$ and $\sum_{s=1}^{d}\beta_s m_s$.

Furthermore, depending on the number of mutant cell types (k), the state space can be (l, m_1), (l, m_1, m_2), $\cdots$ $(l, m_1, m_2 \ \cdots \ m_d)$.

Let us denote by $(l_t, m_{1t}, m_{2t} \ \cdots \ m_{dt})$ the process state after the tth event (death or division); $Y_t = (l_t, m_{1t})$ (for $k = 1$); $Y_t = (l_t, m_{1t}, m_{2t})$ (for $k = 2$); $Y_t = (l_t, m_{1t}, m_{2t}; \cdots ; m_{dt})$ (for $k = d$).

- For $k = 1$

Table 1: Description of the parameters and variables

Parameter	Description
λ	Non-mutant cell division rate
μ	Non-mutant cell death race
γ_s	Non-mutant cell probability of changing background (mutation) to cell type s after division
α_s	Mutant cell type s division rate
β_s	Mutant cell type s death rate
l	Total number of disease cells
m_s	Number of mutant cells type s

The transition probabilities are mathematically expressed as follows:

$$P_r[Y_{t+1} = (l+1, m_1) \mid Y_t = (l, m_1)] = \frac{\lambda(1-\gamma_1)(l-m_1)}{\Phi_{l,m_1}}, \tag{8}$$

$$P_r[Y_{t+1} = (l+1, m_1+1) \mid Y_t = (l, m_1)] = \frac{\lambda\gamma_1(l-m_1)+\alpha_1 m_1}{\Phi_{l,m_1}}, \tag{9}$$

$$P_r[Y_{t+1} = (l-1, m_1) \mid Y_t = (l, m_1)] = \frac{\mu(l-m_1)}{\Phi_{l,m_1}}, \tag{10}$$

$$P_r[Y_{t+1} = (l-1, m_1-1) \mid Y_t = (l, m_1)] = \frac{\beta_1 m_1}{\Phi_{l,m_1}}, \tag{11}$$

with $\Phi_{l,m_1} = (l-m_1)(\lambda+\mu) + m_1(\alpha_1+\beta_1)$, the total sum of the rates, ensuring that the overall probability is normalized to 1.

- For $k = 2$

The transition probabilities are expressed as follows:

$$P_r[Y_{t+1} = (l+1, m_1, m_2) \mid Y_t = (l, m_1, m_2)] = \frac{\lambda(1-\gamma_1-\gamma_2)(l-m_1-m_2)}{\Phi_{l,m_1,m_2}}, \tag{12}$$

$$P_r[Y_{t+1} = (l+1, m_1+1, m_2) \mid Y_t = (l, m_1, m_2)] = \frac{\lambda\gamma_1(l-m_1-m_2)+\alpha_1 m_1}{\Phi_{l,m_1,m_2}}, \tag{13}$$

$$P_r[Y_{t+1} = (l+1, m_1, m_2+1) \mid Y_t = (l, m_1, m_2)] = \frac{\lambda\gamma_2(l-m_1-m_2)+\alpha_2 m_2}{\Phi_{l,m_1,m_2}}, \tag{14}$$

$$P_r[Y_{t+1} = (l-1, m_1, m_2) \mid Y_t = (l, m_1, m_2)] = \frac{\mu(l-m_1-m_2)}{\Phi_{l,m_1,m_2}}, \tag{15}$$

$$P_r[Y_{t+1} = (l-1, m_1-1, m_2) \mid Y_t = (l, m_1, m_2)] = \frac{\beta_1 m_1}{\Phi_{l,m_1,m_2}}, \tag{16}$$

$$P_r[Y_{t+1} = (l-1, m_1, m_2-1) \mid Y_t = (l, m_1, m_2)] = \frac{\beta_2 m_2}{\Phi_{l,m_1,m_2}}, \tag{17}$$

with $\Phi_{l,m_1,m_2} = (\lambda+\mu)(l-m_1-m_2) + (\alpha_1+\beta_1)m_1 + (\alpha_2+\beta_2)m_2$.

- For $k \geq 2$

The transition probabilities are expressed as follows:

$$P_r[Y_{t+1} = (l+1, m_1, \ldots, m_d) \mid Y_t = (l, m_1, \ldots, m_d)] = \frac{\lambda(1-\sum_{s=1}^{d}\gamma_s)(l-\sum_{s=1}^{d}m_s)}{\Phi_{l,m_1,m_2,\ldots,m_d}}, \tag{18}$$

$$P_r[Y_{t+1} = (l+1, m_1+1, \ldots, m_d) \mid Y_t = (l, m_1, \ldots, m_d)] = \frac{\lambda\gamma_1(l-\sum_{s=1}^{d}m_s)+m_1\alpha_1}{\Phi_{l,m_1,\ldots,m_d}}, \tag{19}$$

$$P_r[Y_{t+1} = (l+1, m_1, m_2+1, \ldots, m_d) \mid Y_t = (l, m_1, \ldots, m_d)] = \frac{\lambda\gamma_2(l-\sum_{s=1}^{d}m_s)+m_2\alpha_2}{\Phi_{l,m_1,\ldots,m_d}}, \tag{20}$$

$$\ldots\ldots\ldots\ldots\ldots\ldots$$

$$P_r[Y_{t+1} = (l+1, m_1, \ldots, m_d+1) \mid Y_t = (l, m_1, \ldots, m_d)] = \frac{\lambda\gamma_d(l-\sum_{s=1}^{d}m_s)+m_d\alpha_d}{\Phi_{l,m_1,\ldots,m_d}}, \tag{21}$$

$$P_r[Y_{t+1} = (l-1, m_1, \ldots, m_d) \mid Y_t = (l, m_1, \ldots, m_d)] = \frac{\mu(l - \sum_{s=1}^{d} m_s)}{\Phi_{l,m_1,\ldots,m_d}}, \tag{22}$$

$$P_r[Y_{t+1} = (l-1, m_1-1, \ldots, m_d) \mid Y_t = (l, m_1, \ldots, m_d)] = \frac{m_1\beta_1}{\Phi_{l,m_1,\ldots,m_d}}, \tag{23}$$

$$P_r[Y_{t+1} = (l-1, m_1, m_2-1, \ldots, m_d) \mid Y_t = (l, m_1, \ldots, m_d)] = \frac{m_2\beta_2}{\Phi_{l,m_1,\ldots,m_d}}, \tag{24}$$

$$\ldots\ldots\ldots\ldots\ldots\ldots$$

$$P_r[Y_{t+1} = (l-1, m_1, \ldots, m_d-1) \mid Y_t = (l, m_1, \ldots, m_d)] = \frac{m_d\beta_d}{\Phi_{l,m_1,\ldots,m_d}}, \tag{25}$$

with $\Phi_{l,m_1,\ldots,m_d} = (\lambda + \mu)(l - \sum_{s=1}^{d} m_s) + \sum_{s=1}^{d}(\alpha_s + \beta_s)m_s$.

3.2 Probability mass function

Let us define w, w_1, w_2,…, w_d, and x as follows:

$$w = l_{(t+1)} - l_t, \tag{26}$$

$$w_1 = m_{1(t+1)} - m_{1(t)}, \tag{27}$$

$$w_2 = m_{2(t+1)} - m_{2(t)}, \tag{28}$$

$$\ldots\ldots\ldots$$

$$w_d = m_{d(t+1)} - m_{d(t)}, \tag{29}$$

$$x = w + w_1 + w_2 \cdots + w_d. \tag{30}$$

Four different values are possible for x:

1: When the total number of disease cells increases due to the non-mutant cells; 2: when the total number of disease cells increases due to one of the mutant cell types; −1: when the total number of disease cells decreases due to the non-mutant cells; −2: when the total number of disease cells decreases due to one of the mutant cell types.

Using the transition probabilities and considering those four possibilities, the probability mass function can be derived as follows:

$$\begin{aligned} P_r(X = x) = &\left(\frac{\lambda(1 - \sum_{s=1}^{d}\gamma_s)(l - \sum_{s=1}^{d} m_s)}{\Phi_{l,m_1,\ldots,m_d}}\right)^{\mathbb{1}_1(x)} \left(\frac{\lambda(\sum_{s=1}^{d}\gamma_s)(l - \sum_{s=1}^{d} m_s) + \sum_{s=1}^{d}\alpha_s m_s}{\Phi_{l,m_1,\ldots,m_d}}\right)^{\mathbb{1}_2(x)} \\ &\times \left(\frac{\mu(l - \sum_{s=1}^{d} m_s)}{\Phi_{l,m_1,\ldots,m_d}}\right)^{\mathbb{1}_{-1}(x)} \left(\frac{\sum_{s=1}^{d}\beta_s m_s}{\Phi_{l,m_1,\ldots,m_d}}\right)^{\mathbb{1}_{-2}(x)}, \end{aligned} \tag{31}$$

with $x \in \{1, 2, -1, -2\}$ and $\mathbb{1}_i(x)$ is an indicator function defined as: $\mathbb{1}_i(x) = 1$ if $i = x$, and $\mathbb{1}_i(x) = 0$ if not.

4 Parameter estimation

The MLE method can be used to estimate the parameters. To do so, the probability mass function (equation (31)) is needed.

- For $k = 1$

The probability mass function (pmf) is given by the following equation:

$$P_r(X = x) = \left(\frac{\lambda(1-\gamma_1)(l-m_1)}{\Phi_{l,m_1}}\right)^{\mathbb{1}_1(x)}\left(\frac{\lambda\gamma_1(l-m_1)+\alpha_1 m_1}{\Phi_{l,m_1}}\right)^{\mathbb{1}_2(x)}\left(\frac{\mu(l-m_1)}{\Phi_{l,m_1}}\right)^{\mathbb{1}_{-1}(x)}\left(\frac{\beta_1 m_1}{\Phi_{l,m_1}}\right)^{\mathbb{1}_{-2}(x)} \tag{32}$$

with $x \in \{1, 2, -1, -2\}$ and $\Phi_{l,m_1} = (l-m_1)(\lambda+\mu) + m_1(\alpha_1+\beta_1)$.

Assuming a dataset of N observations, the likelihood function is defined as follows:

$$\begin{aligned} L(\lambda, \gamma_1, \alpha_1, \mu, \beta_1|X) = \prod_{n=1}^{N}\Bigg[&\left(\frac{\lambda(1-\gamma_1)(l_n-m_{1n})}{\Phi_{l_n,m_{1n}}}\right)^{\mathbb{1}_1(x)}\left(\frac{\lambda\gamma_1(l_n-m_{1n})+\alpha_1 m_{1n}}{\Phi_{l_n,m_{1n}}}\right)^{\mathbb{1}_2(x)} \\ &\times\left(\frac{\mu(l_n-m_{1n})}{\Phi_{l_n,m_{1n}}}\right)^{\mathbb{1}_{-1}(x)}\left(\frac{\beta_1 m_{1n}}{\Phi_{l_n,m_{1n}}}\right)^{\mathbb{1}_{-2}(x)}\Bigg]. \end{aligned} \tag{33}$$

Let us state $\log L(\lambda, \gamma_1, \alpha_1, \mu, \beta_1|X) = \log L$. The log-likelihood function is

$$\begin{aligned} \log L = \sum_{n=1}^{N}\Bigg[&\mathbb{1}_1(x_n)\log\left(\frac{\lambda(1-\gamma_1)(l_n-m_{1n})}{\Phi_{l_n,m_{1n}}}\right) + \mathbb{1}_2(x_n)\log\left(\frac{\lambda\gamma_1(l_n-m_{1n})+\alpha_1 m_{1n}}{\Phi_{l_n,m_{1n}}}\right) \\ &+ \mathbb{1}_{-1}(x_n)\log\left(\frac{\mu(l_n-m_{1n})}{\Phi_{l_n,m_{1n}}}\right) + \mathbb{1}_{-2}(x_n)\log\left(\frac{\beta_1 m_{1n}}{\Phi_{l_n,m_{1n}}}\right)\Bigg]. \end{aligned} \tag{34}$$

