

Received 26 March 2024, accepted 24 April 2024, date of publication 30 April 2024, date of current version 7 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3395544

## RESEARCH ARTICLE

# A Stochastic Epidemiological SIRD-V Model With LSM-EKF Algorithm for Forecasting and Monitoring the Spread of COVID-19 Pandemic: Real Data

ABDENNOUR SEBBAGH<sup>1</sup>, CHEMESSE ENNEHAR BENCHERIE<sup>1</sup>, AND SIHEM KECHIDA

Laboratoire d'Automatique et Informatique de Guelma (LAIG), Université 8 Mai 1945 Guelma, Guelma 24000, Algeria

Corresponding author: Abdennour Sebbagh (sebbagh.abdennour@univ-guelma.dz)

**ABSTRACT** Predictive algorithms for the coronavirus epidemic are indispensable tools for monitoring the dynamic spread of COVID-19 and for implementing intervention and preparedness measures to mitigate the outbreak. Many of the existing mathematical models used for epidemic analysis are deterministic in nature, which may not fully capture the complex dynamics of disease transmission. In this paper, we introduce a novel stochastic predictive algorithm known as the LSM-EKF-SIRD-V algorithm. This algorithm combines a SIRD-V model, which accounts for susceptible, infected, recovered, deceased, and vaccinated cases, with the Least Square Method (LSM) and an Extended Kalman Filter (EKF). It provides daily dynamic predictions of the system's parameters and is employed to analyze the COVID-19 disease profile in Algeria from January 29, 2021, to October 2, 2022. The primary goal of this approach is to create a decision-support system that empowers governments and health authorities with future pandemic statistics. This information enables them to adapt and optimize hospitalization resources, allowing for more effective intervention and preparedness measures to control the spread of the pandemic. Simulation results demonstrate the effectiveness of the proposed algorithm in accurately predicting the future dynamics of coronavirus spread based on historical and current case data.

**INDEX TERMS** Coronavirus, COVID-19, SIRD-V model, extended Kalman filter, LSM-EKF-SIRD-V algorithm.

## I. INTRODUCTION

The coronavirus, or COVID-19, is an extremely contagious respiratory virus that originated in Hubei province, China [1], [2]. Within a very brief period, it spread to numerous countries worldwide, causing a significant impact on the affected regions. The World Health Organization (WHO) officially declared the COVID-19 epidemic a “public health emergency of international concern” on January 30, 2020, and subsequently escalated it to the status of a pandemic on March 11, 2020.

The associate editor coordinating the review of this manuscript and approving it for publication was Zhengmao Li<sup>1</sup>.

Since the onset of the COVID-19 pandemic, various efforts have been made to monitor the progression of the viral spread and forecast the course of this infectious disease. These endeavors share a common goal of providing crucial insights to inform the decisions of public health authorities. The approaches developed can be categorized based on the number of compartments studied and are broadly divided into two groups: those employing artificial intelligence tools [3], [4], [5] and those relying on classical epidemiological models.

The second category, which involves classical epidemiological models, stands out for the mathematical equations used. These equations may encompass differential equations,

difference equations, or even fractional calculus, depending on the specific requirements of modeling biological or medical phenomena [6], [7], [8].

Regarding the application of mathematical models, initial research efforts have primarily concentrated on epidemiological studies, focusing on the transmission dynamics and forecasting of the disease in the most heavily impacted countries across the globe. Many of these studies have employed various forms of mathematical models, often utilizing different compartments, such as susceptible-exposed-infectious-recovered-death (and even more compartments like protected and treated). These models are drawn from well-established frameworks found in the literature, including the widely recognized SIR, SIRD, SIIRD, PSIRD, and more. These models serve as tools for understanding and analyzing the spread and transmission dynamics of the disease [9], [10].

In subsequent phases of the research, these strategies have evolved to include the consideration of isolated or hospitalized populations within the modeling framework [11], [12].

More recently, researchers have delved into extended epidemic models that account for vaccinated individuals. In studies like [13], [14], and [15], authors examine the impact of vaccination campaigns and shed light on the development of the epidemic in various countries, with particular emphasis on developing nations. They use models such as the deterministic SEIR model or the SEIR model formulated with white noise and time delays. Additionally, Poonia et al. propose an enhanced SEIR model for predicting the severity of COVID-19 in vaccinated populations [16]. By combining the generalized SEIR model with the Wells–Riley model, the results obtained offer insights into the effectiveness of preventive strategies, including vaccination, face mask usage, and nucleic acid testing, particularly in the context of mass gatherings such as large sports events [17].

In the realm of epidemiological modeling, the basic SIR model with vaccination has gained widespread usage in the literature due to its simplicity. Numerous approaches have been developed, presenting different versions of the SIRV model to monitor transmission and predict the trends of COVID-19 and other infectious diseases [18], [19], [20].

In [20], an age-structured SIR model was employed to quantify the effects of vaccine hesitancy in the United States during the Delta wave of the pandemic.

A significant model, known as the SIRDV model, has proven to be efficient in the study of severe infectious diseases like Ebola and COVID-19 [7], [15]. This model incorporates both vaccination and deceased individuals in its formulation and is a crucial tool for analyzing epidemiological characteristics.

While many previous works on vaccination campaigns focused on a single dose, some recent approaches have taken multiple vaccinations into account, breaking down vaccinated individuals into different categories, such as vaccination1 (those who received the first dose) and vaccination2 (those

who received the second dose) [21], [22]. Similarly, the authors of [23] formulated the rollout of vaccination for two different vaccines as an optimization problem with the primary goal of achieving herd immunity rapidly.

As evident from the papers discussed, many issues are addressed using a deterministic framework of compartmental models. While deterministic models are advantageous for their simplicity and ease of analysis, they may not be the most appropriate choice when dealing with uncertainties or random variations, especially in cases where the phenomenon under study exhibits stochastic properties, as is often the case with the spread of diseases [24], [25].

This recognition of the stochastic nature of disease spread has prompted researchers to approach COVID-19-related issues from a stochastic perspective, employing various modeling techniques. These include regression models, agent-based models, and traditional models like SIR, SEIRD, etc. These models are often combined with a range of tools, such as the Kalman filter [26], [27] and its suboptimal estimators like the Ensemble Kalman Filter (EnKF) [28], [29], Unscented Kalman Filter (UKF), Extended Kalman Filter (EKF) [30], [31], [32], particle filters [33], Bayesian approaches [34], stochastic models [35]. These stochastic modeling techniques provide a more accurate representation of the inherent randomness and uncertainties in the spread of diseases like COVID-19.

The stochastic approaches have evolved to include the consideration of vaccinated individuals to better understand the dynamics of COVID-19 and assess the effects of vaccination.

The current research endeavors to investigate the influence of vaccination on the transmission dynamics of COVID-19. This study is a follow-up to a previously published work [30], which introduced the well-established concept of the Kalman filter used in target tracking to predict the spread of diseases. The goal here is to combine the Least Square Method (LSM) with the Extended Kalman Filter (EKF) within the framework of a SIRDV model. This combined approach aims to forecast and monitor the dynamics of compartmental classes in the COVID-19 pandemic in Algeria.

Up to this point, there have been only a handful of contributions discussing the epidemiological profile in Algeria, primarily focusing on modeling and prediction studies [30], [36], [37]. Notably, none of these prior works have considered the impact of vaccination.

In this context, we introduce an estimation approach based on the Extended Kalman Filter (EKF) for an extended SIRD model, which incorporates five distinct classes: susceptible, infected, recovered, deceased, and vaccinated. The primary objective of this paper is to delve into the transmission dynamics of COVID-19, using data spanning from January 29, 2021, to October 2, 2022, in the specific context of Algeria.

The contributions and innovations of this study can be summarized as follows:

1. The introduction of a stochastic aspect achieved by combining the Least Square Method (LSM) and the

Extended Kalman Filter (EKF) within the framework of the epidemiological SIRD-V model.

2. A daily optimization and prediction of various key parameters, including the infection rate  $\alpha(k)$ , recovery rate  $\beta(k)$ , death rate  $\gamma(k)$ , and vaccination rate  $\mu(k)$  is accomplished through the application of LSM.
3. Utilization of the epidemiological SIRD-V model in conjunction with the Extended Kalman Filter (EKF) to enable the simultaneous prediction of daily model states, daily transmission rates, and all parameters related to COVID-19.

These contributions provide a comprehensive framework for monitoring and forecasting the dynamics of COVID-19, considering both stochastic aspects and essential parameters in the context of Algeria.

