Understanding and monitoring the evolution of the Covid-19 epidemic from medical emergency calls: the example of the Paris area

Stéphane Gaubert, Marianne Akian, Xavier Allamigeon, Marin Boyet,

Baptiste Colin, Théotime Grohens, Laurent Massoulié, David Parsons,

Frederic Adnet, Érick Chanzy, et al.

▶ To cite this version:

HAL Id: hal-02648075
https://hal.inria.fr/hal-02648075v2

Submitted on 12 Jun 2020

Understanding and monitoring the evolution of the Covid-19 epidemic from medical emergency calls: the example of the Paris area

Stéphane Gaubert[1,2], Marianne Akian[1,2], Xavier Allamigeon[1,2], Marin Boyet[1,2]
Baptiste Colin[1,2], Théotime Grohens[1,3], Laurent Massoulié[1,4,5], David P. Parsons[1]
Frédéric Adnet[6,7], Érick Chanzy[6], Laurent Goix[6], Frédéric Lapostolle[6,7]
Éric Lecarpentier[6], Christophe Leroy[6], Thomas Loeb[6], Jean-Sébastien Marx[6]
Caroline Télion[6], Laurent Tréluyer[6] and Pierre Carli[6,8]

[1] INRIA

[2] CMAP, École polytechnique, IP Paris, CNRS

[3] Université de Lyon, CNRS, INSA-Lyon, LIRIS, UMR5205

[4] ENS, CNRS, PSL University

[5] Microsoft Research-INRIA Joint Centre

[6] AP-HP

[7] Université Paris XIII, Bobigny

[8] Université Paris-Descartes, Paris

Emails: [1]: Prenom.Nom@inria.fr [6]: Prenom.Nom@aphp.fr

June 10, 2020

Abstract. We portray the evolution of the Covid-19 epidemic during the crisis of March-April 2020 in the Paris area, by analyzing the medical emergency calls received by the EMS of the four central departments of this area (Centre 15 of SAMU 75, 92, 93 and 94). Our study reveals strong dissimilarities between these departments. We show that the logarithm of each epidemic observable can be approximated by a piecewise linear function of time. This allows us to distinguish the different phases of the epidemic, and to identify the delay between sanitary measures and their influence on the load of EMS. This also leads to an algorithm, allowing one to detect epidemic resurgences. We rely on a transport PDE epidemiological model, and we use methods from Perron-Frobenius theory and tropical geometry.

Comprendre et surveiller l'évolution de l'épidémie de Covid-19 à partir des appels au numéro 15: l'exemple de l'agglomération parisienne

Résumé. Nous décrivons l'évolution de l'épidémie de Covid-19 dans l'agglomération parisienne, pendant la crise de Mars-Avril 2020, en analysant les appels d'urgence au numéro 15 traités par les SAMU des quatre départements centraux de l'agglomération (75, 92, 93 et 94). Notre étude révèle de fortes disparités entres ces départements. Nous montrons que le logarithme de toute observable épidémique peut être approché par

1

une fonction du temps linéaire par morceaux. Cela nous permet d'identifier les différentes phases d'évolution de l'épidémie, et aussi d'évaluer le délai entre la prise de mesures sanitaires et leur effet sur la sollicitation de l'aide médicale urgente. Nous en déduisons un algorithme permettant de détecter une resurgence éventuelle de l'épidémie. Notre approche s'appuie sur un modèle d'EDP de transport de l'évolution épidémique, ainsi que sur des méthodes de théorie de Perron-Frobenius et de géométrie tropicale.

1 Introduction

The outbreak of Covid-19 in France has put the national Emergency Medical System (EMS), the *SAMU*, in the front line. In the *Île-de-France* region, one most affected by the epidemic, the SAMU centers of Paris and its inner suburbs experienced a major increase in the number of calls received and of the number of ambulance dispatches for Covid-19 patients.

We show that indicators based on EMS calls and vehicle dispatches allow to analyze the evolution of the epidemic. In particular, we show that EMS calls are early signals, allowing one to anticipate vehicle dispatch. We provide a method of short term prediction of the evolution of the epidemic, based on mathematical modeling. This leads to *early detection and early alarm mechanisms* allowing one either to confirm that certain sanitary measures are strong enough to contain the epidemic, or to detect its resurgence. These mechanisms rely on simple data generally available in EMS: numbers of patient records tagged as Covid-19, and among these, numbers of records resulting in medical advice, ambulance dispatch, or Mobile Intensive Care Unit dispatch. We also provide a comparative description of the evolution of the epidemic in the four central departments of the Paris area, showing spatial dissimilarities, including a strong variation of the doubling time, depending on the department.

Our approach relies on several mathematical tools in an essential way. Indeed, the Covid-19 epidemic has unprecedented characteristics, and, given the lack of experience of similar epidemics, one needs to rely on mathematical models. We use transport PDE to represent the dynamics of Covid-19 epidemic. Transport PDE capture epidemics with a significant time interval between contamination and the start of the infectious phase (in contrast, ODE models without time delays allow instantaneous transitions from contamination to the infectious phase). In the early stage of the epidemic, in which the majority of the population is susceptible, this dynamics becomes approximately linear and order preserving. Then, it can be analyzed by methods of Perron–Frobenius theory. Our main theoretical result shows that the logarithm of epidemic observables can be approximated by a piecewise linear map, with as many pieces as there are phases of the epidemic (i.e., periods with different contamination conditions), see Theorem 1. This methods allows us to identify, the phases of the epidemic evolution, and also to evaluate the time interval between sanitary measures and their impact on epidemic observables, like vehicle dispatch. The idea of piecewise linear approximation and of "log glasses", a key ingredient of the present approach, arises from tropical geometry.

The present work started on March 13th, and led to the algorithm presented here. A preliminary version of this algorithm was used, on March 20th, to forecast the epidemic wave, anticipating that the peak load of SAMU (which occurred around March 27th) would be different depending on the department of the Paris area. We subsequently applied our method to provide Assistance Publique – Hôpitaux de Paris (AP-HP), on April 5th, with an early report, quantifying the efficiency of the lockdown measures from the estimation of the contraction rate of the epidemic in the different departments. This algorithm is now deployed operationally in the four SAMU of AP-HP. This work may be quickly reproduced in any EMS.

Although it was developed for Covid-19 and for EMS calls, the present monitoring method is generic. It may also apply to other medical indicators, see Section 3.4, and to other epidemics, for instance, influenza.