To find the values of the parameters that maximize the log-likelihood function (equation (33)), there is a need for the derivative of $\log L$ for each parameter. This gives five equations (one for each parameter $\lambda, \gamma_1, \alpha_1, \mu, \beta_1$).

$$\begin{aligned} \frac{\partial \log L}{\partial \lambda} = \sum_{n=1}^{N}\Bigg[&\frac{\mathbb{1}_1(x_n)\Phi_{l_n,m_{1n}}}{\lambda(1-\gamma_1)(l_n-m_{1n})}\left(\frac{(1-\gamma_1)(l_n-m_{1n})}{\Phi_{l_n,m_{1n}}} - \frac{\lambda(1-\gamma_1)(l_n-m_{1n})^2}{\Phi^2_{l_n,m_{1n}}}\right) \\ &+ \mathbb{1}_2(x_n)\left(\gamma_1(l_n-m_{1n}) - \frac{\alpha_1 m_{1n}\,(l_n-m_{1n})}{\Phi^2_{l_n,m_{1n}}}\right)\left(\lambda\gamma_1(l_n-m_{1n}) + \frac{\alpha_1 m_{1n}}{\Phi_{l_n,m_{1n}}}\right)^{-1} \\ &- \frac{\mathbb{1}_{-1}(x_n)\,(l_n-m_{1n})}{\Phi_{l_n,m_{1n}}} - \frac{\mathbb{1}_{-2}(x_n)\,(l_n-m_{1n})}{\Phi_{l_n,m_{1n}}}\Bigg], \end{aligned} \tag{35}$$

$$\begin{aligned} \frac{\partial \log L}{\partial \mu} = \sum_{n=1}^{N}\Bigg[&-\frac{\mathbb{1}_1(x_n)\,(l_n-m_{1n})}{\Phi_{l_n,m_{1n}}} - \frac{\mathbb{1}_2(x_n)\alpha_1 m_{1n}(l_n-m_{1n})}{\Phi^2_{l_n,m_{1n}}}\left(\lambda\gamma_1(l_n-m_{1n}) + \frac{\alpha_1 m_{1n}}{\Phi_{l_n,m_{1n}}}\right)^{-1} \\ &+ \frac{\mathbb{1}_{-1}(x_n)\Phi_{l_n,m_{1n}}}{\mu(l_n-m_{1n})}\left(\frac{l_n-m_{1n}}{\Phi_{l_n,m_{1n}}} - \frac{\mu(l_n-m_{1n})^2}{\Phi^2_{l_n,m_{1n}}}\right) - \frac{\mathbb{1}_{-2}(x_n)(l_n-m_{1n})}{\Phi_{l_n,m_{1n}}}\Bigg], \end{aligned} \tag{36}$$

$$\frac{\partial \log L}{\partial \gamma_1} = \sum_{n=1}^{N}\left[-\frac{\mathbb{1}_1(x_n)}{1-\gamma_1} + \mathbb{1}_2(x_n)\lambda(l_n-m_{1n})\left(\lambda\gamma_1(l_n-m_{1n}) + \frac{\alpha_1 m_{1n}}{\Phi_{l_n,m_{1n}}}\right)^{-1}\right], \tag{37}$$

$$\begin{aligned} \frac{\partial \log L}{\partial \alpha_1} = \sum_{n=1}^{N}\Bigg[&-\frac{\mathbb{1}_1(x_n)m_{1n}}{\Phi_{l_n,m_{1n}}} + \mathbb{1}_2(x_n)\left(\frac{m_{1n}}{\Phi_{l_n,m_{1n}}} - \frac{\alpha_1 {m_{1n}}^2}{\Phi^2_{l_n,m_{1n}}}\right)\left(\lambda\gamma_1(l_n-m_{1n}) + \frac{\alpha_1 m_{1n}}{\Phi_{l_n,m_{1n}}}\right)^{-1} \\ &- \frac{\mathbb{1}_{-1}(x_n)m_{1n}}{\Phi_{l_n,m_{1n}}} - \frac{\mathbb{1}_{-2}(x_n)m_{1n}}{\Phi_{l_n,m_{1n}}}\Bigg], \end{aligned} \tag{38}$$

$$\begin{aligned} \frac{\partial \log L}{\partial \beta_1} = \sum_{n=1}^{N}\Bigg[&-\frac{\mathbb{1}_1(x_n)m_{1n}}{\Phi_{l_n,m_{1n}}} - \frac{\mathbb{1}_2(x_n)\alpha_1 {m_{1n}}^2}{\Phi^2_{l_n,m_{1n}}}\left(\lambda\gamma_1(l_n-m_{1n}) + \frac{\alpha_1 m_{1n}}{\Phi_{l_n,m_{1n}}}\right)^{-1} \\ &- \frac{\mathbb{1}_{-1}(x_n)m_{1n}}{\Phi_{l_n,m_{1n}}} + \frac{\mathbb{1}_{-2}(x_n)\Phi_{l_n,m_{1n}}}{\beta_1 m_{1n}}\left(\frac{m_{1n}}{\Phi_{l_n,m_{1n}}} - \frac{\beta_1 {m_{1n}}^2}{\Phi^2_{l_n,m_{1n}}}\right)\Bigg]. \end{aligned} \tag{39}$$

- For $k = 2$

The probability mass function (pmf) is

$$P_r(X = x) = \left(\frac{\lambda(1-\gamma_1-\gamma_2)(l-m_1-m_2)}{\Phi_{l,m_1,m_2}}\right)^{\mathbb{1}_1(x)}\left(\frac{\lambda(\gamma_1+\gamma_2)(l-m_1-m_2)+\alpha_1 m_1+\alpha_2 m_2}{\Phi_{l,m_1,m_2}}\right)^{\mathbb{1}_2(x)} \times \left(\frac{\mu(l-m_1-m_2)}{\Phi_{l,m_1,m_2}}\right)^{\mathbb{1}_{-1}(x)}\left(\frac{\beta_1 m_1+\beta_2 m_2}{\Phi_{l,m_1,m_2}}\right)^{\mathbb{1}_{-2}(x)}, \tag{40}$$

with $x \in \{1, 2, -1, -2\}$ and $\Phi_{l,m_1,m_2} = (\lambda+\mu)(l-m_1-m_2) + (\alpha_1+\beta_1)m_1 + (\alpha_2+\beta_2)m_2$.

Assuming a dataset of N observations, the likelihood function is

$$\begin{aligned} L(\lambda, \gamma_1, \gamma_2, \alpha_1, \alpha_2, \mu, \beta_1, \beta_2 | X) &= \prod_{n=1}^{N}\left[\left(\frac{\lambda(1-\gamma_1-\gamma_2)(l_n-m_{1n}-m_{2n})}{\Phi_{l_n,m_{1n},m_{2n}}}\right)^{\mathbb{1}_1(x_n)}\right. \\ &\times \left(\frac{\lambda(\gamma_1+\gamma_2)(l_n-m_{1n}-m_{2n})+\alpha_1 m_{1n}+\alpha_2 m_{2n}}{\Phi_{l_n,m_{1n},m_{2n}}}\right)^{\mathbb{1}_2(x_n)} \\ &\left.\times \left(\frac{\mu(l_n-m_{1n}-m_{2n})}{\Phi_{l_n,m_{1n},m_{2n}}}\right)^{\mathbb{1}_{-1}(x_n)}\left(\frac{\beta_1 m_{1n}+\beta_2 m_{2n}}{\Phi_{l_n,m_{1n},m_{2n}}}\right)^{\mathbb{1}_{-2}(x_n)}\right]. \end{aligned} \tag{41}$$

The log-likelihood function is

$$\begin{aligned} \log L = \sum_{n=1}^{N}\Bigg[&\mathbb{1}_1(x_n)\log\left(\frac{\lambda(1-\gamma_1-\gamma_2)(l_n-m_{1n}-m_{2n})}{\Phi_{l_n,m_{1n},m_{2n}}}\right) \\ &+ \mathbb{1}_2(x_n)\log\left(\frac{\lambda(\gamma_1+\gamma_2)(l_n-m_{1n}-m_{2n})+\alpha_1 m_{1n}+\alpha_2 m_{2n}}{\Phi_{l_n,m_{1n},m_{2n}}}\right) \\ &+ \mathbb{1}_{-1}(x_n)\log\left(\frac{\mu(l_n-m_{1n}-m_{2n})}{\Phi_{l_n,m_{1n},m_{2n}}}\right) + \mathbb{1}_{-2}(x_n)\log\left(\frac{\beta_1 m_{1n}+\beta_2 m_{2n}}{\Phi_{l_n,m_{1n},m_{2n}}}\right)\Bigg]. \end{aligned} \tag{42}$$

The first derivatives with respect to the parameters are given by

$$\begin{aligned} \frac{\partial \log L}{\partial \lambda} = \sum_{n=1}^{N}\Bigg[&\frac{\mathbb{1}_1(x_n)\Phi_{l_n,m_{1n},m_{2n}}}{\lambda(1-\gamma_1-\gamma_2)(l_n-m_{1n}-m_{2n})} \\ &\times \left(\frac{(1-\gamma_1-\gamma_2)(l_n-m_{1n}-m_{2n})}{\Phi_{l_n,m_{1n},m_{2n}}} - \frac{\lambda(1-\gamma_1-\gamma_2)(l_n-m_{1n}-m_{2n})^2}{\Phi^2_{l_n,m_{1n},m_{2n}}}\right) \\ &+ \left(\frac{\mathbb{1}_2(x_n)\Phi_{l_n,m_{1n},m_{2n}}}{\lambda(\gamma_1+\gamma_2)(l_n-m_{1n}-m_{2n})+\alpha_1 m_{1n}+\alpha_2 m_{2n}}\right) \\ &\times \left(\frac{(\gamma_1+\gamma_2)(l_n-m_{1n}-m_{2n})}{\Phi_{l_n,m_{1n},m_{2n}}} - \frac{\lambda(\gamma_1+\gamma_2)(l_n-m_{1n}-m_{2n})+\alpha_1 m_{1n}+\alpha_2 m_{2n}}{\Phi^2_{l_n,m_{1n},m_{2n}}}\right) \\ &- \frac{\mathbb{1}_{-1}(x_n)(l_n-m_{1n}-m_{2n})}{\Phi_{l_n,m_{1n},m_{2n}}} - \frac{\mathbb{1}_{-2}(x_n)(l_n-m_{1n}-m_{2n})}{\Phi^2_{l_n,m_{1n},m_{2n}}}\Bigg], \end{aligned} \tag{43}$$

$$\begin{aligned} \frac{\partial \log L}{\partial \mu} = \sum_{n=1}^{N}\Bigg[&-\frac{\mathbb{1}_1(x_n)(l_n-m_{1n}-m_{2n})}{\Phi_{l_n,m_{1n},m_{2n}}} - \frac{\mathbb{1}_2(x_n)(l_n-m_{1n}-m_{2n})}{\Phi_{l_n,m_{1n},m_{2n}}} \\ &+ \frac{\mathbb{1}_{-1}(x_n)\Phi_{l_n,m_{1n},m_{2n}}}{\mu(l_n-m_{1n}-m_{2n})}\left(\frac{l_n-m_{1n}-m_{2n}}{\Phi_{l_n,m_{1n},m_{2n}}} - \frac{\mu(l_n-m_{1n}-m_{2n})^2}{(\Phi_{l_n,m_{1n},m_{2n}})^2}\right) \\ &- \frac{\mathbb{1}_{-2}(x_n)(l_n-m_{1n}-m_{2n})}{\Phi_{l_n,m_{1n},m_{2n}}}\Bigg], \end{aligned} \tag{44}$$

$$\frac{\partial \log L}{\partial \gamma_1} = \sum_{n=1}^{N}\left[-\frac{\mathbb{1}_1(x_n)}{1-\gamma_1-\gamma_2} + \frac{\mathbb{1}_2(x_n)\lambda(l_n - m_{1n} - m_{2n})}{\lambda(\gamma_1+\gamma_2)(l_n - m_{1n} - m_{2n}) + \alpha_1 m_{1n} + \alpha_2 m_{2n}}\right], \tag{45}$$