The rest of this paper is organized into three sections. Section II is dedicated to firstly the problem formulation and the description of the chosen epidemiological model with the least square method (LSM), secondly, it details the data assimilation methods, more precisely the EKF algorithm used in the context of this work. In the next section, we discuss the application of this algorithm and present simulation results to verify this study. Finally, the last section recapitulates the concluding remarks of this study and suggests a purpose for future work.

## II. MATERIALS AND METHODS

### A. EPIDEMIOLOGICAL MODEL

In this paper, we present an extended version of the SIRD epidemiological compartment model, through, which we examine the spread of COVID-19 and account for daily and periodic intervention in terms of restrictions and vaccination campaigns in Algeria. This extended version is named the SIRD-V model and consists of the susceptible population  $S(t)$ , the infected population  $I(t)$ , the recovered population  $R(t)$ , the deceased population  $D(t)$  and the vaccinated population  $V(t)$ , such that:

$$S(t) + I(t) + R(t) + D(t) + V(t) = N \quad (1)$$

where  $N$  represents the total number of Algerian populations, the schematic of a SIRD-V model is shown in fig 1. and it is governed by the following set of nonlinear equations:

$$\frac{dS(t)}{dt} = -\frac{\alpha(t)}{N}S(t)I(t) - \mu(t)S(t) \quad (2)$$

$$\frac{dI(t)}{dt} = +\frac{\alpha(t)}{N}S(t)I(t) - \beta(t)I(t) - \gamma(t)I(t) \quad (3)$$

$$\frac{dR(t)}{dt} = \beta(t)I(t) \quad (4)$$

$$\frac{dD(t)}{dt} = \gamma(t)I(t) \quad (5)$$

$$\frac{dV(t)}{dt} = \mu(t)S(t) \quad (6)$$

$\alpha(t)$ ,  $\beta(t)$ ,  $\gamma(t)$  and  $\mu(t)$  are the daily infection, daily recovery, daily death and daily vaccination rates, which are optimized by the least square method (LSM) as follow:

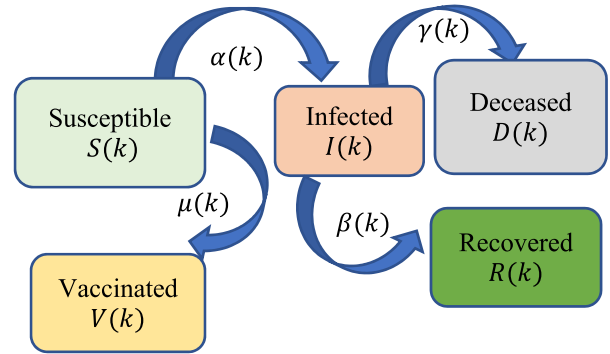


FIGURE 1. Scheme of SIRD-V epidemiological model.

In discrete time and if we suppose that  $S = N$ , then:

$$\alpha(k) = \frac{\sum_{j=1}^k I(j) \cdot \Delta I(j) + \sum_{j=1}^k I(j) \cdot \Delta R(j) + \sum_{j=1}^k I(j) \cdot \Delta D(j)}{\sum_{j=1}^k I^2(j)} \quad (7)$$

If  $S \neq N$ , then (8)–(11), as shown at the bottom of the next page:

$I(j)$  is the total currently infected at time  $j$  (day)

$S(j)$  is the susceptible population at time  $j$  (day)

$\Delta I(j) = I(j) - I(j-1)$  is daily currently infected at time  $j$  (day)

$\Delta R(j) = R(j) - R(j-1)$  is daily new recovered at time  $j$  (day)

$\Delta D(j) = D(j) - D(j-1)$  is daily new deceased at time  $j$  (day)

$\Delta V(j) = V(j) - V(j-1)$  is daily new vaccinated at time  $j$  (day)

The nonlinear discrete SIRD-V model is given by:

$$X(k+1) = \begin{pmatrix} x_1(k) - \frac{\alpha(k)}{N}x_1(k) \cdot x_2(k) - \mu(k) \cdot x_1(k) \\ x_2(k) + \frac{\alpha(k)}{N}x_1(k) \cdot x_2(k) - \beta(k)x_2(k) - \gamma(k)x_2(k) \\ x_3(k) + \beta(k)x_2(k) \\ x_4(k) + \gamma(k)x_2(k) \\ x_5(k) + \mu(k)x_1(k) \end{pmatrix} + V_k \quad (12)$$

where

$$X(k) = \begin{pmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \\ x_5(k) \end{pmatrix} = \begin{pmatrix} S(k) \\ I(k) \\ R(k) \\ D(k) \\ V(k) \end{pmatrix}$$

$V_k$  is a zero-mean white noise with covariance  $Q_V$ .

The Jacobian matrix of this model is obtained as (13) and (14), as shown at the bottom of the next page, where  $\hat{\alpha}(k)$ ,  $\hat{\beta}(k)$  and  $\hat{\gamma}(k)$  are given in [30].

We suppose that the measurement equation is given daily by:

$$Y_{k+1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ x_4(k+1) \\ x_5(k+1) \end{pmatrix} + W_k$$

$$Y_{k+1} = CX_{k+1} + W_k \tag{15}$$

With

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$W_k$  is a zero-mean white noise with covariance  $\sum_W$ .

**B. EXTENDED KALMAN FILTER**

Kalman Filter can be used in estimation of the state  $X_k \in \mathbb{R}^{n_x}$  where posterior PDF is Gaussian in every time step and the process is governed by linear difference equation. But in many cases this PDF is Not-Gaussian and we need to use different approach such as an approximate grid-based method or extended Kalman Filter. These methods are also labeled as sub-optimal algorithms [38], [39].

Again, let  $X_k \in \mathbb{R}^{n_x}$  be the state sequence, but in opposite to the previous case, the process is governed by the nonlinear

difference equation.

$$X_{k+1} = f(X_k) + V_k \tag{16}$$

with measurement  $Y_k \in \mathbb{R}^{n_y}$ :

$$Y_{k+1} = h(X_{k+1}) + W_k \tag{17}$$

where  $V_k$  and  $W_k$  represent process and measurement noise vectors with a zero mean and covariances  $Q_k$  and  $\sum_k$  respectively. Function  $f$  can be used to compute a state in time step  $k$  from the previous estimate and function  $h$  can be used to compute the predicted measurement from the predicted state.

Extended Kalman Filter is based upon an approximation of the Bayes' rule using linearization. Again, as Kalman Filter, its extended version works also in two phases: prediction and update. A predict stage can be described using following equations:

$$\hat{X}_{k+1|k} = f(\hat{X}_{k|k}) + V_k \tag{18}$$

where  $\hat{X}_{k+1|k}$  is the predicted state estimate at time  $k+1$  given measurements up to time  $k$  and

$$P_{k+1|k} = \hat{F}_k P_{k|k} \hat{F}_k^T + Q_k \tag{19}$$

where  $P_{k+1|k}$  is the error covariance matrix

$$\hat{Y}_{k+1} = h(\hat{X}_{k+1|k}) + W_k \tag{20}$$

$$\alpha(k) = N \cdot \frac{\left[ \sum_{j=1}^k I^2(j) \sum_{j=1}^k (S(j) \cdot I(j) \cdot \Delta I(j)) \right]}{\sum_{j=1}^k I^2(j) \sum_{j=1}^k (I^2(j) \cdot S^2(j))} + N \frac{\left( \sum_{j=1}^k I^2(j) \cdot S(j) \right) \left( \sum_{j=1}^k (I(j) \cdot \Delta R(j)) + \sum_{j=1}^k (I(j) \cdot \Delta D(j)) \right)}{\sum_{j=1}^k I^2(j) \sum_{j=1}^k (I^2(j) \cdot S^2(j))} \tag{8}$$

