Continuous-time Hybrid Markov/semi-Markov Model with Sojourn Time Approach in the Spread of Infectious Diseases

Faihatuz Zuhairoh, Dedi Rosadi, and Adhitya Ronnie Effendie

Abstract—The semi-Markov model is another name for the continuous-time Markov model. This model exponentially disperses the amount of time spent on each sojourn. This work aims to develop a model of infectious disease transmission using a continuous-time hybrid Markov/semi-Markov model. The model will combine assumptions from the Markov and semi-Markov models. The sojourn time in the semi-Markov model may have an exponential or a Weibull distribution, depending on the circumstances. A hybrid Markov/semi-Markov model can consist of both types of sojourn time distributions. In general, there are two parameterizations in the semi-Markov model: the transition intensity function and the sojourn time distribution, each of which has a different probabilistic and inferential perspective. This paper uses data on COVID-19 cases in DIY, Indonesia, from March 15, 2020, to July 31, 2020. This research uses four states: susceptible, infected, recovered, and deceased, where the sojourn time in the susceptible state is in the Weibull distribution while the sojourn time in other states has an Exponential distribution. This is because, at the beginning of the spread of COVID-19, few cases were found, so the distribution of sojourn time in each state tends to be constant except for the sojourn time in susceptible states. Under the semi-Markov model, the Weibull distribution leads to a dynamic probability with a higher degree of decline and a slight difference. In the final section, a comparison is made of the Markov, semi-Markov, and hybrid Markov/semi-Markov models. The hybrid Markov/semi-Markov model shows the best results with the smallest AIC value. Next, make a prediction equation for the SIRD model assuming a hybrid Markov/semi-Markov, which gives a MAPE $< 20\%$. This means that the model's ability to predict COVID-19 cases is good.

Index Terms—hybrid Markov/semi-Markov, sojourn time, transition intensity, continuous-time, epidemic model.

I. INTRODUCTION

IN epidemiology, the transmission of infectious diseases can be modelled using deterministic and stochastic models. The multi-state Markov model helps describe how an individual can move between states in continuous-time. The primary objective of modeling infectious diseases is to analyze the spread and provide the government with solutions for controlling the spread. As happened at the end of 2019, the coronavirus outbreak spread worldwide to cause a pandemic. Infected people reach hundreds of millions of people and cause the death of around six million people worldwide.

The multi-state model has been applied to various health conditions, including screening for abdominal aortic aneurysms [1], problems after lung transplantation [2], heart transplant problems [3], [4], liver cancer [5], infeksi HIV infection and AIDS [6], [7], complications of diabetes [8], [9], and breast cancer screening [10].

The duration of time spent in state k before transitioning to state l is entirely dependent on state k, which is a property of the Markov model. In addition, the sojourn time distribution is memoryless, which means that it does not take into account the length of conceivable lengths in a given state. The distribution of sojourn time is Exponential distribution. However, such assumptions may result in arbitrary constraints in practice. The semi-Markov model, which can be viewed as an extension of the Markov model is an alternate model capable of addressing this issue. Markov and semi-Markov models have both benefits and drawbacks. The primary advantage of the Markov model is its apparent simplicity. This makes interpreting and comprehending Markov models more accessible when simulating the spread of infectious diseases. In comparison, the semi-Markov model allows the length of time to be in a specific state due to a broader distribution of sojourn times. This justifies developing a hybrid model that combines both assumptions.

The contribution of this study is the combination of semi-Markov and Markov models that are applied to the continuous-time infectious disease epidemic model with four states, namely susceptible, infected, recovered, and deceased. In this study, vaccinated states were not included due to limited research data, where COVID-19 data in several countries did not include individuals infected with COVID-19 who had been vaccinated. This study employs the Exponential and Weibull distributions to model sojourn time. Then a theorem is formulated to predict the number of COVID-19 cases in the SIRD epidemic model under the continuous-time hybrid Markov/semi-Markov assumption.

Previous work of ours on COVID-19 has covered topics such as the use of transition intensity for the determination of reproduction numbers [11], [12], the implementation of a discrete-time Markov model for the spread of the virus [13]–[15], and the prediction of the COVID-19 space using a Richards curve model for the single-wave [16], [17]. The extension of the Richards curve model for the multi-wave [18] case. In addition, insight into the discrete time Markov chain model contributes to the use of the [19] continuous time Markov chain model.

F. Zuhairoh is an assistant professor from the Study Program of Mathematics Education, STKIP YPUP Makassar, Indonesia (e-mail: fzuhairoh@stkip.ypup.ac.id).

D. Rosadi is a professor from the Department of Mathematics, Universitas Gadjah Mada, Indonesia (corresponding author to provide phone: (0274)552243, Fax. (0274) 555131, e-mail: dedirosadi@ugm.ac.id).

A. R. Effendie is an associate professor from the Department of Mathematics, Universitas Gadjah Mada, Indonesia (e-mail: adhityaronnie@ugm.ac.id).

The remaining parts of the manuscript are laid out in the following order. Section II discusses the sojourn time approach, the sojourn time distribution consisting of Exponential and Weibull distributions. Section III discusses the hybrid Markov/semi-Markov model and several advantages model, including the ability to capture a fixed effect duration, which is advantageous when minimizing the number of parameters to be estimated. In addition, we also develop theorems based on the hybrid Markov/semi-Markov model. Section IV applies the SIRD epidemic model with four states: susceptible, infected, recovered, and deceased. COVID-19 case data from the DIY Special Region in Indonesia is analyzed to see if the model strictly adheres to Markov assumptions. If not, then the semi-Markov assumption will be used for each transition using test statistics to produce transition probabilities. Lastly, section V contains the conclusion.