This paper is a crisis report, giving a unified picture of a work done jointly by a team of physicians of the SAMU of AP-HP and applied mathematicians from INRIA and École polytechnique. Medical, epidemiological, and mathematical aspects are intricated in this work. We received help from several physicians, researchers and engineers, not listed as authors, and also help from several organizations. They are thanked in the acknowledgments section.

This paper should be understood as an announce. The results will be subsequently developed in several papers, with different subsets of coauthors. It is intended to be read both by a medical and a mathematical

audience. The first part of the paper, up to Section 6 included, and the conclusion, are intended to a broad audience. Mathematical tools are presented in Section 7, Section 8, Section 9 and in the appendix.

The present work shows the epidemiological significance of the calls received by the EMS, it focuses on the mathematical modeling aspects, on the description of the evolution of the epidemy in the Paris area, and on prediction algorithms. The current work[1] with an intersecting set of authors, is coordinated with the present one. It focuses on medical aspects. It makes a case study of the Covid-19 crisis of March-April 2020, in Paris, considering the EMS and the hospital services in a unified perspective. It shows that the calls received by SAMU are early predictors of the future load on ICU.

2 Context

The mission of the SAMU centers is to provide an appropriate response to calls to the number 15, the French toll-free phone number dedicated to medical emergencies. This service is based on the medical regulation of emergency calls, in the sense that for each patient, a physician decides which response is most appropriate. Thus, depending on the evaluation over the phone of the severity of the case and the circumstances, the response may be a medical advice, a home visit by a general practitioner, the dispatch of a team of EMTs (Emergency Medical Technicians) of either a first aid association or the Fire brigade, or an ambulance of a private company. A Mobile Intensive Care Unit (MICU), staffed by a physician, a nurse and an EMT, is sent to the scene as a second or a first tier, when a life threatening problem is suspected. The role of the SAMU in the management of disasters or mass casualties has been described elsewhere [18, 4]. The city of Paris and its inner suburbs are covered by 4 departmental SAMU Center-15 : Paris (75), Hauts-de-Seine (92), Seine Saint-Denis (93), and Val-de-Marne (94), see the map on Figure 3. They serve a population of 6.77 million inhabitants. These four Center-15 are part of the public hospital administration, AP-HP (Assistance Publique – Hôpitaux de Paris). They operate identically and use the same computerized call management system. Since the outbreak of the Covid-19 epidemic, the French government instructed the public that anyone with signs of respiratory infection or fever should not go directly to the hospital emergency room to limit overcrowding, but should call number 15 for orientation. To comply with the recommendations of the health care authorities, the four Center-15 applied the same procedures: after medical call regulation, only patients with signs of severity or significant risk factors were transported by EMTs and ambulances to hospitals, either to Emergency Room (ER) or newly created Covid-19 Units. The cases presenting a life-threatening emergency, mostly respiratory distress, were managed by an MICU team and then admitted directly in Intensive Care Unit (ICU). All other cases were advised to stay at home and isolate themselves. When necessary, these patients were also eligible for a home visit by a general practitioner or a consultation appointment the following days.

In order to maintain a rapid response when a major increase in the number of calls was observed, the four Center-15 implemented specific procedures. Switchboard operators and medical staff was reinforced, and for calls related to Covid-19 an interactive voice server —triaging the calls to dedicated computer stations— was developed. Patient evaluation and management were improved by introducing video consultation, sending of instruction using SMS, giving the patient the option to be called back. Prehospital EMT teams were also significantly reinforced by first aid volunteers, and additional MICU were created. Since January 20th 2020 all calls and patient records related to Covid-19 were flagged in the information system of Center-15 and a daily automated activity report was produced.

3 Methods

In this section, we describe the methods used in this work, in a way adapted to a general audience. Mathematical developments appear in Sections 7 to 9 and in the appendix.

---

[1]COVID19 APHP-Universities-INRIA-INSERM, Emergency calls are early indicators of ICU bed requirement during the COVID-19 epidemic, medRxiv:2020.06.02.20117499, June 2020.
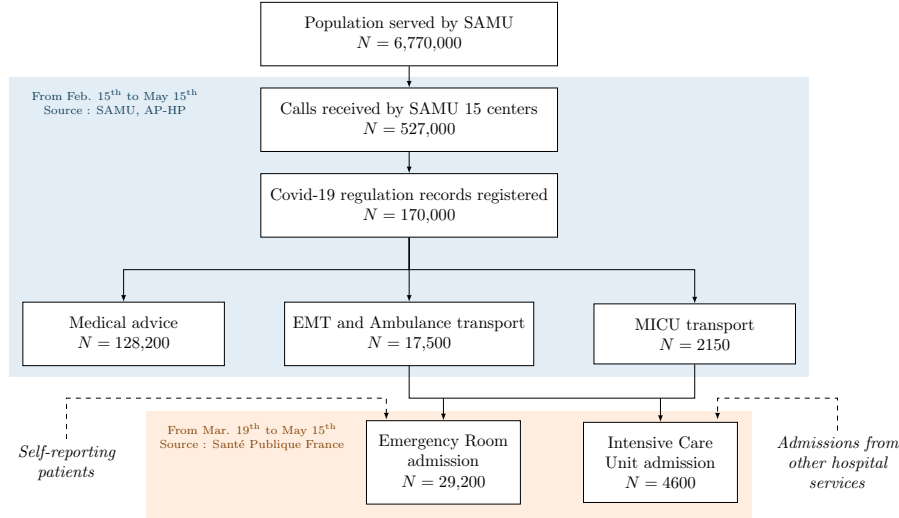
Figure 1: Flowchart: from calls to Center 15 to admission in hospital units. The numbers are summed over the departments 75, 92, 93 and 94 of the Paris area.

3.1 Classification of calls

In order to develop a mathematical analysis of the evolution of the epidemic, we classified the calls tagged as Covid-19 in three categories, according to the decision taken:

Class 1: calls resulting in the dispatch of a Mobile Intensive Care Unit;

Class 2: calls resulting in the dispatch of an ambulance staffed with EMT;

Class 3: calls resulting in no dispatch decision. Such calls correspond to different forms of medical advice (recommendation to consult a GP, specific instructions to the patient, etc.).

We shall denote by $Y_{\mathrm{MICU}}(t)$ (resp. $Y_{\mathrm{EMT}}(t)$ and $Y_{\mathrm{adv}}(t)$) the number of MICU transports (resp. the number of ambulances transport and the number of medical advices) on day t, for patients tagged with suspicion of Covid-19. We shall call these functions of time the *observables*, in contrast with $C(t)$, the actual number of new contaminations on day t, which cannot be measured. We developed a piece of software that computes these observables by analyzing the medical decisions associated with the patient records, made accessible daily by AP-HP.