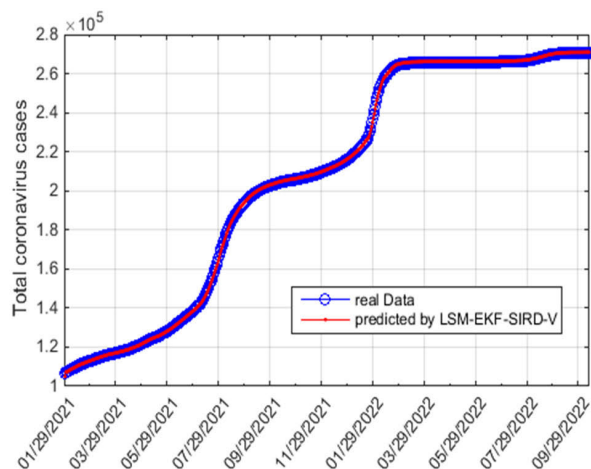
$$\beta(k) = \frac{\sum_{j=1}^k I(j) \cdot \Delta R(j)}{\sum_{j=1}^k I^2(j)} \tag{9}$$

$$\gamma(k) = \frac{\sum_{j=1}^k I(j) \cdot \Delta D(j)}{\sum_{j=1}^k I^2(j)} \tag{10}$$

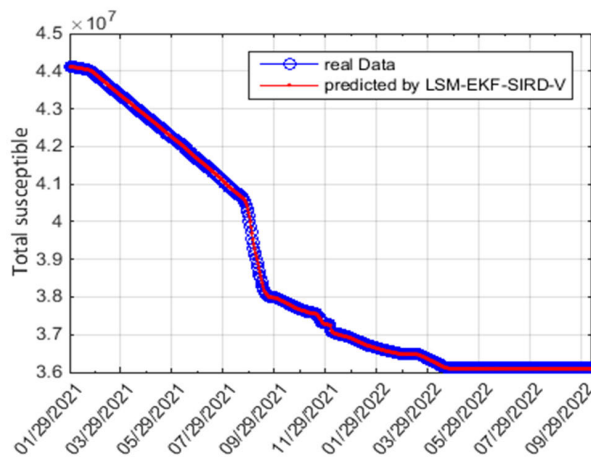
$$\mu(k) = \frac{\sum_{j=1}^k S(j) \cdot \Delta V(j)}{\sum_{j=1}^k S^2(j)} \tag{11}$$

$$F(X_k) = \frac{\partial f}{\partial X_k} = \begin{pmatrix} 1 - \frac{\hat{\alpha}(k)}{N} x_2(k) - \hat{\mu}(k) & -\frac{\hat{\alpha}(k)}{N} x_1(k) & 0 & 0 & 0 \\ \frac{\hat{\alpha}(k)}{N} x_2(k) & 1 - \hat{\beta}(k) - \hat{\gamma}(k) + \frac{\hat{\alpha}(k)}{N} x_1(k) & 0 & 0 & 0 \\ 0 & \hat{\beta}(k) & 1 & 0 & 0 \\ 0 & \hat{\gamma}(k) & 0 & 1 & 0 \\ \hat{\mu}(k) & 0 & 0 & 0 & 1 \end{pmatrix} \tag{13}$$

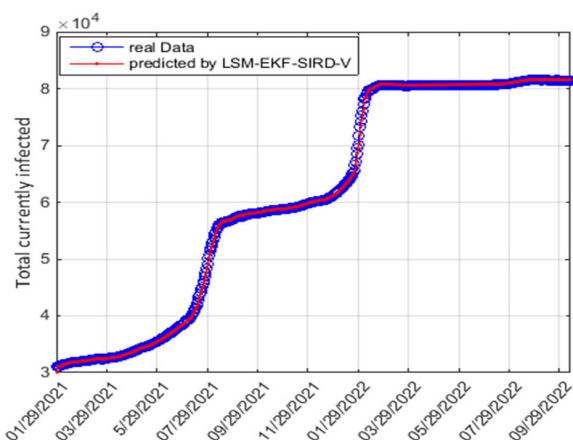
$$\hat{\mu}(k) = \frac{\text{predicted daily new vaccinated}}{\text{estimate of total susceptible}} = \frac{x_5(k+1/k) - x_5(k/k-1)}{x_1(k/k-1)} \tag{14}$$



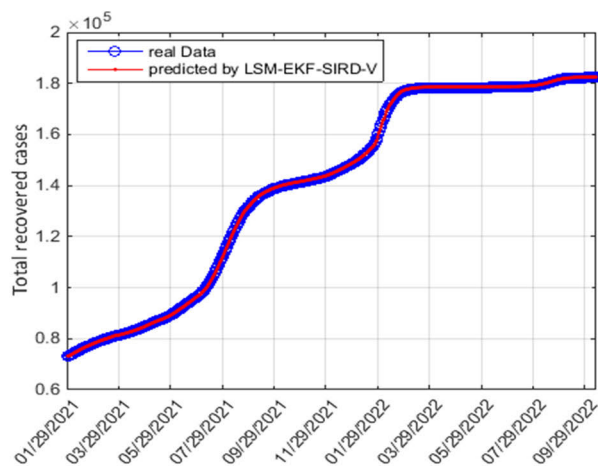
(a) Time step  $k$  (day)



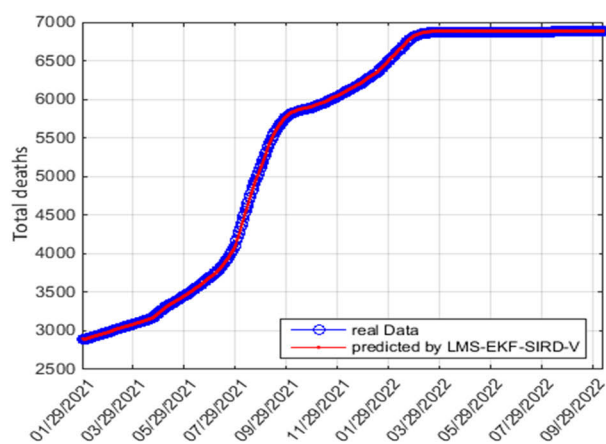
(b) Time step  $k$  (day)



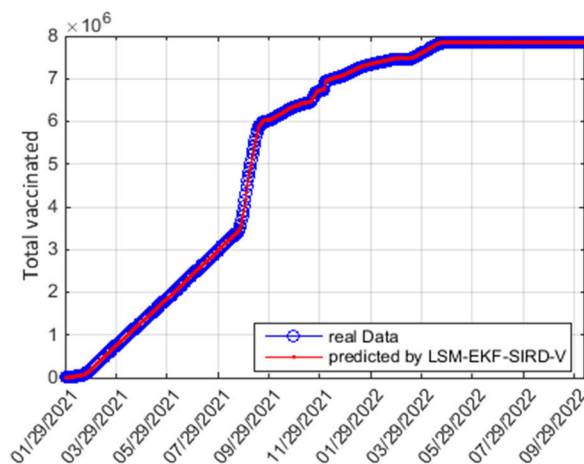
(c) Time step  $k$  (day)



(d) Time step  $k$  (day)