$$\frac{\partial \log L}{\partial \gamma_2} = \sum_{n=1}^{N}\left[-\frac{\mathbb{1}_1(x_n)}{1-\gamma_1-\gamma_2} + \frac{\mathbb{1}_2(x_n)\lambda(l_n - m_{1n} - m_{2n})}{\lambda(\gamma_1+\gamma_2)(l_n - m_{1n} - m_{2n}) + \alpha_1 m_{1n} + \alpha_2 m_{2n}}\right], \tag{46}$$

$$\begin{aligned}\frac{\partial \log L}{\partial \alpha_1} = \sum_{n=1}^{N}\Bigg[&-\frac{\mathbb{1}_1(x_n)m_{1n}}{\Phi_{l_n,m_{1n},m_{2n}}} + \left(\frac{\mathbb{1}_2(x_n)\Phi_{l_n,m_{1n},m_{2n}}}{\lambda(\gamma_1+\gamma_2)(l_n - m_{1n} - m_{2n}) + \alpha_1 m_{1n} + \alpha_2 m_{2n}}\right)\\ &\times\left(\frac{m_{1n}}{\Phi_{l_n,m_{1n},m_{2n}}} - \frac{(\lambda(\gamma_1+\gamma_2)(l_n - m_{1n} - m_{2n}) + \alpha_1 m_{1n} + \alpha_2 m_{2n})m_{1n}}{\Phi^2_{l_n,m_{1n},m_{2n}}}\right)\\ &- \frac{\mathbb{1}_{-1}(x_n)m_{1n}}{\Phi_{l_n,m_{1n},m_{2n}}} - \frac{\mathbb{1}_{-2}(x_n)m_{1n}}{\Phi_{l_n,m_{1n},m_{2n}}}\Bigg],\end{aligned} \tag{47}$$

$$\begin{aligned}\frac{\partial \log L}{\partial \alpha_2} = \sum_{n=1}^{N}\Bigg[&-\frac{\mathbb{1}_1(x_n)m_{2n}}{\Phi_{l_n,m_{1n},m_{2n}}} + \left(\frac{\mathbb{1}_2(x_n)\Phi_{l_n,m_{1n},m_{2n}}}{\lambda(\gamma_1+\gamma_2)(l_n - m_{1n} - m_{2n}) + \alpha_1 m_{1n} + \alpha_2 m_{2n}}\right)\\ &\times\left(\frac{m_{2n}}{\Phi_{l_n,m_{1n},m_{2n}}} - \frac{(\lambda(\gamma_1+\gamma_2)(l_n - m_{1n} - m_{2n}) + \alpha_1 m_{1n} + \alpha_2 m_{2n})m_{2n}}{\Phi^2_{l_n,m_{1n},m_{2n}}}\right)\\ &- \frac{\mathbb{1}_{-1}(x_n)m_{2n}}{\Phi_{l_n,m_{1n},m_{2n}}} - \frac{\mathbb{1}_{-2}(x_n)m_{2n}}{\Phi_{l_n,m_{1n},m_{2n}}}\Bigg],\end{aligned} \tag{48}$$

$$\begin{aligned}\frac{\partial \log L}{\partial \beta_1} = \sum_{n=1}^{N}\Bigg[&-\frac{\mathbb{1}_1(x_n)m_{1n}}{\Phi_{l_n,m_{1n},m_{2n}}} - \frac{\mathbb{1}_2(x_n)m_{1n}}{\Phi_{l_n,m_{1n},m_{2n}}} - \frac{\mathbb{1}_{-1}(x_n)m_{1n}}{\Phi_{l_n,m_{1n},m_{2n}}}\\ &+ \frac{\mathbb{1}_{-2}(x_n)\Phi_{l_n,m_{1n},m_{2n}}}{\beta_1 m_{1n} + \beta_2 m_{2n}} \times\left(\frac{m_{1n}}{\Phi_{l_n,m_{1n},m_{2n}}} - \frac{(\beta_1 m_{1n} + \beta_2 m_{2n})m_{1n}}{(\Phi_{l_n,m_{1n},m_{2n}})^2}\right)\Bigg],\end{aligned} \tag{49}$$

$$\begin{aligned}\frac{\partial \log L}{\partial \beta_2} = \sum_{n=1}^{N}\Bigg[&-\frac{\mathbb{1}_1(z_n)m_{2n}}{\Phi_{l_n,m_{1n},m_{2n}}} - \frac{\mathbb{1}_2(z_n)m_{2n}}{\Phi_{l_n,m_{1n},m_{2n}}} - \frac{\mathbb{1}_{-1}(z_n)m_{2n}}{\Phi_{l_n,m_{1n},m_{2n}}}\\ &+ \frac{\mathbb{1}_{-2}(z_n)\Phi_{l_n,m_{1n},m_{2n}}}{\beta_1 m_{1n} + \beta_2 m_{2n}}\left(\frac{m_{2n}}{\Phi_{l_n,m_{1n},m_{2n}}} - \frac{(\beta_1 m_{1n} + \beta_2 m_{2n})m_{2n}}{\Phi^2_{l_n,m_{1n},m_{2n}}}\right)\Bigg],\end{aligned} \tag{50}$$

• For $k \geq 2$

The probability mass function (pmf) is

$$\begin{aligned}P_r(X = x) = &\left(\frac{\lambda(1 - \sum_{s=1}^{d}\gamma_s)(l - \sum_{s=1}^{d} m_s)}{\Phi_{l,m_1,\ldots,m_d}}\right)^{\mathbb{1}_1(x)}\left(\frac{\lambda(\sum_{s=1}^{d}\gamma_s)(l - \sum_{s=1}^{d} m_s) + \sum_{s=1}^{d}\alpha_s m_s}{\Phi_{l,m_1,\ldots,m_d}}\right)^{\mathbb{1}_2(x)}\\ &\times\left(\frac{\mu(l - \sum_{s=1}^{d} m_s)}{\Phi_{l,m_1,\ldots,m_d}}\right)^{\mathbb{1}_{-1}(x)}\left(\frac{\sum_{s=1}^{d}\beta_s m_s}{\Phi_{l,m_1,\ldots,m_d}}\right)^{\mathbb{1}_{-2}(x)},\end{aligned} \tag{51}$$

with $x \in \{1, 2, -1, -2\}$ and $\Phi_{l,m_1,\ldots,m_d} = (\lambda + \mu)(l - \sum_{s=1}^{d} m_s) + \sum_{s=1}^{d}(\alpha_s + \beta_s)m_s$. The likelihood function is

$$
\begin{aligned}
&L(\lambda, \mu, \gamma_1, \alpha_1, \beta_1, \cdots \gamma_d, \alpha_d, \beta_d | X) \\
&= \prod_{n=1}^{N} \left[\left(\frac{\lambda(1 - \sum_{s=1}^{d} \gamma_s)(l_n - \sum_{s=1}^{d} m_{sn})}{\Phi_{l_n, m_{1n}, \cdots, m_{dn}}} \right)^{\mathbb{1}_1(x_n)} \right. \\
&\quad \times \left(\frac{\lambda(\sum_{s=1}^{d} \gamma_s)(l_n \sum_{s=1}^{d} m_{sn}) + \sum_{s=1}^{d} \alpha_s m_{sn}}{\Phi_{l_n, m_{1n}, \cdots, m_{dn}}} \right)^{\mathbb{1}_2(x_n)} \\
&\quad \left. \times \left(\frac{\mu(l_n - \sum_{s=1}^{d} m_{sn})}{\Phi_{l_n, m_{1n}, \ldots, m_{dn}}} \right)^{\mathbb{1}_{-1}(x_n)} \left(\frac{\sum_{s=1}^{d} \beta_s m_{sn}}{\Phi_{l_n, m_{1n}, \cdots, m_{dn}}} \right)^{\mathbb{1}_{-2}(x_n)} \right].
\end{aligned} \tag{52}
$$

The log-likelihood function is

$$
\begin{aligned}
\log L = \sum_{n=1}^{N} \Bigg[& \mathbb{1}_1(x_n) \log \left(\frac{\lambda(1 - \sum_{s=1}^{d} \gamma_s)(l_n - \sum_{s=1}^{d} m_{sn})}{\Phi_{l_n, m_{1n}, \cdots, m_{dn}}} \right) \\
& + \mathbb{1}_2(x_n) \log \left(\frac{\lambda(\sum_{s=1}^{d} \gamma_s)(l_n \sum_{s=1}^{d} m_{sn}) + \sum_{s=1}^{d} \alpha_s m_{sn}}{\Phi_{l_n, m_{1n}, \cdots, m_{dn}}} \right) \\
& + \mathbb{1}_{-1}(x_n) \log \left(\frac{\mu(l_n - \sum_{s=1}^{d} m_{sn})}{\Phi_{l_n, m_{1n}, \ldots, m_{dn}}} \right) + \mathbb{1}_{-2}(x_n) \log \left(\frac{\sum_{s=1}^{d} \beta_s m_{sn}}{\Phi_{l_n, m_{1n}, \cdots, m_{dn}}} \right) \Bigg].
\end{aligned} \tag{53}
$$

The first derivatives are

$$
\begin{aligned}
\frac{\partial \log L}{\partial \lambda} = \sum_{n=1}^{N} \Bigg[& \frac{\mathbb{1}_1(x_n) \Phi_{l_n, m_{1n}, \cdots, m_{dn}}}{\lambda(1 - \sum_{s=1}^{d} \gamma_s)(l_n - \sum_{s=1}^{d} m_{sn})} \left(\frac{(1 - \sum_{s=1}^{d} \gamma_s)(l_n - \sum_{s=1}^{d} m_{sn})}{\Phi_{l_n, m_{1n}, \ldots, m_{dn}}} \right. \\
& \left. - \frac{\lambda(1 - \sum_{s=1}^{d} \gamma_s)(l_n - \sum_{s=1}^{d} m_{sn})^2}{\Phi^2_{l_n, m_{1n}, \ldots, m_{dn}}} \right) + \frac{\mathbb{1}_2(x_n) \Phi_{l_n, m_{1n}, \ldots, m_{dn}}}{\lambda \sum_{s=1}^{d} \gamma_s l_n \sum_{s=1}^{d} m_{sn} + \sum_{s=1}^{d} \alpha_s m_{sn}} \\
& \times \left(\frac{\sum_{s=1}^{d} \gamma_s l_n \sum_{s=1}^{d} m_{sn}}{\Phi_{l_n, m_{1n}, \cdots, m_{dn}}} - \frac{(\lambda \sum_{s=1}^{d} \gamma_s l_n \sum_{s=1}^{d} m_{sn} + \sum_{s=1}^{d} \alpha_s m_s)(l_n - \sum_{s=1}^{d} m_{sn})}{\Phi^2_{l_n, m_{1n}, \ldots, m_{dn}}} \right) \\
& - \frac{\mathbb{1}_{-1}(x_n)(l_n - \sum_{s=1}^{d} m_{sn})}{\Phi_{l_n, m_{1n}, \ldots, m_{dn}}} - \frac{\mathbb{1}_{-2}(x_n)(l_n - \sum_{s=1}^{d} m_{sn})}{\Phi_{l_n, m_{1n}, \ldots, m_{dn}}} \Bigg],
\end{aligned} \tag{54}
$$

$$
\begin{aligned}
\frac{\partial \log L}{\partial \mu} = \sum_{n=1}^{N} \Bigg[& - \frac{\mathbb{1}_1(x_n)(l_n - \sum_{s=1}^{d} m_{sn})}{\Phi_{l_n, m_{1n}, \ldots, m_{dn}}} - \frac{\mathbb{1}_2(x_n)(l_n - \sum_{s=1}^{d} m_{sn})}{\Phi_{l_n, m_{1n}, \ldots, m_{dn}}} + \frac{\mathbb{1}_{-1}(x_n) \Phi_{l_n, m_{1n}, \ldots, m_{dn}}}{\mu(l_n - \sum_{s=1}^{d} m_{sn})} \\
& \times \left(\frac{l_n - \sum_{s=1}^{d} m_{sn}}{\Phi_{l_n, m_{1n}, \ldots, m_{dn}}} - \frac{\mu(l_n - \sum_{s=1}^{d} m_{sn})^2}{\Phi^2_{l_n, m_{1n}, \ldots, m_{dn}}} \right) - \frac{\mathbb{1}_{-2}(x_n)(l_n - \sum_{s=1}^{d} m_{sn})}{\Phi_{l_n, m_{1n}, \ldots, m_{dn}}} \Bigg],
\end{aligned} \tag{55}
$$

