A System Dynamics model to predict the impact of COVID-19 in Spain

M.T. Sanz [♭1], A. Caselles [♮], J. C. Micó[♯] and C. Soler[♯♯]

(♭) Departament de Didàctica de la Matemàtica,
Universitat de València, Av. dels Tarongers, 4, 46022 València, Spain,
(♮) IASCYS member, Departament de Matemàtica Aplicada,
Universitat de València (retired), Dr. Moliner 50, 46100 Burjassot, Spain,
(♯)Institut Universitari de Matemàtica Multidisciplinar,
Universitat Politècnica de Valéncia, Camí de Vera s/n, 46022 Valéncia, Spain,
(♯♯)Institut Valencià de Cultura,
Generalitat Valenciana, Polígon Accés Ademús s/n, 46100 Burjassot, Spain.

1 Introduction

On 21 January 2020, the World Health Organization published its first report about the situation of the world with the Novel Coronavirus (2019-nCoV), a new human infectious disease. A total number of confirmed cases in that day was 282 and this disease was only present in China, Japan, Republic of Korea and Thailand. Nowadays, this organization has published its 94 report and it reflects that COVID-19 is around the world (Figure 1), the actual cases number is 2544792 and there are 175694 deaths.

Figure 1: Cases of COVID-19 on 23 April 2020. Obtained from `https://www.who.int`

[1]e-mail: m.teresa.sanz@uv.es

The scientific community concentrates its efforts on obtaining predictive models that can be an instrument to help governments to establish sanitary and social measures [1] or to identify cases as fast as possible [2]. In Spain, the Spanish Mathematics Committee calls on mathematicians and statisticians to participate in the construction of a meta-predictor to provide the authorities with information on the short-term behaviour of important variables in the spread of the COVID-19. The method uses the predictions from different models or algorithms that are provided by the participating researchers (`https://covid19.citic.udc.es/`), and build optimized combinations of these, disaggregated by Autonomous Communities. Thus, in order to help our Government in this fight, the DinSisCovid19 group is created, the members of which are the authors of this paper.

2 The model

The model presented in this research is an extension of the simplest model of the epidemic outbreak evolution, the SIR [3]. According to this model, the population can be divided into three different groups: a) susceptible (S), individuals without immunity to the infectious agent but they can be infected if exposed to the infectious agent; b) infected (I), individuals who are infected at any given time and can transmit the infection to susceptible population; and c) recovered (R), individuals who are immune to infection, and consequently the transmission does not affect when they are in contact with other individuals. This model provides a coupled system of three differential equations for the cited variables.

$$\frac{dS(t)}{dt} = -\beta \cdot S(t) \cdot I(t), \tag{1}$$

$$\frac{dI(t)}{dt} = \beta \cdot S(t) \cdot I(t) - \gamma \cdot I(t), \tag{2}$$

$$\frac{dN(t)}{dt} = \gamma \cdot I(t). \tag{3}$$

Being β the transmission rate, γ the recovery rate and $1/\gamma$ the average recovery period of the epidemic. These equations have been generalized in this paper in order to increase their predictive power. First, the population of the territory, $N(t)$, is considered. It is calculated as the initial population minus the deaths flow (4),

$$\frac{dR(t)}{dt} = -kf \cdot \left(\frac{E(t)}{\tau f} + S(t) \cdot E(t) \cdot \frac{ka}{\tau a} + S(t) \cdot R(t) \cdot \frac{kb}{\tau b} \right), \tag{4}$$

where, kf is the deaths rate of the sick population ($E(t)$); τf is the sick population deaths delay; and $S(t) \cdot E(t) \cdot ka/\tau a + S(t) \cdot R(t) \cdot kb/\tau b$, refers to those infected at each time step, i.e., the infection flow. The latter expression is used to calculate the infected population $I(t)$ that is modified in (2) through (5).

$$\frac{dI(t)}{dt} = S(t) \cdot \left(E(t) \cdot \frac{ka}{\tau a} + R(t) \cdot \frac{kb}{\tau b} \right). \tag{5}$$