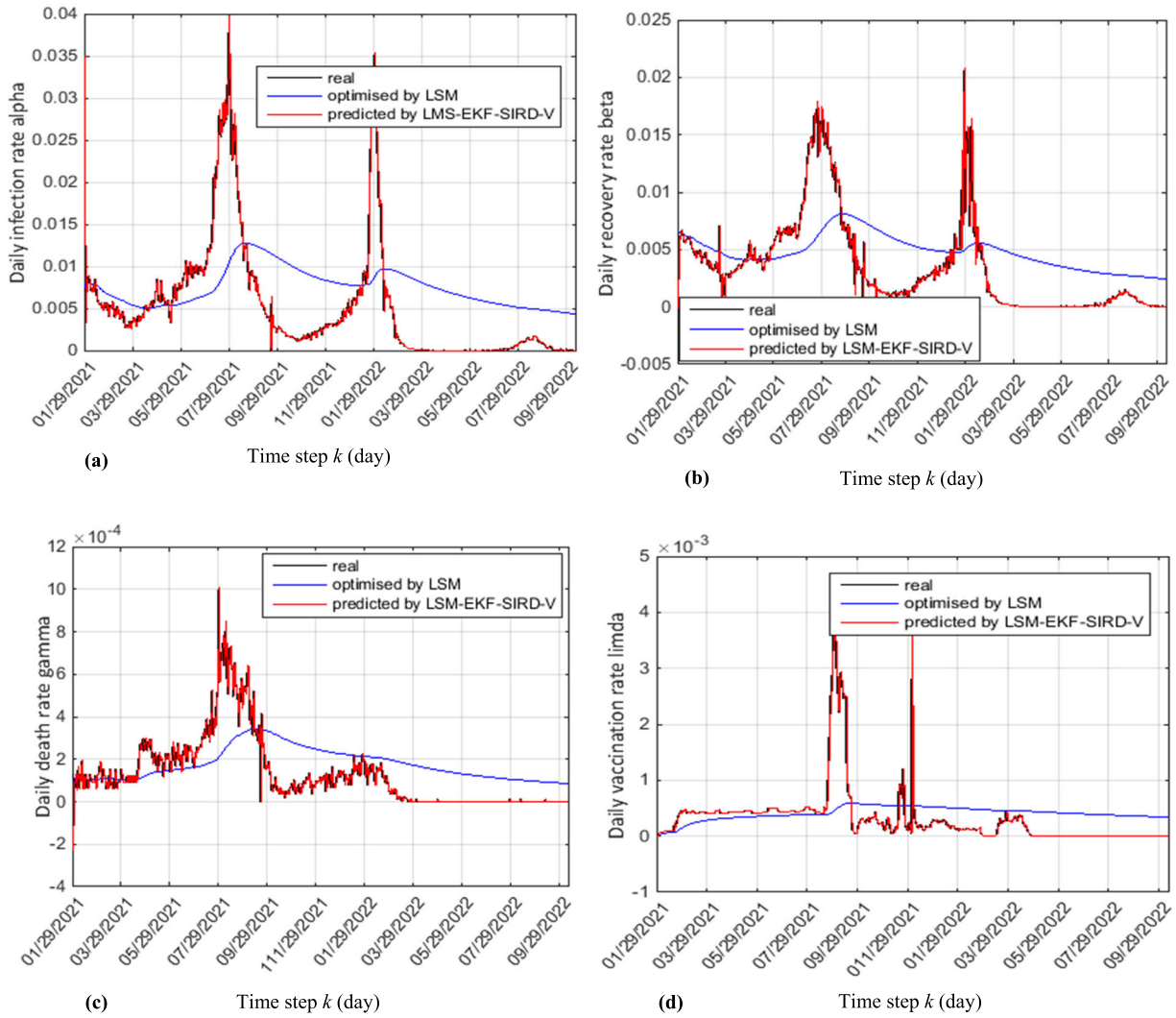


(e) Time step  $k$  (day)



(f) Time step  $k$  (day)

**FIGURE 2.** Real and predicted trajectories, (a) total coronavirus cases, (b) total susceptible, (c) total currently infected, (d) total recovered cases, (e) total deceased and (f) total vaccinated.



**FIGURE 3.** Real and predicted trajectories, (a) daily infection rate  $\alpha(k)$ , (b) daily recovery rate  $\beta(k)$ , (c) daily death rate  $\gamma(k)$  and (d) daily vaccination rate  $\mu(k)$ .

$\hat{Y}_{k+1}$  is the predicted measurement at time  $k + 1$  state transition and observation matrices are defined by the following

$$\hat{F}_k = \left. \frac{\partial f}{\partial X} \right|_{X_{k|k}} \quad (21)$$

$$\hat{H}_{k+1} = \left. \frac{\partial h}{\partial X} \right|_{X_{k+1|k}} \quad (22)$$

and  $S_{k+1}$  is the innovation covariance defined by:

$$S_{k+1} = \hat{H}_{k+1} P_{k+1|k} \hat{H}_{k+1}^T + \sum_k \quad (23)$$

$$K_{k+1} = P_{k+1|k} \hat{H}_{k+1}^T S_{k+1}^{-1} \quad (24)$$

where  $K_{k+1}$  is the Kalman gain,

Update stage can be described with the following equations:

$$\tilde{y}_{k+1} = Y_{k+1} - \hat{Y}_{k+1} \quad (25)$$

where  $\tilde{y}_{k+1}$  is innovation term,

$$\hat{X}_{k+1|k+1} = \hat{X}_{k+1|k} + K_{k+1} \tilde{y}_{k+1} \quad (26)$$

is update state estimate and

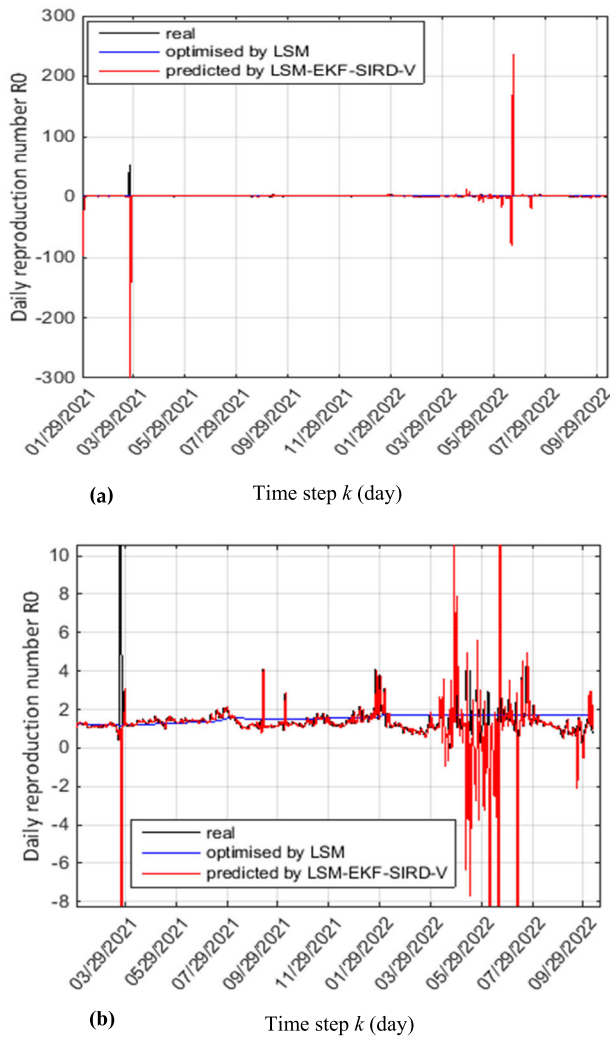
$$P_{k+1|k+1} = (I - K_{k+1} \hat{H}_{k+1}) P_{k+1|k} \quad (27)$$

is update estimate covariance.

Equations for the extended Kalman filter shown above utilize first term in a Taylor expansion of the non-linear function. Utilizing higher order terms is possible, but computational complexity prohibited their use.

### III. RESULTS AND DISCUSSION

Algeria was among the early adopters in the WHO Africa region to kickstart its COVID-19 vaccination campaign. The government has allocated substantial human and material resources to provide its population with safe and effective vaccines. The aim is to mitigate the spread of the



**FIGURE 4.** Real and predicted trajectories, (a) daily basic reproduction number  $R_0$  and (b) daily basic reproduction number (zoom).

virus across the country, including all its cities (Wilayas). These vaccination efforts are crucial in the fight against COVID-19 and in achieving herd immunity to protect public health.

The initial phase of the vaccination campaign in Algeria was initiated on January 30, 2021, starting in the city (Wilaya) of Blida. Blida was the first national hotspot of the pandemic in Algeria. During this phase, the Russian Vaccine Sputnik-V was administered to the population as part of the vaccination campaign. This marked the beginning of a critical effort to protect individuals and communities from the spread of COVID-19.

As of September 4th, 2022, Algeria had administered three additional types of vaccines, namely AstraZeneca, Sinovac, and Sinopharm. The total number of vaccine doses administered in the country had reached approximately 15.27 million. This widespread vaccination effort aims to provide immunity and protection against COVID-19 for a significant portion of the population.

To implement the LSM-EKF-SIRD-V algorithm for predicting the dynamics of COVID-19 spread with vaccination, we utilize real data obtained from the Algerian Health Ministry and the World Health Organization (WHO). This data spans from January 29, 2021, to October 02, 2022, and serves as the foundation for our analysis and predictions regarding the transmission of the coronavirus in Algeria.

