

# From the hospital scale to nationwide: observability and identification of models for the COVID-19 epidemic waves

Emeric Scharbarg<sup>a,b</sup>, Claude H. Moog<sup>a</sup>, Nicolas Mauduit<sup>b</sup>, Claudia Califano<sup>c</sup>

<sup>a</sup>*Emeric Scharbarg and Claude H. Moog are with LS2N, UMR CNRS 6004, BP 92101, 44321 Nantes Cedex 3, France. emeric.scharbarg@ls2n.fr*

<sup>b</sup>*Nicolas Mauduit is with Nantes University Hospital, 5 allée de l'île Gloriette 44093 Nantes Cedex 1, France. nicolas.mauduit@chu-nantes.fr*

<sup>c</sup>*Claudia Califano is with DIAG, Università di Roma, via Ariosto 25, 00184 Roma, Italy. claudia.califano@uniroma1.it*

---

## Abstract

Two mathematical models of the COVID-19 dynamics are considered as the health system in some country consists in a network of regional hospital centers. The first *macroscopic* model for the virus dynamics at the level of the general population of the country is derived from a standard SIR model. The second *local* model refers to a single node of the health system network, i.e. it models the flows of patients with a smaller granularity at the level of a regional hospital care center for COVID-19 infected patients. Daily (low cost) data are easily collected at this level, and are worked out for a *fast evaluation* of the local health status thanks to control systems methods.

Precisely, the identifiability of the parameters of the hospital model is proven and thanks to the availability of clinical data, essential characteristics of the local health status are identified. Those parameters are meaningful not only to alert on some increase of the infection, but also to assess the efficiency of the therapy and health policy.

*Keywords:* Epidemiology, Covid-19, identifiability, observability, identification.

---

## 1. Introduction

COVID-19 is an aerial virus which strickens humans through a respiratory infection, see Zhu et al. [2020]. It belongs to the Coronavirus family which was discovered in the 60's and has already infected humans through SARS and MERS. SARS is an atypical pneumonia which appeared for the first time in 202 in China as described in Ksiazek et al. [2003]. MERS which appeared in 2012 in China too (see Zaki et al. [2012]), is very similar to SARS but with a higher mortality. COVID-19 was declared to WHO (the World Health Organization) at the end of 2019 from cases in Wuhan, China, see Zhu et al. [2020]. On the 20th of February 2020 the declared positive cases were 76,000 with almost 2,500 deaths, mainly in China. At the same time, there were about 50 declared positive cases in Northern Italy, while in France there were less than 30 declared positive cases. The health status was far from being homogeneous in those or other countries. The WHO declared on March 11th, 2020, that COVID-19 was a pandemic on the website WHO [2020]. On June 1st, 2020, the declared positive cases

around the world were 6.15 million with 375,000 deaths, with almost 233,000 positive cases and 33,000 deaths in Italy and 152,091 positive cases and 28,833 deaths in France. On September 21st 2020, the confirmed cases globally amount to 30,909,405 with 958,754 deaths communicated to WHO by national authorities. The total number of cases in France on September 21st is 420,855 with 31,109 deaths, while in Italy the total number of cases is 296,569 with 35,692 deaths (website WHO [2020b]).

Mathematical models are fundamental to understand and to predict the mechanisms of the spread of an epidemic. The most popular and widely used is the SIR model in Kermack and McKendrick [1927] for human-to-human transmission. For the Covid-19 pandemic, many models have been built to explore the epidemic at the scale of a country like the SIDARTHE model which is an upgrade of the SIR model and is found in Giordano et al. [2020]. Recall that SIDARTHE stands for a Susceptible, Infected, Diagnosed, Ailing, Recognized, Threatened, Healed and Extinct model. The SIDARTHE model has highlighted that restrictive social-distancing measures need to be combined with widespread testing and contact tracing in Giordano et al. [2020]. Gervetz et al. [2020] incorporate explicit social distancing via separate compartments for susceptible and asymptomatic individuals in the SIR model, while in Di Giamberardino et al. [2020] infected people are split into infected with low viral load, undiagnosed ones, diagnosed ones and in quarantine ones. Alternatively the SARS-CoV-2 dynamics is also described at a within-host level in Vargas et al. [2020].

In this paper, a new continuous-time macroscopic model, valid at the scale of a country, is introduced and it is argued that suitable delays are mandatory to reflect the dynamics of the infection while maintaining a certain simplicity in the modelling. An observability analysis is processed for further insight of this model.

Since the health status is far from being homogeneous over a full country and since the health system consists in a network of major regional hospitals, it is worth to take advantage of the availability and agility of local hospital data rather than just merging them in some centralized information system.

At this point, the mathematical modelling of the local population flows at the level of one single hospital appears to be definitely relevant. Such models may be interconnected to include the transfer of COVID-19 patients from one hospital to another to distribute the pressure at some peaks of the infection, as done in Italy with transfers from the north to the south, or in France with transfers from the east to the west and even to Germany.

The identification of the parameters of the second local model is agile, fast and relevant not only to alert on some increase of the infection, but also to evaluate the flows of severely infected patients, and thus to assess the efficiency of the therapy and the local health policy.

Data from Nantes University Hospital were collected on a daily basis for about 6 months from March 16th to September 17th 2020, and are used here to identify essential characteristics of the local health conditions. These data include the daily values of patients in conventional care, intensive care, patients which died for COVID-19 and patients which were discharged from the hospital because completely recovered or partially recovered but in the condition of continuing their treatment at home. The originality of the data from Nantes university hospital lies in the fact that the lockdown

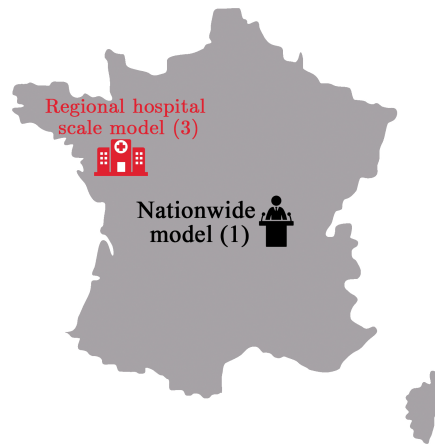


Figure 1: Epidemiologic models for two different scales. (In dark grey : the map of France , in red : Nantes University Hospital).

started early, the emergency services were not saturated and they focus in a population of patients with mild or severe Covid-19 cases. The macroscopic and local models are illustrated in Figure 1.