II. Sojourn Time Approach

As already established in the semi-Markov process, a Y_w, can be formed from the state and the respective jump durations using the pair $\{Z_n, W_n\}_{n \geq 0}$. The parameters supporting this formation consist of the transition probabilities of the embedded chain, $p_{kl} = \Pr(Z_n = l | Z_{n-1} = k)$, and its distribution from the sojourn time for each transition $k \to l$, provided in $F_{kl}(w) = \Pr(H_n \leq w | Z_{n-1} = k, Z_n = l), w \geq 0$.

Embedded chain transition probabilities are frequently represented as a stochastic matrix, $\mathbf{P} = [p_{kl}]$ (where $p_{kk} = 0$ is set). Moreover, given that the distribution of sojourn time is continuous, this distribution is frequently represented in several ways, such as survival and density functions.

The survival function is

$$
\begin{aligned}
S_{kl}(w) &= \Pr(H_n > w | Z_{n-1} = k, Z_n = l) \\
&= 1 - F_{kl}(w)
\end{aligned} \tag{1}
$$

($S_{kl}(w)$ is a descending function, where $S_{kl}(0) = 1$ and $\lim_{w \to +\infty} S_{kl}(w) = 0$). The density function of the sojourn time is

$$
f_{kl}(w) = \lim_{\Delta w \to 0} \frac{\Pr(H_n \in (w, w + \Delta w) | Z_{n-1} = k, Z_n = l)}{\Delta w} \tag{2}
$$

The density function of each distribution used for continuous-time is presented as follows [20].

1) Weibull distribution ($Y \sim \text{WEI}(y; \lambda, r)$)

$$
f(y) = \begin{cases} \frac{r}{\lambda} \left(\frac{y}{\lambda}\right)^{r-1} e^{-\left(\frac{y}{\lambda}\right)^r}, & y > 0 \\ 0, & y \text{ otherwise} \end{cases} \tag{3}
$$

If $r < 1$ the monotonic hazard level goes down; if $r = 1$ the hazard level is constant (Exponential distribution), and if $r > 1$ the monotonic hazard level increases [21].

2) Exponential distribution ($Y \sim \text{EXP}(y; \lambda)$)

$$
f(y) = \begin{cases} \frac{1}{\lambda} e^{-\frac{1}{\lambda} y}, & y > 0 \\ 0, & y \text{ otherwise} \end{cases} \tag{4}
$$

The continuous sojourn time distribution uses two distributions: the Exponential distribution and the Weibull distribution. If there are two states, namely k and l, the hazard function which is often considered as the probability of a transition occurring at a specific interval $g = (w, w + \Delta w)$ without any transition before the time w, is written as

$$
\begin{aligned}
\mu_{kl}(t) &= \lim_{\Delta w \to 0} \frac{\Pr(H_n \in g | Z_{n-1} = k, Z_n = l, H_n > w)}{\Delta w} \\
&= \lim_{\Delta w \to 0} \frac{\Pr(H_n \in g | Z_{n-1} = k, Z_n = l)}{\Delta w \Pr(H_n > w | Z_{n-1} = k, Z_n = l)} \\
&= \frac{f_{kl}(w)}{S_{kl}(w)}
\end{aligned} \tag{5}
$$

The duration spent in state k before changing to state l is represented by a survival function, denoted as

$$
S_{kl}(t) = \Pr(H_n > w | Z_{n-1} = k, Z_n = l)
$$

So to show the survival function of the sojourn time in state k at the time, w can be written as

$$
\begin{aligned}
S_k(w) &= \Pr(H_n > w | Z_{n-1} = k) \\
&= \Pr(Z_n = l | Z_{n-1} = k) \frac{\Pr(H_n > w | Z_{n-1} = k)}{\Pr(Z_n = l | Z_{n-1} = k)} \\
&= p_{kl} S_{kl}(w) \\
&= \sum_{k \neq l} p_{kl} S_{kl}(w)
\end{aligned} \tag{6}
$$

The relationship between the sojourn time approach and the transition intensity function approach according to [21] can be obtained using conditional probabilities. The relationship is described as follows

$$
\begin{aligned}
\tilde{\mu}_{kl}(w) &= \lim_{\Delta w \to 0} \frac{\Pr(H_n \in g, Z_n = l | Z_{n-1} = k, H_n > w)}{\Delta w} \\
&= \lim_{\Delta w \to 0} \frac{\Pr(H_n \in g, Z_n = l | Z_{n-1} = k)}{\Delta w \Pr(H_n > w | Z_n = k)} \\
&= \frac{p_{kl} f_{kl}(w)}{S_k(w)} \\
&= \frac{p_{kl} S_{kl}(w)}{S_k(w)} \frac{f_{kl}(w)}{S_{kl}(w)} \\
&= \frac{p_{kl} S_{kl}(w)}{S_k(w)} \mu_{kl}(w)
\end{aligned} \tag{7}
$$

A. The Sojourn time is Exponential distributed

Equation (4) can be used to determine the survival function and the hazard function of the Exponential distribution as follows.

$$
\begin{aligned}
S_{kl}(w) &= 1 - F_{kl}(w) \\
&= 1 - \int_0^w f_{kl}(y) \, dy \\
&= 1 - \int_0^w \frac{1}{\lambda} e^{-\frac{1}{\lambda} y} \, dw \\
&= 1 - \left(e^{-\frac{1}{\lambda} y} |_0^w\right) \\
&= e^{-\frac{1}{\lambda} w}
\end{aligned} \tag{8}
$$

