



OPEN

EKF-SIRD model algorithm for predicting the coronavirus (COVID-19) spreading dynamics

Abdenmour Sebbagh & Sihem Kechida

In this paper, we study the Covid 19 disease profile in the Algerian territory since February 25, 2020 to February 13, 2021. The idea is to develop a decision support system allowing public health decision and policy-makers to have future statistics (the daily prediction of parameters) of the pandemic; and also encourage citizens for conducting health protocols. Many studies applied traditional epidemic models or machine learning models to forecast the evolution of coronavirus epidemic, but the use of such models alone to make the prediction will be less precise. For this purpose, we assume that the spread of the coronavirus is a moving target described by an epidemic model. On the basis of a SIRD model (Susceptible-Infection-Recovery- Death), we applied the EKF algorithm to predict daily all parameters. These predicted parameters will be much beneficial to hospital managers for updating the available means of hospitalization (beds, oxygen concentrator, etc.) in order to reduce the mortality rate and the infected. Simulations carried out reveal that the EKF seems to be more efficient according to the obtained results.

Since its appearance in late December 2019, the new COVID-19 epidemic has spread rapidly across the world. The first cases of COVID-19 were reported in Wuhan, China and this disease then dilated to Europe, North and South America affecting the most of developed countries such as Italy, France, USA, etc. where sporadic cases have been imported via returning travelers from China.

While it has long seemed spared, or almost, by Covid-19, the African continent is not immune to this coronavirus epidemic, it is now affected like the rest of the world, even if the number of deaths remains very limited. A sudden acceleration in the number of cases was observed in July and August and then the contaminations slowed down again. At the start of the year, we are witnessing a "new wave" very visible in the North of the continent, and observable in several large countries of the East and the South, while the health authorities are getting organized for the arrival of the first doses of vaccine.

As of February 13, 2021, the virus had spread to most African countries, with more than 3 734 227 confirmed cases and more than 97 863 reported deaths, including Algeria with 112 461 cases and 2 970 deaths¹.

For the purpose of control and prevention from the spread of this outbreak of coronavirus, Algerian authorities have implemented various containment measures since March 28, 2020, including traffic restrictions, contact tracing, mandatory face masks in public spaces, entry or exit screening, quarantine and awareness campaigns.

The current outbreak of coronavirus disease (COVID-19) is declared as Public Health Emergency of International concern and a pandemic by the World Health Organization (WHO). This alarming situation has prompted scientists to indulge in studies concerning the transmission dynamics and forecasting of the virus to the most affected countries in the world such as China^{2,3}, Italy⁴, France⁵, India⁶ and then to other countries⁷⁻¹³, etc.

These works focus on epidemiological studies whose the main objective is to develop strategies to fight against the spread of the coronavirus and provide guidance to control its transmission dynamic^{14,15}. A considerable number of strategies require or involve mathematical models dedicated for studying infection diseases such as SIRD, SIR or SEIR models in different context¹⁶⁻¹⁸ (analysis, forecasting the spread and prediction).

The most of these works are intended to the modeling of transmission dynamics with the aim to predict the trend of the epidemic and control the outbreak evolution. In this context, authors of¹⁹⁻²³ proposed mathematical models translating the transmission dynamics of COVID-19 to forecast the number of active cases or to estimate the total number of infected and deaths⁷, while those of² develop a strategy based on SIR model to estimate the actual number of people infected and to deduce the IFR (Infected Fatality Ratio). The forecast of future COVID-19 cases has discussed in²⁴ using regression analysis.

Laboratoire d'Automatique et Informatique de Guelma (LAIG), Université 8 mai 1945 Guelma, Bp: 401, 24000 Guelma, Algeria. email: sebbagh.abdenmour@univ-guelma.dz

The estimation of infection, mortality and recovery rates and the basic reproduction number (R_0) are provided in³ using a SIRD model. Afterwards, Dhillon and all study the trend analysis of mortality and recovery rate considering scenario of most affected countries and Indian States⁶. To mitigate disease transmission, mathematical models introducing a quarantine measures are formulated by Liu and all in² and Mandal and all in²⁵. Other research works establish the prediction of epidemic peak under the impact of lockdown using an improvised compartment mathematical model²⁶ (SEIR or SEIRD) i.e., Susceptible (S)-Exposed (E)-Infected (I)-Recovered (R)-Death (D) while, in^{27,28}, authors study the forecast of the spread tendency of the COVID-19 through an improved SEIR model. Others methods such as fractional concept, optimization algorithms, Artificial Neural Network... are introduced sometimes for study the growth of cumulative confirmed and cured people and sometimes for formulate the prediction problem as an optimization framework²⁹ or to estimate the COVID-19 cases⁸.

Additionally, and for containing the epidemic spread in African countries, research works are being conducted in the top infected countries through studies modeling and forecasting of COVID-19. Among which, there have been some comparative studies between the African countries including Algeria^{30–32}.

However, there are a little peer reviewed papers about epidemiological profile in Algerian territory; these research studies consider traditional epidemic models (SIR, SI, SEIR, ...) dedicated to historical data analysis for forecasting the incidence and /or estimation of parameters^{33–37}.

In all these developed methodologies, the authors consider mathematical models whose parameters are estimated over a limited period of time. The model once defined is applied in different studies of COVID'19 evolution without taking into account the update of the model parameters and the various measures taken by those responsible.

In this work, we project the engineering techniques used in targets tracking on epidemiology assuming that the spread of the coronavirus is a moving target described by an epidemic model. The idea is to investigate the Kalman filter on SIRD model with the goal to predict the spreading of the Covid 19 and to effectively manage the burden of COVID-19 pandemic in Algeria.

This study shows the disease profile in the Algerian territory since February 25, 2020 to February 13, 2021. Here we are fascinated in applying the extended Kalman filter (EKF) using an epidemic SIRD model to provide a daily prediction of infection, mortality and recovery rates and the basic reproduction number (R_0).

In addition, these data are much beneficial to hospital managers and public health decision-makers for updating the available means of hospitalization (beds, oxygen concentrator, etc.) in order to reduce the mortality rate and the infected.

The rest of this paper is organized in 4 sections. “**Problem formulation**” Section is dedicated the problem formulation and the description of chosen model. The next section details Bayesian approach and more precisely the EKF algorithm used in the context of this work. The application of this technique and the simulation results are discussed in “**Simulation results**” section and finally, the last section recapitulates concluding remarks of this study and to suggest some outlooks for future works.

Problem formulation

In the literature, the works carried out in the epidemiology study use mathematical models each stratify the dynamics of individuals. The choice of these individuals depends on the problem formulation. The most of the epidemic models for human-to-human transmission rely on the susceptible-infected-recovered (SIR) structure, considered as a fundamental model widely used to delineate various infectious diseases. SIRD model is the standard famous SIR model incorporating an additional compartment: Death class (D). Other structures have emerged to monitor the dynamics of others compartments (classes), such as quarantined susceptible individuals, asymptomatic infectious individuals, isolated infected individuals, exposed individuals, etc.^{3,21,22,26,37}

For the SIRD model, the population N is divided into sub-population: susceptible (S), infected (I), recovered (R) and deceased (D) for all time k , i.e., $N = S + I + R + D$.