Control systems theoretic tools have shown their efficiency to give a new insight for various biomedical systems including for instance HIV infection in Chang et al. [2014]. It is argued that news solutions to cope with the COVID-19 pandemic may also benefit of those engineering science tools.

The outline of the paper is as follows. A new infection model involving delays is derived from the standard SIR model in Section 2. To fit to the mainstream of the current literature, this model is described in continuous time. The observability analysis of this continuous time-delay model is processed in Section 3. A subsystem, consisting in a node of the global health system network, is extracted and further detailed in Section 4 as it models the dynamics inside a hospital center. Since such a model is obtained from the daily data available from a University hospital, this second local model is directly designed in discrete-time. In Section 5 a discussion on the identification of the parameters characterizing the local model and their interpretation is carried out. Conclusions are pointed out in Section 6.

## 2. The dynamics in the general population

The evolution of the illness is described through the following dynamical model which represents a modification of the well-known SIR model introduced in 1927 in Kermack and McKendrick [1927] to describe the diffusion of an epidemic disease by considering the evolution of three classes of people (compartments) the Susceptible individuals, those who can become infected, the Infected, those who spread the disease around, and the Recovered, those who recovered from the disease. With respect to the SIR model, hereafter we split the Infected people into two compartments due to the

specificity of the disease, and we consider also the compartment of dying people. In particular apart the high transmission rate, other two aspects were immediately pointed out by the physicians which did strongly influence the diffusion of the disease and the medical resources: first it was estimated that a large delay of time (10 to 14 days) is present between the moment in which a person becomes infected and can infect, and the instant in which symptoms become evident and the person is isolated and sent to quarantine. Secondly it took a long time for many patients to recover from the disease. Some of them were in hospital and in intensive care for a big amount of time (many more than one month) thus maintaining the resources unavailable to help other patients. To highlight these two aspects we then refer in the following to a time delay system. As it will be shown in the next section the delay introduced will have an important role on the observability properties of the dynamics and thus cannot be neglected. Such a dynamics can then be described by the following delay-differential equations:

$$\left\{ \begin{array}{l} \dot{I}_q = \varepsilon\beta SI_a - \alpha I_q - \gamma_q I_q + \eta I_a(t - \tau) \\ \dot{R} = \gamma_q I_q(t - \ell_q \tau) + \gamma_a I_a(t - \ell_a \tau) \\ \dot{D}_q = \alpha I_q \\ \dot{S} = -\beta SI_a \\ \dot{I}_a = (1 - \varepsilon)\beta SI_a - \gamma_a I_a - \eta I_a(t - \tau) \end{array} \right. \quad (1)$$

- $I_q$  are the infected patients aware of their disease and who are thus in quarantine. They include all hospitalized patients, but not only.
- $R$  is the sub-population which has recovered from the infection.
- $D_q$  denotes the cumulative number of patients who deceased from the infection and were already identified so they are part of the  $I_q$  population.
- $S$  is the amount of naive individuals among a given population which are susceptible to become infected. The  $S$  population is infected by the  $I_a$  infected individuals which are not in quarantine.
- $I_a$  are asymptomatic infected people which represent the main source of infection and who spread out the infection among the general population.
- $\beta SI_a$  is the amount of newly infected individuals per time unit. This term splits into two parts, a smaller part of very sensitive people which affects the dynamics of  $I_q$  and the majority of newly infected individuals will increase the number  $I_a$ .
- $\alpha$  denotes the death rate due to the infection and mainly affects the  $I_q$  population.
- $\gamma_q I_q(t - \ell_q \tau)$  is the amount of patients in quarantine who recover from the disease.
- $\gamma_a I_a(t - \ell_a \tau)$  is the amount of other infected individuals among the general population who recover from the disease.

- $\eta I_a(t - \tau)$  is the amount of unaware infected individuals who become aware of their infection and go into quarantine. This phenomenon occurs with some delay  $\tau$  which is typically evaluated from 2 to 3 weeks.

**Remark.** As already highlighted,  $D_q$  represents the COVID deaths from the infected and in quarantine compartment. There may be of course unknown COVID deaths which can be described by an additional dynamics of the form

$$\dot{D}_a = \alpha_a I_a$$

so that the total number of deaths due to COVID is actually given by  $D = D_a + D_q$ .

$D_a$  cannot be measured, and had an important role essentially at the beginning of the pandemic when the number of detected COVID cases was much lower than the real one. Due to the high mortality caused by COVID in that period, a rough estimation could be obtained by comparing the month death rate of a single region with the corresponding one in the current year.  $\square$

**Remark.** As already underlined, the dynamics (1) is affected by two different kind of delays: the first one characterizes the amount of time that passes between the moment a person is infected and can infect, and the moment the person becomes aware of the illness and is put to quarantine. This delay is intrinsic of the disease and can only be marginally affected. The second delay which characterizes the dynamics (1) is instead linked to the time infected people recover from the disease. This delay instead has drastically changed from the starting of the pandemic and is mainly due to a better knowledge of the disease and the therapies needed by the patients which allows now to recover faster from the disease.  $\square$

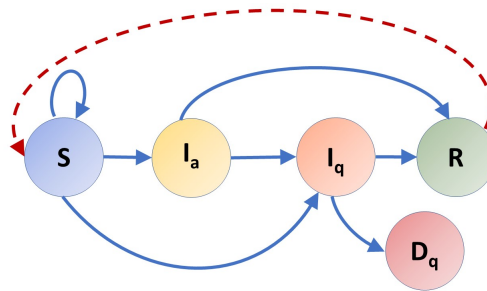


Figure 2: Flow Diagram associated to the evolution of the disease in a population described by eq. (1)

In Figure 2 a representation of the propagation of the disease is given. Susceptible people ( $S$ ) remain susceptible in time or become infected. Once infected they will move to the compartment of people in quarantine ( $I_q$ ) or infected asymptomatic people ( $I_a$ ), depending whether they have symptoms or not. From the compartment  $I_a$  people will become symptomatic and move to  $I_q$  or recover and move to  $R$ . Those in quarantine will recover and move to  $R$  or die due to the disease ( $D_q$ ). Finally as it has been recently put in evidence people who recover may be infected again, so they should be counted in

the susceptible people. The model described by equation (1) has from this point of view a short term validity. It is possible to take into account the possibility of reinfection of recovered people. In this case, the dynamics (1) would be modified by setting

$$\dot{S} = -\beta SI_a + \nu R$$

The coefficient  $\nu$  seems however to be very small at this stage, and is neglected in this study.