So we can also get the hazard function from the following Exponential distribution.

$$
\begin{aligned}
\mu_{kl}(w) &= \frac{f_{kl}(w)}{S_{kl}(w)} \\
&= \frac{\frac{1}{\lambda}e^{-\frac{1}{\lambda}w}}{e^{-\frac{1}{\lambda}w}} = \frac{1}{\lambda}
\end{aligned} \tag{9}
$$

The following transition intensity function can be obtained using the relationship in the Equation (7).

$$
\tilde{\mu}_{kl}(w) = p_{kl}\frac{\frac{1}{\lambda_{kl}}e^{-\frac{1}{\lambda_{kl}}w}}{\sum_{m\neq k} p_{km}e^{-\frac{1}{\lambda_{km}}w}} \tag{10}
$$

B. The Sojourn time is Weibull distributed

Equation (3) can be used to determine the survival and hazard functions of the Weibull distribution where the scale and shape parameters for the transition from k to l are λ_{kl} and η_{kl}, is obtained as follows.

$$
\begin{aligned}
S_{kl}(w) &= 1 - F_{kl}(w) \\
&= 1 - \int_0^w f_{kl}(y)\,\mathrm{d}y \\
&= 1 - \int_0^w \frac{\eta}{\lambda}\left(\frac{y}{\lambda}\right)^{\eta-1} e^{-\left(\frac{y}{\lambda}\right)^\eta}\,\mathrm{d}y \\
&= 1 - \int_0^w \frac{\eta\left(\frac{y}{\lambda}\right)^{\eta-1} e^{-\left(\frac{y}{\lambda}\right)^\eta}}{\lambda}\,\mathrm{d}y \\
&= 1 - \int_0^w \frac{\eta y^{\eta-1} e^{-\left(\frac{y}{\lambda}\right)^\eta}}{\lambda^\eta}\,\mathrm{d}y
\end{aligned} \tag{11}
$$

to do $\int_0^w \frac{\eta y^{\eta-1} e^{-\left(\frac{y}{\lambda}\right)^\eta}}{\lambda^\eta}\,\mathrm{d}y$, for example $u = -\frac{y^\eta}{\lambda^\eta}$, then $\frac{\mathrm{d}u}{\mathrm{d}y} = -\frac{\eta y^{\eta-1}}{\lambda^\eta}$ or $\mathrm{d}y = -\frac{\lambda^\eta y^{1-\eta}}{\eta}\,\mathrm{d}u$ so that the following results are obtained.

$$
\begin{aligned}
\int_0^w e^{-\left(\frac{y}{\lambda}\right)^\eta}\frac{\eta y^{\eta-1}}{\lambda^\eta}\,\mathrm{d}y &= \int_0^w e^{-u}\frac{\eta y^{\eta-1}}{\lambda^\eta}\left(-\frac{\lambda^\eta y^{1-\eta}}{\eta}\,\mathrm{d}u\right) \\
&= -\int_0^w e^{-u}\,\mathrm{d}u
\end{aligned} \tag{12}
$$

Equation (12) can be substituted for the survival function in Equation (11) to obtain the following form.

$$
\begin{aligned}
S_{kl}(w) &= 1 - \int_0^w \frac{\eta y^{\eta-1} e^{-\left(\frac{y}{\lambda}\right)^\eta}}{\lambda^\eta}\,\mathrm{d}y \\
&= 1 - \int_0^w e^{-u}\,\mathrm{d}u \\
&= 1 + e^{-u}\big|_0^w \\
&= 1 + \left(e^{\left(\frac{w}{\lambda}\right)^\eta} - 1\right) \\
&= e^{-\left(\frac{w}{\lambda}\right)^\eta}
\end{aligned} \tag{13}
$$

So that the hazard function can also be obtained from the following Weibull distribution.

$$
\begin{aligned}
\mu_{kl}(w) &= \frac{f_{kl}(w)}{S_{kl}(w)} \\
&= \frac{\frac{\eta}{\lambda}\left(\frac{w}{\lambda}\right)^{\eta-1} e^{-\left(\frac{w}{\lambda}\right)^\eta}}{e^{-\left(\frac{w}{\lambda}\right)^\eta}} \\
&= \frac{\eta}{\lambda}\left(\frac{w}{\lambda}\right)^{\eta-1}
\end{aligned} \tag{14}
$$

The following transition intensity function can be obtained using the relationship in the Equation (7).

$$
\begin{aligned}
\tilde{\mu}_{kl}(w) &= p_{kl}\frac{e^{-(w/\lambda_{kl})^{\eta_{kl}}}}{\sum_{\eta\neq k} p_{km}e^{-(w/\lambda_{km})^{\eta_{km}}}}\frac{\eta_{kl}}{\lambda_{kl}}\left(\frac{w}{\lambda_{kl}}\right)^{\eta_{kl}-1} \\
&= K_{kl}(w)w^{\eta_{kl}-1}
\end{aligned} \tag{15}
$$

III. Hybrid Markov/Semi-Markov Epidemic Model

Both the semi-Markov and the Markov models have their positive and negative aspects to consider, primarily because the Markov model is more superficial. This makes it easier to interpret and comprehend conventional Markov models when simulating the spread of infectious diseases. Due to a broader distribution of sojourn times, the semi-Markov model simultaneously permits quantification of the time spent in a given state. This provides the context for the authors' development of a hybrid model that incorporates Markov and semi-Markov assumptions into a multistate model. In addition, the estimations contain a limited number of factors and are not too demanding in terms of the availability of data.