3.2 Mathematical properties of the observables

To analyze these observables, we shall rely on a mathematical model.

A standard approach represents the evolution of an epidemic by an ordinary differential equation (SEIR ODE), representing the evolution of the population in four compartments: "susceptible" (S), "exposed" but not yet infectious (E), "infectious" (I), and finally, "removed" from the contamination chain (R), either by recovery or death. A refinement of the SEIR model splits the S and E compartments in sub-compartments corresponding to different age classes. It includes a contact matrix, providing differentiated age-dependent infectiosity rates [27]. Another refinement includes additional compartments, representing, for instance, patients at hospital [13].

In contrast with such ODE models, we use a partial differential equation (PDE), i.e., an infinite dimensional dynamical system, described in Section 7. Our approach is inspired by the PDE model of Kermack

and McKendrick [22]. We use PDE, rather than ODE, to take into account the presence of *delays* in the contamination process: the median incubation period of Covid-19 is estimated of 5.1 days, with a 95% confidence interval of 4.5-5.8 and an heavy tail [25], in line with other human coronaviruses having also long incubation times, like SARS [35] and MERS [34]. (This may be compared with a median incubation time of 1.4 days [95% CI, 1.3–1.6] for the toxigenic Cholera [2], or with an interval of 36 hours between infection by pneumonic plague and first symptoms in Brown Norway rats, with rapid letality, 2-4 days after infection [1].) ODE models assume exponentially distributed transitions times from one compartment from another. This entails that the interval elapsed between contamination and the time an individual becomes infectious can be arbitrarily small, so ODE models are more adapted to epidemics with a short incubation time. Using PDE, as is done in Section 7, takes delays into account.

We shall limit here our analysis to the early stage of the epidemic, assuming that the population that has been infected is much smaller than the susceptible population. This approximation is reasonable at least in the initial part of the epidemic, according to the study [32] which gives an estimate of 5.7% for the proportion of the population in France that has been infected prior to May 11th, 2020. Then, the dynamics becomes linear and order-preserving. The latter property entails that the observables are an increasing function of the size of the initial population that is either exposed or infected.

Results of Perron–Frobenius and of Krein-Rutman theory, which we recall in Section 7, entail that, *if the sanitary measures stay unchanged*, there is a rate λ, such that the number of newly contaminated individuals at day t grows as $C(t) \simeq K_C \exp(\lambda t)$, as $t \to \infty$, where K_C is a positive constant.

The number $\delta := (\log 2)/\lambda$, when it is positive, represents the *doubling time*: every δ days, the number of new contaminations per day doubles. When δ is negative, the epidemic is in a phase of exponential decay. Then, the opposite of δ yields the time after which the number of new contaminations per day is cut by half. For the analysis which follows, it is essential to consider, instead of $C(t)$, its logarithm, $\log C(t) \simeq \log K_C + \lambda t$. The exponential growth or decay of $C(t)$ corresponds to a linear growth or decay of the logarithm.

We shall also see in Section 7 that all the epidemiological observables evolve with the same rate. E.g., assuming that all the patients transported by MICU were contaminated τ_{MICU} days before the transport, and that a proportion π_{MICU} of the contaminated individuals will require MICU transport, we arrive at $Y_{\text{MICU}}(t) = \pi_{\text{MICU}} C(t - \tau_{\text{MICU}})$, and so $\log Y_{\text{MICU}}(t) \simeq \log \pi_{\text{MICU}} + \log K_C + \lambda(t - \tau_{\text{MICU}})$. Similar formulæ apply to Y_{EMT} and Y_{adv}, and to other observables based for instance on ICU admissions or deceases. A finer model of observables, taking into account a distribution of times τ_{MICU}, instead of a single value, is presented in Section 7.3.

For the analysis which follows, we shall keep in mind that *the logarithm of all the observables is asymptotically linear as $t \to \infty$*, and that *the rate, λ, is independent of the observable*.

3.3   Piecewise linear approximation of the logarithm of the observables

When the sanitary measures change, for instance, when lockdown is established, the rate λ changes. So, the logarithm of the observables cannot be approximated any more by a linear function. However, a general result, stated as Theorem 1 below, shows that this logarithm can be approximated by a *piecewise linear function* with as many linear pieces as there are phases of sanitary policy. This result stems from the order preserving and linear character of the epidemiological dynamics, and so, it holds for a broad class of epidemiological models; several examples of such models are discussed in Section 7.

In the Paris area, there are three relevant sanitary phases to consider from February to May, 2020: initial growth (no restrictions); "stade 2" (stage 2) starting on Feb. 29th (prevention measures), and then lockdown from March 17th to May 11th. Sanitary phases are further described in Section 5.1.

Since the number ν of sanitary phases is known (here $\nu = 3$), we can infer the different values of λ attached to each of these phases, by computing the best piecewise linear approximation, $\mathcal{L}(t)$ with at most ν pieces of the logarithm of an observable $Y(t)$. To compute a robust approximation, we minimize the ℓ_1 norm, $\sum_t |\mathcal{L}(t) - \log Y(t)|$, where the sum is taken over the days t in which the data are available. Finding the best approximation $\mathcal{L}$ is a difficult optimization problem, for the objective function is both non-smooth and non-convex. Methods to solve this problem are discussed in Appendix A.

5

3.4 Epidemic alarms based on doubling times

To construct epidemic alarms, we shall compute a linear fit, $\mathcal{L}(t) = \alpha + \beta t$, to the variables $\log Y(t)$, where Y is an epidemic observable. The principle is to trigger an alarm when the doubling time becomes positive, or equivalently, when the slope β becomes positive.

Assuming that values of $Y(t)$ are known over a temporal window, there are simple ready-to-use methods for computing estimates $\hat{\beta}$ for the slope β. We can also determine the probability p^+ that the slope is positive. These methods are detailed in Section 9. On their basis, we propose the following **alarm raising mechanism, allowing one to deploy a gradual response**.

This mechanism relies on the two following observables, Y_{adv}, the number of calls resulting in medical advice, and $Y_{\mathrm{disp}} := Y_{\mathrm{EMT}} + Y_{\mathrm{MICU}}$, the number of dispatched vehicles. The consolidation of the observables Y_{EMT} and Y_{MICU} is justified, because the two time series both correspond to the stage of aggravation, albeit with different degrees, and so they evolve more or less at the same time.