The discrete nonlinear SIRD model is given by:

$$S(k+1) = S(k) - \frac{\alpha(k)}{N} S(k)I(k) \quad (1)$$

$$I(k+1) = I(k) + \frac{\alpha(k)}{N} S(k)I(k) - \beta(k)I(k) - \gamma(k)I(k) \quad (2)$$

$$R(k+1) = R(k) + \beta(k)I(k) \quad (3)$$

$$D(k+1) = D(k) + \gamma(k)I(k) \quad (4)$$

where $\alpha(k)$, $\beta(k)$ and $\gamma(k)$ are the daily infection, daily recovery and daily death rates respectively, see Fig. 1, note that, these rates are optimized daily using the least square method (LSM) as follows:

If we accept that $S = N$, then:

$$\alpha(k) = \frac{\sum_{j=1}^k I(j) \cdot \Delta S(j)}{\sum_{j=1}^k I^2(j)} \quad (5)$$

If $S \neq N$, then:

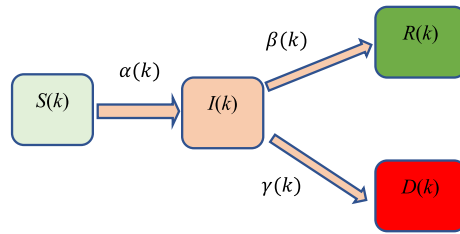


Figure 1. Flow between the populations of SIRD model.

$$\alpha(k) = N \cdot \frac{\sum_{j=1}^k S(j) \cdot I(j) \cdot \Delta S(j)}{\sum_{j=1}^k I^2(j) \cdot S^2(j)} \tag{6}$$

$$\beta(k) = \frac{\sum_{j=1}^k I(j) \cdot \Delta R(j)}{\sum_{j=1}^k I^2(j)} \tag{7}$$

$$\gamma(k) = \frac{\sum_{j=1}^k I(j) \cdot \Delta D(j)}{\sum_{j=1}^k I^2(j)} \tag{8}$$

$I(j)$ is the total currently infected in the time j (day).
 $\Delta S(j) = S(j) - S(j - 1)$ is daily new coronavirus cases at time j
 $\Delta R(j) = R(j) - R(j - 1)$ is daily new recovered at time j
 $\Delta D(j) = D(j) - D(j - 1)$ is daily new deceased at time j
 We suppose that:

$$x_1(k) = S(k), x_2(k) = I(k), x_3(k) = R(k), \text{ and } x_4(k) = D(k)$$

Then the SIRD model becomes:

$$\begin{pmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ x_4(k+1) \end{pmatrix} = \begin{pmatrix} f_1(X_k) \\ f_2(X_k) \\ f_3(X_k) \\ f_4(X_k) \end{pmatrix} = \begin{pmatrix} x_2(k) + \frac{\alpha(k)}{N} x_1(k) \cdot x_2(k) - \beta(k)x_2(k) - \gamma(k)x_2(k) \\ x_3(k) + \beta(k)x_2(k) \\ x_4(k) + \gamma(k)x_2(k) \end{pmatrix} + V_k \tag{9}$$

X_k is the state vector including susceptible (S), infected (I), recovered (R) and deceased (D), defined as:

$$X_k = (S \ I \ R \ D)^T$$

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

The Jacobian matrix of this model is obtained as:

$$F(X_k) = \frac{\partial f}{\partial X_k} = \begin{pmatrix} 1 - \frac{\hat{\alpha}(k)}{N} x_2(k) & -\frac{\hat{\alpha}(k)}{N} x_1(k) & 0 & 0 \\ \frac{\hat{\alpha}(k)}{N} x_2(k) & 1 - \hat{\beta}(k) - \hat{\gamma}(k) + \frac{\hat{\alpha}(k)}{N} x_1(k) & 0 & 0 \\ 0 & \hat{\beta}(k) & 1 & 0 \\ 0 & \hat{\gamma}(k) & 0 & 1 \end{pmatrix} \tag{10}$$

where $\hat{\alpha}(k)$, $\hat{\beta}(k)$ and $\hat{\gamma}(k)$ are the predicted daily infection, predicted daily recovery and predicted daily death rates respectively and are calculated as:

$$\hat{\alpha}(k) = \frac{\text{predicted daily new cases}}{\text{estimate of total currently infected}} = \frac{[x_2(k+1/k) - x_2(k)] + [x_3(k+1/k) - x_3(k)] + [x_4(k+1/k) - x_4(k)]}{x_2(k)} \tag{11}$$

$$\hat{\beta}(k) = \frac{\text{predicted daily new recovered}}{\text{estimate of total currently infected}} = \frac{[x_3(k+1/k) - x_3(k)]}{x_2(k)} \tag{12}$$

$$\hat{\gamma}(k) = \frac{\text{predicted daily new deceased}}{\text{estimate of total currently infected}} = \frac{[x_4(k+1/k) - x_4(k)]}{x_2(k)} \tag{13}$$

The predicted daily new cases = the predicted daily new currently infected + the predicted daily new recovered + the predicted daily new deceased.

We suppose that the measurement equation is given daily by:

$$Y_{k+1} = \begin{pmatrix} y_1(k+1) \\ y_2(k+1) \\ y_3(k+1) \end{pmatrix} = \begin{pmatrix} I(k+1) \\ R(k+1) \\ D(k+1) \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ x_4(k+1) \end{pmatrix} + W_k \quad (14)$$

$$Y_{k+1} = CX_{k+1} + W_k$$

with

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

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

Bayesian filtering

In Bayesian approach we attempt to construct the posterior PDF of the state given all measurements. All available information is used to form such PDF. So, this PDF represents complete solution.

Let $X_k, k \in \mathbb{N}$, be the state sequence:

$$X_k = f_k(X_{k-1}, u_{k-1}, V_{k-1}) \quad (15)$$

where f_k is in generally nonlinear function of the previous state $X_{k-1} \in \mathbb{R}^{n_x}$, $V_{k-1} \in \mathbb{N}^{n_v}$ is state noise, $u_{k-1} \in \mathbb{R}^{n_u}$ is known input, n_x, n_v et n_u are dimensions of the state, process and input noise vectors.

let Y_k be the measurement:

$$Y_k = h_k(X_k, W_k) \quad (16)$$

where $Y_k \in \mathbb{R}^{n_y}$, h_k is in generally non-linear measurements function, $W_k \in \mathbb{N}^{n_w}$ is measurement noise, n_y and n_w are dimensions of the measurement and measurement noise vectors.

We want to find estimate of the X_k based on all available measurements at time k (marked as $Y_{1:k}$) by constructing the posterior PDF $p(X_k, Y_{1:k})$. It is assumed, that initial PDF $p(X_0|Y_0) \equiv p(X_0)$ is available. Posterior PDF can be obtained recursively in two stages, namely prediction and update. Suppose that required PDF $p(X_{k-1}|Y_{1:k-1})$ at time step $k-1$ is available. Then using the system model, it is possible to obtain the prior PDF of the state at the time step k ^{38,39}:

$$p(X_k|Y_{1:k-1}) = \int p(X_k|X_{k-1})p(X_{k-1}|Y_{1:k-1})dX_{k-1} \quad (17)$$

Prediction step usually deforms, spreads state PDF due to noise. Measurement Y_k is available at time step k , so it can be used to update the prior. Using Bayes' rule, we obtain:

$$p(X_k|Y_{1:k}) = \frac{p(Y_k|X_k)p(X_k|Y_{1:k-1})}{p(Y_k|Y_{1:k-1})} \quad (18)$$

where the normalizing constant is:

$$p(Y_k|Y_{1:k-1}) = \int p(Y_k|X_k)p(X_k|Y_{1:k-1})dX_k \quad (19)$$

In the update Eq. (19), the measurement Y_k is used to modify the predicted prior from the previous time step to obtain PDF of the state. Equations (17) and (18) theoretically allow optimal Bayesian solution. But it is only conceptual solution and integrals in these equations are intractable. Solution exists in some restricted cases such as Kalman Filter.