The steps for implementing the continuous time hybrid Markov/semi-Markov epidemic model are described as follows.

1) Use the chi-square test described by [22] to determine whether or not the transition from state k to state l can be considered a Markov transition. A Markov test must determine if the transition probability from the current state to the subsequent state is independent of the previous state. Written as $\Pr(m|l,k)$, which is the probability of transitioning from state k to state $(k \leq l \leq m)$, this is the probability of transitioning to state k. If the Markov property is satisfied, $p_{klm} = p_{lm}$.

   A test to determine whether a Markov chain is second order or not is defined as follows by [22]:

$$
\begin{aligned}
H_0 &: \quad p_{1lm} = p_{2lm} = \cdots = p_{klm} = p_{lm}, \\
&\quad\quad l, m = 1, 2, \ldots, d. \\
H_1 &: \quad \text{is a second-order Markov chain}
\end{aligned}
$$

   The following is the value of the chi-square test statistic for the null hypothesis:

$$
\chi^2 = \sum_{k,l,m} n_{kl}\frac{(\hat{p}_{klm} - \hat{p}_{lm})^2}{\hat{p}_{lm}} \tag{16}
$$

   where

$$
\widehat{p}_{klm} = \frac{n_{klm}}{\sum_{l=1}^d n_{klo}} \quad \text{and} \quad \widehat{p}_{lm} = \frac{\sum_{l=1}^d n_{klm}}{\sum_{k=1}^d \sum_{o=1}^d n_{klo}}
$$

   d is the number of states used. The degree of freedom used is $d(d-1)^2$. If $\chi^2 > \chi^2_{tab}$ then H_0 is rejected. It may be seen from this that the multi-state model does not conform to the Markov model. Hypothesis testing can also use the p-value. If p-value is less than the given significance level, then H_0 is rejected.

2) Examining whether or not the semi-Markov model holds true at the level of the distribution of sojourn

times h_{kl}. In continuous-time, utilizing the Equation (15), where $K_{kl}(w)$ represents all components of $\tilde{\mu}_{kl}(w)$ omitting $w_{kl}^{\eta}-1$. The form $K_{kl}(w)$ can be used to determine if the approximation of the transition intensity function is of the Weibull type. A Weibull-type representation will follow if $K_{kl}(w)$ is independent of w.

3) The transition from state k to l can be considered a Markov chain if the sojourn time of h_{kl} is exponentially distributed for continuous-time.
4) Finds the transition probability of each transition between states.

Theorem 1: If there is a susceptible vector at time w ($\mathbf{S}(w)$), an infected vector at time w ($\mathbf{I}(w)$), a cured vector at time w ($\mathbf{R}(w)$) and vectors that die at time w ($\mathbf{D}(w)$), The following is how the prediction equation for an individual who has been infected with an infectious disease at time $w + 1$ for a semi-Markov system can be generated.

$$
\begin{aligned}
\mathbf{I}(w+1) \quad = \quad & \sum_{k=1}^{m} n_{i\mathbf{I}}(w)p_{k\mathbf{I}} + \sum_{l=1}^{m}(n_{\mathbf{I}l}(w+1)p_{\mathbf{I}} - \\
& n_{\mathbf{I}l}(w+1)p_{\mathbf{I}l})
\end{aligned} \quad (17)
$$

Proof: The SIRD epidemic model includes four states. Therefore, it can be expressed as follows.

$$
\mathbf{S}(w) + \mathbf{I}(w) + \mathbf{R}(w) + \mathbf{D}(w) = \mathbf{N}
$$

A change from a susceptible condition will increase in the infected state. In the meantime, it will reduce if a shift to a condition of recuperation or death occurs. Both claims can be expressed as

$$
\begin{aligned}
\mathbf{I}(w+1) \quad = \quad & \mathbf{S}(w+1) - \mathbf{S}(w) + \mathbf{I}(w) + \mathbf{R}(w+1) + \\
& \mathbf{D}(w+1)
\end{aligned}
$$

Suppose we are aware of how many people were present in each state at time w. In that situation, the value of the Binomial random variable $n_{kl}(w)$ is determined by the parameters $n_k(w)$ and p_{kl}, which leads to the predicted value presented in the next paragraph.

$$
\bar{n}_{kl}(w) = \bar{n}_k(w-1)p_{kl}
$$

Similarly, the expected value of $n_{i,m+1}(w)$ reflects the number of people who make their way out of the diseased state, either by recovering or passing away.

$$
\begin{aligned}
\bar{n}_{k,m+1}(w) \quad &= \quad \bar{n}_k(w-1)v_{kl} \\
&= \quad \bar{n}_k(w-1)(r_{kl} + d_{kl})
\end{aligned}
$$

In the meantime, $n_{0l}(w+1)$ reflects the number of persons from vulnerable states who have recently become infected, and has a value of $S(w+1)s_l$ as its predicted value. The resulting equation for making predictions looks like this.

$$
n_l(w+1) = \sum_{k=1}^{d} n_{kl}(w) + n_{0l}(w+1) - v_{kl}(w+1) \quad (18)
$$

Given $R(w+1), D(w+1), n_{0l}(w+1)$ with expected values $R(w+1)r_l$ and $D(w+1)d_l$. The expected value from the

equation (18) is obtained.