The infected population come not only from the sick population but also from the recovered population ($R(t)$), and both interact with the susceptible population ($S(t)$) through the ratio of interaction rate and its corresponding delay, $ka/\tau a$ y $kb/\tau b$, respectively. Regarding the susceptible population, (1) is transformed into (6). In (1), a fixed initial susceptible population is changed through the infection flow. In the present model, we start from total population that is becoming susceptible due to the susceptibility rate (kq). It is calculated as the new susceptible population, $kq \cdot N(t)$, minus the infection flow.

$$\frac{dS(t)}{dt} = kq \cdot N(t) - S(t) \cdot \left(E(t) \cdot \frac{ka}{\tau a} + R(t) \cdot \frac{kb}{\tau b} \right). \tag{6}$$

Regarding the recovered population, equation (2) becomes (7). It is defined as the recovered rate, kr, multiplied by the sum of the infection flow and the sick population afected by the recovered population delay (τr).

$$\frac{dR(t)}{dt} = kr \cdot \left(S(t) \cdot E(t) \cdot \frac{ka}{\tau a} + S(t) \cdot R(t) \cdot \frac{kb}{\tau b} + \frac{E(t)}{\tau r} \right). \tag{7}$$

The deaths flow allows calculating the total deaths of the system (8).

$$\frac{dD(t)}{dt} = kf \cdot \left(S(t) \cdot E(t) \cdot \frac{ka}{\tau a} + S(t) \cdot R(t) \cdot \frac{kb}{\tau b} + \frac{E(t)}{\tau f} \right). \tag{8}$$

Finally, the sick population is calculated through the infected flow minus deaths and recovered flow.

$$\begin{aligned} \frac{dE(t)}{dt} = & \left(S(t) \cdot (E(t) \cdot \frac{ka}{\tau a} + S(t) \cdot R(t) \cdot \frac{kb}{\tau b} \right) - \\ & kf \cdot \left(S(t) \cdot E(t) \cdot \frac{ka}{\tau a} + S(t) \cdot R(t) \cdot \frac{kb}{\tau b} + \frac{E(t)}{\tau f} \right) - \\ & kr \cdot \left(S(t) \cdot E(t) \cdot \frac{ka}{\tau a} + S(t) \cdot R(t) \cdot \frac{kb}{\tau b} + \frac{E(t)}{\tau r} \right). \end{aligned} \tag{9}$$

It should be noted that two additional variables are added to the model: hospitalized (H) and ICU (U), both are included by the collaboration that our research group carries out with the Spanish Mathematics Committee (CEMAT).

$$\frac{dH(t)}{dt} = kh \cdot \left(S(t) \cdot E(t) \cdot \frac{ka}{\tau a} + S(t) \cdot R(t) \cdot \frac{kb}{\tau b} + \frac{E(t)}{\tau h} \right), \tag{10}$$

$$\frac{dU(t)}{dt} = ku \cdot \left(S(t) \cdot E(t) \cdot \frac{ka}{\tau a} + S(t) \cdot R(t) \cdot \frac{kb}{\tau b} + \frac{E(t)}{\tau u} \right). \tag{11}$$

Where kh is the hospitalization rate of the sick population and ku is the ICU rate of the hospitalized population. Moreover, τh and τu represent the delay of the flow of both variables with respect to the sick and hospitalized population, respectively.

3 Results and discussion

For the calibration of the model, the genetic algorithm (hereinafter GA) implemented in Sigem [1] is used. With this tool, rates and delays are variables that the model calculates from an assumed initial data defined on the chromosome. The genetic algorithm is provided with a minimum and maximum oscillation of these variables, as well as a percentage that corresponds to the corresponding search window. In addition, the sum of the absolute values of the differences between the infected, deceased, recovered, hospitalized and admitted to the ICU calculated by the model and the actual trend of each variable is defined as an objective variable to be minimized by the GA. The aforementioned real trend is temporarily fitted through Gaussian combinations.

These adjustments are the ones that allow extrapolations of the variables contained in the chromosome to be able to make future predictions. The real data are obtained from `https://covid19.isciii.es/` and the time period here studied corresponds to the one from Januray 31 to April 22, 2020.