$$
\frac{\partial \log L}{\partial \gamma_s} = \sum_{n=1}^{N} \left[- \frac{\mathbb{1}_1(x_n) d}{1 - \sum_{s=1}^{d} \gamma_s} + \frac{\mathbb{1}_2(x_n) \lambda d l_n \sum_{s=1}^{d} m_{sn}}{\lambda \sum_{s=1}^{d} \gamma_s l\_n \sum_{s=1}^{d} m_{sn} + \sum_{s=1}^{d} \alpha_s m_{sn}} \right], \tag{56}
$$

$$\frac{\partial \log L}{\partial \alpha_s} = \sum_{n=1}^{N}\Bigg[-\frac{\mathbb{1}_1(x_n)\sum_{s=1}^{d} m_{sn}}{\Phi_{l_n,m_{1n},\cdots,m_{dn}}} + \frac{\mathbb{1}_2(x_n)\Phi_{l_n,m_{1n},\cdots,m_{dn}}}{\lambda\sum_{s=1}^{d}\gamma_s l_n\sum_{s=1}^{d} m_{sn} + \sum_{s=1}^{d}\alpha_s m_{sn}}\Bigg(\frac{\sum_{s=1}^{d} m_{sn}}{\Phi_{l_n,m_{1n},\cdots,m_{dn}}} - \frac{\left(\lambda\sum_{s=1}^{d}\gamma_s l_n\sum_{s=1}^{d} m_{sn} + \sum_{s=1}^{d}\alpha_s m_{sn}\right)\sum_{s=1}^{d} m_{sn}}{\Phi^2_{l_n,m_{1n},\cdots,m_{dn}}}\Bigg) - \frac{\mathbb{1}_{-1}(x_n)\sum_{s=1}^{d} m_{sn}}{\Phi_{l_n,m_{1n},\cdots,m_{dn}}} - \frac{\mathbb{1}_{-2}(x_n)\sum_{s=1}^{d} m_{sn}}{\Phi_{l_n,m_{1n},\cdots,m_{dn}}}\Bigg], \tag{57}$$

$$\frac{\partial \log L}{\partial \beta_s} = \sum_{n=1}^{N}\Bigg[-\frac{\mathbb{1}_1(x_n)\sum_{s=1}^{d} m_{sn}}{\Phi_{l_n,m_{1n},\cdots,m_{dn}}} - \frac{\mathbb{1}_2(x_n)\sum_{s=1}^{d} m_{sn}}{\Phi_{l_n,m_{1n},\cdots,m_{dn}}} - \frac{\mathbb{1}_{-1}(x_n)\sum_{s=1}^{d} m_s}{\Phi_{l_n,m_{1n},\cdots,m_{dn}}} + \frac{\mathbb{1}_{-2}(x_n)\Phi_{l_n,m_{1n},\cdots,m_{dn}}}{\sum_{s=1}^{d}\beta_s m_{sn}}\Bigg(\frac{\sum_{s=1}^{d} m_{sn}}{\Phi_{l_n,m_{1n},\cdots,m_{dn}}} - \frac{\sum_{s=1}^{d}\beta_s m_{sn}\sum_{s=1}^{d} m_{sn}}{\Phi^2_{l_n,m_{1n},\cdots,m_{dn}}}\Bigg)\Bigg]. \tag{58}$$

The search for analytical solutions using the aforementioned equations would be quite complex. In practice, this optimization needs to be solved numerically rather than analytically. The limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) algorithm, a Quasi-Newton method, has been chosen in this study for various reasons. Quasi-Newton methods are widely recognized for their effectiveness in nonlinear optimization, are frequently incorporated into various software libraries, and prove particularly advantageous when the computation of the Hessian matrix is challenging [12]. Among the quasi-Newton methods, the BFGS method stands out as the most popular and effective [5,35]. The L-BFGS algorithm, an extension of the BFGS method, addresses the computational challenges associated with a high number of parameters [16].

For the general case (k > =2), we want to estimate $\theta = (\lambda, \mu, \gamma_1, \alpha_1, \beta_1, \ldots, \gamma_d, \alpha_d, \beta_d)$, the vector of the parameters, by maximizing the log-likelihood function $\log L$:

$$\theta_{\mathrm{MLE}} = \arg\max_{\theta}\Bigg\{\sum_{n=1}^{N}\Bigg[\mathbb{1}_1(x_n)\log\left(\frac{\lambda(1-\sum_{s=1}^{d}\gamma_s)(l_n-\sum_{s=1}^{d} m_{sn})}{\Phi_{l_n,m_{1n},\cdots,m_{dn}}}\right) + \mathbb{1}_2(x_n)\log\left(\frac{\lambda(\sum_{s=1}^{d}\gamma_s)(l_n\sum_{s=1}^{d} m_{sn}) + \sum_{s=1}^{d}\alpha_s m_{sn}}{\Phi_{l_n,m_{1n},\cdots,m_{dn}}}\right) + \mathbb{1}_{-1}(x_n)\log\left(\frac{\mu(l_n-\sum_{s=1}^{d} m_{sn})}{\Phi_{l_n,m_{1n},\ldots,m_{dn}}}\right) + \mathbb{1}_{-2}(x_n)\log\left(\frac{\sum_{s=1}^{d}\beta_s m_{sn}}{\Phi_{l_n,m_{1n},\cdots,m_{dn}}}\right)\Bigg]\Bigg\}. \tag{59}$$

The BFGS algorithm applied to that optimization problem (equation (59)) is as follows:

Initial step : Let $\varepsilon > 0$ be the stopping criteria. State initial values θ_1 and a symmetric positive definite matrix Q_1.

Step 1: If

$$\Bigg\|\nabla\Bigg\{\sum_{n=1}^{N}\Bigg[\mathbb{1}_1(x_n)\log\left(\frac{\lambda(1-\sum_{s=1}^{d}\gamma_s)(l_n-\sum_{s=1}^{d} m_{sn})}{\Phi_{l_n,m_{1n},\cdots,m_{dn}}}\right) + \mathbb{1}_2(x_n)\log\left(\frac{\lambda(\sum_{s=1}^{d}\gamma_s)(l_n\sum_{s=1}^{d} m_{sn}) + \sum_{s=1}^{d}\alpha_s m_{sn}}{\Phi_{l_n,m_{1n},\cdots,m_{dn}}}\right) + \mathbb{1}_{-1}(x_n)\log\left(\frac{\mu(l_n-\sum_{s=1}^{d} m_{sn})}{\Phi_{l_n,m_{1n},\ldots,m_{dn}}}\right) + \mathbb{1}_{-2}(x_n)\log\left(\frac{\sum_{s=1}^{d}\beta_s m_{sn}}{\Phi_{l_n,m_{1n},\cdots,m_{dn}}}\right)\Bigg]\Bigg\}\Bigg\| < \varepsilon, \tag{60}$$

stop.

If not, state

$$
\begin{aligned}
d_k = Q_k \nabla \log \Bigg\{ \sum_{n=1}^{N} \Bigg[& \mathbb{1}_1(x_n) \log \left(\frac{\lambda(1 - \sum_{s=1}^{d} \gamma_s)(l_n - \sum_{s=1}^{d} m_{sn})}{\Phi_{l_n, m_{1n}, \cdots, m_{dn}}} \right) \\
& + \mathbb{1}_2(x_n) \log \left(\frac{\lambda(\sum_{s=1}^{d} \gamma_s)(l_n \sum_{s=1}^{d} m_{sn}) + \sum_{s=1}^{d} \alpha_s m_{sn}}{\Phi_{l_n, m_{1n}, \cdots, m_{dn}}} \right) \\
& + \mathbb{1}_{-1}(x_n) \log \left(\frac{\mu(l_n - \sum_{s=1}^{d} m_{sn})}{\Phi_{l_n, m_{1n}, \ldots, m_{dn}}} \right) + \mathbb{1}_{-2}(x_n) \log \left(\frac{\sum_{s=1}^{d} \beta_s m_{sn}}{\Phi_{l_n, m_{1n}, \cdots, m_{dn}}} \right) \Bigg] \Bigg\},
\end{aligned} \tag{61}
$$

choose $\alpha_k > 0$ and compute $\theta_{k+1} = \theta_k - \alpha_k d_k$, with ∇, the gradient.

Step 2: Compute Q_{k+1}

$$
Q_{k+1} = Q_k + \frac{a_k a_k^t}{a_k b_k^t} + \frac{b_k^t Q_k b_k a_k a_k^t}{(a_k b_k^t)^2} - \frac{a_k b_k^t Q_k + Q_k b_k a_k^t}{a_k^t b_k}, \tag{62}
$$

with

$$
a_k = \alpha_k d_k \equiv \theta_{k+1} - \theta_k, \tag{63}
$$

$$
\begin{aligned}
b_k = & \nabla_{k+1} \Bigg\{ \sum_{n=1}^{N} \Bigg[\mathbb{1}_1(x_n) \log \left(\frac{\lambda(1 - \sum_{s=1}^{d} \gamma_s)(l_n - \sum_{s=1}^{d} m_{sn})}{\Phi_{l_n, m_{1n}, \cdots, m_{dn}}} \right) \\
& + \mathbb{1}_2(x_n) \log \left(\frac{\lambda(\sum_{s=1}^{d} \gamma_s)(l_n \sum_{s=1}^{d} m_{sn}) + \sum_{s=1}^{d} \alpha_s m_{sn}}{\Phi_{l_n, m_{1n}, \cdots, m_{dn}}} \right) \\
& + \mathbb{1}_{-1}(x_n) \log \left(\frac{\mu(l_n - \sum_{s=1}^{d} m_{sn})}{\Phi_{l_n, m_{1n}, \ldots, m_{dn}}} \right) + \mathbb{1}_{-2}(x_n) \log \left(\frac{\sum_{s=1}^{d} \beta_s m_{sn}}{\Phi_{l_n, m_{1n}, \cdots, m_{dn}}} \right) \Bigg] \Bigg\} \\
& - \nabla_k \Bigg\{ \sum_{n=1}^{N} \Bigg[\mathbb{1}_1(x_n) \log \left(\frac{\lambda(1 - \sum_{s=1}^{d} \gamma_s)(l_n - \sum_{s=1}^{d} m_{sn})}{\Phi_{l_n, m_{1n}, \cdots, m_{dn}}} \right) \\
& + \mathbb{1}_2(x_n) \log \left(\frac{\lambda(\sum_{s=1}^{d} \gamma_s)(l_n \sum_{s=1}^{d} m_{sn}) + \sum_{s=1}^{d} \alpha_s m_{sn}}{\Phi_{l_n, m_{1n}, \cdots, m_{dn}}} \right) \\
& + \mathbb{1}_{-1}(x_n) \log \left(\frac{\mu(l_n - \sum_{s=1}^{d} m_{sn})}{\Phi_{l_n, m_{1n}, \ldots, m_{dn}}} \right) + \mathbb{1}_{-2}(x_n) \log \left(\frac{\sum_{s=1}^{d} \beta_s m_{sn}}{\Phi_{l_n, m_{1n}, \cdots, m_{dn}}} \right) \Bigg] \Bigg\}.
\end{aligned} \tag{64}
$$

We then replace k with $k + 1$ and return to **Step 1**.