First define a temporal window of days t over which the linear fit $\mathcal{L}_{\mathrm{adv}}(t) = \alpha_{\mathrm{adv}} + \beta_{\mathrm{adv}}t$ to $\log Y_{\mathrm{adv}}(t)$ is made. By default we consider the last ten days prior to the current day. Similarly, we compute a linear fit $\mathcal{L}_{\mathrm{disp}}(t) = \alpha_{\mathrm{disp}} + \beta_{\mathrm{disp}}t$ to $\log Y_{\mathrm{disp}}(t)$ over the same time window.

Our algorithm will generate both a *warning* and *alarms*. A warning is a mere incentive to be careful. An unjustified warning is bothersome but generally harmless, so we accept a high probability of false positive for warnings. An alarm may imply some actions, so we wish to avoid false alarms. For this reason, we shall consider two different probability thresholds, $\vartheta_{\mathrm{alarm}}$ and $\vartheta_{\mathrm{warn}}$, say $\vartheta_{\mathrm{alarm}} = 75\%$ and $\vartheta_{\mathrm{warn}} = 25\%$. With this setting, we will be warned as soon as the probability of the undesirable event is $\geqslant 25\%$, and we will be alarmed when the same probability becomes $\geqslant 75\%$. Of course, these thresholds can be changed, depending on the risk level deemed to be acceptable. We shall denote by p^+_{adv} the probability that the slope β_{adv} is positive, and by p^+_{disp} the probability that β_{disp} is positive. These probabilities are evaluated on the basis of statistical assumptions detailed in Section 9.

1. A **warning** is provided when $p^+_{\mathrm{adv}} \geqslant \vartheta_{\mathrm{warn}}$, meaning that the probability that the slope β_{adv} of the curve of the logarithm of the *calls for medical advice* over the corresponding time window be positive is at least $\vartheta_{\mathrm{warn}}$. This should be interpreted as a mere warning of epidemic risk: choosing $\vartheta_{\mathrm{warn}}$ as above, the odds are at least 25% that the epidemic is growing.

2. This warning is subsequently transformed into an **alarm** when $p^+_{\mathrm{adv}} \geqslant \vartheta_{\mathrm{alarm}}$. Choosing $\vartheta_{\mathrm{alarm}}$ as above, the odds that the epidemic is growing are now at least 75%.

3. Such an alarm is then subsequently transformed into a **confirmed alarm** if we still have $p^+_{\mathrm{adv}} \geqslant \vartheta_{\mathrm{alarm}}$, and if, in addition, $p^+_{\mathrm{disp}} \geqslant \vartheta_{\mathrm{alarm}}$, meaning that the probability that the slope of the logarithm of the curve of ambulances and MICU dispatches be positive is now above $\vartheta_{\mathrm{alarm}}$. Again, this estimate is defined in terms of a time window over which β_{disp} is estimated. We use the same default values of ten days and $\vartheta_{\mathrm{alarm}}$ as above.

As shown in Section 4, the indicators based on vehicle dispatch are by far less noisy than the indicators based on calls for medical advices, but their evolution is delayed. This is the rationale for using medical advice for an early warning and early alarm, and then vehicle dispatch for confirmation.

Instead of considering the probability p^+, we could consider the upper and lower bounds of a confidence interval $[\beta^-_{\epsilon}, \beta^+_{\epsilon}]$ for the estimated slope β, with a probability threshold ϵ. Then we may, trigger a warning when $\beta^+_{\epsilon} \geqslant 0$, and an alarm when $\beta^-_{\epsilon} \geqslant 0$. This leads to an essentially equivalent mechanism. We prefer the algorithm above as it allows to interpret the thresholds in terms of false positives and false negatives.

Given the severity of the risk implied by Covid-19, it may be desirable to complete the previous alarm, based only on tail probabilities of the slope, by a different type of alarm, based on a threshold of doubling time, D. The alarm will be triggered if the odds that the doubling time be positive and smaller than D are at least one half. An indicative value of D might be 14 days: a doubling of the number of arrivals of Covid-19 patients in hospital services every 14 days may be quite challenging, justifying an alarm, and the slope corresponding to this doubling time seems significant enough to avoid false alarms. Again, the value of

D can be changed arbitrarily depending on the acceptable level of risk. Moreover, this other type of alarm can still be implemented in two stages: early alarm, with the medical advice signal, and then confirmed alarm, with the vehicle dispatch signal.

In addition, Section 9 provides more sophisticated ready-to-use methods for obtaining sharper confidence intervals or probabilities for the slope β, resulting in more precise alarm mechanisms, when different time series are available. We require, however, that these series correspond to events occurring approximately at the same stage in the pathology unfolding. Here, we used the trivial aggregator, $Y_{\text{disp}} = Y_{\text{EMT}} + Y_{\text{MICU}}$. There is an optimal way to mix different series to minimize the variance of the composite estimator, explained in Section 9.

This methodology is generic. It could thus also apply to obtain a sharper confidence interval for the early indicator by combining its estimate $\hat{\beta}_{\text{adv}}$ with that of other time series associated with signals that correspond to the same stage in pathology unfolding. Specifically, the count $Y_{\text{GP}}(t)$ of patients consulting general practitioners for recently developed Covid-19 symptoms, if available, provides such a signal. A linear fit to $\log Y_{\text{GP}}(t)$ would then yield an estimate $\hat{\beta}_{\text{GP}}$ which can be combined with $\hat{\beta}_{\text{adv}}$ to refine the corresponding confidence interval. In this way, we can mix several early but noisy indicators to get an early but less noisy consolidated indicator.

4   Results – data analysis

4.1   Key figures and graphs

From February 15th to May 15th, we counted a total of 170,166 patient files tagged with a suspicion of Covid-19, distributed as follows in the different departments: 53,646 in Dep. 75; 36,721 in Dep. 92; 49,703 in Dep. 93; and 30,096 in Dep. 94.

The flow of calls to the SAMU of the Paris area, and its impact on ER and ICU, is shown on Figure 1. The data concerning the ER and the ICU are taken from the governmental website SPF (Santé Publique France) [15], it is available only from March 19th.

On Figure 2, we represent, in logarithmic ordinates, the numbers of events of different types, summed over the four departments of the Paris area (75, 92, 93 and 94): (i) the number of patients calling the SAMU (including patients not calling for Covid-19 suspicion); (ii) the number of calls tagged as Covid-19 not resulting in a vehicle dispatch (i.e., as discussed in §3.1, all kinds of medical advices); (iii) the number of calls tagged as Covid-19 resulting in an ambulance or MICU dispatch,

We obtained the data (i) by analyzing the phone operator log files. Since a patient may call the Center 15 several times, we eliminated multiple calls to count unique patients. To compute data (ii) and (iii), we developed a software to analyze the "medical decision" field of the regulation records.

