Infectious Disease Modelling 8 (2023) 341–355

Contents lists available at ScienceDirect

Infectious Disease Modelling

journal homepage: www.keaipublishing.com/idm

Early estimation of the number of hidden HIV infected subjects: An extended Kalman filter approach★

Paolo Di Giamberardino, Daniela Iacoviello*

Dept. of Computer, Control and Management Engineering, Sapienza, University of Rome, Via Ariosto 25, 00185, Rome, Italy

ARTICLE INFO

Article history:
Received 12 January 2023
Received in revised form 16 February 2023
Accepted 7 March 2023
Available online 11 March 2023
Handling Editor: Dr Yijun Lou

Keywords:
Epidemic modeling
Infection spread
HIV-AIDS
Extended Kalman filter

ABSTRACT

In the last decades several epidemic emergencies have been affecting the world, influencing the social relationships, the economics and the habits. In particular, starting in the early '80, the Acquired Immunodeficiency Syndrome, AIDS, is representing one of the most worrying sanitary emergency, that has caused up to now more than 25 million of dead patients. The infection is caused by the Human Immunodeficiency Virus, HIV, that may be transmitted by body fluids; therefore with wise behaviours the epidemic spread could rapidly be contained. This sanitary emergency is peculiar for the long incubation time: it can reach even 10 years, a long period in which the individual can unconsciously infect other subjects. The identification of the number of infected unaware people, mandatory to define suitable containment measures, is here obtained by using the extended Kalman filter applied to a noisy model in which, reasonably, only the number of infected diagnosed patients is available. Numerical simulations and real data analysis support the effectiveness of the approach.

1. Introduction

In the last years, mathematical modeling has shown its power in describing the evolution of an epidemic spread, predicting the number of possible patients and the effects of containment measures. Usually, compartmental models are adopted, in which the population is partitioned into different groups according to specific conditions with respect to the infection evolution.

Starting from the basic SIR models (Di Giamberardino & Iacoviello, 2017; Ledzewicz & Schattler, 2011, pp. 981–990), with the compartments of Susceptible, Infected and Recovered individuals, improvements and enrichments are proposed to better classify the subjects, so getting, for example, SEIR models (Khan et al., 2015; Yang et al., 2020) adding the Exposed individuals compartment, SIRC models (Casagrandi, Bolzoni, Levin, & Andreasen, 2006; Iacoviello & Stasio, 2013) with the Cross–immune persons group, SIS models (Sanatkar, White, Natarajan, Scoglio, & Garrett, 2016), where the Cross–immune individuals are directly reconsidered as Susceptible ones, or SIRD models (Borri, Palumbo, Papa, & Possieri, 2021; Villaverde & Jones, 2020) which explicitly introduce the Dead individuals.

★ Peer review under responsibility of KeAi Communications Co., Ltd.
This work was supported by Sapienza Grant n.RM120172A824EB2C.

* Corresponding author.
E-mail addresses: daniela.iacoviello@uniroma1.it, iacoviello@diag.uniroma1.it (D. Iacoviello).

https://doi.org/10.1016/j.idm.2023.03.001

When dealing with particular epidemics, additional compartments are introduced to include specific health conditions or behavioural attitudes; with reference to three of the most dangerous virus spread emergencies currently present, it is possible to refer to (Attaullah & Sohaib, 2020; Ayele, Goufo, & Mugisha, 2021; Boonyaprasorn, Ngiamsunthorn, & Sethaput, 2016; Di Giamberardino, Compagnucci, De Giorgi, & Iacoviello, 2019; Di Giamberardino & Iacoviello, 2018a, 2018b; Sun, Nishiura, & Xiao, 2020) for the HIV-AIDS emergency, to (Assefa et al., 2022; Contreras, Villavicencio, Median-Ortiz, Biron-Lattes, & Olivera-Nappa, 2020; Di Giamberardino et al., 2021a, 2021b; Di Giamberardino & Iacoviello, 2021) when dealing with COVID-19, or, for the measles epidemic spread, to (Adewale, Olopade, Ajao, & Adeniran, 2016; Momoh, Ibrahim, Uwanta, & Manga, 2013; Pang, Ruan, Liu, Zhao, & Zhang, 2015).

In particular, among the pandemics, since 1982 the Acquired Immunodeficiency Syndrome, AIDS, is representing one of the most worrying sanitary emergency, that has caused up to now more than 25 million of dead subjects. Only in 2020, in the world 1.7 million of new diagnosis have been registered with about 680000 dead patients. It has been noted that since 2015 there has been a dangerous increase in the number of patients that become aware of the infection in an advanced stage.

The infection is due to the Human Immunodeficiency Virus, HIV; it infects cells of the immune system, making the person more susceptible to infections. The HIV can be transmitted only by some body fluids: blood, semen, pre-seminal fluid, rectal and vaginal fluid, and breast milk. According to the World Health Organization (WHO), there are several ways to prevent the HIV transmission: practice safe sexual behaviour, get tested for HIV, avoid injecting drugs or always use sterile needles and syringes, and ensure that any blood products to be used are tested regularly. Moreover, if one has HIV start antiretroviral therapy as soon as possible for his health and to prevent HIV transmission. The peculiarity of HIV-AIDS pandemic is that with proper precautions, the spread could stop immediately making realistic the goal of WHO of ending AIDS by 2030.

The HIV infection degenerates in AIDS that can be reached even after 10–15 years from the infection. Depending on its rate of progression, the following situations are possible, (Zeller, McCain, & Swanson, 1996): rapid progression and intermediate progression corresponding to the development of AIDS after three years or slower (between three and ten years from the infection). A third advisable condition is possible, the long term non progressions (LTNP), regarding less than 5% of the population: in this case the HIV infected patients still preserve an high number of CD4 and CD8 T-cell counts.

Such a number is an important indicator to classify the infected patients; in fact, from literature (Okoye & Picker, 2014), it is established that a subject is classified as a non infected patient if his CD4 T-cells count is in the range of 800–1050 per mm^3. Moreover, the ratio between the number of CD4 and CD8 cells is an alert for a possible HIV infection since in that case the lymphocytes CD4 are destroyed more rapidly with respect to the lymphocytes CD8 (Kumar, 2013).

As a result of recent advances in access to antiretroviral therapy (ART), HIV-positive people now live longer and healthier lives. In addition, it has been confirmed that ART prevents onward transmission of HIV; currently, almost 27.5 million people in the worlds are receiving the antiretroviral treatment (World Health Organization, 2023).

Mathematical modeling of HIV-AIDS epidemic has been developed in two main directions, at a cellular and at the population levels. In the first framework, the evolution of the number of CD4 T-cells is studied, (Althaus & Boer, 2011; Boonyaprasorn et al., 2016).

In the description at population level, the community is partitioned into homogeneous groups with respect to some characteristics (Anderson, 1988; Papa et al., 2018; Sun et al., 2020; Tuckwell & Corfec, 2016). This is the approach adopted in this paper, in which two main characteristics of HIV-AIDS are stressed: the strongly-behaviour dependent character of this epidemic spread and the presence of a significant number of asymptomatic infected patients. This latter fact suggests the introduction of a compartment containing individuals strongly involved in the epidemic spread whose number is not known being unaware of their status. Moreover, following the illness progression, they could possibly need future medical assistance, so involving the public healthcare system with consequent costs. As for the recent COVID-19 emergency, the number of infected asymptomatic subjects is an important information that, if available, could help in tuning the control actions. What typically happens, especially in a not-well informed society, is that most of the individuals in a population think to be not infected by the HIV, so avoiding a safe behaviour in their social relationships. The behavioural aspect, that strongly influences the HIV spread, is considered in the modelling adopted herein. Regarding the number of asymptomatic individuals, its knowledge could be very useful both for trying to control the actual spread and to forecast the required medical effort.

In this paper, an approach for the estimation of the number of asymptomatic individuals is proposed on the basis of a suitable mathematical model. The approach must take into account the fact that both the model itself and the data collected from monitoring the epidemic status suffer from the presence of uncertainties and errors, modelled as noise.