5 Simulation study

For this simulation study, three main scenarios have been considered: one type of mutant cells, two types of mutant cells, and five types of mutant cells with sample sizes of 30, 100, and 500. Monte-Carlo experiments with 1,000 replications have been conducted. For each parameter, we compute the mean, the bias, and the mean-squared error (MSE) defined as

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}\theta_i, \tag{65}$$

$$\text{Bias} = \frac{1}{n}\sum_{i=1}^{n}(\theta_i - \theta), \tag{66}$$

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\theta_i - \theta)^2, \tag{67}$$

where θ is the true value vector.

Practically, we sort the data based on the variable Z such that from index 1 to N_1, the value of $Z = 1$, from $N_1 + 1$ to N_2, $Z = 2$, from $N_2 + 1$ to N_3, $Z = -1$ and $N_3 + 1$ to N, $Z = -2$. Then, Z has the following distribution (Table 2).

Table 2: Conceptual distribution of the variable Z for a dataset of N observations

Z	1	2	1	2	*Total*
Number of observations	N_1	$N_2 - N_1$	$N_3 - N_2$	$N - N_3$	N

The log-likelihood function can be simplified as follows:

$$\begin{aligned}\log L = &\sum_{n=1}^{N_1}\log\left(\frac{\lambda(1-\sum_{s=1}^{d}\gamma_s)(l_n-\sum_{s=1}^{d}m_{sn})}{\Phi_{l_n,m_{1n},\cdots,m_{dn}}}\right)\\ &+\sum_{n=N_1+1}^{N_2}\log\left(\frac{\lambda(\sum_{s=1}^{d}\gamma_s)(l_n\sum_{s=1}^{d}m_{sn})+\sum_{s=1}^{d}\alpha_s m_{sn}}{\Phi_{l_n,m_{1n},\cdots,m_{dn}}}\right)\\ &+\sum_{n=N_2+1}^{N_3}\log\left(\frac{\mu(l_n-\sum_{s=1}^{d}m_{sn})}{\Phi_{l_n,m_{1n},\ldots,m_{dn}}}\right)+\sum_{n=N_3+1}^{N}\log\left(\frac{\sum_{s=1}^{d}\beta_s m_{sn}}{\Phi_{l_n,m_{1n},\cdots,m_{dn}}}\right).\end{aligned} \tag{68}$$

This formula (equation (68)) is easier to compute compared to the previous one (equation (53)).

5.1 Scenario 1: One type of mutant cells ($k = 1$)

For λ, the mean estimates exhibit a gradual convergence toward the true value as sample size increases. The bias decreases from 0.0968 to 0.0007 as the sample size expands from 30 to 500, indicating a diminishing systematic error in estimation. Moreover, the MSE decreases substantially from 0.0450 to 0.0091, highlighting the improved accuracy of estimation with larger sample sizes. Similar trends can be observed for γ_1, α_1, μ, and β_1. As the sample size grows, biases decrease, indicating a reduction in systematic errors, and MSE values decline, reflecting enhanced precision in estimation (Table 3).

Table 3: Parameter estimation statistics for $k = 1$

Parameter	True value	Sample size 30			Sample size 100			Sample size 500		
		Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
λ	0.198	0.2948	0.0968	0.0450	0.2065	0.0085	0.0147	0.1987	0.0007	0.0091
γ_1	0.486	0.5262	0.0402	0.0458	0.4813	0.0047	0.0472	0.4859	0.0001	0.0418
α_1	0.172	0.1935	0.0215	0.0269	0.1726	0.0006	0.0227	0.1720	0.0000	0.0221
μ	0.224	0.2656	0.0416	0.0276	0.2256	0.0016	0.0247	0.2241	0.0001	0.0266
β_1	0.414	0.2909	0.1231	0.0420	0.3981	0.0159	0.0253	0.4138	0.0002	0.0245

5.2 Scenario 2: Two types of mutant cells ($k = 2$)

For λ, μ, γ_1, α_1, β_1, γ_2, α_2, and β_2, we observe similar patterns to those in the previous table. Specifically, biases tend to decrease as sample size increases, indicating improved accuracy in estimation. Moreover, MSE values generally decrease with larger sample sizes, suggesting enhanced precision. Notably, for parameters such as α_1, α_2, β_1, and β_2, biases are relatively high for smaller sample sizes (30) but decrease substantially as the sample size increases, indicating the influence of sample size on the accuracy of estimation (Table 4).

Table 4: Parameter estimation statistics for $k = 2$

Parameter	True value	Sample size 30			Sample size 100			Sample size 500		
		Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
λ	0.145	0.1818	0.0368	0.0188	0.1422	0.0028	0.0064	0.1448	0.0002	0.0036
μ	0.154	0.1620	0.0080	0.0082	0.1477	0.0063	0.0078	0.1584	0.0044	0.0087
γ_1	0.217	0.2449	0.0279	0.0145	0.2149	0.0021	0.0146	0.2169	0.0001	0.0139
α_1	0.113	0.1870	0.0740	0.0232	0.1312	0.0182	0.0125	0.1126	0.0004	0.0087
β_1	0.275	0.2283	0.0467	0.0207	0.2517	0.0233	0.0143	0.2747	0.0003	0.0074
γ_2	0.27	0.2955	0.0255	0.0133	0.2662	0.0038	0.0132	0.2705	0.0005	0.0119
α_2	0.11	0.1854	0.0754	0.0234	0.1306	0.0206	0.0124	0.1103	0.0003	0.0083
β_2	0.274	0.2253	0.0487	0.0208	0.2579	0.0161	0.0141	0.2738	0.0002	0.0065

5.3 Scenario 3: Five types of mutant cells ($k = 5$)

Across all parameters, biases tend to decrease as sample size increases, indicating improved accuracy in estimation. Moreover, MSE values generally decrease with larger sample sizes, suggesting enhanced precision. For each parameter, there are notable differences in biases and MSE values across different sample sizes. Generally, larger sample sizes lead to smaller biases and MSE values, indicating more accurate and precise estimation. Notably, parameters such as α_i and β_i exhibit relatively high biases and MSE values for smaller sample sizes (30), but these decrease substantially as the sample size increases, highlighting the influence of sample size on the accuracy of estimation (Table 5).

Table 5: Parameter estimation statistics for $k = 5$

Parameter	True value	Sample size 30			Sample size 100			Sample size 500		
		Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
λ	0.059	0.0359	0.0231	0.0015	0.0353	0.0237	0.0012	0.0585	0.0005	0.0008
μ	0.066	0.0441	0.0219	0.0015	0.0455	0.0205	0.0013	0.0662	0.0002	0.0015
γ_1	0.095	0.0972	0.0022	0.0026	0.0927	0.0023	0.0025	0.0951	0.0001	0.0024
α_1	0.066	0.1182	0.0522	0.0074	0.0802	0.0142	0.0040	0.0662	0.0002	0.0027
β_1	0.117	0.1204	0.0034	0.0046	0.1098	0.0072	0.0046	0.1173	0.0003	0.0038
γ_2	0.095	0.0972	0.0022	0.0026	0.0928	0.0022	0.0025	0.0951	0.0001	0.0024
α_2	0.062	0.1199	0.0579	0.0077	0.0794	0.0174	0.0039	0.0623	0.0003	0.0023
β_2	0.117	0.1102	0.0068	0.0047	0.1149	0.0021	0.0048	0.1170	0.0000	0.0036
γ_3	0.095	0.0972	0.0022	0.0026	0.0928	0.0022	0.0025	0.0951	0.0001	0.0024
α_3	0.07	0.1226	0.0526	0.0071	0.0819	0.0119	0.0039	0.0699	0.0001	0.0029
β_3	0.12	0.1229	0.0029	0.0047	0.1098	0.0102	0.0048	0.1195	0.0005	0.0037
γ_4	0.095	0.0972	0.0022	0.0026	0.0930	0.0020	0.0025	0.0951	0.0001	0.0024
α_4	0.07	0.1203	0.0503	0.0074	0.0749	0.0049	0.0032	0.0700	0.0000	0.0029
β_4	0.12	0.1201	0.0001	0.0046	0.1118	0.0082	0.0047	0.1202	0.0002	0.0039
γ_5	0.095	0.0974	0.0024	0.0026	0.0928	0.0022	0.0025	0.0951	0.0001	0.0024
α_5	0.066	0.1147	0.0487	0.0070	0.0777	0.0117	0.0038	0.0657	0.0003	0.0028
β_5	0.119	0.1208	0.0018	0.0044	0.1117	0.0073	0.0050	0.1186	0.0004	0.0036

6 Real-life data applications

In this section, we showcase the practicality and utility of the discrete Markov model on dynamic state space using real-world COVID-19 data from the California Department of Public Health (CDPH). The prevalence of circulating SARS-CoV-2 variants in California is being determined by the CDPH through the analysis of data from CDPH Genomic Surveillance and California reportable disease information exchange, the department's communicable disease reporting and surveillance system. Over time, viruses undergo mutations, leading to the emergence and disappearance of various variants. Some variants become widespread and persist, while others are transient. Across specialized laboratories statewide, a fraction of all positive COVID-19 tests have their genomes sequenced to identify circulating variants. Eight main variants have been followed daily (from January 1st, 2021, to November 30, 2023): alpha, beta, delta, epsilon, gamma, lambda, mu, and omicron

The dataset is available through this link: https://data.chhs.ca.gov/sk/dataset/covid-19-variant-data/resource/d7f9acfa-b113-4cbc-9abc-91e707efc08a

The data exploration has been done based on progression analysis (through line plots), violin plots, and decomposition (trend, seasonality, residuals).

6.1 Progression of each SARS-CoV-2 variant over time

Figure 1 show the progression of each SARS-CoV-2 variant over time.

Each variant, excluding omicron, demonstrates distinct temporal dynamics characterized by a heightened incidence in 2021, featuring significant peaks in the first half and a subsequent substantial decline in the latter half of the year. Post-2021, variants like alpha, beta, delta, epsilon, gamma, lambda, and mu have virtually vanished from the epidemiological landscape. Conversely, the omicron variant has emerged as a prominent factor from 2022 onward, showcasing an evolution marked by pronounced fluctuations in subsequent days.

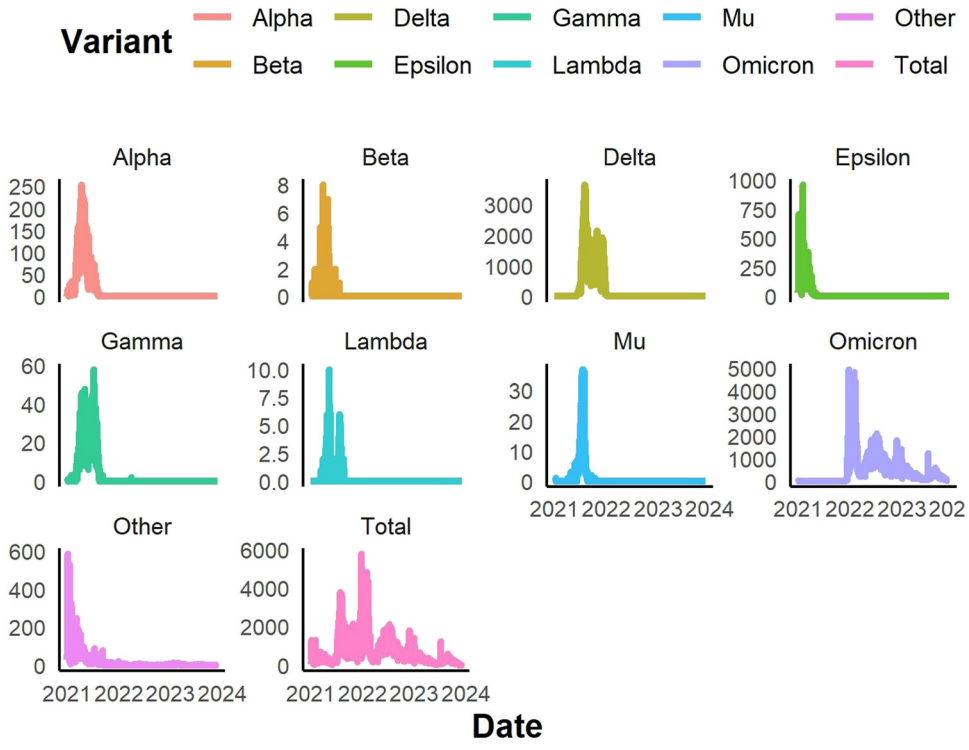

Figure 1: Progression of the variants over time.