In the next Section, the dynamics (1) is further analyzed with respect to its observability properties since this kind of study allows to have an estimation of the state variables. The subsystem (2) consisting by  $I_q$ ,  $R$  and  $D_q$  is then further discussed in Section 4: a group of people who are aware of their infection define the flow of admissions in a local hospital and are split into two populations, the patients admitted in conventional hospitalization and the patients admitted in intensive care.

### 3. Analysis of observability

The health status of a given population has to be assessed from its number  $I_q$  of at-tested COVID-19 cases. It is important to guess the real number of infected individuals, and this is the purpose of observability from the measurement of  $I_q$ .

Considering then as output of the dynamics (1),  $y = I_q$  we will see hereafter that successive time differentiations of the measurement  $y$  will involve  $I_q$ ,  $I_a$  and  $S$ . Thus, the dynamics (1) with the only output  $y = I_q$  is not fully observable and at most  $I_a$  and  $S$  can be estimated in addition to the measurement  $I_q$ . On the other hand also the number of deaths due to COVID-19 in the  $I_q$  population can be measured, while the number of recovered people cannot clearly be estimated. In the present section we will thus investigate the observability properties of the given system starting from the measurement of  $y = I_q$ , and we will show that given the fact that  $D_q$  is measured and  $R$  cannot be estimated, the problem reduces to the study of the observability properties of a subsystem of order 3.

Now, due to the presence of delays, which in this context are considered constant, we can take the differential representation of the dynamics in order to study its behaviour, using the approach introduced in Califano et al. [2020]. To take into account the link between the delayed variables, the backward shift operator  $\delta$  has to be considered, see Xia et al. [2002]. Let us denote by  $\mathcal{H}$  the field of causal meromorphic functions  $f(x(t), \dots, x(t - s\tau), u(t), \dots, u(t - s\tau))$ , with  $s \in \mathbb{N}$ . Given a function  $\gamma(x(t), \dots, x(t - j\tau)) \in \mathcal{H}$ ,  $\gamma(-1)$  denotes the function shifted by  $\tau$ , that is  $\gamma(-1) := \gamma(x(t - \tau), \dots, x(t - j\tau - \tau))$ . Let  $dx(t)$  denote the differential of  $x$ . Then, thanks to the back shift operator  $\delta$ ,  $dx(t - s\tau) = \delta^s dx$ . Accordingly, given the function

$$y(t) = h(x(t), \dots, x(t - s\tau))$$

its differential form  $dy(t) = \sum_{j=0}^s \frac{\partial h}{\partial x(t-j\tau)} dx(t-j\tau)$  can be written in concise form as

$$dy = \left[ \frac{\partial h}{\partial x(t)} + \dots + \frac{\partial h}{\partial x(t-s\tau)} \delta^s \right] dx.$$

Given  $a(\cdot), f(\cdot) \in \mathcal{K}$ :

$$\delta[a(\cdot)df(\cdot)] = a(-1)df(-1).$$

Finally  $\mathcal{K}[\delta]$  is the (left) ring of non commutative polynomials in  $\delta$  with coefficients in  $\mathcal{K}$ . A general module spanned by the differentials of functions in  $\mathcal{K}$  is then defined over the ring  $\mathcal{K}[\delta]$ , as in Xia et al. [2002]. Then, in this framework setting  $(x_1, x_2, x_3, x_4, x_5) = (I_q, R, D_q, S, I_a)$ , and assuming without loss of generalities  $\ell_1$  and  $\ell_2$  integers, our system is characterized by the differential representation

$$\begin{aligned} d\dot{x} &= \begin{pmatrix} -(\alpha + \gamma_q) & 0 & 0 & \varepsilon\beta x_5 & \varepsilon\beta x_4 + \eta\delta \\ \gamma_q\delta^{\ell_q} & 0 & 0 & 0 & \gamma_a\delta^{\ell_a} \\ \alpha & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\beta x_5 & -\beta x_4 \\ 0 & 0 & 0 & (1-\varepsilon)\beta x_5 & -(\gamma_a + \eta\delta) + (1-\varepsilon)\beta x_4 \end{pmatrix} dx \\ dy &= (1 \ 0 \ 0 \ 0 \ 0) dx. \end{aligned}$$

Compute the observability matrix  $\frac{\partial(y, \dot{y}, \ddot{y})}{\partial x}$  which is given by

$$\begin{pmatrix} dy \\ d\dot{y} \\ d\ddot{y} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -(\alpha + \gamma_q) & 0 & 0 & \varepsilon\beta x_5 & \varepsilon\beta x_4 + \eta\delta \\ (\alpha + \gamma_q)^2 & 0 & 0 & L_1 + \eta(1-\varepsilon)\beta x_5(-1)\delta & L_{20} - L_{21}\delta - \eta^2\delta^2 \end{pmatrix} dx$$

with

$$L_1 = \varepsilon\beta x_5[-\beta x_5 - \alpha - \gamma_q - \gamma_a + 2(1-\varepsilon)\beta x_4] - \varepsilon\beta\eta x_5(-1)$$

and

$$\begin{aligned} L_{20} &= ((1-\varepsilon)\beta x_4 - \gamma_a - 2\beta x_5 - \alpha - \gamma_q)\varepsilon\beta x_4 \\ L_{21} &= [\alpha + \gamma_q + \varepsilon\beta x_4 - (1-\varepsilon)\beta x_4(-1) + \gamma_a]\eta \end{aligned}$$