Figure 2 shows the validation of the model's deterministic formulation through the fitting to reality degree of infected, deceased, recovered, hospitalized, admitted to the ICU and sick population. The results show a great reliability of the model, being the determination coefficients higher than 0.99, a good graphic overlap betweeen real and simulated data, and a good relative error, that in all cases, is less than 5%.

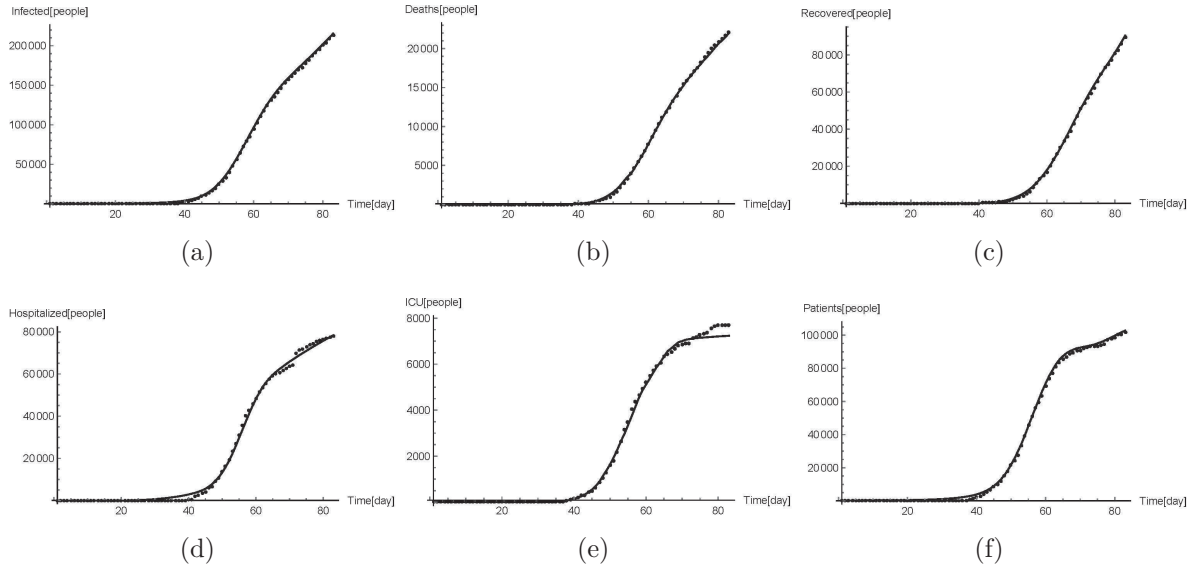

Figure 2: Validation of the model's deterministic formulation in the period from January 31 to April 22, 2020. a) Infected population. R^2= 0.999877; b) Deceased population. R^2= 0.999779; c) Recovered population. R^2= 0.999709; d) Hospitalized population. R^2= 0.998582; e) ICU population. R^2= 0.998089; f) Sick population. R^2= 0.99972. Real data: dots. GA simulated data: line.

The application of the model has been the prediction of the infected flow, deceased flow, recovered flow, hospitalized flow, admitted to the ICU flow, as well as the sicking flow in the period from April 23 to May 14, 2020. For this, the chromosome variables are exptrapolated

into the mentioned period, and they become input variables of the model designed for this work.

Figure 3 presents past and future evolution of the first five previously mentioned variables. It can be seen that by the measure of confinement introduced by the Government, the infection flow has been reduced. However, with the application of massive tests to asymptomatic people, a slight rise in the number of registered cases is expected.