It is noteworthy that, despite the nearly eradicated prevalence of the aforementioned variants, others persist with sustained presence. While their incidence is relatively modest from 2022 onward, their endurance underscores the intricacies of the epidemiological scenario. This observation also underscores the imperative for continual surveillance to evaluate the trajectory of SARS-CoV-2 variants and adjust public health strategies accordingly.

6.2 Violin plots of the SARS-CoV-2 variants

The violin plots (Figure 2) show various distributions for each of the variants. Only the omicron variant is showing a clear box plot, underlying the high randomness of the disease progression over time.

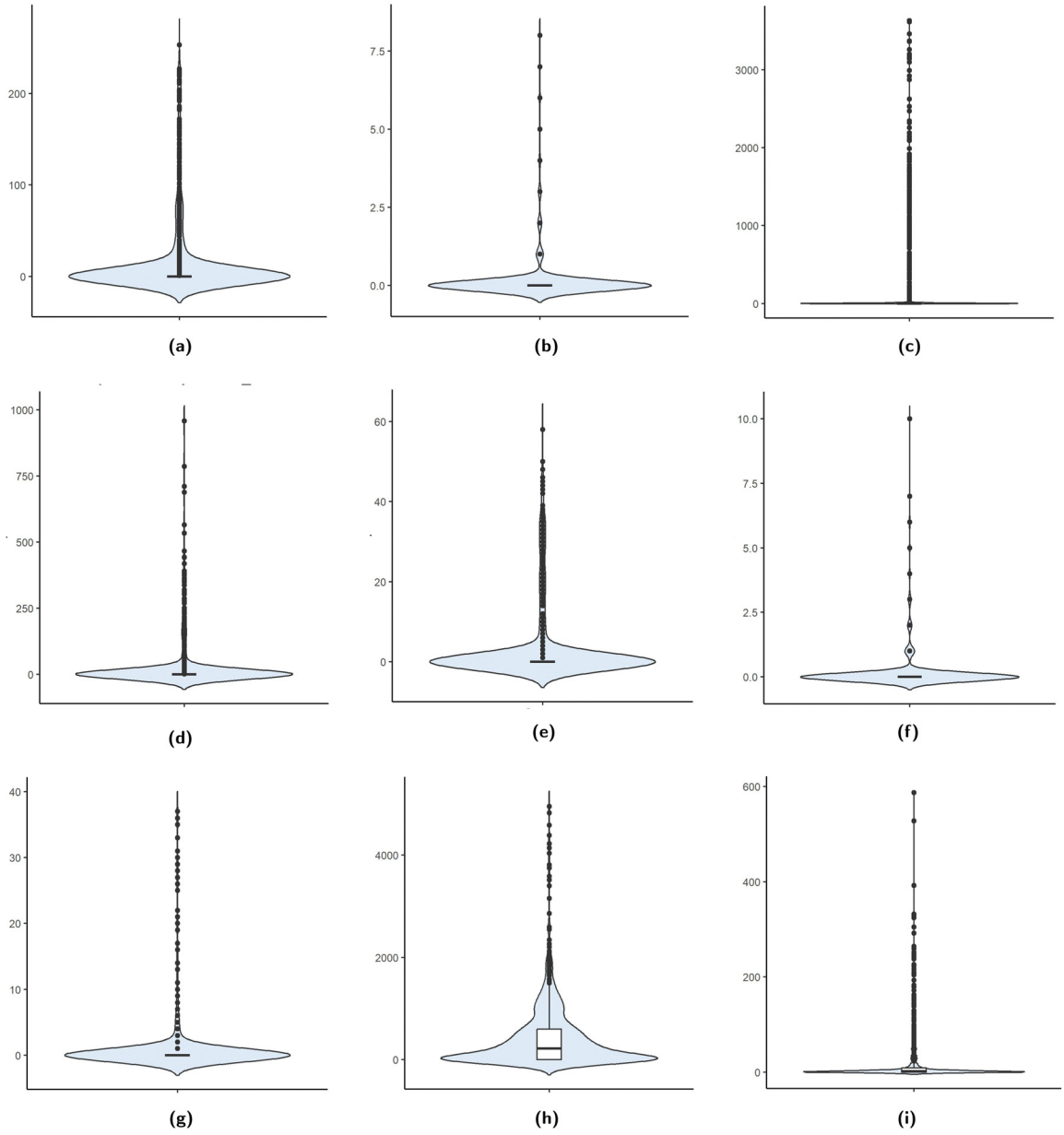

Figure 2: Violin plots for each type of variant: (a) violin plot for alpha variant, (b) violin plot for beta variant, (c) violin plot for delta variant, (d) violin plot for epsilon variant, (e) violin plot for gamma variant, (f) violin plot for lambda variant, (g) violin plot for mu variant, (h) violin plot for omicron variant, (i) violin plot for other variants.

6.3 Decomposition of the SARS-CoV-2 variants

In Figures 3 and 4), the data are decomposed into three key components: trend, seasonality, and random effects for each variant. Diverse trends and seasonality are observed in the data depending on the variant. But for none of them, there is no clear trend, meaning that there is no consistent and sustained upward or downward movement in the data across any variant. That indicates that the data points exhibit significant fluctuations that are not readily explained by underlying trends or seasonality. Then, the data have high random variations over time and outliers.

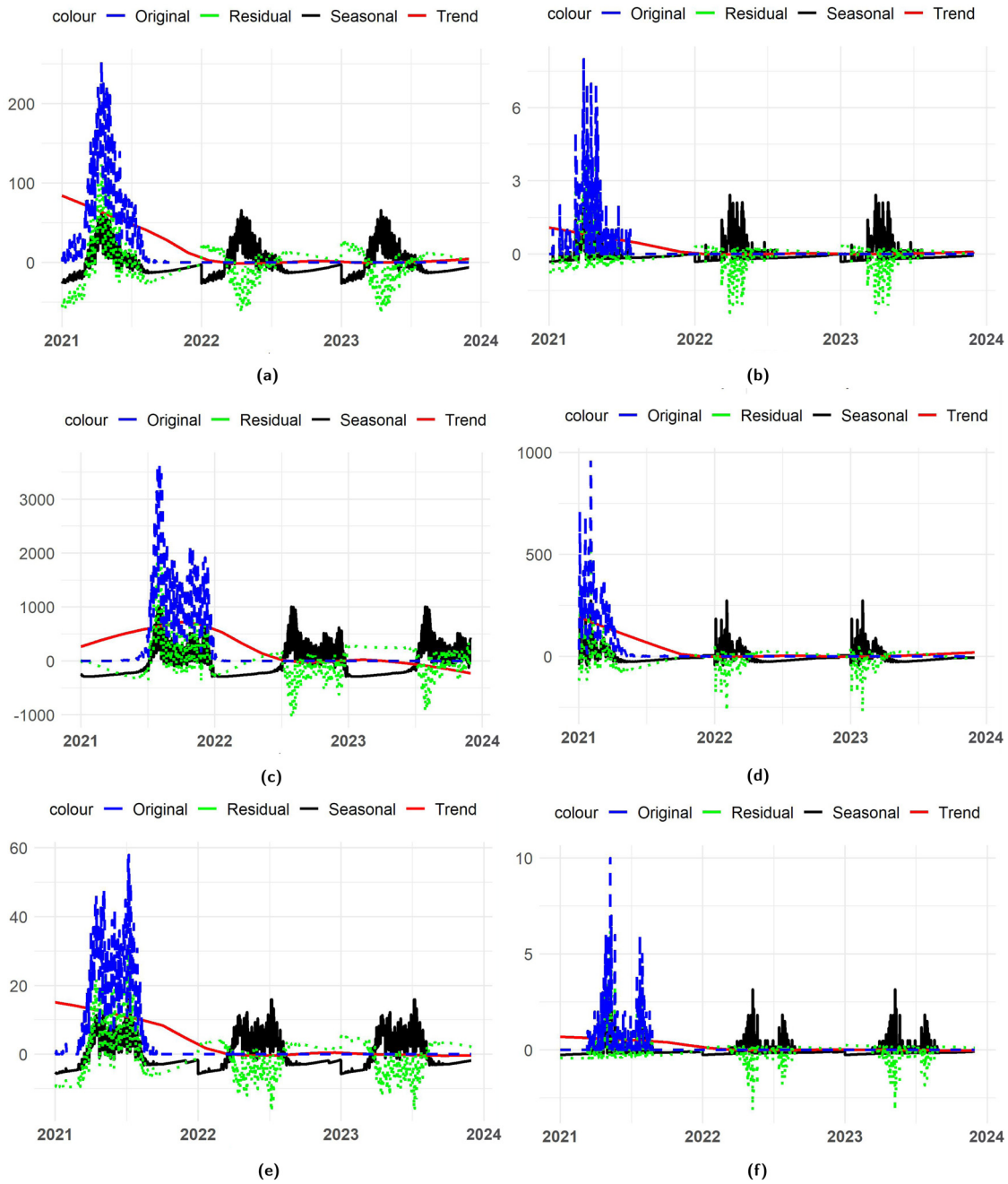

Figure 3: Decomposition plots for alpha, beta, delta, epsilon, gamma, and lambda variants: (a) alpha variant decomposition, (b) beta variant decomposition, (c) delta variant decomposition, (d) epsilon variant decomposition, (e) gamma variant decomposition, and (f) lambda variant decomposition.

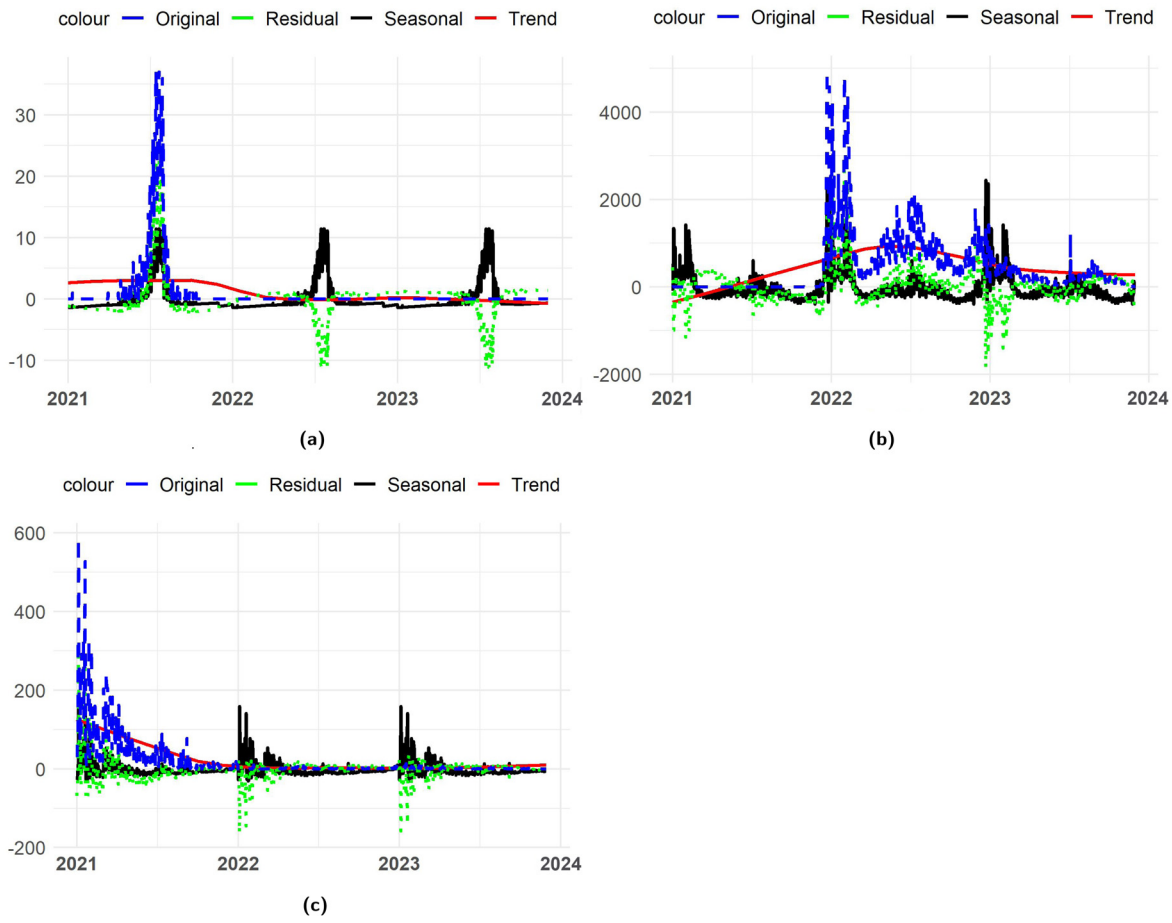

Figure 4: Decomposition plots for mu, omicron, and other variants: (a) mu variant decomposition, (b) omicron variant decomposition, and (c) other variants decomposition.