The real initial vector of coronavirus is given by:

$$X(1) = (44112498 \ 31047 \ 72956 \ 28440)^T$$

The mean vector and covariance matrix initialization of the EKF according to a Gaussian law are:

$$\hat{X}(1|1) = (44113000 \ 31100 \ 73050 \ 2890 \ 0)^T$$

$$P_{1|1} = \begin{pmatrix} 10^{+3} & 0 & 0 & 0 & 0 \\ 0 & 10^{+3} & 0 & 0 & 0 \\ 0 & 0 & 10^{+3} & 0 & 0 \\ 0 & 0 & 0 & 10^{+3} & 0 \\ 0 & 0 & 0 & 0 & 10^{+3} \end{pmatrix}$$

The process noise is zero mean, white with covariance

$$Q_k = \begin{pmatrix} 10^{+3} & 0 & 0 & 0 & 0 \\ 0 & 10^{+3} & 0 & 0 & 0 \\ 0 & 0 & 10^{+3} & 0 & 0 \\ 0 & 0 & 0 & 10^{+3} & 0 \\ 0 & 0 & 0 & 0 & 10^{+3} \end{pmatrix}$$

The measurement noise is also zero mean, white, independent of the process noise, with covariance

$$\Sigma_k = \begin{pmatrix} 10^{-2} & 0 & 0 & 0 & 0 \\ 0 & 10^{-2} & 0 & 0 & 0 \\ 0 & 0 & 10^{-2} & 0 & 0 \\ 0 & 0 & 0 & 10^{-2} & 0 \\ 0 & 0 & 0 & 0 & 10^{-2} \end{pmatrix}$$

Fig 2: (a), (b), (c), (d), (e) and (f) shows the real and predicted trajectories of total coronavirus cases, total susceptible, total currently infected, total recovered, total deceased and total vaccinated respectively.

It's clear that the real and predicted trajectories are almost identical which means that our LSM-EKF-SIRD-V algorithm is correctly predicted the daily evolution of coronavirus spread.

$\alpha(k), \beta(k), \gamma(k)$  and  $\mu(k)$  are the daily infection, daily recovery, daily death and daily vaccination rates respectively, and are optimized firstly by the least square method (LSM) according to (7), (8), (9), (10) and (11) and secondly predicted by LSM-EKF-SIRD-V algorithm as shown in Figure 3: a, b, c and d.

To predict correctly SIRD-V model states and all coronavirus parameters, it must predict correctly  $\alpha(k), \beta(k), \gamma(k)$  and  $\mu(k)$  rates, good predicted rates will certainly give us good predicted states

Fig 3. demonstrates the LSM-EKF-SIRD-V algorithm's ability to make accurate daily predictions of key rates related to COVID-19.

**Algorithm 1** LSM-EKF-SIRD-V Algorithm

• **Initialization:**

$\hat{X}_{1|1} = E(\tilde{X}_1)$ ,  $P_{1|1} = E(\tilde{X}_1 - \hat{X}_{1|1})(\tilde{X}_1 - \hat{X}_{1|1})^T$ ,  $(\alpha(1), \beta(1), \gamma(1) \text{ and } \mu(1))$   
 for  $k = 1 : \text{enddo}$

• **prediction step:**

- $\hat{X}_{k+1|k} = f(\hat{X}_{k|k}) + V_k \rightarrow \text{SIRD-V model predicted state}$
- Calculate the predicted  $(\hat{\alpha}(k), \hat{\beta}(k), \hat{\gamma}(k) \text{ and } \hat{\mu}(k))$
- Calculate  $\hat{F}_k = \left. \frac{\partial f}{\partial X} \right|_{X_{k|k}}$  using the predicted  $(\hat{\alpha}(k), \hat{\beta}(k), \hat{\gamma}(k) \text{ and } \hat{\mu}(k))$
- $P_{k+1|k} = \hat{F}_k P_{k+1|k} \hat{F}_k^T + Q_k \rightarrow \text{predicted covariance}$
- $\hat{Y}_{k+1} = h(\hat{X}_{k+1|k}) + W_k \rightarrow \text{predicted measurement}$
- $S_{k+1} = \hat{H}_{k+1} P_{k+1|k} \hat{H}_{k+1}^T + \sum_k$  where  $\hat{H}_{k+1} = \left. \frac{\partial h}{\partial X} \right|_{X_{k+1|k}}$

• **Update step:**

when a measurement has arrived:  $(Y_{k+1})$

- $\tilde{y}_{k+1} = Y_{k+1} - \hat{Y}_{k+1}$  innovation term
- $\hat{X}_{k+1|k+1} = \hat{X}_{k+1|k} + K_{k+1} \tilde{y}_{k+1}$  SIRD-V model estimate state
- $P_{k+1|k+1} = (I - K_{k+1} \hat{H}_{k+1}) P_{k+1|k}$  estimate covariance
- Optimization of  $(\alpha(k+1), \beta(k+1), \gamma(k+1) \text{ and } \mu(k+1))$  by LSM from real DATA for using in the next step

$$\alpha(k+1) = N \cdot \frac{\left[ \sum_{j=1}^{k+1} I^2(j) \sum_{j=1}^{k+1} (S(j) \cdot I(j) \cdot \Delta I(j)) \right] + \left( \sum_{j=1}^{k+1} I^2(j) \cdot S(j) \right) \left( \sum_{j=1}^{k+1} (I(j) \cdot \Delta R(j)) + \sum_{j=1}^{k+1} (I(j) \cdot \Delta D(j)) \right)}{\sum_{j=1}^{k+1} I^2(j) \sum_{j=1}^{k+1} (I^2(j) \cdot S^2(j))}$$

$$\beta(k+1) = \frac{\sum_{j=1}^{k+1} I(j) \cdot \Delta R(j)}{\sum_{j=1}^{k+1} I^2(j)}$$

$$\gamma(k+1) = \frac{\sum_{j=1}^{k+1} I(j) \cdot \Delta D(j)}{\sum_{j=1}^{k+1} I^2(j)}$$

$$\mu(k+1) = \frac{\sum_{j=1}^{k+1} S(j) \cdot \Delta V(j)}{\sum_{j=1}^{k+1} S^2(j)}$$

**End**

Fig 4. reveals the algorithm’s capacity to predict the daily basic reproduction number  $R_0$ . This prediction is based on the daily values of  $\alpha(k)$ ,  $\beta(k)$ , and  $\gamma(k)$  and is a valuable metric for assessing the transmission dynamics and potential impact of control measures on the spread of the virus, given by the equation:

$$R_0(k) = \frac{\alpha(k)}{\beta(k) + \gamma(k)} \tag{28}$$

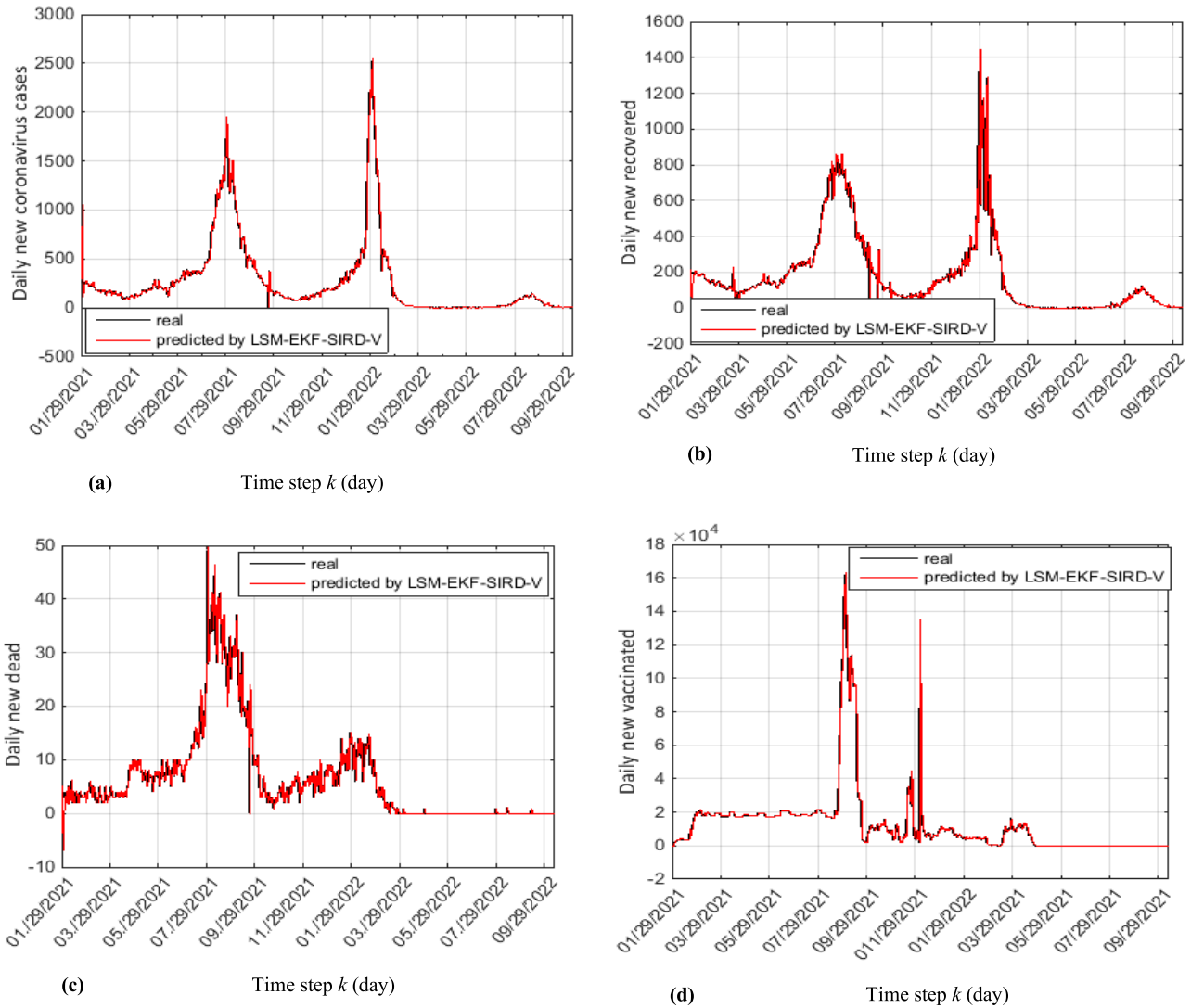
From January 29, 2021, to July 29, 2021, we observed a gradual increase in the reproduction number ( $R_0$ ) from 1 to 2. However, from July 29, 2021, to April 15, 2022, this number began to decrease. The reduction in the reproduction number during the latter period can be attributed to a combination of containment measures taken by the government. These measures included lockdowns, traffic restrictions, contact tracing, mandatory face masks in public spaces, and most significantly, the vaccination campaign. It is evident