The problem has been already faced in literature: a suitable filtering approach can give satisfactory results, as in (Wang, Tse, & Wang, 2022) for epidemic spread over networks. A different solution can be represented by the extended Kalman filter (EKF); it is an improvement of the Kalman filter for nonlinear systems, linearizing dynamically, the nonlinear system model, making, therefore, possible to apply the classical Kalman estimation theory to obtain online state estimation. An early example of the use of a Kalman filter for a HIV compartmental model is (Cazelles & Chau, 1997), aimed at assessing the change in the dynamical evolution of the epidemic by estimating the model parameters. In (Hooshmand, Shariflan, & Shariflan, 2021, pp. 193–197), an extended Kalman particle filter is used to estimate the amount of healthy $CD4^+$ T cells, and the infected ones; the aim was to reduce the rate of increase of the virus. In (Zhu et al., 2021) the EKF is used, referring to COVID-19, to estimate both the model parameters and the transmission state of the virus spread. In (Lal, Huang, & Li, 2021) the ensemble Kalman filter is applied for parameters estimation referring to COVID-19; a simple SIRD model was assumed (Susceptible–Infected–Removed–Dead compartments) with unknown parameters, such as the contact and the recovery rates, as well as the death rate due to COVID-19.

In this paper, as a starting description the model proposed in (Di Giamberardino et al., 2019) is considered, using the compartmental modeling with susceptible and infected subjects; in particular, the category of susceptible subjects is split depending on the more or less wise attitude of the individuals with respect to the possibility of infection. Moreover, the category of infected patients is divided into three compartments: the infected individuals that still don't know to have the infection and therefore can, unconsciously, spread the virus; the category of tested patients in the pre-AIDS condition; the category of patients in the AIDS state.

The paper is organized as follows: in Section 2 a HIV-AIDS discrete noisy time model is proposed and analysed in the deterministic case; successively, the extended Kalman filter is designed and implemented. Numerical results are discussed in Section 4, considering as reference population the italian region. Nevertheless the conclusions, outlined in Section 5, along with future works, can be applied in general cases.

2. Materials and methods

The characteristics of the HIV–AIDS spread emphasize, more than for any other epidemic diseases, the behavioural aspects of each individual; in fact, the awareness of the risks and the consequent wise behaviours would strongly decrease the spread of the HIV virus, reducing the new infections. This observation has suggested in (Di Giamberardino et al., 2019) the split of the healthy subjects into two categories, the ones that with their behaviour try to keep themselves safe from the infection and the ones that, for a lack of information or by unconsciousness, are more exposed to contagion. Another important peculiarity of the HIV infection is that an infected subject could be almost asymptomatic for 10–15 years with two main consequences: the first for the epidemic spread, due to the unconscious possibility of infecting people, and the second for the personal health and dangerousness of the infection evolution because of the delay at the beginning of a proper therapy.

2.1. The mathematical model

Based on the above observations, the model proposed in (Di Giamberardino et al., 2019) is here considered, with the population split into five compartments:

S_1: the compartment of susceptible subjects, composed of the healthy part of the population that can underestimate the risks of the infection;
S_2: compartment of susceptible individuals that give great attention to prevention, thus avoiding potential dangerous conditions;
I: compartment of infected patients still not aware of the infection;
P: compartment of patients with a positive HIV diagnosis;
A: compartment of infected subjects with an AIDS diagnosis.

Depending on his behaviour, an healthy subject is in the S_1 or S_2 compartment. If he is in the former he can be infected and transit in the I compartment, still thinking to be healthy and in the dangerous condition of infecting other subjects and delaying possible medications. As said, he can live without sympthoms up to 10 years and then receive a diagnosis of positiveness to the HIV or even to AIDS. Three controls are introduced, representing the possible actions that can be performed to prevent the contagion or to treat the infected individuals: the informative and testing campaigns, denoted by u_1 and u_2, respectively, for the prevention, and u_3 represents the therapy. The first two act directly on the categories S_1, S_2 and I, whereas the third action influences the evolutions of the subjects of the infected compartments P and A.

It is chosen a modeling in the discrete domain for a realistic description; in fact, the HIV–AIDS epidemic has a slow dynamics, and its data are generally collected and described referring to a long time interval, usually with annual unity of time.

In the model, stochastic elements are introduced to take into account all the uncertainties in the parameters estimation as well as in the relationships among the state variables introduced.

The complete noisy controlled discrete system proposed is described by the following difference equations:

$$S_1(k+1) = S_1(k) + Z - \beta\frac{S_1(k)I(k)}{N(k)} - dS_1(k) + \gamma S_2(k) - r_1 S_1(k)u_1(k) + w_1(k) \tag{1}$$

$$S_2(k+1) = S_2(k) - dS_2(k) - \gamma S_2(k) + r_1 S_1(k)u_1(k) + w_2(k) \tag{2}$$

$$I(k+1) = I(k) + \beta\frac{S_1(k)I(k)}{N(k)} - dI(k) - \delta I(k) - r_2\frac{I(k)}{N(k)}u_2(k) + w_3(k) \tag{3}$$

$$P(k+1) = P(k) - dP(k) - \alpha_1 P(k) + \varepsilon\delta I(k) + r_2\varphi\frac{I(k)}{N(k)}u_2(k) + r_3 P(k)u_3(k) + w_4(k) \tag{4}$$

$$A(k+1) = A(k) - dA(k) + \alpha_1 P(k) + (1-\varepsilon)\delta I(k) - \alpha A(k) + r_2(1-\varphi)\frac{I(k)}{N(k)}u_2(k) - r_3 P(k)u_3(k) + w_5(k), \tag{5}$$

where $N(k) = S_1(k) + S_2(k) + I(k)$ is the total number of subjects potentially involved in risky contacts. It is assumed that $N(k) > 0\ \forall k$ under the practical consideration that it is not possible to have, at a certain time, all the population composed by diagnosed infected individuals; actually, the realistic hypothesis that $S_1(k) + S_2(k) \gg P(k) + A(k)$ is assumed.

The parameters introduced in the model have the following meaning: $d \in [0, 1]$ represents the fraction of subjects in all the compartments that die every unit of time for causes different from the AIDS pathology; β is the contact rate related with the possibility to become infected after a dangerous contact; Z is the number of new incomers (as newborns and immigrants), in the unit of time, assumed with the behavioural characteristics of the compartment S_1 only; $\gamma \in [0, 1]$ is the fraction of individuals in S_2 in the unit of time that, for some reasons, relax their attention and can behave like the ones in S_1 compartment; $\delta \in [0, 1]$ is the fraction of subjects that, in the unit of time, discover their infective condition for the appearance of some symptoms; according to the level of infection discovered, a fraction ε transits from the compartment I to P, and the remaining fraction $(1-\varepsilon)$ moves to A; $\alpha_1 \in [0, 1]$ is the fraction of individuals that, in the unit of time, goes in the A compartment coming from the compartment P because of a progression of the infection; $\alpha \in [0, 1]$ represents the fraction of extra death for the patients in the A compartment due to the direct infection. The constants r_i, $i = 1, 2, 3$ are related with the success and the effectiveness of the control actions; $\varphi \in [0, 1]$ is the probability that a subject in the I compartment receives, after testing (i.e. with the application of control u_2) a diagnosis of HIV positiveness and, therefore, transfer in the P class, while $(1-\varphi)$ holds for AIDS positivity and for transfer to A.

As said, the system considered is affected by noise:

$$W(k) = (w_{S_1}(k) \quad w_{S_2}(k) \quad w_I(k) \quad w_P(k) \quad w_A(k))^T \tag{6}$$

that is assumed, for simplicity, gaussian and non correlated with diagonal covariance matrix:

$$Q = \begin{pmatrix} q_{S1}^2 & 0 & 0 & 0 & 0 \\ 0 & q_{S2}^2 & 0 & 0 & 0 \\ 0 & 0 & q_I^2 & 0 & 0 \\ 0 & 0 & 0 & q_P^2 & 0 \\ 0 & 0 & 0 & 0 & q_A^2 \end{pmatrix}.$$

In Fig. 1 the block diagram of the adopted model is shown.