6.4 Parameter estimated values

Using the L-BFGS algorithm, the estimated values of the parameters are computed and presented in Table 6. It is important to note that α_1 and β_1, α_2 and β_2, α_3 and β_3, α_4 and β_4, α_5 and β_5, α_6 and β_6, α_7 and β_7, α_8 and β_8 are,

Table 6: Parameter estimated values

Parameter	Estimated values	Parameter	Estimated values
λ	0.3099	β_4	0.3112
μ	0.8931	γ_5	0.0694
γ_1	0.0694	α_5	0.0787
α_1	0.1095	β_5	0.2990
β_1	0.2837	γ_6	0.0694
γ_2	0.0694	α_6	0.0777
α_2	0.0779	β_6	0.3007
γ_3	0.0694	γ_7	0.0694
α_3	0.0040	α_7	0.0769
β_3	0.0087	β_7	0.3000
γ_4	0.0694	γ_8	0.0694
α_4	0.1190	α_8	0.0026
β_2	0.3006	β_8	0.0062

respectively, the division and death rates for the SARS-CoV-2 variant alpha, beta, delta, epsilon, gamma, lambda, mu and omicron; λ and μ are, respectively, the division and death rate for the other category variant; $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6, \gamma_7$, and γ_8 are, respectively, the other category background changing (after division) to variant alpha, beta, delta, epsilon, gamma, lambda, mu, and omicron.

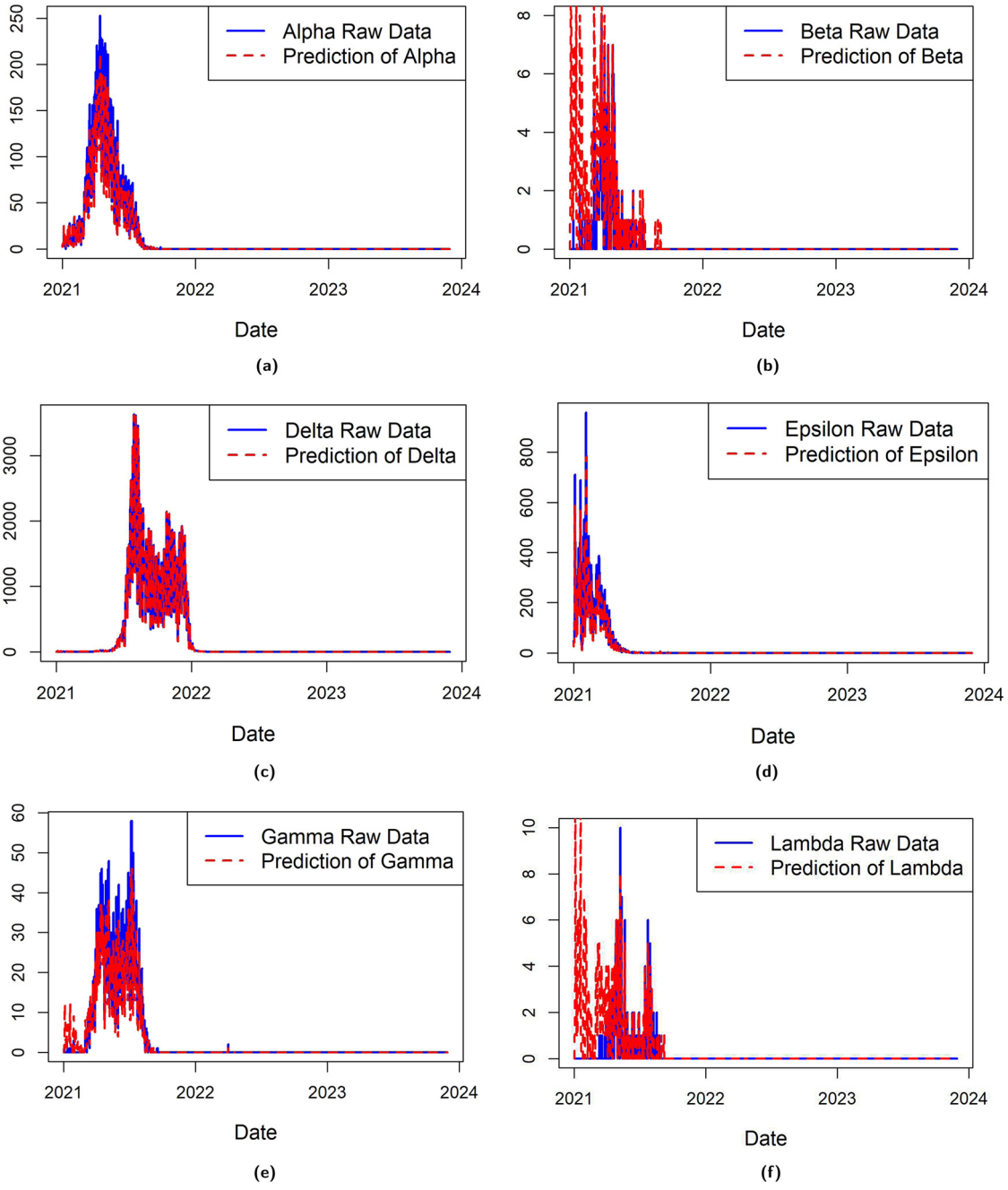

Figure 5: Prediction plots for alpha, beta, delta, epsilon, gamma, and lambda variants: (a) alpha prediction, (b) beta prediction, (c) delta prediction, (d) epsilon prediction, (e) gamma prediction, and (f) lambda prediction.

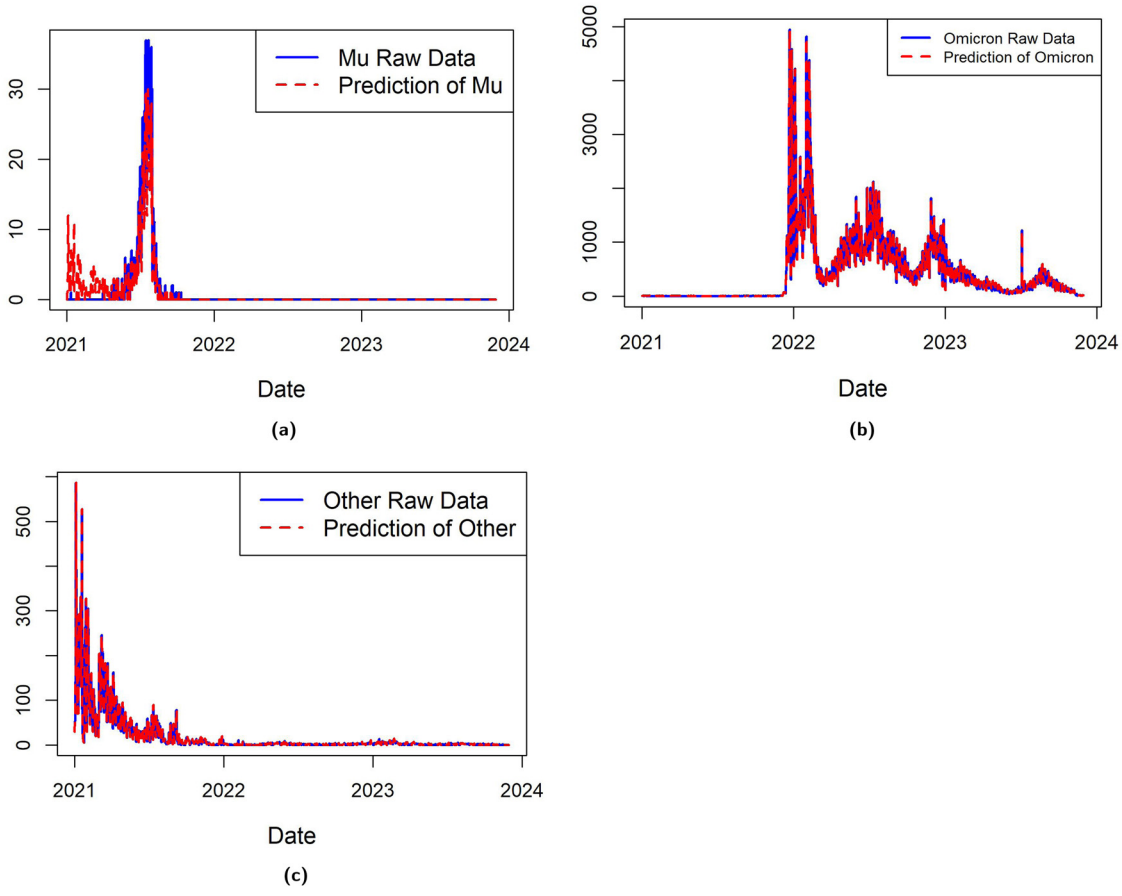

Figure 6: Prediction plots for mu, omicron, and other variants: (a) mu prediction, (b) omicron prediction, and (c) other variants prediction.

6.5 Prediction of the SARS-CoV-2 variants

Based on the estimated values of the parameters, the progression of the disease can be followed. For the SARS-CoV-2 type s variant, we have the following progression formula:

$$m_{s,t+1} = (1 + \alpha_s - \beta_s)m_{s,t} + \lambda\gamma_s\left[l - \sum_{s=1}^{d} m_{s,t}\right] + \varepsilon_s. \tag{69}$$

For each of the variants, namely alpha, beta, delta, epsilon, gamma, lambda, mu, omicron, and others, a comprehensive analysis has been conducted, delving into the predictive values that have been calculated and represented within the confines of a unified graph (Figures 5 and 6). This graphical representation offers a profound insight into the relationship between the predicted values and the corresponding raw data. Remarkably, the predictive values seamlessly align with the raw data over time, elucidating the robustness and efficacy of the predictive model. Upon closer inspection, discernible patterns emerge, unveiling nuanced distinctions among the variants. Notably, the disparities between the raw data and the predicted values manifest more prominently for the beta and lambda variants in contrast to their counterparts. Furthermore, unlike the decomposition illustrated in Figures 3 and 4, where inconsistencies may arise, the predictive model adeptly captures and reflects the intricate temporal variations with precision and fidelity. This underscores the model's robustness in discerning underlying patterns and extrapolating future trajectories.

7 Discussion

The outcomes of the simulation study reveal a significant trend: an increase in sample size across all examined scenarios consistently leads to a decrease in both bias and MSEs. This discovery carries substantial implications for statistical inference and data analysis. The decrease in bias indicates an improvement in estimation accuracy, with larger sample sizes resulting in estimates that closely align with true population parameters. This reduction in bias reflects the enhanced precision and reliability of estimates derived from larger samples. Furthermore, the diminishing MSEs suggest a decrease in variability around the estimates, highlighting the increased stability and robustness of the estimators as the sample size grows. Additionally, the observed trend signifies an enhancement in statistical power, as larger sample sizes make statistical tests more sensitive to detecting true effects or differences.