that the vaccination operation played a crucial role in reducing the rate of infection, contributing to the overall control of the pandemic. This underscores the importance of vaccination in managing and mitigating the spread of COVID-19.

Indeed, when the value of the reproduction number  $R_0$  consistently falls between 1 and 2 over the entire observed range, it suggests that the situation was not particularly alarming.  $R_0$  values in this range indicate that the spread of the virus is relatively controlled and not experiencing rapid exponential growth.

This suggests that the implemented measures, including containment strategies and vaccination, have been effective in managing the COVID-19 situation and keeping it within manageable limits. However, it’s important to remain vigilant and continue these measures to prevent potential resurgences and ensure public health and safety.





**FIGURE 5.** Real and predicted trajectories, (a) daily new coronavirus cases, (b) daily new recovered, (c) daily new dead and (d) daily new vaccinated.

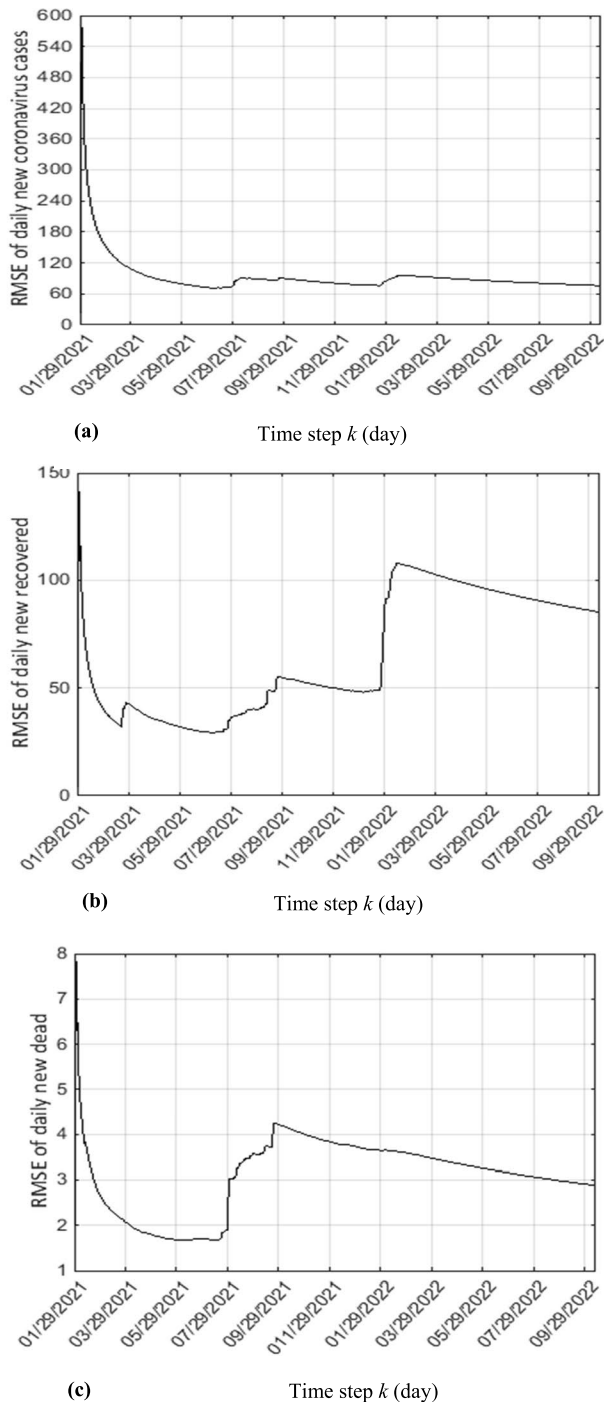
The real and predicted trajectories of daily new coronavirus cases, daily new deceased, daily new recovered, and daily new vaccinated, as presented in fig 5: (a), (b), (c), and (d), illustrate that the proposed algorithm has accurately predicted these daily new quantities. These results reflect the effectiveness and accuracy of our algorithm in forecasting and modeling the dynamics of COVID-19, including the impact of vaccination. Such precise predictions are valuable for public health decision-makers and authorities in managing the pandemic effectively.

The Root Mean Square Errors (RMSEs) illustrated in fig 6: a, b and c are derived from 100 Monte Carlo runs. These runs are conducted based on a specific equation (29) that we are likely referring to within our research or analysis. The RMSE is a common measure of the accuracy of a model’s predictions

compared to the actual data, and conducting multiple Monte Carlo runs helps assess the performance and robustness of a predictive algorithm.

$$RMSE(x(j)) = \sqrt{\frac{\sum_{k=1}^j (x_{real}(k) - x_{predicted}(k))^2}{j}} \quad j = 1 \text{ to } N \quad (29)$$

The quality of the prediction results is evidenced by the smaller Root Mean Square Errors (RMSEs) as illustrated in fig 6: (a), (b), and (c). Smaller RMSE values indicate that the predictions made by this algorithm closely align with the actual data, highlighting the effectiveness and accuracy of this predictive approach in modeling and forecasting COVID-19 dynamics with vaccination.



**FIGURE 6.** RMS error of: (a) daily new coronavirus cases, (b) daily new recovered and (c) daily new dead.

#### IV. CONCLUSION

This study, which spans from January 29, 2021, to October 2, 2022, aims to provide valuable insights to the government and health authorities. By investigating the outbreak of COVID-19 in Algeria during this period, we are contributing essential information that can inform and guide future measures and decisions for managing and controlling the

coronavirus pandemic in the country. This research plays a vital role in helping authorities make data-driven and informed choices to protect public health and mitigate the impact of the virus.

This paper introduces a novel predictive algorithm designed to monitor and manage the dynamics of coronavirus spread. The algorithm is a combination of the Least Square Method (LSM) and the Extended Kalman Filter (EKF), incorporating the epidemiological SIRD-V model. The algorithm brings several novelties to the field:

1. Daily optimization of critical rates: The algorithm optimizes daily values for key rates, including the infection rate ( $\alpha(k)$ ), recovery rate ( $\beta(k)$ ), death rate ( $\gamma(k)$ ), and vaccination rate ( $\mu(k)$ ) using the LSM.
2. Comprehensive predictions: Leveraging the epidemiological SIRD-V model and EKF, the algorithm simultaneously forecasts daily model states, daily transmission rates, and all parameters related to COVID-19. This approach treats the pandemic as a dynamic, evolving entity, allowing for more accurate predictions and control measures.

These innovations offer a robust tool for tracking and controlling the dynamics of coronavirus spread and enable more effective decision-making for public health authorities.

The simulation results indicate that the proposed algorithm is highly effective in predicting the parameters related to the spread of the coronavirus. This predictive capability equips health authorities and policy-makers with valuable insights into the future statistics of the pandemic. By providing accurate forecasts of COVID-19 dynamics, the algorithm offers an essential tool for making informed decisions and implementing effective measures to control and manage the ongoing pandemic.