Clearly  $dy^{(2+j)}$ , for any  $j \geq 0$  does not depend on  $dx_2$  and  $dx_3$  which proves that the whole system cannot be weakly, regularly or strongly observable. We may however be interested in studying the reduced system defined by the variables  $(x_1, x_4, x_5) = (I_q, S, I_a)$  and given by

$$\begin{cases} \dot{x}_1 &= \varepsilon\beta x_4 x_5 - \alpha x_1 - \gamma_q x_1 + \eta x_5(t - \tau) \\ \dot{x}_4 &= -\beta x_4 x_5 \\ \dot{x}_5 &= (1-\varepsilon)\beta x_4 x_5 - \gamma_a x_5 - \eta x_5(t - \tau) \end{cases} \quad (2)$$

Using the previous computations, one gets that the associated observability matrix is

$$\hat{\mathcal{O}}(x, \delta) = \begin{pmatrix} 1 & 0 & 0 \\ -(\alpha + \gamma_q) & \varepsilon\beta x_5 & \varepsilon\beta x_4 + \eta\delta \\ (\alpha + \gamma_q)^2 & L_1 + \eta(1-\varepsilon)\beta x_5(-1)\delta & L_{20} - L_{21}\delta - \eta^2\delta^2 \end{pmatrix}.$$

This subsystem will then be weakly observable if  $\hat{\mathcal{O}}(x, \delta)$  has full rank over  $\mathcal{K}(\delta)$ . It will be strongly observable if it is also unimodular. If it is weakly observable but not

strongly we will have to check if it is regularly observable, see Califano et al. [2020], that is if it is possible to reconstruct the state of the system by using also derivatives of higher order of the output function<sup>1</sup>.

We may essentially distinguish three cases based on the values of the parameters  $\varepsilon$  and  $\eta$  which are discussed hereafter and are represented in Table 1.

- **First Case**  $\varepsilon = 0$ .

This case corresponds to the situation in which there is an important delay between the time people get infected and the time they become aware and get into quarantine. As a consequence, necessarily  $\eta \neq 0$  and the observability matrix associated to the third order subsystem (2) will be

$$\hat{\mathcal{O}}(x, \delta) = \begin{pmatrix} 1 & 0 & 0 \\ -\alpha - \gamma_q & 0 & \eta \delta \\ (\alpha + \gamma_q)^2 & \eta \beta x_5 (-1) \delta & -(\alpha + \gamma_q + \gamma_a - \beta x_4 (-1)) \eta \delta - \eta^2 \delta^2 \end{pmatrix}$$

The subsystem (2) will be then weakly observable for  $\eta \neq 0$ ,  $\beta \neq 0$ ,  $x_5(-1) \neq 0$ . It is easily verified that it will neither be strongly nor regularly observable.

- **Second Case**  $\varepsilon \neq 0$  and  $\eta = 0$ .

In this case there is no structural delay between the two classes  $I_q$  and  $I_a$ . People get infected and after a negligible time are moved to quarantine. The delay characterizes only the large amount of time that ill people need to get recovered. In this situation, for the subsystem (2) one gets

$$\hat{\mathcal{O}}(x, \delta) = \begin{pmatrix} 1 & 0 & 0 \\ -\alpha - \gamma_q & \varepsilon \beta x_5 & \varepsilon \beta x_4 \\ (\alpha + \gamma_q)^2 & \varepsilon \beta x_5 (-\varepsilon \beta x_5 - \alpha - \gamma_q - \gamma_a + 2(1 - \varepsilon) \beta x_4) & L_{20} \end{pmatrix}$$

No delay affects the observability matrix. The subsystem (2) will be strongly observable if and only if the matrix has full rank, which happens if and only if

$$\varepsilon^2 \beta^2 x_4 x_5 \{(\varepsilon - 2) \beta x_5 - (1 - \varepsilon) \beta x_4\} \neq 0$$

---

<sup>1</sup>The different notions of observability are peculiar of time delay systems. **Strong observability** can be tested by verifying the unimodularity of the observability matrix and allows to express the state of the system at time  $t$  as a function of the input and output and their derivatives up to order  $n - 1$  eventually delayed. Weak and regular observability are much weaker notions and have different implications. To have a flavour of these implications we give hereafter two examples to highlight the differences: **Regular Observability:** Consider the dynamics  $\dot{x}(t) = x(t - \tau)u(t)$  with output  $y(t) = x(t) + x(t - \tau)$ . The observability matrix is  $\mathcal{O}(x, \delta) = 1 + \delta$ , which has full rank over  $\mathcal{K}[\delta]$  but is not unimodular. Nevertheless, the state of the system at time  $t$  can still be written as a function of the input and output and their derivatives, but requires higher order derivatives. In case of the example, we get  $x(t) = y(t) - \frac{\dot{y}(t) - \dot{y}(t - \tau)}{u(t) - u(t - \tau)}$ , which is valid whenever  $u(t) \neq u(t - \tau)$ . **Weak Observability:** Consider the dynamics  $\dot{x}(t) = u(t)$  with output  $y(t) = x(t) + x(t - \tau)$ . The observability matrix is still  $\mathcal{O}(x, \delta) = 1 + \delta$ . In this case however the system is neither strongly nor regularly observable, but is said to be weakly observable. In this case only an implicit relation can be written down involving different delayed values of the state, of the derivatives of the input and of the output.



	weakly observable	regularly observable	strongly observable
$\varepsilon = 0$	for $\eta\beta x_5(-1) \neq 0$	$\times$	$\times$
$\varepsilon \neq 0, \eta = 0$	for $\beta x_4 x_5 \neq 0$	for $\beta x_4 x_5 \neq 0$	for $\beta x_4 x_5 \neq 0$
$\varepsilon \neq 0, \eta \neq 0$	for $\beta x_5 \neq 0$	$\times$	$\times$

Table 1: Observability properties of subsystem (2)

that is whenever  $\beta \neq 0, x_4 \neq 0, x_5 \neq 0$ . The last condition in fact

$$x_5 \neq \frac{1-\varepsilon}{\varepsilon-2}x_4$$

is always satisfied since  $x_5 \geq 0$ .