Moreover, the utilization of the Markov model on dynamic state space with COVID-19 data has produced promising results, notably demonstrating a close alignment between predicted values and observed raw data. This convergence between predicted and actual values holds significant implications for comprehending the dynamics of the pandemic and guiding decision-making processes. The agreement between predicted values and raw data underscores the model's effectiveness in capturing the underlying dynamics of COVID-19 progression, instilling confidence in its predictive capabilities for accurate forecasting of key epidemiological metrics.

However, additional validation efforts are indispensable to ensure the model's robustness and generalizability across a wide range of scenarios. By employing larger datasets and extended timeframes, the model's performance can be assessed under varying conditions. This comprehensive evaluation will not only strengthen the model's credibility but also provide valuable insights into its limitations and potential areas for improvement. Moreover, incorporating additional variables such as demographic information, environmental factors, and vaccination rates is a crucial step toward enhancing the model's capacity to reflect the complex dynamics of real-world disease progression.

Finally, expanding the model to account for the reversibility of mutations could provide deeper insights into disease evolution. Mutations can lead to the emergence of new variants with altered transmission dynamics, virulence, or immune escape properties. This enhanced model can contribute to a more comprehensive understanding of the disease's evolution and facilitate the development of effective countermeasures against emerging variants.

8 Conclusion

This study developed a dynamic state-space Markov model of genetic disorders and infectious diseases with mutations, as well as an approach to parameter estimation based on the L-BFGS algorithm. The simulation scenarios underlined the performance of the model: as sample size increased, the model demonstrably improved its ability to estimate parameters, as evidenced by consistently decreasing biases and MSEs. These simulation results established a strong foundation for the model's application in real-world settings with complex disease dynamics.

Applying the model to real-world COVID-19 variant data from California revealed distinct temporal patterns for each variant. Despite the data's complexities, the model displayed remarkable accuracy in predicting future prevalence trajectories for most variants, closely aligning with observed data points. This underscores its robustness and potential for real-time monitoring and analysis. The accurate prediction of future prevalence trends could empower timely resource allocation, targeted vaccination campaigns, and effective containment measures, ultimately contributing to enhanced pandemic management and preparedness.

However, further validation with larger datasets and longer timeframes is crucial to solidify the model's generalizability. Moreover, incorporating additional factors such as demographics, environmental influences, and vaccination rates could enhance its ability to capture the real-world complexities of disease dynamics. Furthermore, the model can be extended taking into account the reversibility of mutations.

Acknowledgements: We confirm that this article is an original work and is not currently being assessed by any other publication.

Funding information: This research received financial support from the African Union (AU), through Pan African University, Institute of Basic Sciences, Technology, and Innovation (PAUSTI).

Author contributions: Mouhamadou Djima Baranon: conceptualization, methodology, writing – original draft, investigation, software, writing – review, and editing. Patrick Guge Oloo Weke: conceptualization, supervision, validation, methodology, writing – review. Judicaël Alladatin: conceptualization, supervision, validation, and writing – review. Boni Maxime Ale: conceptualization, supervision, validation, and writing – review.

Conflict of interest: The authors have no conflicts of interest to disclose.

Ethical approval: This research does not require ethical approval.

Data availability statement: The data used to conduct this research are openly available through this link https://data.chhs.ca.gov/sk/dataset/covid-19-variant-data/resource/d7f9acfa-b113-4cbc-9abc-91e707efc08a.

References

[1] Alberts, B. (2017). *Molecular biology of the cell.* New York, USA: Garland Science.
[2] Antle, C. E., & Bain, L. J. (1969). A property of maximum likelihood estimators of location and scale parameters. *Siam Review*, *11*(2), 251–253.
[3] Blazer, D. G., & Hernandez, L. M.. (2006). *Genes, behavior, and the social environment: Moving beyond the nature/nurture debate*. Washington DC, USA: The National Academies Press.
[4] Brock, T. D., Madigan, M. T., Martinko, J. M., & Parker, J. (2003). *Brock biology of microorganisms.* Upper Saddle River (NJ): Prentice-Hall.
[5] Brownlee, J. (2021). A gentle introduction to the BFGS optimization algorithm. *Tutorial on Optimization*. https://machinelearningmastery.com/bfgs-optimization-in-python/ (accessed on 19 May 2021).
[6] Craig, B. A., Fryback, D. G., Klein, R., & Klein, B. E. (1999). A bayesian approach to modelling the natural history of a chronic condition from observations with intervention. *Statistics in Medicine*, *18*(11), 1355–1371.
[7] Data, M. C., Komorowski, M., & Raffa, J. (2016). Markov models and cost effectiveness analysis: Applications in medical research. *Secondary analysis of electronic health records* (pp. 351–367). New York, USA: Springer.
[8] Divoli, A., Mendonça, E. A., Evans, J. A., & Rzhetsky, A. (2011). Conflicting biomedical assumptions for mathematical modeling: The case of cancer metastasis. *PLoS Computational Biology*, *7*(10), e1002132.
[9] Drabo, E. F., & Padula, W. V. (2023). Introduction to Markov modeling. *Handbook of Applied Health Economics in Vaccines* (p. 264). England: Oxford University Press.
[10] Gallager, R. G. (1996). Markov processes with countable state spaces. In *Discrete Stochastic Processes* (pp. 187–222). New York, USA: Springer.
[11] Goss, C. (2014). Genetic disorders. *JEMS: a Journal of Emergency Medical Services*, *39*(2), 64–71.
[12] Griva, I., Nash, S. G., & Sofer, A. (2008). *Linear and Nonlinear Optimization* 2nd Edition. SIAM.
[13] Halevy, T., & Urbach, A. (2014). Comparing ESC and IPSC-based models for human genetic disorders. *Journal of Clinical Medicine*, *3*(4), 1146–1162.
[14] Hallen-Adams, H. E., & Suhr, M. J. (2017). Fungi in the healthy human gastrointestinal tract. *Virulence*, *8*(3), 352–358.
[15] Howard, R. A. (1960). *Dynamic programming and Markov processes*. USA: Technology Press of Massachusetts Institute of Technology.
[16] Ian, G., Yoshua, B., & Aaron, C. (2017). *Deep learning: Adaptive computation and machine learning*. USA: MIT Press.
[17] Ingram, D., & Stan, G.-B. (2023). Modelling genetic stability in engineered cell populations. *Nature Communications*, *14*(1), 3471.
[18] Jackson, C. H., Sharples, L. D., Thompson, S. G., Duffy, S. W., & Couto, E. (2003). Multistate Markov models for disease progression with classification error. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *52*(2), 193–209.
[19] Jin, J., Wu, X., Yin, J., Li, M., Shen, J., Li, J., ..., Wen, Q., et al. (2019). Identification of genetic mutations in cancer: Challenge and opportunity in the new era of targeted therapy. *Frontiers in Oncology*, *9*, 263.
[20] Khayatkhoei, M., & AbdAlmageed, W. (2023). *Emergent asymmetry of precision and recall for measuring fidelity and diversity of generative models in high dimensions*. arXiv: http://arXiv.org/abs/arXiv:2306.09618.
[21] Köhler, J. R., Hube, B., Puccia, R., Casadevall, A., & Perfect, J. R. (2017). Fungi that infect humans. *Microbiology Spectrum*, *5*(3), 5–3.

[22] Kordnoori, S., Mostafaei, H., Kordnoori, S., & Ostadrahimi, M. (2020). Testing the semi Markov model using Monte Carlo simulation method for predicting the network traffic. *Pakistan Journal of Statistics and Operation Research*, *16*(4), 713–720.
[23] Lee, S. Y., Nielsen, J., & Stephanopoulos, G. (2016). *Industrial Biotechnology: Products and Processes.* New Jersey, USA: John Wiley & Sons.
[24] Lillacci, G., & Khammash, M. (2010). Parameter estimation and model selection in computational biology. *PLoS Computational Biology*, *6*(3), e1000696.
[25] Liu, X., & Stechlinski, P. (2017). Infectious disease modeling. *A Hybrid System Approach.* Cham: Springer.
[26] Matsuno, K. (1975). Ergodicity of observable and ergodic hypothesis in markovian kinetics. *Journal of Mathematical Physics*, *16*(3), 604–608.
[27] Meira-Machado, L., de Uña-Álvarez, J., Cadarso-Suárez, C., and Andersen, P. K. (2009). Multi-state models for the analysis of time-to-event data. *Statistical Methods in Medical Research*, *18*(2), 195–222.
[28] Michel, B., Boubakri, H., Baharoglu, Z., LeMasson, M., & Lestini, R. (2007). Recombination proteins and rescue of arrested replication forks. *DNA Repair*, *6*(7), 967–980.
[29] Myers, D. S., Wallin, L., & Wikström, P. (2017). An introduction to Markov chains and their applications within finance. *MVE220 Financial Risk: Reading Project*, 26.
[30] Naeem, M. F., Oh, S. J., Uh, Y., Choi, Y., & Yoo, J. (2020). Reliable fidelity and diversity metrics for generative models. In *International Conference on Machine Learning* (pp. 7176–7185). PMLR.
[31] Newton, P. K., Mason, J., Bethel, K., Bazhenova, L. A., Nieva, J., & Kuhn, P. (2012). A stochastic Markov chain model to describe lung cancer growth and metastasis. *PloS One*, *7*(4), e34637.
[32] Norris, J. R. (1998). *Markov chains.* Number 2. Cambridge, United Kingdom: Cambridge University Press.
[33] Ollagnier, J. M. (2007). *Ergodic theory and statistical mechanics* (Vol. 1115). New York, USA: Springer.
[34] Purnell, D. W. (2006). *Discriminative and Bayesian techniques for hidden Markov model speech recognition systems* (PhD thesis). South Africa: University of Pretoria.
[35] Robinso, S. M., Mikosch, T. V., & Resnick, S. I. (2006). *Springer series in operations research and financial engineering*. New York, USA: Springer.
[36] Sanjuán, R., & Domingo-Calap, P. (2016). Mechanisms of viral mutation. *Cellular and Molecular Life Sciences*, *73*, 4433–4448.
[37] Sarker, A., Fisher, P., Gaudio, J. E., & Annaswamy, A. M. (2023). Accurate parameter estimation for safety-critical systems with unmodeled dynamics. *Artificial Intelligence*, *316*, 103857.
[38] Schmid-Hempel, P. (2009). Immune defence, parasite evasion strategies and their relevance for macroscopic phenomena such as virulence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1513), 85–98.
[39] Schneider, M., Johnson, J. R., Krogan, N. J., & Chanda, S. K. (2016). The virus-host interactome: Knowing the players to understand the game. In *Viral Pathogenesis* (pp. 157–167). Elsevier.
[40] Schwardt, L., & Preez, J. D. (2000). Efficient mixed-order hidden Markov model inference. In *Sixth International Conference on Spoken Language Processing*. Citeseer.
[41] Self, S. G., & Liang, K.-Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under non-standard conditions. *Journal of the American Statistical Association*, *82*(398), 605–610.
[42] Seneta, E. (2016). *Markov chains as models in statistical mechanics*. USA: Institute of Mathematical Statistics.
[43] Shivahare, R., Dubey, S., & McGwire, B. S. (2023). The tug of war between parasites survival and host immunity. *Frontiers in Immunology*, *14*, 1234191.
[44] Silhavy, T. J., Kahne, D., & Walker, S. (2010). The bacterial cell envelope. *Cold Spring Harbor perspectives in biology*, *2*(5), a000414.
[45] Swarts, F. (2014). *Markov characterization of fading channels*. South Africa: University of Johannesburg.
[46] Umair, M., & Alfadhel, M. (2019). Genetic disorders associated with metal metabolism. *Cells*, *8*(12), 1598.
[47] Vermolen, F., & Pölönen, I. (2020). Uncertainty quantification on a spatial Markov-chain model for the progression of skin cancer. *Journal of Mathematical Biology*, *80*(3), 545–573.
[48] Wu, J., Dhingra, R., Gambhir, M., & Remais, J. V. (2013). Sensitivity analysis of infectious disease models: Methods, advances and their application. *Journal of the Royal Society Interface*, *10*(86), 20121018.