Kalman filter. Kalman filter together with its basic variants are commonly the used tools in statistical signal processing, especially in the context of causal, real-time applications.

There are several approaches in the derivation of the Kalman Filter. We can assume Gaussian distribution of the deriving process and of the initial state. In the next phase, we derive the posterior distribution of the states given the observations, taking the mean of the resulting distributions as the estimation of the state. The second approach combines a recursive weighted least-squares method with special weighting of the previous estimate of the states in the role of additional measurements^{40,41}.

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. But in many cases this PDF is not Gaussian and we need to use different approach such as extended Kalman Filter. This method is also labelled as sub-optimal algorithm^{42,43}.

Extended Kalman filter. Most processes in real life are unfortunately nonlinear, and therefore needs to be linearized before they can be estimated by Kalman filter.

The extended Kalman filter (EKF)^{38,39,44–46}, is the nonlinear genre of the Kalman filter^{41,42} which linearizes about an estimate of the current mean and covariance^{43,47}. The state transition and measurement models for the extended Kalman filter are taken as:

$$X(k+1) = f(X(k)) + V(k) \quad (20)$$

$$Y(k+1) = h(X(k+1)) + W(k) \quad (21)$$

where $V(k)$ is the process noise with zero mean and covariance Q_k , and $W(k)$ is the measurement noise with zero mean and covariance Σ_k .

The functions $f(X(k))$ and $h(X(k+1))$ are used to compute the predicted state from the previous estimate and predicted measurement from the predicted state, respectively. Instead of applying $f(X(k))$ and $h(X(k+1))$ to the covariance directly, a Jacobian matrix is applied which is evaluated with current predicted states at each time step. Extended Kalman Filter is based upon approximation of the Bayes' rule using linearization.

Discrete-time extended Kalman filter's prediction (time update) and correction (measurement update) equations are given by,

- **Prediction (time update)**

Predict stage can be described using following equations:

$$\hat{X}_{k+1|k} = f(\hat{X}_{k|k}) \quad (22)$$

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

$$P_{k+1|k} = \hat{F}_k P_{k|k} \hat{F}_k^T + Q_k \quad (23)$$

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

- **Correction (measurement update)**

Update stage can be described with the following equations:

$$\tilde{y}_{k+1} = Y_{k+1} - h(\hat{X}_{k+1|k}) \quad (24)$$

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

$$S_{k+1} = \hat{H}_{k+1} P_{k+1|k} \hat{H}_{k+1}^T + \sum_k \quad (25)$$

where S_{k+1} is the innovation covariance,

$$K_{k+1} = P_{k+1|k} \hat{H}_{k+1}^T S_{k+1}^{-1} \quad (26)$$

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

$$\hat{X}_{k+1|k+1} = \hat{X}_{k+1|k} + K_{k+1} \tilde{y}_{k+1} \quad (27)$$

is update state estimate and

$$P_{k+1|k+1} = (I - K_{k+1} \hat{H}_{k+1}) P_{k+1|k} \quad (28)$$

is update estimate covariance.

Where the Jacobian for state transition and measurement matrices are defined as:

$$\hat{F}_k = \left. \frac{\partial f}{\partial X} \right|_{X_{k|k}} \quad (29)$$

$$\hat{H}_{k+1} = \left. \frac{\partial h}{\partial X} \right|_{X_{k+1|k}} \quad (30)$$

Figure 2 Shows the EKF-SIRD Algorithm.

Simulation results

For the application of EKF estimator on coronavirus (covid-19) modelled by the SIRD model, we use the real data provided by the Ministry of Algerian health and the WHO, from February 25, 2020 to February 13, 2021 in our daily predictions.

We consider that the spread of coronavirus is a target that begins its movement from the initial vector:

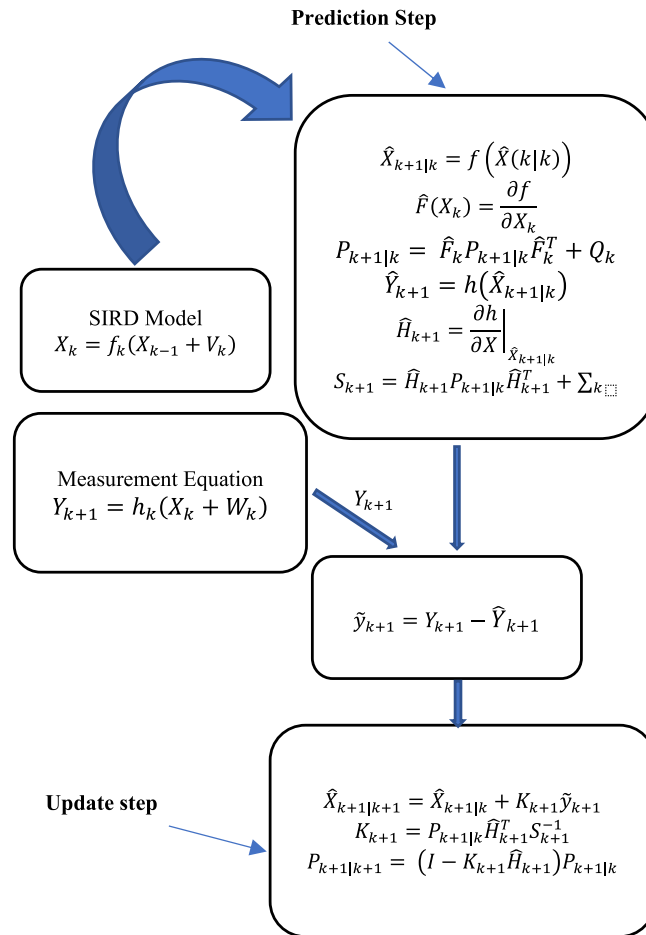


Figure 2. EKF-SIRD Model Algorithm.

$$X(1) = (N - 1 \ 1 \ 0 \ 0)^T$$

where $N = 44219385$ is the Algerian population number. The mean vector and covariance matrix initialization of the EKF according to a Gaussian law are:

$$\hat{X}(1) = (N - 1 \ 1 \ 0 \ 0)^T$$

$$P(1|1) = \begin{pmatrix} 100 & 0 & 0 & 0 \\ 0 & 100 & 0 & 0 \\ 0 & 0 & 100 & 0 \\ 0 & 0 & 0 & 100 \end{pmatrix}$$

The process noise is zero mean, white and with covariance

$$Q = \begin{pmatrix} 100 & 0 & 0 & 0 \\ 0 & 100 & 0 & 0 \\ 0 & 0 & 100 & 0 \\ 0 & 0 & 0 & 100 \end{pmatrix}$$

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

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

The trajectories plotted in Fig. 3a, b, c and d are the real data of Algeria and predicted by EKF of total coronavirus cases, total currently infected, total recovered and total deceased respectively.

We observe that the predicted trajectories by the EKF are superposable on the trajectories of real data, which allowed us to say that the EKF is correctly predicted the evolution of these quantities.