$$
\begin{aligned}
\bar{n}_l(w+1) \quad = \quad & \sum_{k=1}^{d} \bar{n}_k(w)p_{kl} + S(w+1)s_l - \\
& R(w+1)r_l - D(w+1)d_l
\end{aligned} \quad (19)
$$

So that the SIRD epidemic model is obtained

$$
\begin{aligned}
n_{\mathbf{I}}(w+1) \quad = \quad & \sum_{k=1}^{d} n_{k\mathbf{I}}(w)p_{k\mathbf{I}} + \sum_{l=1}^{d}(n_{\mathbf{I}l}(w+1)p_{\mathbf{I}} - \\
& n_{\mathbf{I}l}(w+1)p_{\mathbf{I}l})
\end{aligned}
$$

$\blacksquare$

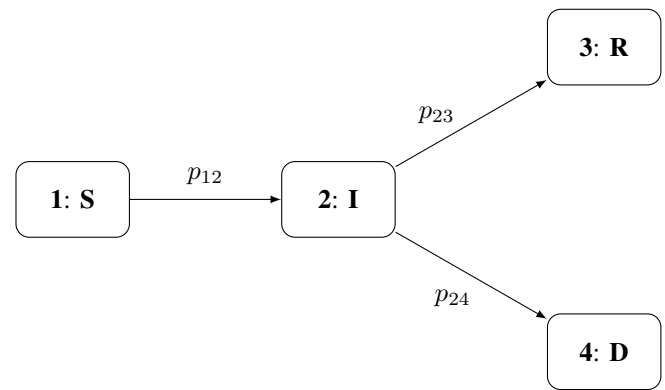## IV. Application of the Continuous-time SIRD Epidemic Model

Fig. 1. SIRD epidemic model.

The first step, use the following hypothesis in conjunction with Equation (16) to determine whether or not the SIRD epidemic model satisfies Markov characteristics.

$$
\begin{aligned}
H_0 \quad &: \quad p_{1lm} = p_{2lm} = p_{3lm} = p_{4lm}, \quad l, m = 1, 2, 3, 4. \\
H_1 \quad &: \quad \text{is a second-order Markov chain}
\end{aligned}
$$

H_0 states that the Markov property is satisfied. Under the null hypothesis, the test statistic χ^2 in this model with four states is distributed χ^2 with degrees of freedom $4(3)^2$ with $\alpha = 0.05$ significance level. The formula for calculating the value of χ^2 is as follows.

$$
\begin{aligned}
\chi^2 \quad &= \quad \sum_{k,l,m} n_{kl} \frac{(\hat{p}_{klm} - \hat{p}_{lm})^2}{\hat{p}_{lm}} \\
&= \quad n_{12}\frac{(\hat{p}_{123} - \hat{p}_{23})^2}{\hat{p}_{23}} + n_{12}\frac{(\hat{p}_{124} - \hat{p}_{24})^2}{\hat{p}_{24}} \\
&= \quad 102.48
\end{aligned}
$$