In future work, we propose to combine this algorithm with an artificial intelligence tool to obtain more precise results.

#### REFERENCES

- [1] C. Huang et al., "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China," *Lancet*, vol. 395, no. 10223, pp. 497–506, Feb. 2020, doi: [10.1016/s0140-6736\(20\)30183-5](https://doi.org/10.1016/s0140-6736(20)30183-5).
- [2] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G. F. Gao, and W. Tan, "A novel coronavirus from patients with pneumonia in China, 2019," *New England J. Med.*, vol. 382, no. 8, pp. 727–733, Feb. 2020, doi: [10.1056/nejmoa2001017](https://doi.org/10.1056/nejmoa2001017).
- [3] P. Wang, X. Zheng, J. Li, and B. Zhu, "Prediction of epidemic trends in COVID-19 with logistic model and machine learning technics," *Chaos, Solitons Fractals*, vol. 139, Oct. 2020, Art. no. 110058, doi: [10.1016/j.chaos.2020.110058](https://doi.org/10.1016/j.chaos.2020.110058).
- [4] A. B. Payedimarri, D. Concina, L. Portinale, M. Canonico, D. Seys, K. Vanhaecht, and M. Panella, "Prediction models for public health containment measures on COVID-19 using artificial intelligence and machine learning: A systematic review," *Int. J. Environ. Res. Public Health*, vol. 18, no. 9, p. 4499, Apr. 2021, doi: [10.3390/ijerph18094499](https://doi.org/10.3390/ijerph18094499).
- [5] A. Bousquet, W. H. Conrad, S. O. Sadat, N. Vardanyan, and Y. Hong, "Deep learning forecasting using time-varying parameters of the SIRD model for COVID-19," *Sci. Rep.*, vol. 12, no. 1, Feb. 2022, doi: [10.1038/s41598-022-06992-0](https://doi.org/10.1038/s41598-022-06992-0).
- [6] R. A. Singh, R. Lal, and R. R. Kotti, "Time-discrete SIR model for COVID-19 in Fiji," *Epidemiology Infection*, vol. 150, p. e75, Jul. 2022, doi: [10.1017/s0950268822000590](https://doi.org/10.1017/s0950268822000590).

- [7] T. W. Tulu, B. Tian, and Z. Wu, "Modeling the effect of quarantine and vaccination on ebola disease," *Adv. Difference Equ.*, vol. 2017, no. 1, p. 178, Dec. 2017, doi: [10.1186/s13662-017-1225-z](https://doi.org/10.1186/s13662-017-1225-z).
- [8] A. S. Khoojine, M. Mahsuli, M. Shadabfar, V. R. Hosseini, and H. Kordestani, "A proposed fractional dynamic system and Monte Carlo-based back analysis for simulating the spreading profile of COVID-19," *Eur. Phys. J. Special Topics*, vol. 231, nos. 18–20, pp. 3427–3437, Dec. 2022, doi: [10.1140/epjs/s11734-022-00538-1](https://doi.org/10.1140/epjs/s11734-022-00538-1).
- [9] K. Sarkar, S. Khajanchi, and J. J. Nieto, "Modeling and forecasting the COVID-19 pandemic in India," *Chaos, Solitons Fractals*, vol. 139, Oct. 2020, Art. no. 110049, doi: [10.1016/j.chaos.2020.110049](https://doi.org/10.1016/j.chaos.2020.110049).
- [10] M. J. Beira and P. J. Sebastião, "A differential equations model-fitting analysis of COVID-19 epidemiological data to explain multi-wave dynamics," *Sci. Rep.*, vol. 11, no. 1, Aug. 2021, Art. no. 16312, doi: [10.1038/s41598-021-95494-6](https://doi.org/10.1038/s41598-021-95494-6).
- [11] S. Khajanchi, K. Sarkar, J. Mondal, K. S. Nisar, and S. F. Abdelwahab, "Mathematical modeling of the COVID-19 pandemic with intervention strategies," *Results Phys.*, vol. 25, Jun. 2021, Art. no. 104285, doi: [10.1016/j.rinp.2021.104285](https://doi.org/10.1016/j.rinp.2021.104285).
- [12] J. Mondal and S. Khajanchi, "Mathematical modeling and optimal intervention strategies of the COVID-19 outbreak," *Nonlinear Dyn.*, vol. 109, no. 1, pp. 177–202, Jul. 2022, doi: [10.1007/s11071-022-07235-7](https://doi.org/10.1007/s11071-022-07235-7).
- [13] O. J. Watson, G. Barnsley, J. Toor, A. B. Hogan, P. Winskill, and A. C. Ghani, "Global impact of the first year of COVID-19 vaccination: A mathematical modelling study," *Lancet Infectious Diseases*, vol. 22, no. 9, pp. 1293–1302, 2022, doi: [10.1016/S1473-3099\(22\)0320-6](https://doi.org/10.1016/S1473-3099(22)0320-6).
- [14] T. Usherwood, Z. LaJoie, and V. Srivastava, "A model and predictions for COVID-19 considering population behavior and vaccination," *Sci. Rep.*, vol. 11, no. 1, p. 12051, Jun. 2021, doi: [10.1038/s41598-021-91514-7](https://doi.org/10.1038/s41598-021-91514-7).
- [15] G. Rinaldi, P. P. Menon, A. Ferrara, W. D. Strain, and C. Edwards, "Epidemiological model based periodic intervention policies for COVID-19 mitigation in the United Kingdom," *Sci. Rep.*, vol. 12, no. 1, p. 15660, Sep. 2022, doi: [10.1038/s41598-022-19630-6](https://doi.org/10.1038/s41598-022-19630-6).
- [16] R. C. Poonia, A. K. J. Saudagar, A. Altameem, M. Alkhatami, M. B. Khan, and M. H. A. Hasanat, "An enhanced SEIR model for prediction of COVID-19 with vaccination effect," *Life*, vol. 12, no. 5, p. 647, Apr. 2022, doi: [10.3390/life12050647](https://doi.org/10.3390/life12050647).
- [17] Z. Liu, H. Zhou, N. Ding, J. Jia, X. Su, H. Ren, X. Hou, W. Zhang, and C. Liu, "Modeling the effects of vaccination, nucleic acid testing, and face mask wearing interventions against COVID-19 in large sports events," *Frontiers Public Health*, vol. 10, Nov. 2022, doi: [10.3389/fpubh.2022.1009152](https://doi.org/10.3389/fpubh.2022.1009152).
- [18] M. Shadabfar, M. Mahsuli, A. S. Khoojine, and V. R. Hosseini, "Time-variant reliability-based prediction of COVID-19 spread using extended SEIVR model and Monte Carlo sampling," *Results Phys.*, vol. 26, Jul. 2021, Art. no. 104364, doi: [10.1016/j.rinp.2021.104364](https://doi.org/10.1016/j.rinp.2021.104364).
- [19] A. Elazzouzi, A. L. Alaoui, M. Tilioua, and A. Tridane, "Global stability analysis for a generalized delayed SIR model with vaccination and treatment," *Adv. Difference Equ.*, vol. 2019, no. 1, p. 532, Dec. 2019, doi: [10.1186/s13662-019-2447-z](https://doi.org/10.1186/s13662-019-2447-z).
- [20] A. de Miguel-Arribas, A. Aleta, and Y. Moreno, "Impact of vaccine hesitancy on secondary COVID-19 outbreaks in the US: An age-structured SIR model," *BMC Infectious Diseases*, vol. 22, no. 1, p. 511, Jun. 2022, doi: [10.1186/s12879-022-07486-0](https://doi.org/10.1186/s12879-022-07486-0).
- [21] G. Sepulveda, A. J. Arenas, and G. González-Parra, "Mathematical modeling of COVID-19 dynamics under two vaccination doses and delay effects," *Mathematics*, vol. 11, no. 2, p. 369, Jan. 2023, doi: [10.3390/math11020369](https://doi.org/10.3390/math11020369).
- [22] Y. Omae, Y. Kakimoto, M. Sasaki, J. Toyotani, K. Hara, Y. Gon, and H. Takahashi, "SIRVVD model-based verification of the effect of first and second doses of COVID-19/SARS-CoV-2 vaccination in Japan," *Math. Biosci. Eng.*, vol. 19, no. 1, pp. 1026–1040, 2021, doi: [10.3934/mbe.2022047](https://doi.org/10.3934/mbe.2022047).
- [23] F. Parino, L. Zino, G. C. Calafiore, and A. Rizzo, "A model predictive control approach to optimally devise a two-dose vaccination rollout: A case study on COVID-19 in Italy," *Int. J. Robust Nonlinear Control*, vol. 33, no. 9, pp. 4808–4823, Jun. 2023, doi: [10.1002/mc.5728](https://doi.org/10.1002/mc.5728).
- [24] F. S. Alshammari, F. T. Akyildiz, M. A. Khan, A. Din, and P. Sunthrayuth, "A stochastic mathematical model for understanding the COVID-19 infection using real data," *Symmetry*, vol. 14, no. 12, p. 2521, Nov. 2022, doi: [10.3390/sym14122521](https://doi.org/10.3390/sym14122521).
- [25] P. Wang and J. Jia, "Stationary distribution of a stochastic SIRD epidemic model of ebola with double saturated incidence rates and vaccination," *Adv. Difference Equ.*, vol. 2019, no. 1, p. 433, Dec. 2019, doi: [10.1186/s13662-019-2352-5](https://doi.org/10.1186/s13662-019-2352-5).
- [26] A. Gómez-Expósito, J. A. Rosendo-Macías, and M. A. González-Cagigal, "Monitoring and tracking the evolution of a viral epidemic through nonlinear Kalman filtering: Application to the COVID-19 case," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 4, pp. 1441–1452, Apr. 2022, doi: [10.1109/JBHI.2021.3063106](https://doi.org/10.1109/JBHI.2021.3063106).
- [27] A. A. H. Ahmadi, M. Naeem, M. Aamir, R. Dewan, S. S. A. Alshqaq, and W. K. Mashwani, "Analysis and forecast of the number of deaths, recovered cases, and confirmed cases from COVID-19 for the top four affected countries using Kalman filter," *Frontiers Phys.*, vol. 9, Aug. 2021, doi: [10.3389/fphy.2021.629320](https://doi.org/10.3389/fphy.2021.629320).
- [28] Q. Sun, T. Miyoshi, and S. Richard, "Analysis of COVID-19 in Japan with extended SEIR model and ensemble Kalman filter," *J. Comput. Appl. Math.*, vol. 419, Feb. 2023, Art. no. 114772, doi: [10.1016/j.cam.2022.114772](https://doi.org/10.1016/j.cam.2022.114772).
- [29] R. Lal, W. Huang, and Z. Li, "An application of the ensemble Kalman filter in epidemiological modelling," *PLoS One*, vol. 16, no. 8, 2021, Art. no. e0256227, doi: [10.1371/journal.pone.0256227](https://doi.org/10.1371/journal.pone.0256227).
- [30] A. Sebbagh and S. Kechida, "EKF-SIRD model algorithm for predicting the coronavirus (COVID-19) spreading dynamics," *Sci. Rep.*, vol. 12, no. 1, p. 13415, Aug. 2022, doi: [10.1038/s41598-022-16496-6](https://doi.org/10.1038/s41598-022-16496-6).
- [31] X. Zhu, B. Gao, Y. Zhong, C. Gu, and K.-S. Choi, "Extended Kalman filter based on stochastic epidemiological model for COVID-19 modelling," *Comput. Biol. Med.*, vol. 137, Oct. 2021, Art. no. 104810, doi: [10.1016/j.compbiomed.2021.104810](https://doi.org/10.1016/j.compbiomed.2021.104810).
- [32] J. Song, H. Xie, B. Gao, Y. Zhong, C. Gu, and K.-S. Choi, "Maximum likelihood-based extended Kalman filter for COVID-19 prediction," *Chaos, Solitons Fractals*, vol. 146, May 2021, Art. no. 110922, doi: [10.1016/j.chaos.2021.110922](https://doi.org/10.1016/j.chaos.2021.110922).
- [33] C. Sun, S. Richard, T. Miyoshi, and N. Tsuzu, "Analysis of COVID-19 spread in Tokyo through an agent-based model with data assimilation," *J. Clin. Med.*, vol. 11, no. 9, p. 2401, Apr. 2022, doi: [10.3390/jcm11092401](https://doi.org/10.3390/jcm11092401).
- [34] A. De Simone and M. Piangerelli, "A Bayesian approach for monitoring epidemics in presence of undetected cases," *Chaos, Solitons Fractals*, vol. 140, Nov. 2020, Art. no. 110167, doi: [10.1016/j.chaos.2020.110167](https://doi.org/10.1016/j.chaos.2020.110167).
- [35] H. Sha, S. Tang, and R. A. Libin, "A discrete stochastic model of the COVID-19 outbreak: Forecast and control," *Math. Biosci. Eng.*, vol. 17, no. 4, pp. 2792–2804, Mar. 2021, doi: [10.3934/mbe.2020153](https://doi.org/10.3934/mbe.2020153).
- [36] S. Bentout, A. Tridane, S. Djilali, and T. M. Touaoula, "Age-structured modeling of COVID-19 epidemic in the USA, UAE and Algeria," *Alexandria Eng. J.*, vol. 60, no. 1, pp. 401–411, Feb. 2021, doi: [10.1016/j.aej.2020.08.053](https://doi.org/10.1016/j.aej.2020.08.053).
- [37] C. M. Leveau, H. A. Aouissi, and F. K. Kebaili, "Spatial diffusion of COVID-19 in Algeria during the third wave," *GeoJournal*, vol. 88, no. 1, pp. 1175–1180, Mar. 2022, doi: [10.1007/s10708-022-10608-5](https://doi.org/10.1007/s10708-022-10608-5).
- [38] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, 2002, doi: [10.1109/78.978374](https://doi.org/10.1109/78.978374).
- [39] D. Fox, J. Hightower, L. Liao, D. Schulz, and G. Borriello, "Bayesian filtering for location estimation," *IEEE Pervasive Comput.*, vol. 2, no. 3, pp. 24–33, Jul. 2003, doi: [10.1109/MPRV.2003.1228524](https://doi.org/10.1109/MPRV.2003.1228524).