- **Third Case**  $\varepsilon \neq 0$  and  $\eta \neq 0$ .

To check it let us use Smith decomposition. We have that

$$\hat{\mathcal{O}}(x, \delta) = T_1^{-1}(x, \delta)S(x, \delta)T_2^{-1}(x, \delta)$$

where

$$T_1(x, \delta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & T_{11} - \frac{\eta}{\varepsilon}\delta & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \eta\delta & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ \alpha + \gamma_q & 1 & 0 \\ 0 & \alpha + \gamma_q & 1 \end{pmatrix}$$

$$T_2(x, \delta) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$S(x, \delta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \varepsilon\beta x_5 & \varepsilon\beta x_4 + \eta\delta \\ 0 & 0 & S_{33}(x, \delta) \end{pmatrix}$$

with

$$T_{11} = \varepsilon\beta x_5 + \gamma_a - 2(1-\varepsilon)\beta x_4 + \eta \frac{x_5(-1)}{x_3}$$

$$S_{33}(x, \delta) = \left( \eta \frac{x_5(-1)}{x_5} - (1-\varepsilon)\beta x_4 - \beta x_5 \right) \varepsilon\beta x_4$$

$$+ \left( \varepsilon\beta x_5 - (2-\varepsilon)\beta x_4 + \eta \frac{x_5(-1)}{x_5} \right) \eta\delta - \frac{\eta^2}{\varepsilon}\delta^2.$$

The subsystem (2) is then weakly observable for  $\beta x_5 \neq 0$ . It cannot be strongly observable. After some tedious computations, it appears that it is not regularly observable either.

#### 4. Data from a regional university hospital

The originality of this section is to consider a smaller granularity with a subsystem of the previous dynamics: a local hospital center in charge of Covid-19 patients. Its input is the number  $A(k)$  of admissions in the hospital on day  $k$ . It is then in principle less than or at most equal to  $I_q$  in model (1), and it is further split into the number  $I(k)$  of patients admitted in intensive care and the number  $C(k)$  of patients admitted in conventional hospitalization on day  $k$ . The model describes the evolution of the dynamics defined by the hospitalized patients  $I(k)$  and  $C(k)$  as well as the number of deaths  $D(k)$  in hospital and the number  $R(k)$  of patients who have recovered from the disease (completely or partially and continue their treatment outside the hospital).

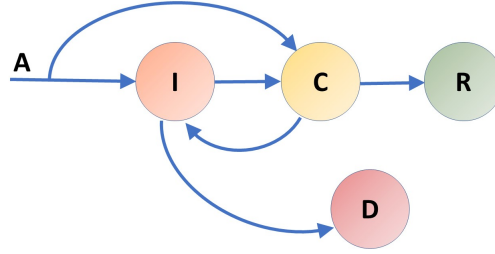


Figure 3: Flow Diagram associated to the management of the COVID-19 patients in the hospital described by eq. (3)

In Figure 3 a representation of the management of the COVID-19 patients in the hospital is highlighted. Each day  $k$  a set of  $(A)$  new patients arrive at the hospital admission centre. Depending on their health status they can be forwarded to the conventional care section  $(C)$  or the intensive care one  $(I)$ . Patients move from intensive care to conventional care and then are released from the hospital  $(R)$  once their health allows. Some of them instead do not survive  $(D)$  once their condition worsens.

As already underlined data from Nantes University Hospital were collected for about 6 months from March 16th to September 17th, 2020. These data include the daily values of  $C(k)$ ,  $I(k)$ ,  $R(k)$  and  $D(k)$  valid at day  $k$ .

Figures 4 and 5 depict the number  $D(k)$  of deaths and of recovered patients  $R(k)$  at day  $k$ , respectively. The value  $k = 0$  corresponds to March 16, 2020.

Due to the format of the daily data, a discrete time model with four parameters is in order as follows.

$$\begin{cases} C(k+1) &= (1 - \gamma - \mu)C(k) + \theta A(k) + \lambda I(k) \\ I(k+1) &= (1 - \alpha - \lambda)I(k) + (1 - \theta)A(k) + \mu C(k) \\ D(k+1) &= D(k) + \alpha I(k) \\ R(k+1) &= R(k) + \gamma C(k) \end{cases} \quad (3)$$

Raw data from the Nantes University Hospital Covid-19 database [2020] include, on a daily basis, the number  $I(k)$  of patients in intensive care, the number  $C(k)$  of patients

in conventional hospitalization, the cumulative number of deaths and the cumulative number of patients which have recovered since day 1. The daily Covid-19 admissions are not directly measured but are computed as