It describes current condition; a subject without specific symptoms is assumed to belong to one of the susceptible classes S_1 and S_2 or, if infection is already present, to the infected class I. Among such asymptomatic individuals, there is no way to distinguish directly between infected and uninfected without a (blood) test. To overcome this problem, in the present paper it is proposed an approach for estimating the number infected individuals $I(t)$, information crucial to evaluate the spread and the severity of the epidemics. It has been said that for susceptible uninfected individuals two possible attitudes are considered: a wise one, avoiding the contagious risky behaviours (S_2), and a risk underestimation approach to dangerous behaviours (S_1). In presence of informative campaign (u_1), an individual can change his attitudes, reducing the risks of infection; if some symptoms arise and/or a testing campaign is applied (u_2), a subject can have the reassurance on his state of

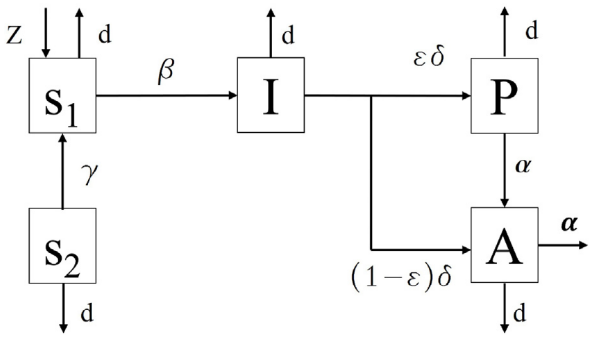

Fig. 1. Block diagram of the adopted HIV-AIDS model (Di Giamberardino et al., 2019).

health, or receive a positive diagnosis. In this case a suitable therapy can start, helping the patient in reaching the Long Term Non Progression State that avoids fatal consequences to the infected individuals.

The proposed model in absence of control and with the presence of infected patients will lead to the progressive emptying of the class S_2 and in a more or less long period most of the healthy individuals would become infected; this mainly depends on the contact rate β and on δ, the rate with which symptoms appear and an infected patient in I becomes aware of having got the infection, and therefore transfers to the class P or A. The control actions are able to interrupt this process; as already noted, from the very beginning of the start of this pandemic (early '80) up to nowadays, the increased knowledge on the spread modalities as well as the new medications suggest that the containment measures u_i, $i = 1, 2, 3$, dependent on time, have slow changes or show almost a piecewise constant evolution, so that they can be considered constant over specific periods.

It is useful to describe the proposed nonlinear noisy system (1)–(5) in a compact way, introducing the state vector

$$X(k) = (\, S_1(k) \quad S_2(k) \quad I(k) \quad P(k) \quad A(k) \,)^T. \tag{7}$$

Let $F(X) = (\, F_1(X) \quad F_2(X) \quad F_3(X) \quad F_4(X) \quad F_5(X) \,)^T$ be the vector function which collects all the uncontrolled and not noisy terms of equations (1)–(5), while the matrix $G(X(k))$ defines the functional relationship between the state $X(k)$ and the control vector $U(k) = (\, u_1(k) \quad u_2(k) \quad u_3(k) \,)^T$:

$$G(X(k)) = \begin{pmatrix} -r_1 S_1(k) & 0 & 0 \\ r_1 S_1(k) & 0 & 0 \\ 0 & -r_2 \dfrac{I(k)}{N(k)} & 0 \\ 0 & r_2 \varphi \dfrac{I(k)}{N(k)} & r_3 P(k) \\ 0 & r_2 (1-\varphi) \dfrac{I(k)}{N(k)} & -r_3 P(k) \end{pmatrix}.$$

Therefore, the discrete nonlinear noisy system can be written in the compact form:

$$X(k+1) = F(X(k)) + G(X(k))U(k) + W(k) = f(X(k), U(k)) + W(k) \tag{8}$$

with the obvious meaning of the vector $f(X(k), U(k))$; note that the relation between the state and the control is affine.

As already noted, the only available and measurable information regards the number of infected patients, that is, in the proposed framework, the number of subjects in the P and A compartments. The remaining individuals are supposed to be healthy, and possibly also wise-healthy, that is in the S_2 class. This justify the following choice regarding the measure equation:

$$Y(k) = H(X(k), V(k)) = P(k) + A(k) + V(k) \tag{9}$$

with V measuring noise with covariance R, uncorrelated with W.

The unawareness of the real epidemic situation implies that the number of people potentially at risk could be higher than the hypothesized one. In particular, the number of infected patients in the compartment I, due to the delayed diagnosis of infection, can be strongly underestimated, with evident problems in the sanitary system organization and, of course, in facing the infection for the unaware individuals.

2.2. Positiveness analysis

The mathematical model (1)–(5) is consistent with the non-negativeness of the state space variables. It is possible to verify the following Proposition.

Proposition 1. *Sufficient conditions for having the non-negative subspace of the state space invariant for the non forced dynamics are:*

i. *the physical assumptions on positiveness of parameters;*
ii. *the physical assumptions on positiveness of controls;*
iii. *the conditions*
   a. $1 - d - \gamma \geq 0$,
   b. $1 - d - \alpha_1 \geq 0$,

c. $1 - d - \alpha \geq 0$,
d. $1 - d - \beta \geq 0$,
e. $1 - d - \delta \geq 0$.

Once the control actions are considered, the additional conditions on control amplitudes are:

f. $u_1(k) \leq \frac{1}{N(k)}(1-d)$,
g. $u_2(k) \leq \frac{N(k)}{r_2}(1-d-\delta+\beta)$,
h. $u_3(k) \leq \frac{\alpha_1}{r_3}$.

Proof: Conditions for non negativity of the state at time $k+1$ for any non-negative values at time k are:

$$S_1(k) + Z - \beta\frac{S_1(k)I(k)}{N(k)} - dS_1(k) + \gamma S_2(k) - r_1 S_1(k)u_1(k) \geq 0,$$
$$S_2(k) - dS_2(k) - \gamma S_2(k) + r_1 S_1(k)u_1(k) \geq 0,$$
$$I(k) + \beta\frac{S_1(k)I(k)}{N(k)} - dI(k) - \delta I(k) - r_2\frac{I(k)}{N(k)}u_2(k) \geq 0,$$
$$P(k) - dP(k) - \alpha_1 P(k) + \varepsilon\delta I(k) + r_2\varphi\frac{I(k)}{N(k)}u_2(k) + r_3 P(k)u_3(k) \geq 0,$$
$$A(k) - dA(k) + \alpha_1 P(k) + (1-\varepsilon)\delta I(k) - \alpha A(k) + r_2(1-\varphi)\frac{I(k)}{N(k)}u_2(k) - r_3 P(k)u_3(k) \geq 0$$

$\forall\ S_1(k), S_2(k), I(k), P(k)$ and $A(k) \geq 0$. In the uncontrolled case, they can be rewritten as

$$\begin{aligned}
Z + \left(1 - d - \beta\frac{I(k)}{N(k)}\right)S_1(k) + \gamma S_2(k) &\geq 0,\\
(1 - d - \gamma)S_2(k) &\geq 0,\\
\left(1 + \beta\frac{S_1(k)}{N(k)} - d - \delta\right)I(k) &\geq 0,\\
(1 - d - \alpha_1)P(k) + \varepsilon\delta I(k) &\geq 0,\\
(1 - d - \alpha)A(k) + \alpha_1 P(k) + (1-\varepsilon)\delta I(k) &\geq 0.
\end{aligned}$$

recalling that

$$\sum_i c_i N_i \geq 0 \qquad \forall N_i \geq 0$$

if and only if $c_i \geq 0\ \forall i$, from the conditions above one can get directly, in addition to the positiveness of all the parameters already stated, the constraints