with $\chi^2_{tab} = \chi^2_{(0.05;36)} = 50.998$. Because the value of $\chi^2 > \chi^2_{tab}$ then H_0 is rejected. The conclusion that can be drawn from this is that the SIRD epidemic model does not meet the conditions set forth by Markov.
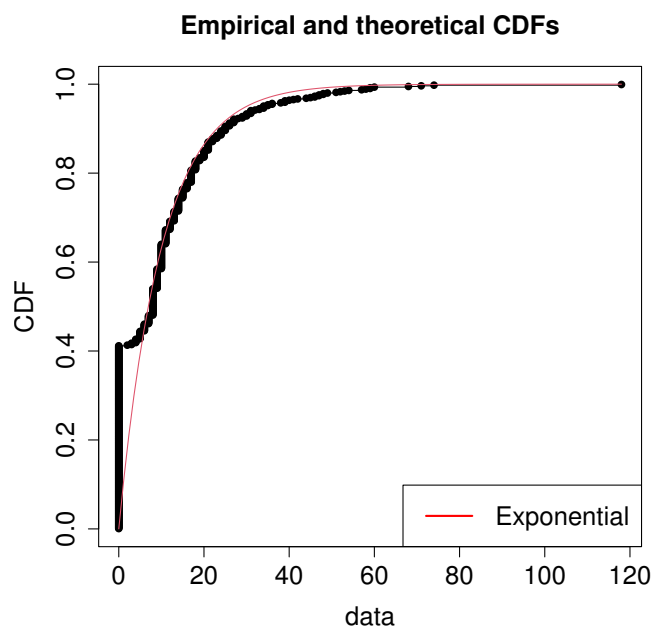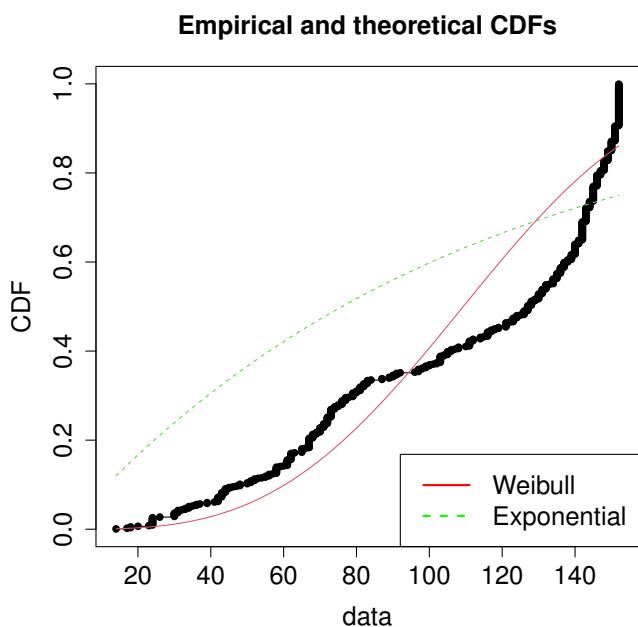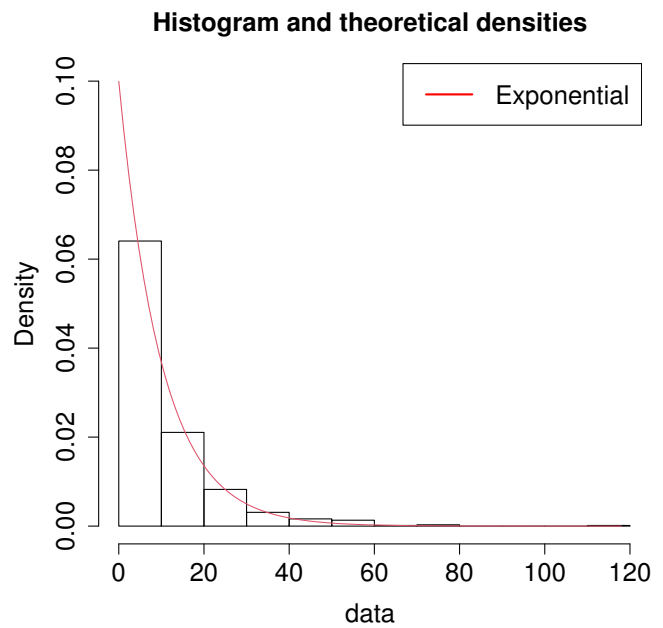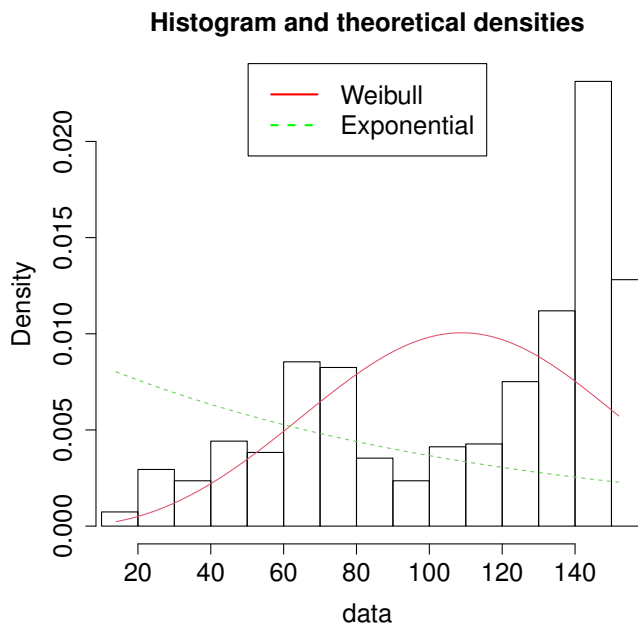
Histogram and theoretical densities

Histogram and theoretical densities

Empirical and theoretical CDFs

Empirical and theoretical CDFs

Fig. 2.   *Goodness-of-fit* plots for fitted distributions of probability of transition from a susceptible to infected state.

Fig. 3.   *Goodness-of-fit* plots for fitted distributions of probability of transition from infected to recovered state.

Second step, test every transition that occurs using the help of Software R, namely *package* "fitdistplus". The fitdistrplus package is a generic package that aims to help fit univariate parametric distributions with censored or uncensored data. The method used is the maximum likelihood estimation. Each transition is made a histogram plot, and then the suitability of the distribution used is seen and tested for the two distributions used in this study, the Exponential and Weibull distributions. Regarding the transition from susceptible to infected, the results in Figure 2 are obtained where the most suitable distribution is the Weibull distribution. As for the other transitions after testing, they all meet the Exponential distribution assumption.

TABLE I
ESTIMATION OF THE TRANSITION PROBABILITY OF EACH MODEL

Model	Transition	Distribution	$\widehat{p}_{ij}$
Markov	$1 \to 2$	Exponential	1.00
	$2 \to 3$	Exponential	0.96
	$2 \to 4$	Exponential	0.04
semi-Markov	$1 \to 2$	Weibull	1.00
	$2 \to 3$	Weibull	0.95
	$2 \to 4$	Weibull	0.05
hybrid Markov/semi-Markov	$1 \to 2$	Weibull	1.00
	$2 \to 3$	Exponential	0.96
	$2 \to 4$	Exponential	0.04

After getting the results of the distribution of each transition between states, the next step is to compare the Akaike