Using logarithmic ordinates is essential on Figure 2, as it allows to visualize on the same graph signals of different orders of magnitude (e.g, there is a ratio of 20 between the peak number of patients calling and the peak number of vehicles dispatched).

The evolution of the number of vehicles dispatched (MICU and ambulances) is shown on Figure 3, for each department of the Paris area (still with logarithmic ordinates).

We provide in Table 1 the doubling times of the number of vehicles dispatched (ambulances and MICU), for the different departments, measured in days (abbreviation "d").

We now draw several conclusions from the previous analysis.

4.2   The increase in the number of calls for medical advice provides an early, but noisy, indicator of the epidemic growth

As shown in Figure 2, the peak of the number of calls for medical advice was on March 13th. However, this date, four days before the lockdown (March 17th), is not consistent with epidemiological modeling. This peak seems rather to be caused by announcements to the population, see the discussion in Section 6.1.
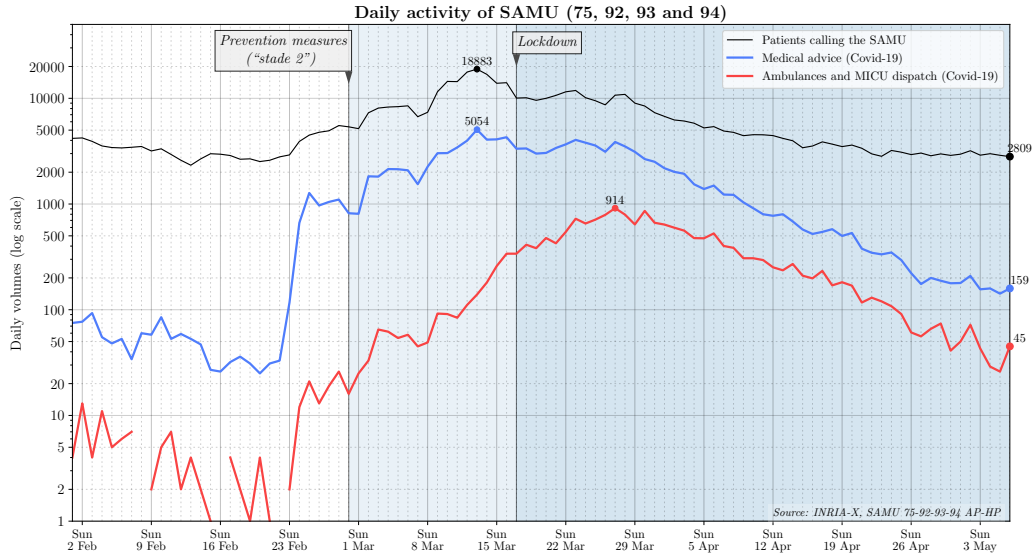
Figure 2: Number of patients calling Center-15, of MICU and ambulances dispatch for Covid-19 suspicion in the Paris area (departments 75, 92, 93 and 94)
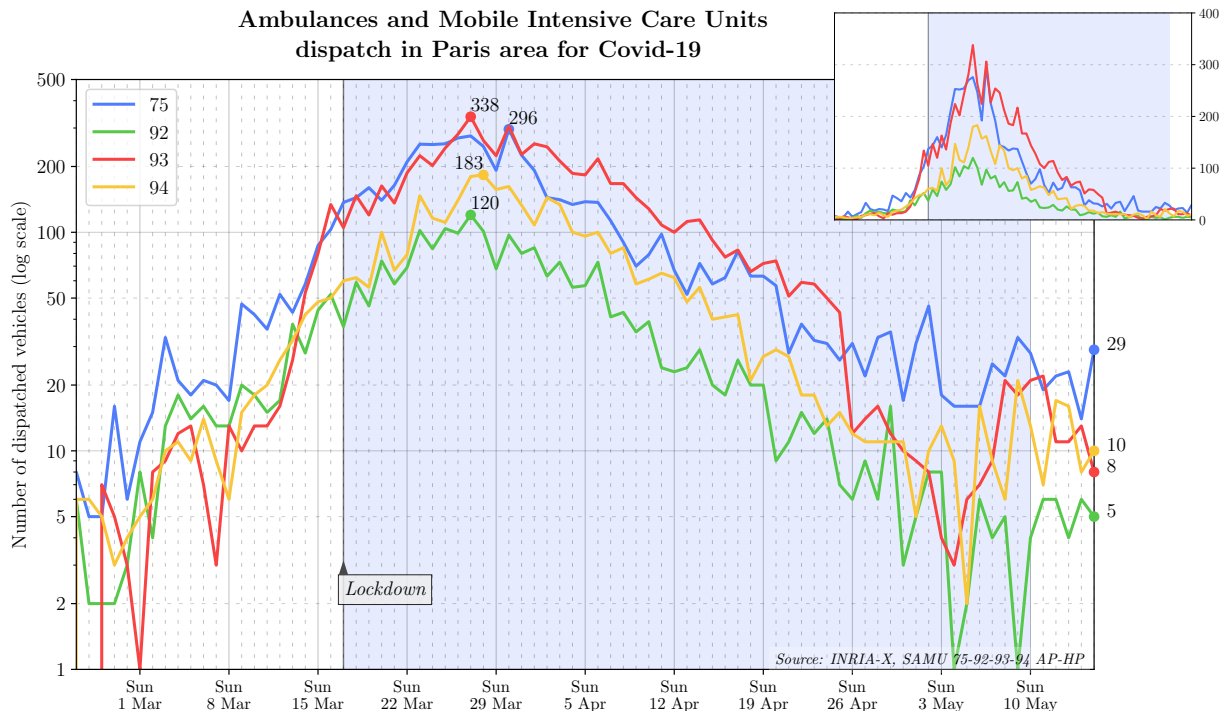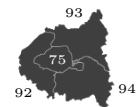

Figure 3: Comparison of the evolution of the epidemic in the different departments of the Paris area: numbers of vehicles dispatch by department. The figure inset displays the same curves in usual linear ordinates to keep in mind the different magnitudes at stake. A map of the Paris area, showing the departments 75, 92, 93, 94, is at the bottom right of the figure.

	Feb. 28th– Mar. 15th	Mar. 15th– Mar. 29th	Mar. 29th– April 24th
75	5.9 d	9.8 d	-9.4 d
92	4.9 d	10.6 d	-8.3 d
93	4.2 d	8.5 d	-10.2 d
94	4.6 d	6.9 d	-7.7 d

Table 1: Doubling time of the number of MICU and ambulances dispatched, for different periods, for each department, obtained by a least squares approximation of the logarithm of this number. The opposite of a negative doubling time yields the halving time.