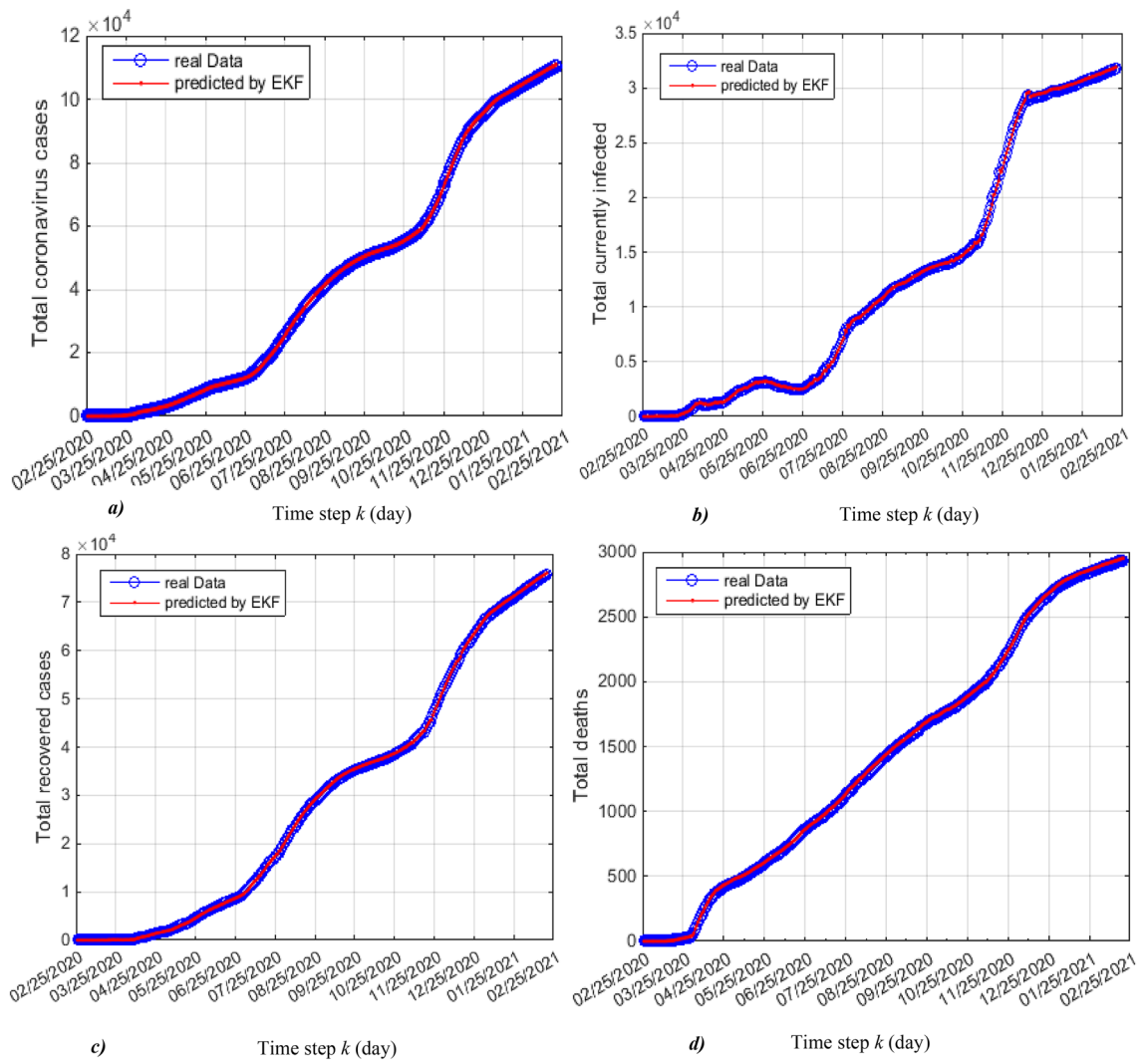


Figure 3. Real and predicted trajectories, (a) total coronavirus cases, (b) total currently infected, (c) total recovered cases and (d) total deceased.

The daily infection rate $\alpha(k)$, daily recovery rate $\beta(k)$ and daily death rate $\gamma(k)$ are optimised by using least square method (LSM) according to Eqs. (5), (6), (7) and (8), and also predicted by the EKF according to Eqs. (11), (12) and (13) as shown in Fig. 4a, b and c.

From these previous predictions (by LMS and by EKF), we can daily predict the basic reproduction number R_0 as shown in Figs. 5 and 6, according to Eq. (31).

$$R_0(k) = \frac{\alpha(k)}{\beta(k) + \gamma(k)} \tag{31}$$

We see that generally, the value of the basic reproduction number between the end of April, 2020 and October 15, 2020 is between 1 and 2 except for some disturbances in July, this comes down to the containments measures taken by the country officials (lockdown), including traffic restrictions, contact tracing, mandatory face masks in public spaces.

From the mid of October, 2020 until the end of December, we see some disturbances in the basic reproduction number because of the appearance of the second coronavirus wave, where the daily new coronavirus number has been increased and reached 1133 cases on November 24, 2020.

From the beginning of January, 2021 until February 13, 2021 the basic reproduction number stabilizes between 1 and 1.5.

Using the predictions by EKF of total currently infected, total recovered and total deceased, we can daily predict the case fatality ratio (CFR), case recovery ratio (CRR) and also case infection ratio (CIR) as shown in Fig. 7a, b and c, according to these equations:

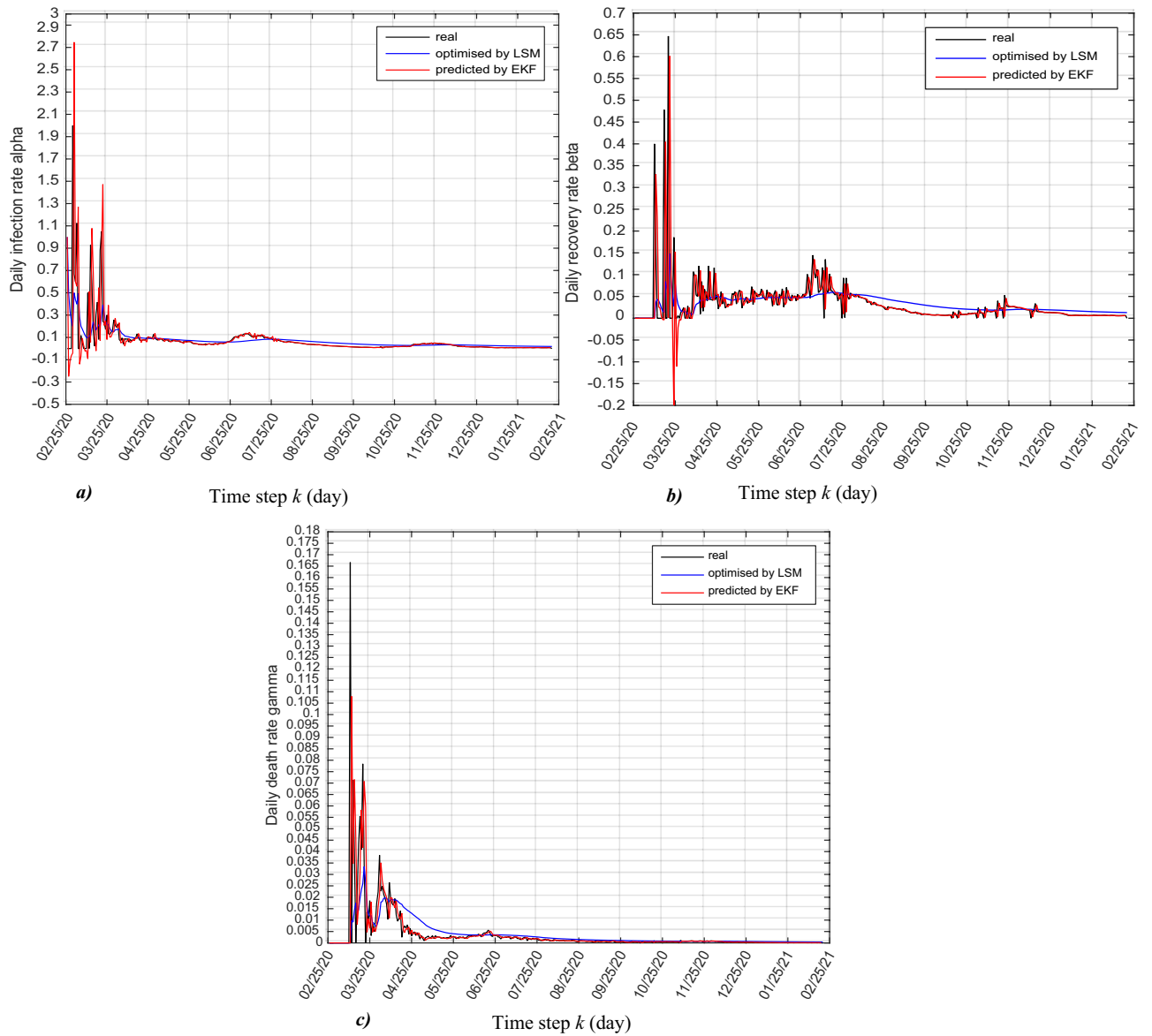


Figure 4. Real and predicted trajectories, (a) daily infection rate $\alpha(k)$, (b) daily recovery rate $\beta(k)$ and (c) daily death rate $\gamma(k)$.

$$CFR(k + 1)\% = \frac{\text{Total deceased}(x_4(k + 1/k) = D(k + 1))}{\text{Total coronavirus cases} = (x_2(k + 1/k) + x_3(k + 1/k) + x_4(k + 1/k))} \cdot 100 \quad (32)$$

$$CRR(k + 1)\% = \frac{\text{Total recovered}(x_3(k + 1/k) = R(k + 1))}{\text{Total coronavirus cases} = (x_2(k + 1/k) + x_3(k + 1/k) + x_4(k + 1/k))} \cdot 100 \quad (33)$$

$$CIR(k + 1)\% = \frac{\text{Total currently infected}(x_2(k + 1/k) = I(k + 1))}{\text{Total coronavirus cases} = (x_2(k + 1/k) + x_3(k + 1/k) + x_4(k + 1/k))} \cdot 100 \quad (34)$$

Figure 8a, b and c show the real and predicted trajectories of daily new coronavirus cases, daily new deceased and daily new recovered, from these trajectories it shown that the EKF is correctly predicted these daily new quantities.

The good results obtained by the application of the EKF on the coronavirus evolution using SIRD model are demonstrated by the smaller RMSEs (Root mean square errors) of daily new coronavirus cases, daily new deceased and daily new recovered, illustrated in Fig. 9a, b and c.

These RMSEs are obtained from 100 Monte Carlo runs given by the equation:

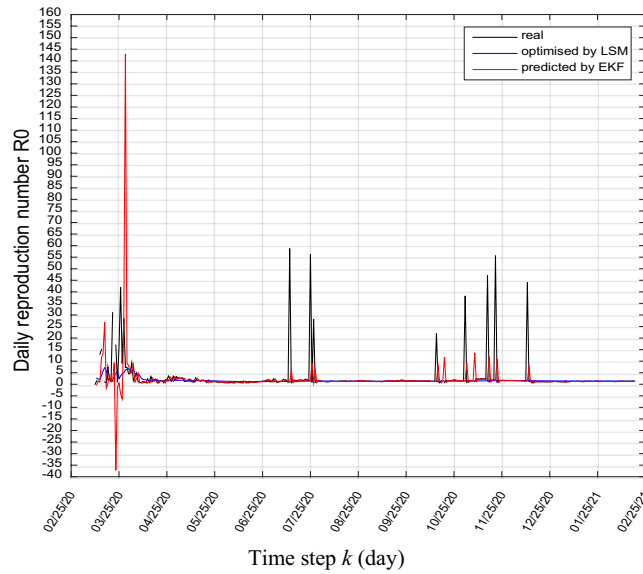


Figure 5. Real and predicted trajectories of daily basic reproduction number R_0 .

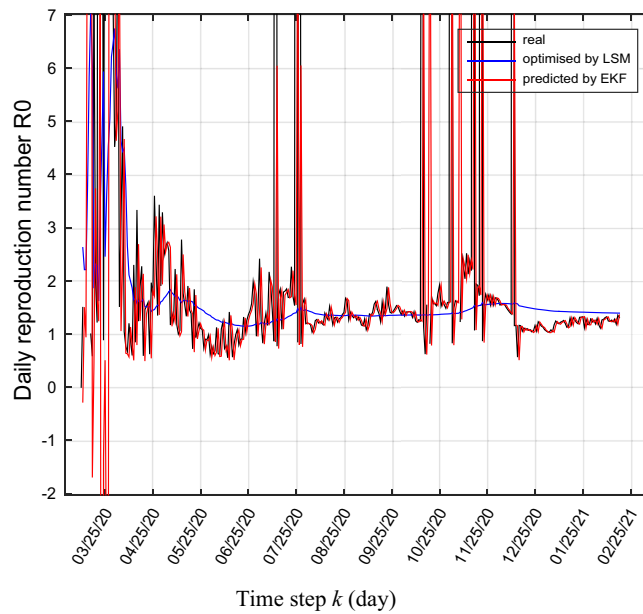


Figure 6. Real and predicted trajectories of daily basic reproduction number R_0 (zoom).

$$RMSE(x(j)) = \sqrt{\frac{\sum_{k=1}^j (x_{real}(k) - x_{predicted}(k))^2}{j}} \quad j = 1, \dots, \dots \tag{35}$$

Conclusion

To track and predict the spread of coronavirus pandemic, we investigated and analysed the outbreak of this Covid-19 disease in Algeria, to help the government and the health ministry take new measures and future decisions to deal with this coronavirus pandemic.

For this, we supposed that the coronavirus epidemic is a target modelled by a nonlinear SIRD model and we apply the engineering technique of target tracking (an EKF algorithm) on the coronavirus spreading to predict daily all parameters i.e., susceptible (S), infected (I), recovered (R) and deceased (D).

The novelty of this work is summed up in two points: the daily updating of the model parameters and the application of the extended Kalman filter on this model, which makes the prediction results more precise and the method more reliable.