Histogram and theoretical densities

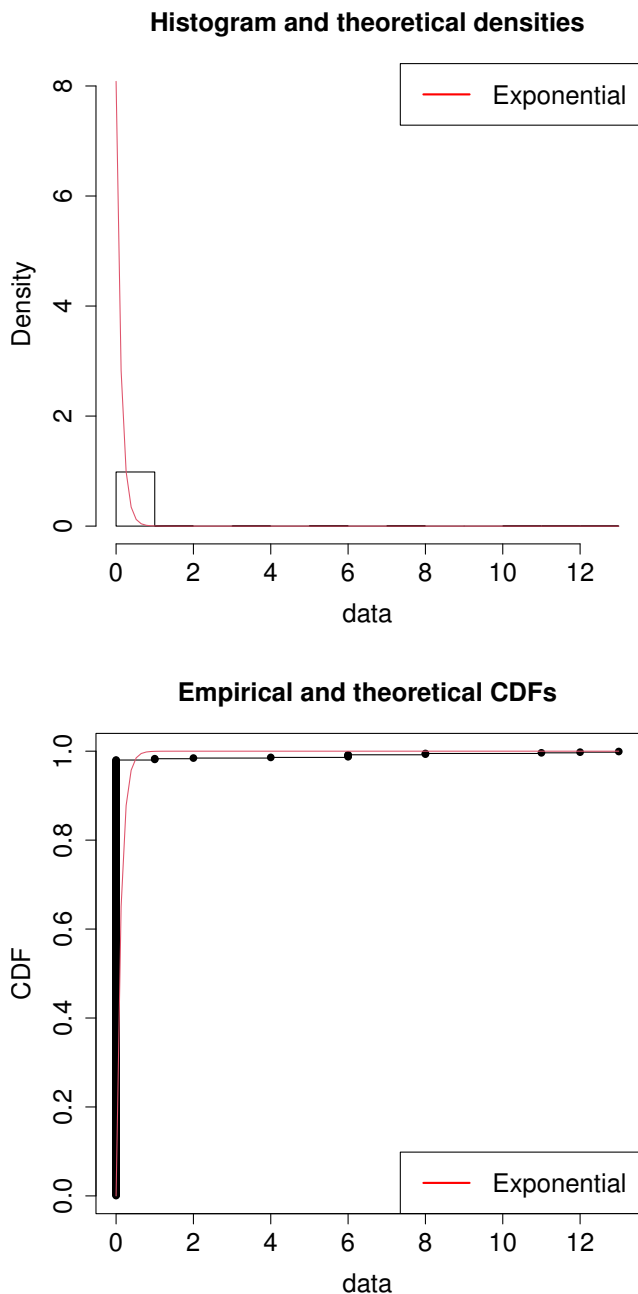## Empirical and theoretical CDFs

Fig. 4.   *Goodness-of-fit* plots for fitted distributions of probability of transition from infected to deceased state.

Information Criterion (AIC) values of each model used. Table II shows that the best model used to make short-term predictions is the hybrid Markov/semi-Markov model, which shows the smallest AIC value, where the sojourn time distribution used is the Weibull distribution and the Exponential distribution.

TABLE II
CONTINUOUS-TIME MODEL AIC VALUES

Model	Distribution	AIC
Markov	Exponential	13521.47
semi-Markov	Weibull	12703.31
Markov/semi-Markov hybrid	Exponential-Weibull	12475.31

$$
\begin{aligned}
\mathbf{I}(w+1) &= \sum_{k=1}^{d} n_{k\mathbf{I}}(w)p_{k\mathbf{I}} + \sum_{l=1}^{d} (n_{\mathbf{I}l}(w+1)p_{\mathbf{I}} - \\
&\quad n_{\mathbf{I}l}(w+1)p_{\mathbf{I}l}) \\
\mathbf{I}(140) &= n_{12}(139)p_{12} + n_{23}(140)p_{22} - n_{23}(140)p_{23} \\
&\quad + n_{24}(140)p_{22} - n_{24}(140)p_{24} \\
&= 27(1.00) + 11(0.32) - 11(0.65) + \\
&\quad 0(0.32) - 0(0.03) \\
&= 23
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbf{I}(w+1) &= \sum_{k=1}^{d} n_{k\mathbf{I}}(w)p_{k\mathbf{I}} + \sum_{l=1}^{d} (n_{\mathbf{I}l}(w+1)p_{\mathbf{I}} - \\
&\quad n_{\mathbf{I}l}(w+1)p_{\mathbf{I}l}) \\
\mathbf{I}(141) &= n_{12}(140)p_{12} + n_{23}(141)p_{22} - n_{23}(141)p_{23} \\
&\quad + n_{24}(141)p_{22} - n_{24}(141)p_{24} \\
&= 24(1.00) + 13(0.32) - 13(0.65) + \\
&\quad 0(0.32) - 0(0.03) \\
&= 20
\end{aligned}
$$

Thus, for a continuous time, the following predictions are obtained for $t = 140$, which corresponds to August 1, 2020, and $t = 141$, which corresponds to August 2, 2020. The forecast results for the next ten days are displayed in Table III.

TABLE III
PREDICTION RESULTS OF THE SIRD EPIDEMIC MODEL

Date	Actual	Prediction	MAPE (%)
2020/08/01	24	23	
2020/08/02	19	20	
2020/08/03	12	14	
2020/08/04	12	10	
2020/08/05	17	10	19.73
2020/08/06	18	15	
2020/08/07	19	16	
2020/08/08	16	17	
2020/08/09	8	12	
2020/08/10	9	7	

V. CONCLUSION

In this work, continuous time is incorporated into the SIRD epidemic model to predict the behavior of individuals who have a positive COVID-19 test. The semi-Markov and Markov assumptions are combined into a multi-state model to create the hybrid Markov/semi-Markov model. In order to get the correct inference for a multi-state model, we first need to do a chi-square test to determine whether or not the model fulfills the Markov assumptions. If it does, we may move on to the following processing step. If this is not the case, further assumptions in the form of a semi-Markov model are necessary.