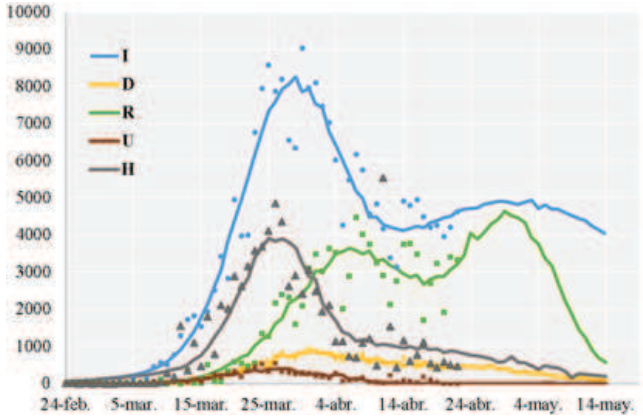

Figure 3: Prediction in a trend scenario in the period April 23 to May 14, 2020. Real data: points. Simulated data: line.

Regarding the sicking flow (Figure 4), it can be seen that for May 14 the trend continues growing, being observed this increase since April 20, that is the date when the relaxation of the containment measures for certain production sectors in the territory took place. In addition, we want to emphasize that the slope of the forecasted curve is the same as the one that occurred in the last period from March 11 to 31, approximately.

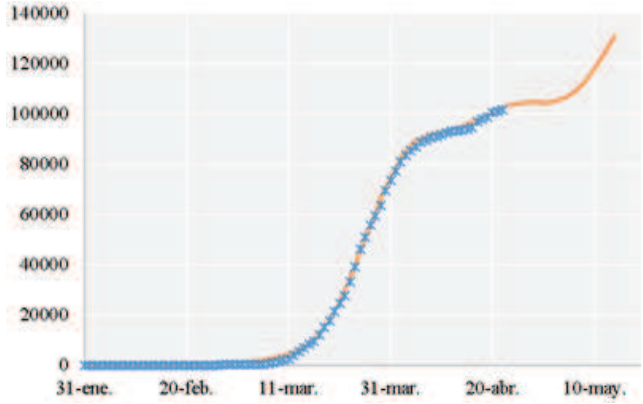

Figure 4: Prediction in a trend scenario of the sicking flow in the period from April 23 to May 14, 2020. Real data: blue blades. Simulated data: orange line.

4 Conclusions and future work

In this report we have analyzed the epidemic data made available to the scientific community by the Gobierno de España (`https://covid19.isciii.es/`) and referred to the 31/01/2020 –

23/04/2020 period. Our results suggest that the containment measures taken by the authorities have reduced the spread of COVID-19 among people in Spain. But the introduction of new rapid tests together with the PCR tests that were carried out from the beginning of the epidemic, have led to an increase in detected cases since April 14. This fact does not indicate an increase in new cases, but rather reflects the actual viral reality that our country had. Likewise, the de-escalation process accelerated by the fear to economic collapse from April 20 also causes an increase in cases. The future work on this model is chalenger. On the one hand, the present model will be compared with models already published by the scientific community. On the other hand, its application to different countries of the world must be done, so that its validity be ratified. Regarding the improvement of the model, the construction of the corresponding stochastic model formulation will allow giving predictions with a confidence interval. Likewise, the definition of the same by using cohorts and sexes could enable us to evaluate the different kinds of population and to develop specific measures in both directions. Finally, a possible study of the past evolution of the parameters included in the cromosome (detected in the calibration process) could teach something about how to influence the future evolucion of the pandemia in Spain.

References

[1] Fanelli, D., Piazza, F., Analysis and forecast of COVID-19 spreading in China, Italy and France *Chaos, Solitons & Fractals*, Volume(134):1–5, 2020.

[2] Rao, A., Vazquez, J., Identification of COVID-19 Can be Quicker through Artificial Intelligence framework using a Mobile Phone-Based Survey in the Populations when Cities/Towns Are Under Quarantine *Infection Control & Hospital Epidemiology*, Volume(41):1–18, 2020.

[3] Kermack, W., McKendrick, A. G., Walker, G., A contribution to the mathematical theory of epidemics *Proceedings of the Royal Society*, Volume(A115):700–721, 1927.

[4] Caselles, A., SIGEM: a realistic models generator expert system. In Trappl (Ed.) *Cybernetics and Systems '88* (pp. 101-108) Dordrecht, Kluwer Academic Publishers, 1988.