$(1 - d - \gamma) \geq 0$,
$(1 - d - \alpha_1) \geq 0$,
$(1 - d - \alpha) \geq 0$,

and

$$\left(1 - d - \beta\frac{I(k)}{N(k)}\right) \geq 0,$$
$$\left(1 - d - \delta + \beta\frac{S_1(k)}{N(k)}\right) \geq 0,$$

for which, since $0 \leq \frac{I(k)}{N(k)} \leq 1$ and $0 \leq \frac{S_1(k)}{N(k)} \leq 1$, the sufficient conditions

$(1 - d - \beta) \geq 0$
$(1 - d - \delta) \geq 0$

are obtained. In the forced case, the conditions involve also the controls; they become

$$\begin{aligned}
\left(1-d-r_1u_1(k)-\beta\frac{I(k)}{N(k)}\right) &\geq 0,\\
\left(1+\beta\frac{S_1(k)}{N(k)}-d-\delta-r_2\frac{1}{N(k)}u_2(k)\right) &\geq 0,\\
(1-d-\alpha_1+r_3u_3(k)) &\geq 0,\\
\left(\varepsilon\delta+r_2\varphi\frac{1}{N(k)}u_2(k)\right) &\geq 0,\\
(\alpha_1-r_3u_3(k)) &\geq 0,\\
\left((1-\varepsilon)\delta+r_2(1-\varphi)\frac{1}{N(k)}u_2(k)\right) &\geq 0,
\end{aligned}$$

from which

$$\begin{aligned}
u_1\left(k\right) &\leq \frac{1}{r_1}\left(1-d-\beta\frac{I(k)}{N(k)}\right)\leq\frac{1}{r_1}(1-d),\\
u_2(k) &\leq \frac{N(k)}{r_2}\left(1-d-\delta+\beta\frac{S_1(k)}{N(k)}\right)\leq\frac{N(k)}{r_2}(1-d-\delta+\beta),\\
u_3(k) &\leq \frac{\alpha_1}{r_3}.
\end{aligned}$$

Moreover, thanks to the positivity of all the parameters involved, from the results stated in Proposition 1 it is possible to verify also that:

Corollary 1. *The conditions a. − e. in Proposition 1 can be additionally bounded as*

a'. $0 \leq 1 - d - \gamma < 1$,
b'. $0 \leq 1 - d - \alpha_1 < 1$,
c'. $0 \leq 1 - d - \alpha < 1$,
d'. $0 \leq 1 - d - \beta < 1$,
e'. $0 \leq 1 - d - \delta < 1$.

2.3. Stability analysis

In the following, an analysis of the proposed model (8) is reported for the uncontrolled case without noise, represented by

$$X(k+1) = F(X(k)). \tag{10}$$

Its equilibrium point $\overline{X}$ must satisfy the condition

$$X(k+1) = X(k) = \overline{X} = (\,\overline{S_1}\quad \overline{S_2}\quad \overline{I}\quad \overline{P}\quad \overline{A}\,)^T \tag{11}$$

corresponding to

$$0 = Z - d\overline{S_1} - \beta\frac{\overline{S_1}\overline{I}}{\overline{N}} + \gamma\overline{S_2}, \tag{12}$$

$$0 = -d\overline{S_2} - \gamma\overline{S_2}, \tag{13}$$

$$0 = -d\overline{I} - \delta\overline{I} + \beta\frac{\overline{S_1}\overline{I}}{\overline{N}}, \tag{14}$$

$$0 = -d\overline{P} - \alpha_1\overline{P} + \varepsilon\delta\overline{I}, \tag{15}$$

$$0 = -d\overline{A} + \alpha_1\overline{P} + (1-\varepsilon)\delta\overline{I} - \alpha\overline{A}. \tag{16}$$

from equation (13) one has $\overline{S_2} = 0$, while (14) gives

$$\bar{I}\left(\beta\frac{\overline{S_1}}{\overline{N}} - d - \delta\right) = 0 \tag{17}$$

from which $\bar{I} = 0$ and

$$\beta\frac{\overline{S_1}}{\overline{N}} - d - \delta = 0 \tag{18}$$

are obtained. The first solution corresponds to the so called *disease free equilibrium point*; it can be fully computed by substitution of $\bar{I} = 0$ in equations (12), (15) and (16), yielding

$$\overline{X}_{DFE} = \begin{pmatrix} \frac{Z}{d} & 0 & 0 & 0 & 0 \end{pmatrix}^T. \tag{19}$$

on the other hand, from (18), the solution

$$\overline{S_1} = \frac{d+\delta}{\beta - d - \delta}\bar{I} \tag{20}$$

is computed, yielding to the so called *endemic equilibrium*. Recalling that the quantities involved, $\overline{S_1}$ and $\bar{I}$, must be non negative, the value (20) is feasible, and then acceptable, if and only if

$$\beta > d + \delta. \tag{21}$$

Equation (15) allows to express $\overline{P}$ as a linear function of $\bar{I}$:

$$\overline{P} = \frac{\delta\varepsilon}{d+\alpha_1}\bar{I} \tag{22}$$

whereas from (16) it can be deduced:

$$\overline{A} = \frac{1}{d+\alpha}\left[\frac{\alpha_1 + \varepsilon\delta}{d+\alpha_1} + (1-\varepsilon)\delta\right]\bar{I}. \tag{23}$$

Note that $\overline{P}$ in (22) and $\overline{A}$ in (23) are always acceptable if $\bar{I} > 0$. By substituting (20), (22) and (23) into (12) it results:

$$\bar{I} = \frac{Z(\beta - d - \delta)}{(d+\delta)(\beta - \delta)} \tag{24}$$

that is acceptable (it is a positive value) once condition (21) is verified. Therefore, denoting by $\bar{I}_{EE}$ the quantity in (24), the endemic equilibrium point is:

$$\overline{X}_{EE} = \begin{pmatrix} \frac{d+\delta}{\beta - d - \delta}\bar{I}_{EE} & 0 & \bar{I}_{EE} & \frac{\delta\varepsilon}{d+\alpha_1}\bar{I}_{EE} & \frac{1}{d+\alpha}\left[\frac{\alpha_1+\varepsilon\delta}{d+\alpha_1} + (1-\varepsilon)\delta\right]\bar{I}_{EE} \end{pmatrix}^T \tag{25}$$

The local condition of stability of the equilibrium points can be established by analysing the Jacobian of the system (10); in particular, evaluating the Jacobian in X_{DFE}, the following matrix is obtained:

$$J_{DFE} = \left(\begin{array}{cc|ccc} 1-d & \gamma & -\beta & 0 & 0 \\ 0 & 1-d-\gamma & 0 & 0 & 0 \\ \hline 0 & 0 & 1-d-\delta+\beta & 0 & 0 \\ 0 & 0 & \epsilon\delta & 1-d-\alpha_1 & 0 \\ 0 & 0 & (1-\epsilon)\delta & \alpha_1 & 1-d-\alpha \end{array}\right).$$

Due to the block structure, its eigenvalues are the elements on the diagonal, $1-d$, $1-d-\gamma$, $1-d-\delta+\beta$, $1-d-\alpha_1$, $1-d-\alpha$ and the stability is obtained once the modulus of all the eigenvalues of the Jacobian is smaller than one.

Thanks to the result in Corollary 1, the only condition still to be checked for the local asymptotic stability of the X_{DFE} is

$$|1 - d - \delta + \beta| < 1. \tag{26}$$

positivity of all the coefficients and the constraints on d and δ implies that the condition for local asymptotic stability of the disease free equilibrium point is

$$\beta < d + \delta \tag{27}$$

holding when condition (21) is violated. This fact can be summarised in the following proposition.