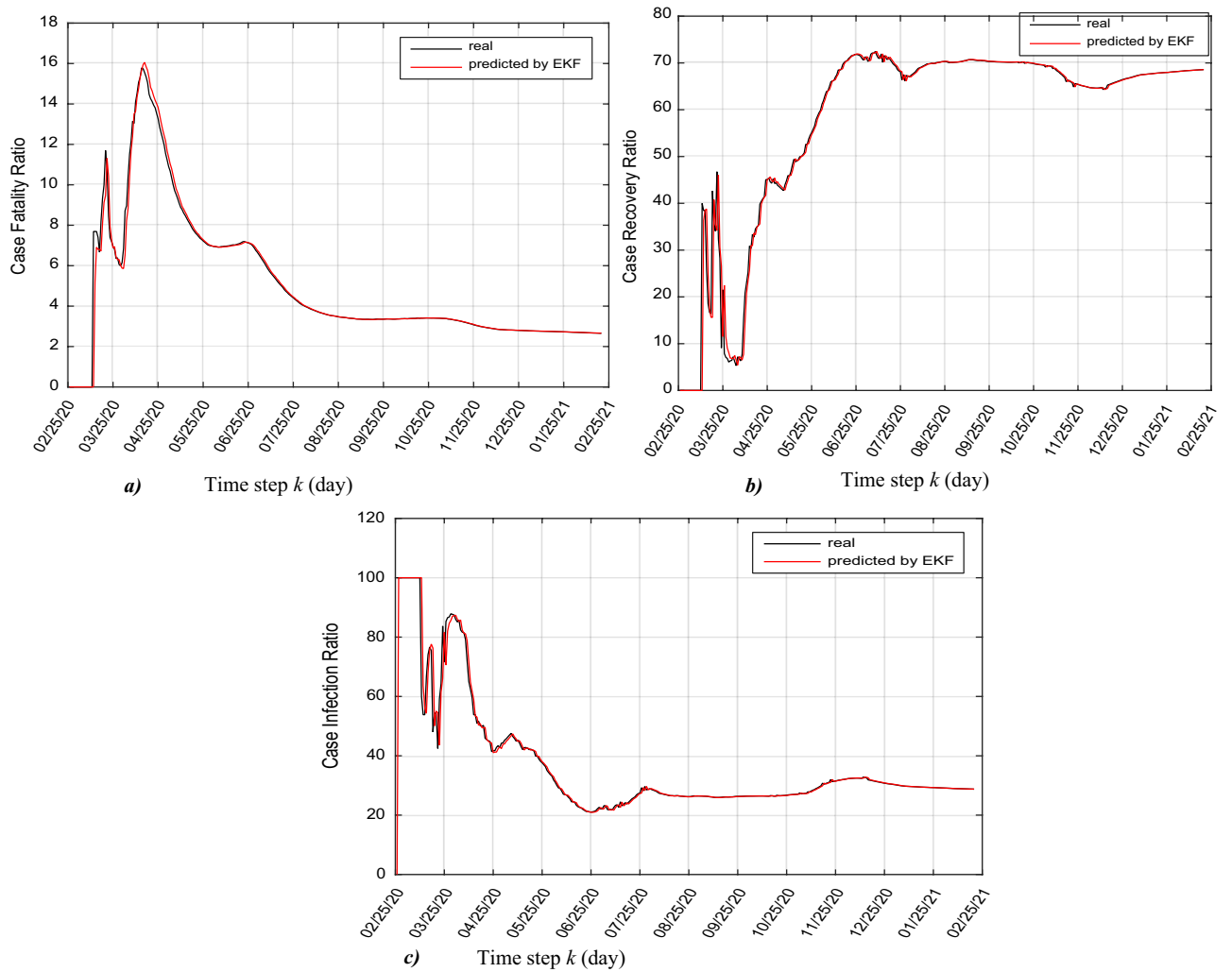


Figure 7. Real and predicted trajectories, (a) case fatality ratio, (b) case recovery ration and, (c) case infection ratio.

The results showed that according to the data provided by the Ministry of Algerian health and the WHO, from February 25, 2020 to February 13, 2021, the EKF algorithm is successfully predicted the daily coronavirus spreading.

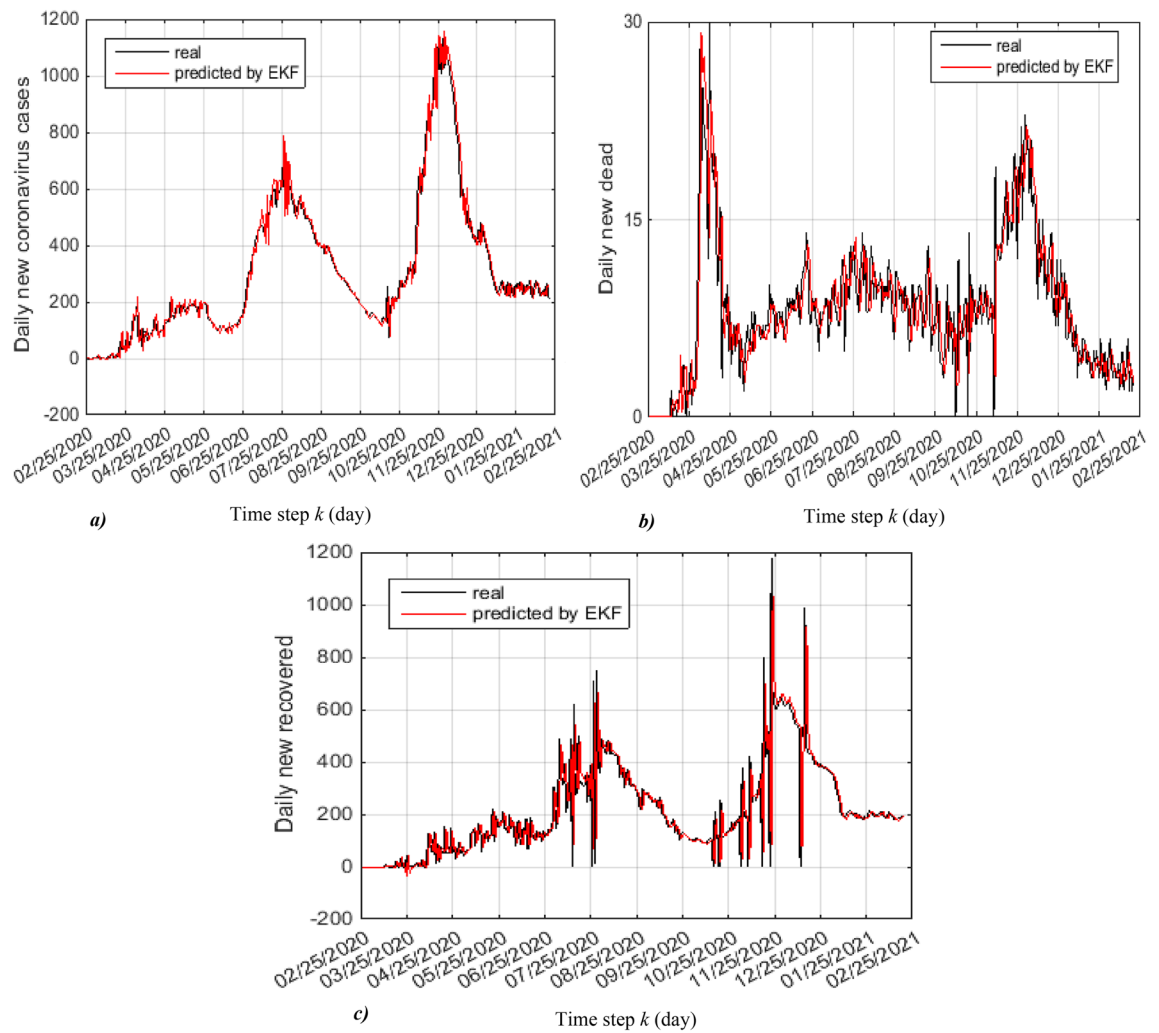


Figure 8. Real and predicted trajectories, (a) daily new coronavirus cases, (b) daily new dead and (c) daily new recovered.