4.3 The epidemic kinetics vary strongly across neighboring departments

In the initial phase of the epidemic (Feb. 28th–March 15th), the doubling time was significantly shorter in the 93 department (4.2 d) than in central Paris (5.9 d). The 93 department, with 1.6M inhabitants, is less populated than central Paris (2.1M inhabitants). Another difference between the departments concerns mobility. Movement from the population from central Paris to smaller towns and cities or to countryside were observed, after March 12th, the date of the first presidential address concerning the Covid-19 crisis.

In order to quantify this mobility, we requested information from Enedis, the company in charge of the electricity distribution network in France, and also from Orange and SFR, two operators of mobile phone networks.

Enedis provided us with an estimation of the departure rates of households, based on a variation of the volume of electricity consumed, aggregated at the level of departments and districts (i.e., *arrondissements*).

SFR provided us with estimates of daily flows from the Paris area to other regions, again aggregated at the scale of the departments or districts, based on mobile phone activity, confirming this decrease of population.

Orange Flux Vision provided us with daily population estimates, at the scale of department, based on mobile phone activity. By March 30th, the population, during the night, was estimated to be 1.6M inhabitants in central Paris, versus 1.35M in the 93.

However, the epidemic peak was higher in the 93 than in the 75 (338 dispatches versus 296). The contraction rate in the period after the peak (March 29th–Apr. 24th) was also smaller in the 93, with a halving time of 10.2 days, to be compared with 9.4 days in the 75. Possible explanations for these strong spatial discrepancies are discussed in Section 6.5.

5 Results – mathematical modeling

5.1 Delay between implementation of sanitary policies and its effect on hospital admissions

We explained in Section 3.3, based on Theorem 1 below, that the logarithm of an epidemic observable can be approached by a piecewise linear map with as many pieces as there are stages of sanitary measures.

So, we look for the best approximation, in the ℓ_1 norm, of the logarithm of the number of vehicles dispatched (ambulances and MICU), by a piecewise linear map with at most three pieces. This best approximation is shown on Figure 4. It is computed by the method of Appendix A.

In order to evaluate the influence of a sanitary measure on the growth of the epidemic, an approach is to compare the date of the measure with the date of the change of slope of the logarithmic curve, consecutive to the measure. This method is expected to be more robust than, for instance, a comparison of peak values, because the best piecewise-linear approximation is obtained by an optimization procedure *taking the whole sequence into account*. Indeed, a local corruption of data will not change significantly the date of change of slope, if the problem is *well conditioned*. This is the case in particular if the difference between consecutive

slopes is sufficiently important. In other words, we can identify in a more robust manner the time of effect of a strong measure than of a mild one.

Let us recall the main changes of sanitary measures in the Paris area, between February and May 2020. We may distinguish the following phases:

- *Initial development of the epidemic*, no general sanitary measures in the Paris area, until Feb 29th, first day of so-called "stade 2" by the authorities (following "stade 1" in which measures intended to prevent the introduction of the virus in France – like quarantine in specific cases– were taken).

- *"Stade 2" (stage 2) measures*: general instructions of social distancing given to the population (e.g., not shaking hands), ban on large gatherings. Moreover, some large companies created crisis committees, and decided to take more restrictive measures than the ones required by the authorities, including for instance banning meetings with more of 10 people, and banning business travels. Restrictive measures in companies were deployed gradually during the work week from March 2nd to March 6th.

- *School closure* on March 16th.

- *Lockdown* on March 17th. The lockdown ended on May 11th, throughout the country.

Hence, we may interpret the variations in the slope in the piecewise linear approximation of the logarithm of the number of ambulances and MICU dispatched, shown on Figure 4, as the effect of sanitary measures. The dates where the slope changes are represented in the figure by dotted lines. Thus, the latest breakpoint of the piecewise linear approximation of the 75 curve (in blue) arises on March 26th, to be compared with March 30th in the 93 (red curve). The dates of breakpoints in the 92 and 94 are intermediate. Given the first strong measure (closing of schools) was taken on March 16th, we may evaluate the delay between a sanitary measure and its effect on the ambulances and MICU dispatch to be between 10 and 14 days. This corresponds to a delay between contamination and occurrence of severe symptoms.

5.2 Construction of statistical indicators of epidemic resurgence based on emergency calls

We implemented the alarm mechanism based on the inference of doubling times described in Section 3.4 and further explained in Section 9. The method is illustrated on Figure 5. Given a time period where the data is known, we perform a linear regression on the number of medical advices and vehicles dispatched.

The light shaded, tubular areas around the curves are based on confidence intervals for the fluctuations of the observed log-counts. The dark-shaded, trapezoidal areas prolongate the tubular areas with straight lines, the slopes of which correspond to confidence intervals for the slope of the linear regression. We display such confidence domains for the last known data in May, performing a 6-day forecast based on the last ten days. In order to validate the method, we also display these domains for older data in March and April, performing for each a 6-day forecast based on a number of past days. For these time-periods, the short-term confidence domains are seen to satisfactorily contain the data of the following days. Observe how the shape of the confidence trapezoids depends on the number of points and the variance of the data used to compute them.

The needle-shaped (or clock hand) indicators depicted in each trapezoidal confidence domain illustrate the alarm mechanism of Section 3.4. For each slope inference, there is a $\vartheta_{warn} = 25\%$ probability that the real-value of the slope we estimated using recent data is greater than the slope of the thin needle on the picture. Likewise there is a $\vartheta_{alarm} = 75\%$ probability that the real slope is greater than the slope of the fat hand. As a result, a warning (resp. an alarm) on the dynamics of the medical advice curve should be triggered as soon as the thin needle (resp. the fat needle) has a positive angle with respect to the horizontal. We have depicted the horizontal with dashed lines to enhance readability.