Proposition 2. *The Disease Free Equilibrium point is locally asymptotically stable if and only if it is the unique equilibrium point.*

2.4. The basic reproduction number

A useful parameter able to yield information on the evolution of the epidemic is the basic reproduction number defined as the number of secondary cases that a unique infected subject can provoke in his infective period in a population of susceptible individuals. As described in (Martcheva, 2015), a possible estimation of this parameter can be obtained by means of the next generation matrix approach. It is useful to reorder the state variable X and the system (10) so that the first three components are the subjects in the compartments I, P and A and the other two the individuals not infected, the ones in S_1 and S_2; thus $\tilde{X} = (I \quad P \quad A \quad S_1 \quad S_2)^T$ and $\tilde{F} = (F_3 \quad F_4 \quad F_5 \quad F_1 \quad F_2)^T$. Linearizing the system in a neighbourhood of the disease free equilibrium point (in the new coordinates $\tilde{X}_{DFE} = \begin{pmatrix} 0 & 0 & 0 & \frac{Z}{d} & 0 \end{pmatrix}^T$), and evaluating the Jacobian $\tilde{J}$ in $\tilde{X}$, the discrete system gives:

$$\xi(k+1) = \tilde{J}\xi(k), \tag{28}$$

$\xi(k)$ being the perturbation; by easy calculations, one gets:

$$\tilde{J}_{DFE} = \left(\begin{array}{ccc|cc} \beta+ & 0 & 0 & 0 & 0 \\ \epsilon\delta & 1-d-\alpha_1 & 0 & 0 & 0 \\ (1-\epsilon)\delta & \alpha_1 & 1-d-\alpha & 0 & 0 \\ \hline -\beta & 0 & 0 & 1-d & \gamma \\ 0 & 0 & 0 & 0 & 1-d-\gamma \end{array}\right) \tag{29}$$

that can be associated to the structure

$$\tilde{J}_{DFE} = \begin{pmatrix} \tilde{J}_1 + \tilde{J}_2 & 0 \\ \tilde{J}_3 & \tilde{J}_4 \end{pmatrix}$$

where $\tilde{J}_1$ and $\tilde{J}_2$ are the sub-matrices of dimension 3×3, 0 the null 3×2 matrix and J_3 and J_4 are easily determined from (29). The sub-matrices $\tilde{J}_1$ and $\tilde{J}_2$ are known as the *fertility matrix* and the *transition matrix* respectively, defined as follows:

$$\tilde{J}_1 = \begin{pmatrix} \beta & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \tilde{J}_2 = \begin{pmatrix} 1-d-\delta & 0 & 0 \\ \varepsilon\delta & 1-d-\alpha_1 & 0 \\ (1-\varepsilon)\delta & \alpha_1 & 1-d-\alpha \end{pmatrix}.$$

The next generation matrix for the proposed discrete model is defined as:

$$\tilde{J}_1 \cdot (\mathbf{I} - \tilde{J}_2)^{-1} = \frac{\begin{pmatrix} \beta(d+\alpha_1)(d+\alpha) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}{(d+\delta)(d+\alpha_1)(d+\alpha)} \tag{30}$$

Where $\mathbf{I}$ is the 3×3 identity matrix. The basic reproduction number R_0 is defined as the spectral radius of $\tilde{J}_1 \cdot (\mathbf{I} - \tilde{J}_2)^{-1}$, that is

$$R_0 = \frac{\beta}{(d+\delta)}. \tag{31}$$

According to its definition, the basic reproduction number (31) is a measure of the spread of the infections since $R_0 < 1$ means that the infected individuals infect a smaller number of persons and, iteratively (asymptotically) the quantity of infected people goes to zero. On other hand, if $R_0 > 1$, each infected person transmit the infection to more than one person and in this case the number of infected, transmission after transmission, grows up.

By comparison of (27) in the previous Subsection 2.3, and (31), and for Proposition 2, it is possible to conclude the following:

Proposition 3. *The reproduction number R_0 is less than* 1 *if and only if the disease free equilibrium is locally asymptotically stable, that is when the endemic equilibrium does not exist.*

The reproduction number decreases as the rate δ increases, meaning that if the rate transfer from the compartment of unaware infected subjects, I, to the one of the P and A increases, the spread starts to decrease. The same happens if β is lowered, implying the suggestion to avoid dangerous conditions. The control actions, not included in the calculation of the basic reproduction number R_0, act exactly in these directions, that is: 1) try to avoid dangerous behaviours, and, therefore, in the proposed model, induce healthy people of compartment S_1 to transfer to compartment S_2, and 2) find, as rapidly as possible, by means of suitable test campaign, new infected people reducing the possibility of new infections and allowing the patients to start the therapy.

3. The extended Kalman filter

Starting from the state and the measure equations, expressions (8) and (9) respectively, it is possible, by using the extended Kalman filter theory, to estimate the state $X(k)$, starting from the available information, that is the number of patients in the HIV and AIDS conditions. The powerful instrument able to deal with this issue in a recursive way, the Kalman filter, provides equations to estimate the state of a system minimizing the mean of the squared error between the state and its estimation. The proposed system being non linear, an extended Kalman filter (EKF) will be used, that is the application of the classical Kalman filter to the system linearized around the current estimate.

The EKF considers two main steps: the *filtering*, or *measurement update*, and the *prediction*, or *time update*.

Let us denote by $\hat{X}(k|k-1)$ the estimation of the state $X(k)$ at time k based on the information available at time $k-1$: it represents the *prediction*; the updated estimation $\hat{X}(k|k)$ is obtained by using the new available measure $Y(k)$ and the last prediction $\hat{X}(k|k-1)$.

The quantity $f_X(\hat{X}(k|k))$ denotes the Jacobian of the function f in (8), with respect to the state X, evaluated in the estimation $\hat{X}(k|k)$; $h_X(\hat{X}(k|k-1))$ is the Jacobian of the measure Y, with respect to the state X, evaluated in the prediction $\hat{X}(k|k-1)$.

An important quantity in the Kalman filter is the *innovation*, that is the difference between the real measure and the measure that would be obtained if, instead of the true state, one used the predicted one:

$$e(k) = Y(k) - h(\hat{X}(k|k-1)). \tag{32}$$

The *a priori estimate error covariance* is:

$$P(k|k-1) = E[(X(k) - \hat{X}(k|k-1))(X(k) - \hat{X}(k|k-1))^T]$$

whereas the *a posteriori estimate error covariance* is indicated by:

$$P(k|k) = E[(X(k) - \hat{X}(k|k))(X(k) - \hat{X}(k|k))^T]. \tag{33}$$

In the *filtering* step (called also *correction*), the estimation is obtained as a correction of the prediction by means of the innovation, suitably weighted by the Kalman gain K:

$$\hat{X}(k|k) = \hat{X}(k|k-1) + K(k)e(k), \tag{34}$$

$$K(k) = P(k|k-1)(h_X(\hat{X}(k|k-1)))^T C_I(k)^{-1}, \tag{35}$$

where $C_I(k)$, the covariance of the innovation process, is given by:

$$C_I(k) = R(k) + (h_X(\hat{X}(k|k-1)))P(k|k-1)(h_X(\hat{X}(k|k-1)))^T.$$

By using the gain K (35), it is possible to write the recursive equation of the *a posteriori* error covariance (33):

$$P(k|k) = P(k|k-1) - K(k)C_I(k)K(k)^T.$$

In the *prediction* of the estimation, $\hat{X}(k+1|k)$ is approximated by

$$\hat{X}(k+1|k) = f(\hat{X}(k|k)) \tag{36}$$

that is by the value of the state that would be obtained if in equation (8), considered in the ideal case of absence of noise, one had the estimation $\hat{X}(k|k)$, instead of the true state $X(k)$.

Finally, the update equation of the *a priori* error covariance $P(k+1|k)$ is obtained by using the approximation:

$$X(K) - \hat{X}(k|k-1) \approx f_X(\hat{X}(k|k)[X(k) - \hat{X}(k|k)]) \tag{37}$$

thus having:

$$P(k+1|k) = f_X(\hat{X}(k|k))P(k|k)(f_X(\hat{X}(k|k)))^T + Q(k) \tag{38}$$

that is the same update equation of the standard Kalman filter.