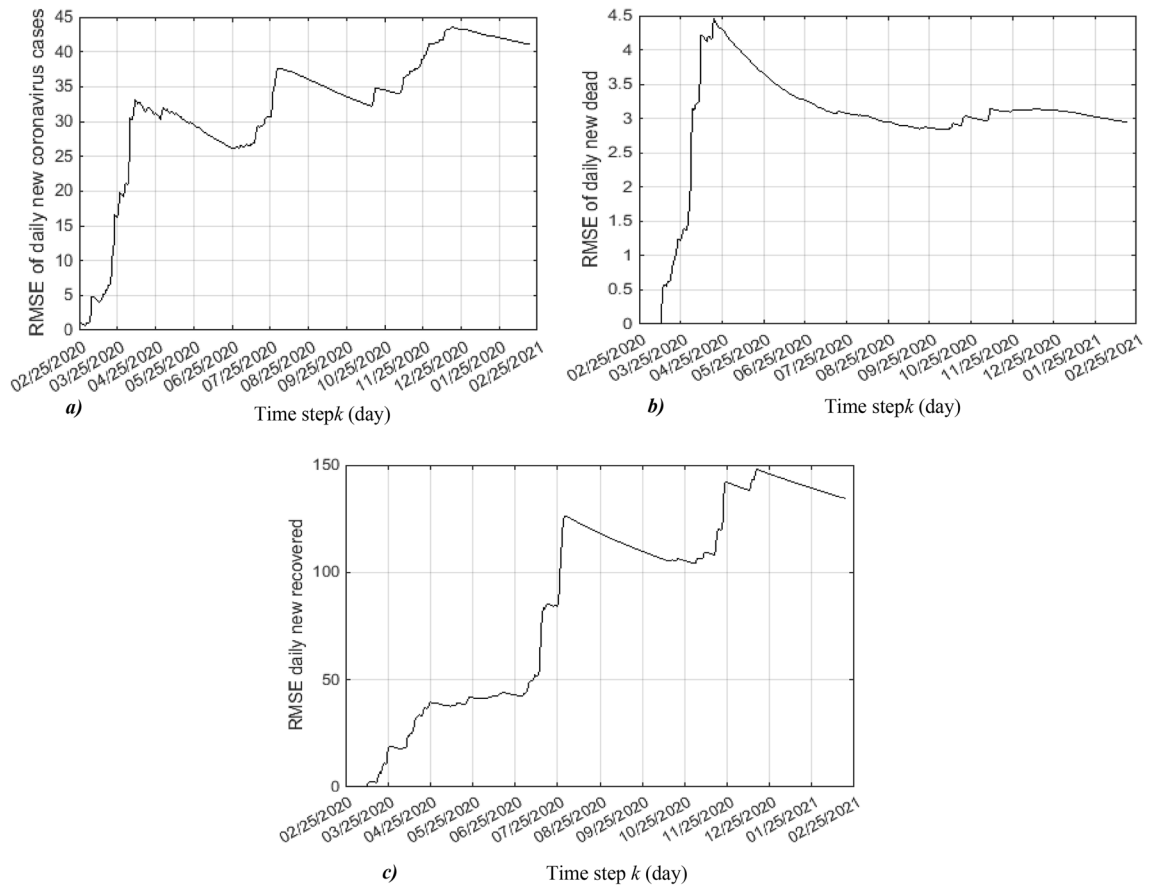


Figure 9. RMS error of: (a) daily new coronavirus cases, (b) daily new dead and (c) daily new recovered.

Data availability

The datasets generated and/or analysed during the current study are available in the [Database Algeria Covid19] https://laig.univ-guelma.dz/sites/laig.univ-guelma.dz/files/Database_%20Algeria_Covid19.xlsx or <https://laig.univ-guelma.dz/fr/node/209>.

Received: 26 July 2021; Accepted: 11 July 2022

Published online: 04 August 2022

References

- <https://www.coronavirus-statistiques.com/stats-continent/coronavirus-nombre-de-cas-afrique>, <https://www.worldometers.info/coronavirus/country/algeria/>.
- Xiuli, L., Geoffrey, H., Shouyang, W., Minghui, Q., Xin, X., Shan, Z., Xuefeng, L. Modelling the situation of COVID-19 and effects of different containment strategies in China with dynamic differential equations and parameters estimation. medRxiv preprint <https://doi.org/10.1101/2020.03.09.20033498> (2020).
- Anastassopoulou, C., Russo, L., Tsakris, A. & Siettos, C. Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PLoS ONE* <https://doi.org/10.1371/journal.pone.0230405> (2020).
- Yongmei, D. & Liyuan, G. An evaluation of COVID-19 in Italy: A data-driven modeling analysis. *Infect. Dis. Model.* **5**, 495–501. <https://doi.org/10.1016/j.idm.2020.06.007> (2020).
- Lionel, R., Etienne, K., Julien, P., Antoine, S. & Samuel, S. Using early data to estimate the actual infection fatality ratio from COVID-19 in France. *Biology* <https://doi.org/10.3390/biology9050097> (2020).
- Preeti, D., Sampurna, K., Chander, S., Usha, R., Laxmi, K. D., Suryakant, Y., Sayeed, U. Case-fatality ratio and recovery rate of COVID-19: Scenario of most affected countries and Indian States. A Situational Analysis Paper for Policy Makers. International Institute for Population Sciences, Mumbai <https://doi.org/10.13140/RG.2.2.25447.68000> (2020).
- Jemy, A. & Mandujano, V. Predicting the number of total COVID-19 cases and deaths in Brazil by the Gompertz model. *Nonlinear Dyn.* **102**, 2951–2957. <https://doi.org/10.1007/s11071-020-06056-w> (2020).
- Torrealba-Rodríguez, O., Conde-Gutiérrez, R. A. & Hernández-Javier, A. L. Modeling and prediction of COVID-19 in Mexico applying mathematical and computational models. *Chaos Solitons Fractals* <https://doi.org/10.1016/j.chaos.2020.109946> (2020).
- Zebin, Z. *et al.* Prediction of the COVID-19 spread in African countries and implications for prevention and control: A case study in South Africa Egypt, Algeria, Nigeria, Senegal and Kenya. *Sci. Total Environ.* **729**, 138959 (2020).
- Issam, D. Modeling Palestinian COVID-19 cumulative confirmed cases: A comparative study. *Infect. Dis. Model.* **5**, 748–754. <https://doi.org/10.1016/j.idm.2020.09.001> (2020).
- Façal, N., Iván, A., Juan, J. N., Cristiana, J. S. & Delfim, F. M. T. Fractional model of COVID-19 applied to Galicia, Spain and Portugal. *Chaos Solitons Fractals* <https://doi.org/10.1016/j.chaos.2021.110652> (2021).