On May 25nd, no warning nor alarm is triggered, since all the needle-indicators are below the positivity threshold, indicating with at least a 75% confidence level that based on the last ten days, the two observed signals are on a decreasing trend. Note that due to the relative stagnation of the medical advice curve in the
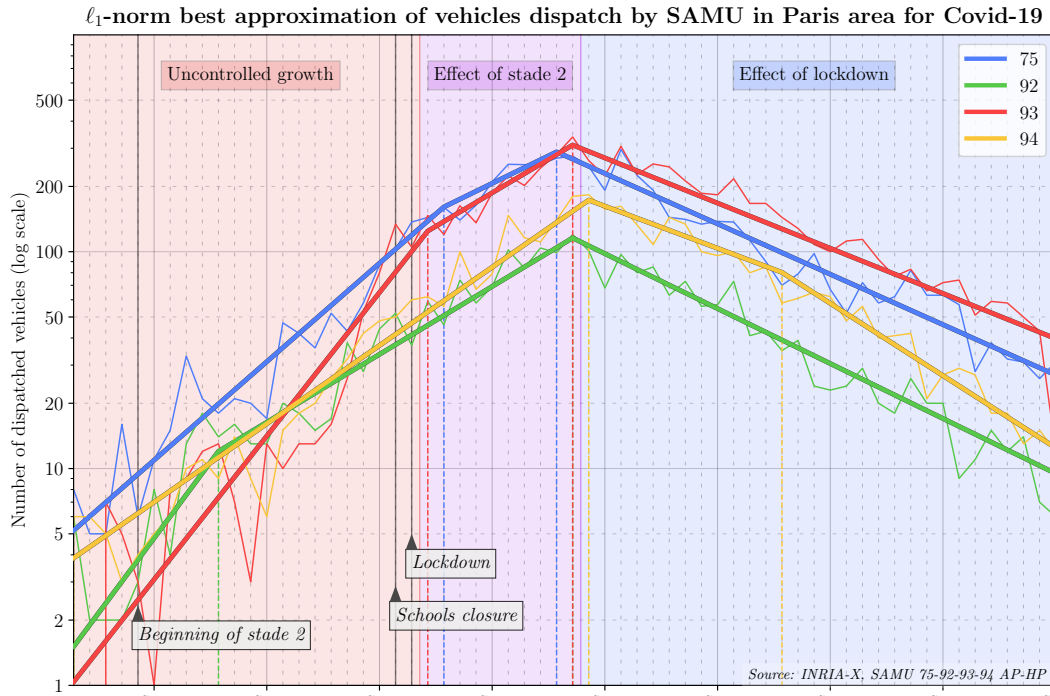
10

Figure 4: Logarithm of the number of ambulances dispatched: the effect of the successive sanitary measures
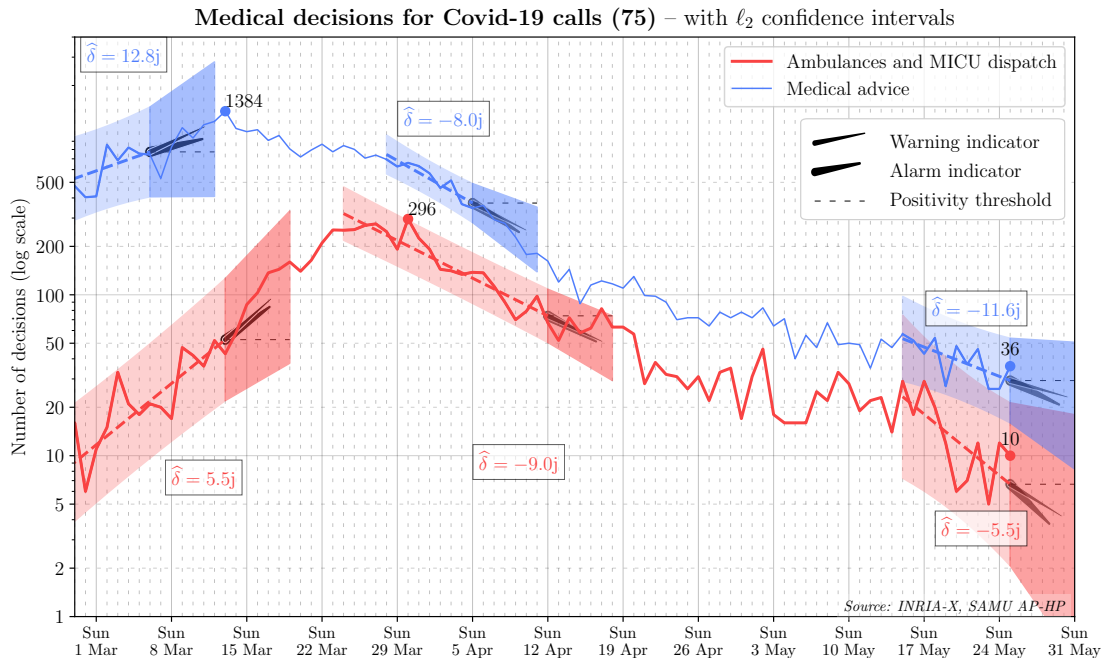

Figure 5: Short term predictor for SAMU 75, with confidence regions and warning and alarms indicators

two first weeks of May, computing our indicators a few days earlier (such as May 22nd) would have raised an warning to arouse vigilance due to the uncertainty on the future trend, but no alarm.

The numbers ϑ_{warn} and ϑ_{alarm} need to be carefully calibrated, and for this, additional data for the forecoming weeks may be helpful.

6 Discussion

6.1 Calls resulting in medical advice are highly influenced by the instructions given to the population

The blue curve on Figures 2 and 5 counts the number of times medical advices was given, of all kinds (calls resulting in recommendations given to the patient but no vehicle dispatched). It is generally associated with early events in the unfolding of the pathology, and in particular occurrence of the first symptoms. An estimate of 5.1 days between the date of contamination and the date of the first symptoms is given in [25], so we may assume that calls for medical advice are made by patients 5-8 days after contamination. We observed in Section 4.2 that the peak in the number of calls resulting in medical advice was on March 13th. Hence, assuming the peak of contaminations was just before lockdown, the peak of the curve of new symptom occurrences should occur *only several days after the lockdown time* (March 17th). This indicates that the curve of calls resulting in a medical advice did not give a reliable picture of the epidemic growth around March 13th. Indeed, this curve is very sensitive to changes in the instructions given to the population and to political announcements, notable examples of which include the following: recommendation to patients to call emergency number 15, instead of going directly to emergency departments (to avoid contamination and overcrowding); – the presidential announcement on March 12th of more restrictive measures to be deployed from March 16th, making the population more aware of the growth of the epidemic.

6.2 The indicators of medical advice given and ambulances and MICU dispatch can be used to monitor the epidemic

Setting aside perturbations due to political announcements or changes in the policy for calling SAMU, the curve of calls for medical advice should be a reliable and early estimate of the curve of ambulances dispatched, which is triggered at a later stage in the unfolding of the pathology when symptom severity increases. Thus, it gives an early signal allowing both SAMU and hospitals to anticipate by several days an increase in load.