In the next Section, the described Kalman filter algorithm is implemented to estimate the evolutions of all the state variables starting from the available measure of the number of diagnosed infected patients $P + A$; note that it is also possible to assume known separately the numbers of patients in P and in A.

4. Numerical results

As pointed out in the Introduction, the HIV-AIDS pandemic has been an emergency since the early '80, when only partial information on the infection and the modalities of diffusion were available. Also, due to inconsistent data collections regarding the infected patients in different countries, the difficulties in tracing the contacts after the discovery of the infection, as well as its strong dependency on the individuals' behaviour makes difficult to study the trend of HIV spread. The availability of a reliable model is a support in understanding the dynamics of the HIV emergency, in particular, the main goal of the paper is to determine an estimation of the number of subjects that are infected but still asymptomatic, corresponding, in the proposed model, to the individuals in compartment I. This information would allow a not delayed action to interrupt the spread by improving containment measures, (Di Giamberardino & Iacoviello, 2018a, 2018b). Moreover, also an estimation of the people in the S_1 and S_2 classes will be available, again allowing a proper tuning of the informative campaign. The tuning of the model parameters will refer to the situation in Italy, but of course the general rationale of the proposed choices could be applied to any country. For the data strictly related to the italian population the *Istituto Nazionale di Statistica* (ISTAT) website is considered (Istituto Nazionale di Statistica, 2023). The unit of time assumed is a year; as far as Z, the number of new incomers in one year is evaluated by adding the number of immigrants (in 2020 about 106503 residency permits) to the newborn (404892). To evaluate the percentage d of dead people (dead individuals in one year over the total population) the data from 2019 are considered to avoid bias due to the COVID−19 pandemic; therefore, $d = \frac{740317}{59.73 \cdot 10^6}$. The parameter δ, related to the natural transfer rate from the I compartment to the P or A ones, is assumed equal to $\frac{1}{10}$, since it has been studied that a subject can live up to $10 - 15$ years with the HIV infection without showing evident symptoms. To estimate the contact rate β, the information on the reproduction number R_0 was used, thanks to their relationship given by (31). This value may be strongly different from one country to the other, see (Williams & Gouws, 2013) for the estimated values in some regions of Africa. In Italy, it can be assumed a value of 2 (Istituto Superiore di Sanità. 2023). From formula (31) and the above estimated parameters d and δ, $\beta = R_0(d + \delta) = 0.2002$ can be determined. In 2019, there have been 2531 new diagnosis of HIV; about $\frac{1}{3}$ of the new patients discover to be positive for some symptoms. According to the proposed model, this number corresponds to: $\delta\varepsilon I(k-1) = \frac{2531}{3}$; at the same time about 571 diagnosis of AIDS have been registered, thus yielding to $\delta(1-\varepsilon)I(k-1) = 571$. From these two expressions it is possible to deduce the estimation of $\varepsilon = 0.59$. The number of dead patients infected by AIDS in 2019 was about 500. To estimate the parameter a denoting the extra percentage of dead people among the ones with AIDS, it has been considered that in Italy there have been about 71591 patients since 1982, of which 46000 died; assuming that about 15000 patients of AIDS are currently living nowadays, the parameter a is equal to 0.033. The parameter a_1, denoting the percentage of patients that transfer in one year from the P compartment to the A one is chosen equal to $\frac{1}{5}$ as a reasonable value with respect to the rapid and strong progression (the most common situations) and the long term non progression of the HIV disease.

The parameters chosen up to now are characteristic of the specific country, in this case Italy. The parameters that will be chosen in the following regard specific aspects of the model such as the natural transfer rate from the compartment S_2 to the S_1, that is γ, or the percentage φ of subjects that, after the control action u_2, transfer from the I compartment to the P one. The variation of these parameters will be studied to determine their influence on the evolution of the state models; as a first choice $\gamma = 0.2$ and $\varphi = 0.8$ are considered, assuming a small amount of subjects that transfer to the compartment of unwise healthy individuals and an high percentage of patients that after the analysis understand to have the infection (and therefore transit to the P class).

In the proposed approach the control actions are fixed, assuming piecewise expressions with increasing values. The rationale of these choices is in the increased effort in the informative and test campaigns, as well as in the improvement in the pharmacological treatments. The containment measures are assumed constant for a fixed number of years, $T = 4$, because it is not reasonable and efficient a continuous change of strategies. In Fig. 2 the proposed choices are shown. All the actions are assumed stronger as time passes, to follow the increased capabilities of intervention as the knowledge of the illness increases and the medical progress makes more drugs available.

The effectiveness of the proposed actions is weighted by r_i, $i = 1, 2, 3$, normalization factors that take into account the different nature of the applied controls, material or human resources; as a possible choice: $r_1 = 0.05$, $r_2 = 5 \cdot 10^3$, and $r_3 = 5 \cdot 10^{-4}$. The order of magnitude of these weights is due to their different action in the dynamical equations: r_1 and r_3 multiply directly the controls (respectively u_1 and u_3) and the number of subjects in one class, S_1 and P respectively, with order of magnitude, for example regarding S_1, of the entire population; on the other hand r_2 multiplies, besides the control u_2, the ratio between the number of infected subjects I and the ones on which the test (u_2) is applied, say the subjects in S_1, S_2 and I, that is a number lower than one.

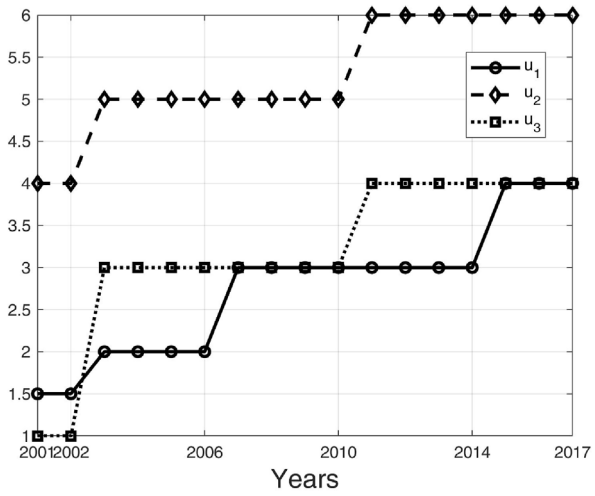

Fig. 2. Trends of the chosen control actions.

Finally, the uncertainties in the dynamical state equations of the model and in the measure equation are chosen in order to stress the partial absence of knowledge on the S_1 and S_2 relationships, as well as the strong indeterminateness on the number of subjects in I, whereas some more information are available on the patients in P and A. Consequently, the following choice is proposed for the matrix Q:

$$Q = \begin{pmatrix} 15 & 0 & 0 & 0 & 0 \\ 0 & 15 & 0 & 0 & 0 \\ 0 & 0 & 300 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 \\ 9 & 0 & 0 & 0 & 5 \end{pmatrix} \tag{39}$$

for the measure noise,a standard deviation equal to 50 is assumed.

In order to test whether the proposed description fits the unique available data (Istituto Superiore di Sanità. 2023; Surveillance Atlas of Infectious Diseases, 2023), the total number of patients with HIV-AIDS from 2001 to 2018 is compared with the output of the proposed model, that is $P + A$, see Fig. 3. In Fig. 4, the trends of the corresponding subjects in the S_1, S_2 and I compartments are shown. Note that the peak in the real data (and in the simulated ones) in the number of infected subjects, that can be seen in Fig. 3, is reasonably preceded by the increased number of infected subjects in the estimated trend of the individuals in I, Fig. 4, supporting the correctness and the effectiveness of the model adopted.The availability of a reliable model with the estimation of the number of infected unaware individuals is useful to study what would have happened with stronger informative and testing campaigns, keeping unchanged the effort of pharmacological therapy. By increasing the effectiveness of the control u_1 and u_2 of 5 times ($r_1 = 0.25$, $r_2 = 25 \cdot 10^3$) the reduction in the number of infected subjects in the compartment I is evident in Fig. 5.The effects can be also quantified by evaluating the total number of infected subjects in the two cases: the difference in the total number of infections is more than 401000. This is due mainly to the informative campaign u_1; in fact, with this stronger action, a high percentage of subjects in the S_1 class transfers to the S_2 one, thus slowing the infection, as shown in Fig. 6.