$$A(k) = C(k+1) - C(k) + I(k+1) - I(k) + D(k+1) - D(k) + R(k+1) - R(k).$$

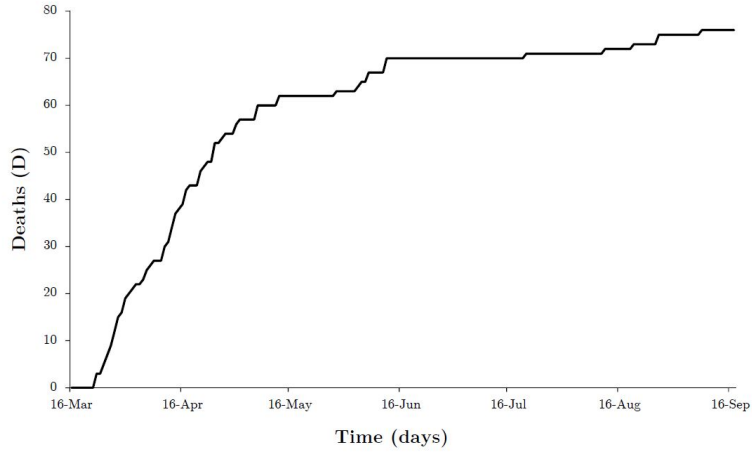


Figure 4: Cumulative deaths over time from Nantes University Hospital Dataset

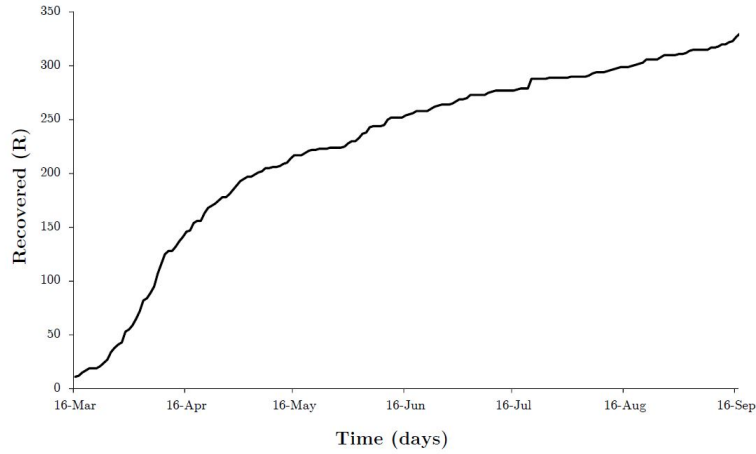


Figure 5: Cumulative number of recovered patients from Nantes University Hospital Dataset

Thus, all those state variables are considered to be measured and the input  $A(k)$  is computed from the measurements over two days.

The new admissions  $A(k)$  are split into  $\theta A(k)$  which increases the number of patients in conventional hospitalization and in  $(1 - \theta)A(k)$  which are directly entering intensive care. A value of  $\theta$  close to 1 is representative of a high level monitoring of the disease by the health care outside hospital. When  $\theta$  is closer to 0, then a major flow of new admissions enter directly into intensive care and, the other way around,  $\lambda I(k)$  is the amount of patients leaving intensive care and entering conventional hospitalization.

The number of patients  $\gamma C(k)$  who recover from the disease is assumed to be proportional to the number of patients in conventional hospitalization.

$\mu C(k)$  denotes the amount of patients who are moved from conventional hospitalization to intensive care.

The number  $\alpha I(k)$  of daily deaths is assumed to be proportional to the number  $I(k)$  of patients in intensive care.

The daily number  $A(k)$  of new admissions highly depends on the social and medical environment, as well as on some political lockdown regulations or meeting restrictions (with some delay). Thus,  $A(k)$  may have a significant variability from one regional hospital center to another. It is proportional to the number of individuals which are susceptible of becoming infected, in some standard SIR population model.

**Remark.** Note that the local scale model (3) models the dynamics of hospitalized patients. Taking the sum of  $C(k) + I(k)$  in (3) over all regional hospitals allows to get the hospitalized patients, nationwide. The latter is a part of the population  $I_q$  in model (1), as not all detected patients are hospitalized. The sum of hospitalized patients and non hospitalized infected people equals  $I_q + I_a$  in model (1). It will be argued next that some parameters that characterize a single hospital may vary with respect to time. These parameters may also differ consistently from one hospital/region to another, depending on several sociological conditions and local regulations.  $\square$

**Remark.** While in the model (1) delays were considered to characterize the evolution of the infection among the population, since they play a fundamental role in particular on the observability of the dynamics, in the local model (3) which refers to the hospitalization of COVID-19 patients, delays could be avoided. On one hand, even if data are available daily, the discrete nature of the model allows to work on a large time window (namely 7 days as it will be discussed in the next section for the identification of the parameters) which is comparable with the delays involved in (1). On the other hand, the eventual presence of a large delay could be handled by splitting the patients in intensive care or in conventional hospitalization in more compartments moving the patients from one compartment to the successive one as the disease evolves, thus avoiding again the use of delays.  $\square$

## 5. Discussion

Note that the infection dynamics described by a standard or modified SIR model, such as the dynamics (1), are external to the hospital and feed the dynamics (3). Thus, the input variable  $A(k)$  of (3), is a percentage of  $I_q$ . The parameters  $\alpha$ ,  $\gamma$ ,  $\theta$ ,  $\mu$  and  $\lambda$  of

model (3) can be identified over a short or longer period so that their evolution can be tracked. The argument in this section is that these parameters are

- easy to derive from standard data from the hospital;
- are representative of the regional health status.

*Identifiability of model (3) parameters*

The death rate  $\alpha$  and the recovery rate  $\gamma$  are easily obtained from the measured state variables, since:

$$\alpha = \frac{D(k+1) - D(k)}{I(k)}, \quad \gamma = \frac{R(k+1) - R(k)}{C(k)}.$$

From Nantes University Hospital Covid-19 database [2020], it is not significant to complete an identification on a daily basis. So, let us identify over an horizon of  $h$  days, which is done rewriting the third and fourth equation of (3) for days  $k, k+1, \dots, k+h$ . Standard computations lead to

$$\alpha = \frac{D(k+h) - D(k)}{I(k) + I(k+1) + \dots + I(k+h-1)} \quad (4)$$

and

$$\gamma = \frac{R(k+h) - R(k)}{C(k) + C(k+1) + \dots + C(k+h-1)}. \quad (5)$$

Note that parameters  $\alpha$  and  $\gamma$  are argued to be time-varying, so that equations (4) and (5) rather compute an average value over  $h$  days.

The parameters  $\theta$  and  $\mu$  are not simultaneously identifiable in this model for constant sequences of  $A(k)$  and  $C(k)$ . Nevertheless, for some input sequence such that  $A(k)C(k+1) \neq A(k+1)C(k)$ , all parameters are identifiable. In other words, the system is generically identifiable, or identifiable for almost all sequence of admissions and for almost all flows of patients.