**ABDENNOUR SEBBAGH** received the Ph.D. degree in automatic control from the Department of Electrotechnical and Automatic Engineering, Université 8 Mai 1945 Guelma, Guelma, Algeria, in 2016.

He is currently an Associate Professor with Université 8 Mai 1945 Guelma. He serves as the Team Research Leader for "Robotics and Systems Modeling" with the LAIG Laboratory. He is the author of several peer-reviewed scientific publications, conference papers, and research projects in different fields. His research interests include automatic control, tracking targets, ATC tracking, traffic control, and forecasting based on epidemiological models.



**CHEMESSE ENNEHAR BENCHERIET** received the Ph.D. degree in signal processing from the Badji Mokhtar University of Annaba, Algeria, in 2007. She is currently a Senior Lecturer (Class A) with the Computer Science Department, Université 8 Mai 1945 Guelma. She is the author of over 40 peer-reviewed scientific publications and conference papers and has contributed to eight national and international research projects in different fields and the author of books entitled

*Détection et Reconnaissance de visages: De l'importance des caractéristiques géométriques faciales et de la couleur de peau* edited by EUE, in 2012, and *Codage et Représentation de l'Information* edited by OPUG, in 2015. Her research interests include machine learning and deep learning, face detection and recognition, fake face detection, skin detection and classification, intelligent transportation systems, and analysis and classification of remote sensing images. She participated as a member of the program and organizing comity in more than 20 national and international conferences. She contributed as a Reviewer in renowned journals, such as *IET Computer Vision*, *IET Biometrics*, *IET Image Processing*, *IET Electronic Letter*, *IEEE ACCESS*, *The Visual Computer* journal (Springer).



**SIHEM KECHIDA** received the Ph.D. degree in industrial automation (engineering control). Holding the position of Full Professor with Guelma University, Algeria, she serves as the Team Research Leader for Diagnosis and Dependability with Laboratoire d'Automatique et Informatique de Guelma (LAIG). Her contributions extend across various domains within industrial automation and control systems. She is the supervisor of many Ph.D. students. She is the Coordinator

of several industrial research projects focusing on remote supervision and reliability of industrial systems. Her work emphasizes fault detection and isolation (FDI) for hybrid dynamical systems and control of transportation systems. In addition to her academic and research responsibilities, she actively engages as an expert in various national and international committees. She collaborates on research activities and serves as a referee for several esteemed national and international journals, contributing to the peer-review process and maintaining the quality of scholarly publications in her field.

...