To emphasize the importance of the two campaigns (the informative and testing ones) three cases are considered: in the first one, it is increased significantly (10 times) the effort for the informative campaign; in the second case, it is increased significantly (again 10 times) the effort for the testing campaign; finally, in the third case, the increment is applied to both the campaigns, up to 10 times. In all the situations, the peak of the number of infected unaware individuals $I(t)$ occurs in the sixth year, as in the real case 2006, assumed as reference. To compare the effects of the chosen strategies in the three cases, two quantities are considered: the difference, in all cases, between the number of infected patients (aware and unaware, people in the I, P, and A compartments) at the end of the control period (year 2017) and the corresponding quantity in the real situation; the second indicator is the difference in the peaks, assuming always as reference the real situation.

In the first case, in 2017 there could have been a minor number of infected patients, about 10682 less than in the real case; with the strategy applied in the second case, the decrease obtained would have been more contained, about 4683 patients, whereas, as obvious, with the strategy hypothesized in the third case, the decrease in the number of infected patients would have been of more than 14200 units. This analysis suggests that the main action should be to induce people to avoid dangerous behaviours (and thus in the proposed model to transfer to the S_2 class). The comparison on the maximum number

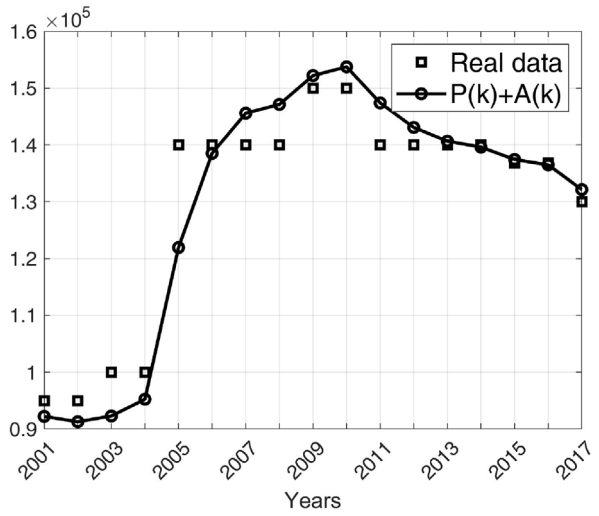

Fig. 3. Comparison between the real data of the total infected patients in Italy and the estimation obtained by applying the EKF approach.

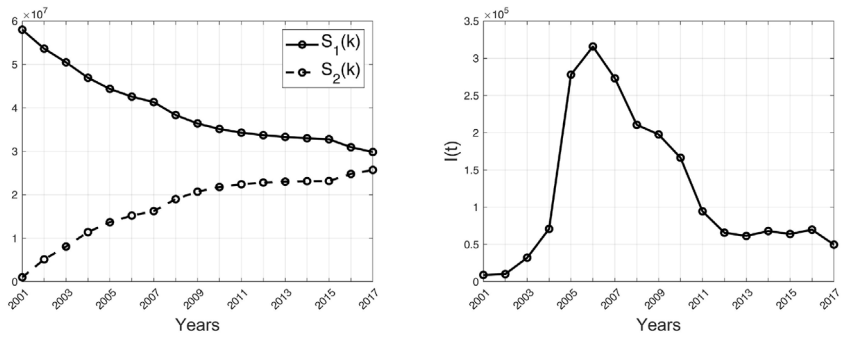

Fig. 4. Trend of the number of individuals in S_1 and S_2 (left) and I (right) in the simulated case relative to Fig. 3.

of infected patients is interesting also to tune the sanitary social effort. In the three cases the results show a decrease in the peak of 10682, 14751 and 85906 patients, respectively. Again, it can be stressed the importance of the informative campaign. The importance of such control is due to the specific characteristics of HIV-AIDS. Unlike other epidemic diseases, it is a *one way* disease, meaning that, at the present knowledge, an infected subject in I (or in P, A) can not go back to the healthy condition in S_1 or S_2. Recent improvements in the research are showing results that bode well for an effective vaccine, (International AIDS Vaccine Initiative, 2023); this would lead to a new scenario, making the goal of bringing down the epidemic by 2030, more reasonable.

5. Conclusions and future work

In this paper a discrete time noisy *stochastic* model of the diffusion of HIV-AIDS is introduced and analysed. The awareness of the infection is the first issue to be pursued to interrupt the spread and to start the antiretroviral therapy. Nevertheless, the absence of symptoms for long time (up to ten years) is real hindrance for the knowledge of the actual number of infected subjects and apply containment measures. In this paper, it is proposed the use of the extended Kalman filter to estimate the number of infected subjects that are still not aware of the infection. The first numerical results appear promising suggesting the following improvements:

- consider a specific population (or category in a population, for example young people) and identify the proposed model taking into account the specificity of the chosen group;
- identify also the possible sources of uncertainties;
- analyse the potentialities of the filtering operation, stressing the operative conditions to determine the limit of validity of the proposed approach;

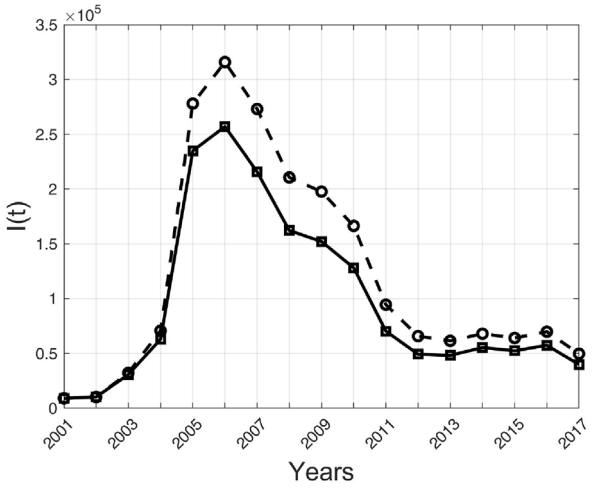

Fig. 5. Comparison between the trend of infected unaware subjects in the case relative to Fig. 3 (dashed line) and if applying stronger informative and testing campaigns (continuous line).

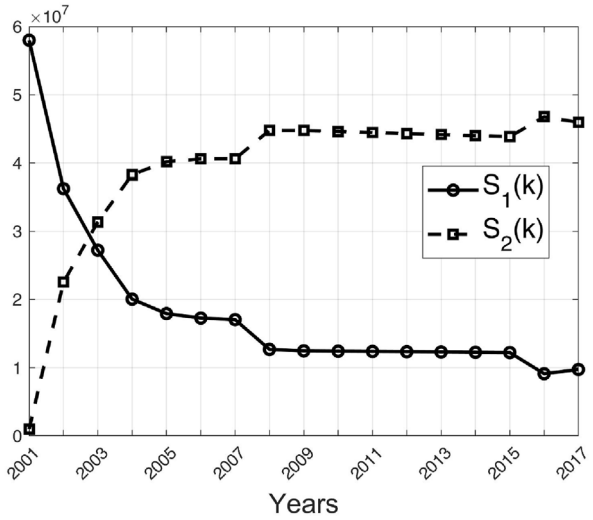

Fig. 6. Trend of the number of individuals in I when stronger control actions u_1 and u_2 are applied.

- introduce in the modeling the relations among populations, referring to contacts among groups of individuals of different countries.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Adewale, S. O., Olopade, I. A., Ajao, S. O., & Adeniran, G. A. (2016). Optimal control analysis of the dynamical spread of measles. *International Journal of research, 4*(5), 169–188.