The identification of  $\theta$  may be processed as follows. Rewrite the first equation of (3) at three different time instants:

$$\begin{cases} C(k+1) &= (1 - \gamma - \mu)C(k) + \lambda I(k) + \theta A(k) \\ C(k+i-1) &= (1 - \gamma - \mu)C(k+i) + \lambda I(k+i) + \theta A(k+i) \\ C(k+j+1) &= (1 - \gamma - \mu)C(k+j) + \lambda I(k+j) + \theta A(k+j) \end{cases} \quad (6)$$

From (6), one can easily eliminate the coefficients  $(1 - \gamma - \mu)$  and  $\lambda$  and after some elementary but tedious computations,  $\theta$  is computed as follows. Rename the following quantities

$$\begin{cases} F_1 &= C(k+1)I(k+j) - C(k+i)I(k+j) \\ F_2 &= C(k+i)I(k+j) - C(k+j)I(k+i) \\ F_3 &= C(k+i+1)I(k+j) - C(k+j+1)I(k+i) \\ F_4 &= C(k+j)I(k) - C_k I(k+j) \\ F_5 &= A(k)I(k+j) - A(k+j)I(k) \\ F_6 &= A(k+j)I(k+j) - A(k+j)I(k+1) \end{cases} \quad (7)$$

so that

$$\theta = \frac{F_1 F_2 + F_3 F_4}{F_2 F_5 + F_4 F_6}. \quad (8)$$

Obviously, there exist some singular input sequences  $A(k)$  which cancel out the denominator of the right-hand side of (8): for instance, when the sequences  $A(k)$  and  $I(k)$  are constant, then formula (8) is inapplicable for the estimation of parameter  $\theta$ .

Following a similar elimination process, it is possible to eliminate parameters  $\lambda$  and  $\alpha$  among (6). Thus  $\mu$  is identifiable for almost all input sequences, though the explicit expression becomes a bit more involved. The generic identifiability of  $\lambda$  is proven following similar lines.

#### *Computation of the death rate $\alpha$ and the recovery rate $\gamma$ from the hospital dataset*

Data from Nantes University Hospital were collected from March 16th to September 17th, 2020 and are available in Nantes University Hospital Covid-19 database [2020]. These data include the daily values of  $C(k)$ ,  $I(k)$ ,  $R(k)$  and  $D(k)$ . In France the lockdown took place from March 17, 2020 to May 11, 2020. When some data were missing or inaccurate then the corresponding periods were skipped.

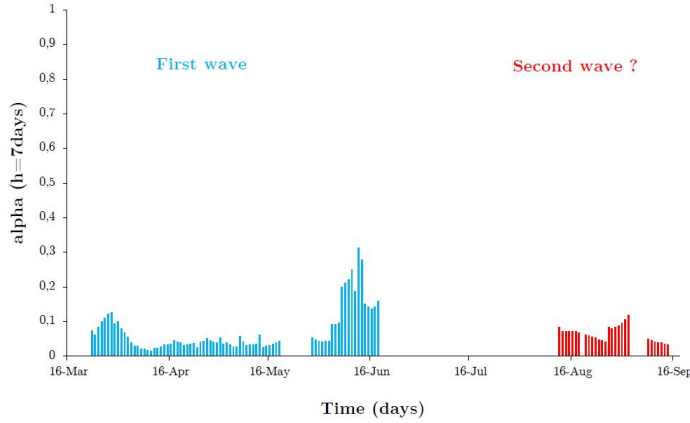


Figure 6: Time varying death rate  $\alpha$  in model (3)

Since  $\alpha$  is computed from the available data over 7 days, its evaluation displayed in Figure 6 starts on the second week. The empty information, instead, from May 20th to 28th corresponds to errors which are due to inaccurate data, such as the number of admissions which were sometimes incorrectly estimated in the hospital data, thus not allowing a correct estimation of  $\alpha$ .

In Figure 6 it is shown that the death rate increased during the early days after the lockdown. Two weeks later it was significantly decreasing and then it stabilized for over two months.

Surprisingly, the death rate  $\alpha$  seems to reach unprecedented levels weeks after the end of the lockdown. A peak of the death rate  $\alpha$  is noticed in early June, about one

month after the end of the lockdown. This paradoxical result is easily interpreted as the general younger population left (intensive care) hospitalization and only elderly patients were remaining in intensive care, sometimes after several months of hospitalization with weak probability of health improvement.

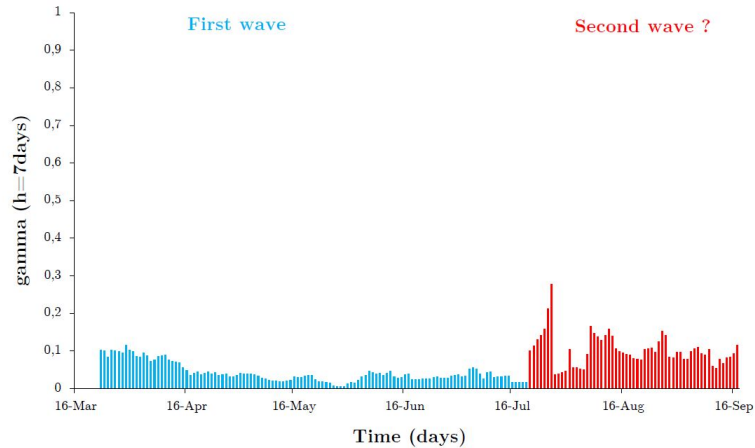


Figure 7: Time varying recovery rate  $\gamma$  in model (3)

From mid-June to mid-August, no patient was in intensive care which yields a zero death rate. This is due to a very low number of new Covid-19 patients admissions. Since the end of July 2020, there are new conventional hospitalizations, followed by intensive care hospitalizations at the beginning of August 2020. This indicates an increase in the circulation of the virus and eventually the beginning of a second epidemic wave. Parameter  $\alpha$  requires a period of 7 days to be estimated, so that the increase in mortality is confirmed on mid-August.

In Figure 7 a higher level of the recovery rate is noticed during the first month of the lockdown, i.e. before the peak of the infection was reached. Once the spread out of the infection was under control, the recovery rate was somehow stabilized.

Population levels were low in July 2020: no patients in intensive care and a decrease in the number of patients in conventional hospitalization from 23 patients at the beginning of the month to 2 patients at the end of the month. Therefore the recovery of one or two patients will have a significant impact on the recovery rate estimate.

Similarly in August 2020, there is an increase in the recovery rate due to a low headcount. These results point to the suspicion of a better patient management by the physicians in connection with an earlier treatment and to a massive screening. Summarizing, a higher recovery rate in August/September is due to a better knowledge of the disease by the physicians.