12. Harun, Y., Aynur, Y., Mustafa, A. T. & Melike, T. Modeling and Forecasting for the number of cases of the COVID-19 pandemic with the Curve Estimation Models, the Box-Jenkins and Exponential Smoothing Methods. *EJMO* 4(2), 160–165 (2020).
13. Osmar, P. N. *et al.* Mathematical model of COVID-19 intervention scenarios for São Paulo—Brazil. *Nat. Commun.* <https://doi.org/10.1038/s41467-020-20687-y> (2021).
14. Calvin, T., Fernando, L., Mark, A. S. & Michael, B. Modeling, state estimation, and optimal control for the US COVID-19 outbreak. *Nat. Sci. Rep.* 10, 10711. <https://doi.org/10.1038/s41598-020-67459-8> (2020).
15. Maira, A., Eduardo, M. O., Joseba, B. V. D., Javier, M. & Nico, S. Modelling COVID 19 in the Basque Country from introduction to control measure response. *Nat. Sci. Rep.* 10, 17306. <https://doi.org/10.1038/s41598-020-74386-1> (2020).
16. Ottar, N. B., Katriona, S., Martin, K. & Naomi, A. Modeling infectious epidemics. *Nat. Methods* 17(5), 455–456. <https://doi.org/10.1038/s41592-020-0822-z> (2020).
17. Ottar, N. B., Katriona, S., Martin, K. & Naomi, A. The SEIRS model for infectious disease dynamics. *Nat. Methods* 17(6), 557–558. <https://doi.org/10.1038/s41592-020-0856-2> (2020).
18. Saulo, B. B. & Daniel, O. C. Modeling and forecasting the early evolution of the Covid-19 pandemic in Brazil. *Nat. Sci. Rep.* 10, 19457. <https://doi.org/10.1038/s41598-020-76257-1> (2020).
19. Subhas, K. & Kankan, S. Forecasting the daily and cumulative number of cases for the COVID-19 pandemic in India. *Chaos* 30, 071101. <https://doi.org/10.1063/5.0016240> (2020).
20. Kankan, S. & Subhas, K. Modeling and forecasting the COVID-19 pandemic in India. *Chaos Solitons Fractals* <https://doi.org/10.1016/j.chaos.2020.110049> (2020).
21. Malavika, B. *et al.* Forecasting COVID-19 epidemic in India and high incidence states using SIR and logistic growth models. *Clin. Epidemiol. Global Health* 9, 26–33. <https://doi.org/10.1016/j.cegh.2020.06.006> (2020).
22. Chatterjee, S., Sarkar, A., Chatterjee, S., Karmakar, M. & Paul, R. Studying the progress of COVID-19 outbreak in India using SIRD model. *Indian J. Phys.* 95(9), 1941–1957. <https://doi.org/10.1007/s12648-020-01766-8> (2021).
23. Yuan, Z. *et al.* Prediction of the COVID-19 outbreak in China based on a new stochastic dynamic model. *Nat. Sci. Rep.* 10, 21522. <https://doi.org/10.1038/s41598-020-76630-0> (2020).
24. Vikas, K. S. & Unnati, N. Modeling and forecasting of Covid-19 growth curve in India. *Trans. Indian Natl. Acad. Eng.* 5, 697–710. <https://doi.org/10.1007/s41403-020-00165-z> (2020).
25. Manotosh, M. *et al.* A model-based study on the dynamics of COVID-19: Prediction and control. *Chaos Solitons Fractals* <https://doi.org/10.1016/j.chaos.2020.109889> (2020).
26. Chatterjee, S., Sarkar, A., Karmakar, M., Chatterjee, S. & Paul, R. SEIRD model to study the asymptomatic growth during COVID-19 pandemic in India. *Indian J. Phys.* 95, 2575–2587. <https://doi.org/10.1007/s12648-020-01928-8> (2021).
27. Vipin, T., Namrata, D. & Nandan, S. B. Mathematical modeling based study and prediction of COVID-19 epidemic dissemination under the impact of lockdown in India. *Front. Phys.* <https://doi.org/10.3389/fphy.2020.586899> (2020).
28. Deshun, S., Li, D., Jianyi, X. & Daping, W. Modeling and forecasting the spread tendency of the COVID-19 in China. *Adv. Differ. Equ.* <https://doi.org/10.1186/s13662-020-02940-2> (2020).
29. Zreiq, R. *et al.* Generalized Richards model for predicting COVID-19 dynamics in Saudi Arabia based on particle swarm optimization Algorithm. *AIMS Public Health* 7(4), 828–843 (2020).
30. Alemayehu, S. A. Modeling and forecasting of COVID-19 new cases in top 10 infected African Countries using regression and time series models. medRxiv preprint, Infectious Diseases <https://doi.org/10.1101/2020.09.23.20200113> (2020)
31. Zebin, Z. *et al.* Prediction of the COVID-19 spread in African countries and implications for prevention and control: A case study in South Africa, Egypt, Algeria, Nigeria, Senegal and Kenya. *Sci. Total Environ.* <https://doi.org/10.1016/j.scitotenv.2020.138959> (2020).
32. Achoki, T. & Alam, U. *et al.* COVID-19 pandemic in the African continent: Forecasts of cumulative cases, new infections, and mortality. medRxiv.preprint <https://doi.org/10.1101/2020.04.09.20059154> (2020)
33. Hamidouche, M. COVID-19 outbreak in Algeria: A mathematical Model to predict cumulative cases. *J. Contemp. Stud. Epidemiol. Public Health* <https://doi.org/10.30935/jconsep/8451> (2020).
34. Balah, B. & Djeddou, M. Forecasting COVID-19 new cases in Algeria using Autoregressive fractionally integrated moving average Models (ARFIMA). medRxiv preprint <https://doi.org/10.1101/2020.05.03.20089615> (2020).
35. Bentout, S., Chekroun, A. & Kuniya, T. Parameter estimation and prediction for coronavirus disease outbreak 2019 (COVID-19) in Algeria. *AIMS Public Health* 7(2), 306–318. <https://doi.org/10.3934/publichealth.2020026> (2020).
36. Belkacem, S. COVID-19 data analysis and forecasting: Algeria and the world. arXiv:2007.09755v2 [stat.AP] (2020).
37. Lounis, M. & Bagal, D. K. Estimation of SIR model's parameters of COVID-19 in Algeria. *Bull Nat Res Cent* 44, 180. <https://doi.org/10.1186/s42269-020-00434-5> (2020).
38. Djouadi, M. S., Sebbagh, A. & Berkani, D. A Nonlinear algorithm for maneuvering target visual-based tracking. In *IEEE Proceedings of the 2nd International Conference on Intell Sens and Infor Proc, ICISIP, Chennai, India* 61–66 (2005).
39. Gordon, N. J., Salmond, D. J. & Smith, A.F.M. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. Radar and Signal Processing. In *IEE Proceedings F*, Vol. 140 107–113 (1993).
40. Gannot, S. & Yeredor, A. The Kalman filter. In *Springer Handbook of Speech Processing* (eds Jacob Benesty, M. *et al.*) 135–160 (Springer, Berlin, Heidelberg, 2008). https://doi.org/10.1007/978-3-540-49127-9_8.
41. Oravec, M., Rozinaj, G. & Beszed, S. M. Detection and recognition of human faces and facial features. In *Speech Audio, Image and Biomedical Signal Processing Using Neural Networks* (eds Prasad, B. & Prasanna, S. M.) 283–301 (Springer, Berlin, Heidelberg, 2008).
42. Arulampalam, M. S., Maskell, S., Gordon, N. & Clapp, T. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Trans. Signal Process.* 50(2), 174–188 (2002).
43. Fox, D., Hightower, J., Liao, L., Schulz, D. & Borriello, G. Bayesian filters for location estimation. *IEEE Pervas. Comput. Mag.* <https://doi.org/10.1109/MPRV.2003.1228524> (2003).
44. Polec, J., Ondrusova, S., Kotuliakova, K. & Karlubikova, T. Hierarchical transform coding using NURBS approximation. In *Proceedings Elmar-2008: 50th International Symposium ELMAR-2008, Zadar, Croatia*, Vol. 1 65–68, ISBN 978-953-7044-09-1 (2008)
45. Gao, Z. W. & Lie, W. N. Video error concealment by using Kalman filtering technique. In *Proceedings of the 2004 International Symposium on Circuits and Systems*, Vol. 2 69–72 ISCAS apos. (2004).
46. Jan, M., Stanislav, M. & Pavol, K. Bayesian filtering techniques: Kalman and extended Kalman filter basics. In *19th IEEE International Conference Radio elektronika, Bratislava, Slovakia* (2009).
47. Sebbagh, A., Djouadi, M. S. & Berkani, D. IMM-UKF algorithm and IMM-EKF algorithm for tracking highly maneuverable target: A comparison. In *ICSIT'05, International Conference on Computer Systems and Information Technology* 527–532, 19–21 July, Algiers, Algeria (2005).

Author contributions

All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-16496-6>.

Correspondence and requests for materials should be addressed to A.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022