Althaus, C. L., & Boer, R. J. (2011). Implications of CTL-mediated killing of HIV-infected cells during the non-productive stage of infection. *PLoS One, 6*(2), Article e16468, 02.

Anderson, R. M. (1988). The role of mathematical models in the study of HIV transmission and the epidemiology of AIDS. *Journal of Acquired Immune Deficiency Syndromes, 1*(3), 241–256.
Assefa, Y., Gilks, C. F., Reid, S., Van de Pas, R., Gete, D. G., & Van Damme, W. (2022). Analysis of the COVID-19 pandemic: Lessons towards a more effective response to public health emergencies. In *Globalization and health*BMC.
Attaullah, A., & Sohaib, M. (2020). Mathematical modeling and numerical simulation of HIV infection model. *Results in Applied Mathematics, 7*, 1–11.
Ayele, T., Goufo, E., & Mugisha, S. (2021). Mathematical modeling of HIV/AIDS with optimal control. *Results in Physics, 26*(104263), 1–14.
Boonyaprasorn, A., Ngiamsunthorn, P. S., & Sethaput, T. (2016). Synergetic control for HIV infection system of CD4+ T cells, International Conference on Control. *Automation and Systems*, 484–488.
Borri, A., Palumbo, P., Papa, F., & Possieri, C. (2021). Optimal design of lock-down and reopening policies for early-stage epidemics through sir-d models. In *Annual reviews in control*Elsevier.
Casagrandi, R., Bolzoni, L., Levin, S. A., & Andreasen, V. (2006). The sirc model and influenza. *A. Mathematical Biosciences, 200*, 152–169.
Cazelles, B., & Chau, P. (1997). Using the Kalman filter and dynamic models to assess the changing HIV/AIDS epidemic. *Mathematical Biosciences, 200*, 131–154.
Contreras, S., Villavicencio, H. A., Median-Ortiz, D., Biron-Lattes, J. P., & Olivera-Nappa, A. (2020). A multigroup SEIRA model for the spread of COVID-19 among heterogeneous populations. In *Caos solitons and fractals*Elsevier.
Di Giamberardino, P., Compagnucci, L., De Giorgi, C., & Iacoviello, D. (2019). Modeling the effects of prevention and early diagnosis on HIV/AIDS infection diffusion. *IEEE Transactions on Systems, Man and Cybernetics: Systems, 49*(10), 2119–2130.
Di Giamberardino, P., & Iacoviello, D. (2017). Optimal control of SIR epidemic model with state dependent switching cost index. *Biomedical Signal Processing and Control, 31*, 377–380.
Di Giamberardino, P., & Iacoviello, D. (2018a). LQ control design for the containment of the HIV/AIDS diffusion. *Control Engineering Practice, 77*, 162–173.
Di Giamberardino, P., & Iacoviello, D. (2018b). HIV infection control: A constructive algorithm for a state-based switching control. *International Journal of Control, Automation and Systems, 16*(3), 1469–1473.
Di Giamberardino, P., & Iacoviello, D. (2021). Evaluation of the effect of different policies in the containment of epidemic spreads for the COVID-19 case. In *Biomedical signal processing and control*Elsevier.
Di Giamberardino, P., Iacoviello, D., Papa, F., & Sinisgalli, C. (2021a). A data-driven model of the COVID-19 spread among interconnected populations: Epidemiological and mobility aspects following the lockdown in Italy. In *Nonlinear dynamics*Elsevier.
Di Giamberardino, P., Iacoviello, D., Papa, F., & Sinisgalli, C. (2021b). Dynamical evolution of COVID-19 in Italy with an evaluation of the size of the asymptomatic infective population. *IEEE Journal of Biomedical and Health Informatics, 25*(4), 1326–1332.
Hooshmand, M., Shariflan, M., & Shariflan, H. (2021). *HIV virus states estimation by extended Kalman particle filter, 29th Iranian conference on electrical engineering*. ICEE.
Iacoviello, D., & Stasio, N. (2013). Optimal control for sirc epidemic outbreak. *Computer Methods and Programs in Biomedicine*, 333–342.
International AIDS vaccine initiative. https://www.iavi.org, (2023).
Istituto Nazionale di Statistica. (2023). https://www.istat.it/en/.
Khan, M., Wahid, A., Islam, S., Khan, I., Shafie, S., & Gul, T. (2015). Stability analysis of an seir epidemic model with non-linear saturated incidence and temporary immunity. *International Journal of Advances in Applied Mathematics and Mechanics, 2*, 1–14.
Kumar, P. (2013). Long term non–progressor (LTNP) HIV infection. *Indian Journal of Medical Research, 138*(3), 291–293.
Lal, R., Huang, W., & Li, Z. (2021). An application of the ensemble kalman filter in epidemiological modelling. *PLoS One, 16*(8), 1–25.
Ledzewicz, U., & Schattler, E. (2011). *On optimal singular controls for a general SIR-model with vaccination and treatment, Discrete and continuous dynamical systems*.
Martcheva, M. (2015). An introduction to mathematical epidemiology. In *Text in applied mathematics* (Vol. 61)Springer.
Momoh, A. A., Ibrahim, M. O., Uwanta, I. J., & Manga, S. B. (2013). Mathematical model for control of measles epidemiology. *International Journal of Pure and Applied Mathematics, 87*(5), 707–718.
Okoye, O., & Picker, L. (2014). CD4+ T cell depletion in HIV infection: Mechanisms of immunological failure. *Immunological Reviews, 254*(1).
Pang, L., Ruan, S., Liu, S., Zhao, Z., & Zhang, X. (2015). Transmission dynamics and optimal control of measles epidemics. *Applied mathematics and computations, 256*, 131–147.
Papa, F., Binda, F., Felici, G., Franzetti, M., Gandolfi, A., Sinisgalli, C., et al. (2018). A simple model of HIV epidemic in Italy: The role of the antiretroviral treatment. In *Math. biosciences and engineering*Elsevier.
Sanatkar, M. R., White, W. N., Natarajan, B., Scoglio, C. M., & Garrett, K. A. (2016). Epidemic threshold of an SIS model in dynamic switching networks. *IEEE Trans. on Systems, Man, and Cybernetics: Systems, 46*(3), 345–355.
Sun, X., Nishiura, H., & Xiao, Y. (2020). Modeling methods for estimating HIV incidence: A mathematical review. *Theoretical Biology and Medical Modelling, 17*(1), 1–14.
Surveillance Atlas of infectious diseases. http://atlas.ecdc.europa.eu/, (2023).
Tuckwell, H. C., & Corfec, E. L. (2016). Stochastic model for early HIV-1 population dynamics. *Journal of Theoretical Biology, 195*.
Villaverde, J. F., & Jones, C. J. (2020). *Estimating and simulating a sird model of covid-19 for many countries, states, and cities*. NBER Working Paper SERIES.
Wang, W., Tse, C. K., & Wang, S. (2022). Generalized correntropy sparse gauss–hermite quadrature filter for epidemic tracking on complex networks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 52*(5), 2770–2778. https://doi.org/10.1109/TSMC.2021.3052538
Williams, B., & Gouws, E. (2013). R_0 and the elimination of HIV in Africa. http://arxiv.org/abs/1304.3720.
World Health Organization. (2023). https://www.who.int/.
Yang, Z., Zeng, Z., Wang, K., Wong, S. S., Liang, W., Zanin, M., et al. (2020). Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. *Journal of Thoracic Disease, 12*(3), 165–174.
Zeller, J. M., McCain, N. L., & Swanson, B. (1996). Immunological and virological markers of HIV-disease progression. *Journal of the Association of Nurses in AIDS Care, 7*, 15–27.
Zhu, X., Gao, B., Zhong, Y., Gu, C., & Choi, K. (2021). Extended Kalman filter based on stochastic epidemiological model for COVID-19 modeling. *Computers in Biology and Medicine, 137*, 1–10.
Istituto Superiore di Sanità.(2023). https://www.salute.gov.it/.