## 6. Conclusion

The main message in this paper is certainly that fast, agile, decentralized and relevant actions can be taken on some decentralized local level both to predict a new wave of the COVID-19 pandemic and to assess the real-time efficiency of treatments or of the health policy.

A local mathematical model of flows of patients will not stand in competition with a standard or upgraded centralized SIR type model but it will give additional insights to the dynamics to help decisions at the local level of a regional hospital.

The information provided by local data can be easily collected from any hospital. This information is rich enough to capture essential indicators about the local health status of the regional population. Not only this information is easy to obtain and does not require the involvement of any national health agency, but it is also more precise in the sense that it gives a smaller granularity picture of the health situation. It is well-known that this situation can be dramatically different from one region to the other; for instance, the Bergamo region in Italy had a very different status when compared with the south of Italy as shown in Alicandro et al. [2020]. Similarly, the Mulhouse region in France was one of the major clusters, see Kuteifan et al. [2020], with a much higher infection rate than western french regions along the Atlantic coast.

Data were collected during and after the lockdown period from the Medical Information Department of Nantes University Hospital Center.

Such a low cost information is able to provide an early alert about a second wave of the infection. In this sense, it is an agile and precise tool able to complement official figures published by central health agencies.

## Acknowledgment

The authors wish to thank the Medical Information Department of Nantes University Hospital for sharing data. Thanks to Dr C. Zhang for his careful reading of a preliminary version of this paper. Thanks to Z.Bondaty for her drawing skills of Figure 1. Thank you to Professor Karim Asehnoune for agreeing to share the hospital data of the Nantes University Hospital on Covid-19. Finally the authors wish to thank the anonymous reviewers for their valuable comments.

## References

- Alicandro G., Remuzzi G., La Vecchia C. (2020), *Italy's first wave of the COVID-19 pandemic has ended: no excess mortality in May*, *The Lancet*, 396 (10253):e27-e28.
- Califano C., and Moog C.H. (2020), *Observability of nonlinear time-delay systems and its application to their state realization*, *IEEE Contr. Sys. Lett.*, 4, pp.803-808.
- Chang H., Moog C.H. and Astolfi A. (2014), *A control systems analysis of HIV prevention model using impulsive input*, *Biomed. Sig. Proces. and Contr.*, 13, pp.123-131.



- Di Giamberardino P., Iacoviello D., Papa F., Sinisgalli C. (2020), *Dynamical evolution of COVID-19 in Italy with an evaluation of the size of the asymptomatic infective population* IEEE J. of Biomed. and Health Inf., doi: 10.1109/JBHI.2020.3009038.
- Gevertz J., Greene J., Hixahuary Sanchez Tapia C., Sontag E.D. (2020), *A novel COVID-19 epidemiological model with explicit susceptible and asymptomatic isolation compartments reveals unexpected consequences of timing social distancing*, medRxiv, Cold Spring Harbor Laboratory Press, doi: 10.1101/2020.05.11.20098335.
- Giordano G., Blanchini F., Erdman D., Bruno R., Colaneri P., Di Filippo A., Di Matteo A., Colaneri M. (2020), *Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy*, Nature Medicine, 26, pp.855–860.
- Kermack W.O. and McKendrick A.G. (1927), *A Contribution to the Mathematical Theory of Epidemics*, Proc. Roy. Soc. Lond., A 115, pp. 700-721.
- Ksiazek T.G., Erdman D., Goldsmith C.S., Zaki S.R., Peret T., Emery S., Tong S., Urbani C., Comer J.A., Lim W., Rollin P.E., Dowell S.F., Ling A.-E., Humphrey C.D., Shieh W.-J., Guarner J., Paddock C.D., Rota P., Fields B., DeRisi J., Yang J.-Y., Cox N., Hughes J.M., LeDuc J.W., Bellini W.J., Anderson L.J. (2003), *A Novel Coronavirus Associated with Severe Acute Respiratory Syndrome*, New England Journal of Medicine, 348 (20), pp.1953-1966.
- Kuteifan K, Pasquier P, Meyer C, Escarment J, Theissen O. (2020), *The outbreak of COVID-19 in Mulhouse : Hospital crisis management and deployment of military hospital during the outbreak of COVID-19 in Mulhouse, France*, Ann Intensive Care, 10(1):59.
- Vargas H., Abelardo E. and Velasco-Hernandez J.X. (2020), *In-host Modelling of COVID-19 Kinetics in Humans*, medRxiv, Cold Spring Harbor Laboratory Press, doi: 10.1101/2020.03.26.20044487
- Xia X., Marquez-Martinez L.A., Zagalak P. and Moog C.H. (2002), *Analysis of non-linear time-delay systems using modules over non-commutative rings*, Automatica, 38, pp.1549-1555.
- Zaki A.M., van Boheemen S., Bestebroer T.M., Osterhaus A.D.M.E., Fouchier R.A.M. (2012), *Isolation of a Novel Coronavirus from a Man with Pneumonia in Saudi Arabia*, New England Journal of Medicine, 367 (19), pp. 1814-1820.
- Zhu N., Zhang D., Wang W., Li X., Yang B., Song J., Zhao X., Huang B., Shi W., Lu R., Niu P., Zhan F., Ma X., Wang D., Xu W., Wu G., Gao G.F., Tan W. (2020), *A Novel Coronavirus from Patients with Pneumonia in China, 2019* New England Journal of Medicine, 382(8), pp.727-733.
- Nantes University Hospital, *Covid19 Epidemiological Database from March 16th to September 17th in 2020*, doi:10.17632/f75kmdmy9r.1 <https://data.mendeley.com/datasets/f75kmdmy9r/draft?a=4d493676-7345-4edc-8f28-c35cffdb9c1>

World Health Organization, Coronavirus Disease 2019 (COVID-19): Situation Report 51, March 11th 2020. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>

World Health Organization, Covid-19 Explorer, (2020), Situation on September 21st 2020, <https://worldhealthorg.shinyapps.io/covid/>