In addition, tests are conducted to determine if each transition satisfies the Markov or semi-Markov assumptions, where it is determined that the sojourn time for the p_{12} transition follows a Weibull distribution. In contrast, the distribution of sojourn time for other transitions is Exponential. In the

final section, COVID-19 case predictions are made using a theorem based on the assumptions of a hybrid Markov/semi-Markov model, yielding a MAPE of $< 20\%$. The hybrid Markov/semi-Markov model simulates the transmission of infectious diseases more accurately than the Markov and semi-Markov models.

REFERENCES

[1] C. Jackson, "Multi-state modelling with R: the msm package," *Journal Of Statistical Software*, vol. 38, no. 8, pp. 1–28, 2011. [Online]. Available: ftp://218.193.60.3/software/rsoft/msm/doc/msm-manual.pdf

[2] C. H. Jackson and L. D. Sharples, "Hidden Markov models for the onset and progression of bronchiolitis obliterans syndrome in lung transplant recipients," *Statistics in Medicine*, vol. 21, no. 1, 2002.

[3] L. D. Sharples, "Use of the gibbs sampler to estimate transition rates between grades of coronary disease following cardiac transplantation," *Statistics in Medicine*, vol. 12, no. 12, pp. 1155–1169, 1993.

[4] J. H. Klotz and L. D. Sharpless, "Estimation for a Markov Heart Transplant Model," *The Statistician*, vol. 43, no. 3, 1994.

[5] R. Kay, "A Markov Model for Analysing Cancer Markers and Disease States in Survival Studies," *Biometrics*, vol. 42, no. 4, 1986.

[6] G. A. Satten and I. M. Longini, "Markov Chains With Measurement Error: Estimating the 'True' Course of a Marker of the Progression of Human Immunodeficiency Virus Disease," *Applied Statistics*, vol. 45, no. 3, 1996.

[7] C. Guihenneuc-Jouyaux, S. Richardson, and I. M. Longini, "Modeling markers of disease progression by a hidden Markov process: Application to characterizing CD4 cell decline," *Biometrics*, vol. 56, no. 3, 2000.

[8] G. Marshall and R. H. Jones, "Multi-state models and diabetic retinopathy," *Statistics in Medicine*, vol. 14, no. 18, 1995.

[9] P. K. Andersen, "Multistate models in survival analysis: A study of nephropathy and mortality in diabetes," *Statistics in Medicine*, vol. 7, no. 6, 1988.

[10] H. H. Chen, S. W. Duffy, and L. Tabar, "A Markov chain method to estimate the tumour progression rate from preclinical to clinical phase, sensitivity and positive predictive value for mammography in breast cancer screening," *Journal of the Royal Statistical Society Series D: The Statistician*, vol. 45, no. 3, 1996.

[11] F. Zuhairoh, D. Rosadi, and A. R. Effendie, "Determination of Basic Reproduction Numbers using Transition Intensities Multi-state SIRD Model for COVID-19 in Indonesia," *Journal of Physics: Conference Series*, vol. 1821, no. 1, p. 012050, 2021.

[12] L. E. Mohamed, M. Derouich, and A. Boutayeb, "A SIAR Model of COVID-19 with Control," *Engineering Letters*, vol. 30, no. 1, pp. 98–107, 2022.

[13] F. Zuhairoh, D. Rosadi, and A. R. Effendie, "Multi-state Discrete-time Markov Chain SVIRS Model on the Spread of COVID-19," *Engineering Letters*, vol. 30, no. 2, pp. 598–608, 2022.

[14] ——, "SVIRD Epidemic Model with Discrete-time Hybrid Semi-Markov," *Communications in Mathematical Biology and Neuroscience*, vol. 29, p. Article ID 29, 2023.

[15] ——, "Multi-state SIRD model with semi-Markov assumptions on the spread of COVID-19 disease based on sojourn time distribution," *AIP Conference Proceedings*, vol. 2614, no. 1, p. 040030, 06 2023. [Online]. Available: https://doi.org/10.1063/5.0125763

[16] F. Zuhairoh and D. Rosadi, "Real-time forecasting of the COVID-19 epidemic using the Richards model in South Sulawesi, Indonesia," *Indonesian Journal of Science and Technology*, vol. 5, no. 3, pp. 456–462, 2020.

[17] ——, "Data-driven Analysis and Prediction of COVID-19 Infection in Southeast Asia using A Phenomenological Model," *Pakistan Journal of Statistics and Operation Research*, vol. 18, no. 1, pp. 59–69, 2022.

[18] ——, "Real-time prediction for multi-wave COVID-19 outbreaks," *Communications for Statistical Applications and Methods*, vol. 29, no. 5, pp. 499–512, 2022.

[19] B. Wu, B. I. G. Maya, and N. Limnios, "Using Semi-Markov Chains to Solve Semi-Markov Processes," *Methodology and Computing in Applied Probability*, 2020.

[20] S. M. Ross, *Introduction to Probability Models*. United States of America: Elsevier Academic Press, 2007.

[21] A. Asanjarani, B. Liquet, and Y. Nazarathy, "Estimation of semi-Markov multi-state models: A comparison of the sojourn times and transition intensities approaches," *International Journal of Biostatistics*, vol. 18, no. 1, pp. 243–262, 2022.

[22] T. W. Anderson and L. A. Goodman, "Statistical inference about Markov chains," *The Annals of Mathematical Statistic*, vol. 28, no. 03, pp. pp. 89–110, 1957.