We can give a rough estimate of this delay by considering the peak dates in Figure 2. Epidemiological modeling indicates that the number of new contaminations grows exponentially until a sanitary measure that is strong enough to contain the epidemic is taken. In the present case, the candidates for such strong measures are the school closing (March 16th) or the lockdown (March 17th). Considering the last day previous to the measures, we may assume that the peak date for new contaminations was on March 15th or March 16th. As mentioned above, according to [25], the time between contamination and first symptoms is estimated to be 5.1 days. Hence, if calls for medical advice were representative of first symptoms, the peak value for these calls would have been between around March 20th or March 21st. The peak of the number of dispatches of ambulances and MICU was on March 27th. This leads to an estimate of 6-7 days for the delay between the curve of the true need for medical advice and the curve of vehicles dispatch.

The alarm mechanism was developed during the crisis, after March 20th. Had it been available before, in the 75, the early warning would have occurred by February 24th (for both Covid-19 indicators on medical advice and vehicle dispatch), alarm would have been been triggered on February 25th based on medical advice, and a confirmation alarm would have occurred on February 27th based on ambulance and MICU dispatch.

Tracking the same signal at a finer spatial resolution (a neighborhood rather than a department) may enable epidemic surveillance during the period following the lifting of measures such as travel bans. Indeed in view of the spatial differentiation in doubling times at the department level, it appears plausible that disparities may also be present at much finer spatial granularities, with resurgences localized to towns or

neighborhoods. Deployment of alarm mechanisms constructed from event counts at finer spatial granularities could be used to identify clusters of resurgence early on, and guide subsequent action.

Seasonal influenza, whose early symptoms can be mistaken for those of Covid-19, can also trigger a linear growth of the logarithms of the number of calls, or vehicle dispatched. However the slope of this logarithmic curve is expected to be shallower, owing to the lower contagiosity of influenza. One could thus distinguish, on the basis of the observed slope, whether one is confronted with seasonal flu or with an outbreak of Covid-19.

6.3 Jumps of the curves of the number of calls may be caused by large clusters or influenced by neighboring countries

The curves of the number of calls for medical advice, and of vehicles dispatched (Figure 2), both jump between February 23rd and 25th. Epidemiogical models are unlikely to produce shocks of this type in the absence of exceptional factors. Several significant epidemic events occurred at nearby dates, including:

1. The development of the Covid-19 epidemic in the north of Italy, a region closely tied to France (the first lockdowns occurred around February 21st in the province of Lodi). The school vacation ended on February 23th in the Paris area. A significant number of parisians went back from Italy during the week-end of February 22nd-23th (end of school vacation).

2. A large Evangelist meeting (Semaine de Carême de l'Eglise La Porte Ouverte Chrétienne de Bourtzwiller, Haut-Rhin), from February 17th to February 21st, identified by Agence Régionale de Santé Grand Est as the source of a cluster phenomenon [12].

The potential influence of the north Italy epidemic was pointed out by Paul-Georges Reuter (private communication). At this stage, the essential factors are not yet known. The influence of mobility on the development of the epidemic in the Paris area will be studied in a further work.

6.4 Patients from different areas tend to call the SAMU at different stages of the pathology

Considering the piecewise linear curves in Figure 4, we note that in the 93, the date of the break point is shifted of 3 days, by comparison with the 75, suggesting that by this time, the patients of 93 were calling Center 15 at a later stage of the evolution of the disease. This hypothesis is confirmed by an examination of the ratio of the number of MICU dispatched over the number of ambulances dispatched. For instance, on March 30th, there were 9 MICU dispatches and 276 other dispatches in the 75, to be compared with 25 MICU dispatches and 262 other dispatches in the 93, i.e., ratios of 3.3% in the 75 and 9.5% in the 93.

In the same way, the first breakpoints of the curves give an indication of the times at which "stade 2" measures influence the epidemic growth. These dates range from March 15th (for 92) to March 22nd (for 75). It may be the case that these dates are dispersed because the change of slope is relatively small, meaning that the effect of stade 2 is mild. Indeed, the milder the slope change, the more the estimation of the corresponding date is sensitive to noise. Another effect which may have perturbed the curves is the important mobility of the population in the 4 departments, between March 12th and March 16th.

6.5 The strong spatial heterogeneity of the evolution of the epidemic may be explained by local conditions

We observed in Section 4.3 that in the initial phase of the epidemic (Feb. 28th–March 15th), the doubling time was significantly shorter in the 93 department, whereas in the contraction phase, the halving time was significantly higher.

One may speculate that the contraction rate in the lockdown phase is influenced by intra-familial contaminations. In this respect, according to a survey of INSEE, the national institute of statistics, the average size of a household is of 2.6 in the 93, versus 1.9 in the 75 (values in 2016 [19]).

We also remark that just after the peak on March 27, and up to April 6, the curve of the department 93 on Figure 3 has the shape of a high, oscillating plateau, decaying more slowly than the curve of the department 75. This may be caused by changes in the nature of the dominant mode of contaminations, intra-familial contamination becoming an essential part of the kinetics during lockdown.

One may also speculate that the blowup rate in the initial phase is higher when the population is more dependent on public transport, or working in jobs with more contamination risk. These aspects will be further studied elsewhere.

After Stade 2 was announced, during the period from March 2$^{\text{nd}}$ to March 6$^{\text{th}}$, a number of large companies took specific measures (e.g., forbidding avoidable small group meetings, enforcing travel restrictions, restricting office access), in addition to the general measures (not shaking hands, forbidding large meetings) enforced by the authorities. This may have led to a decrease of the number of contamination on the workplace, and one explanation for the increase of the doubling time.

7 Epidemiological model based on transport PDE

7.1 Taking delays into account: a transport PDE SEIR model

We now introduce a multi-compartment transport PDE model, representing the dynamics of Covid-19. As explained in Section 3.2, in contrast to ODE models, that assume that the transition time from a compartment to the next one has an exponential distribution, PDE models capture *transition delays* bounded away from zero, an essential feature of Covid-19. An interest of this PDE model also lies in its unifying character: it includes as special cases, or as variations, SEIR ODE models that have been considered [27, 13].

We shall keep the traditional decomposition of individuals in compartments, "susceptible" (S), "exposed" (E), "infectious" (I), and "removed" from the contamination chain (R), as explained in Section 3.2, but the state variables attached to the E and I compartments will take the time elapsed in the compartment into account, and thus, will be infinite dimensional.

For all $t \geqslant 0$, we denote by $n_E(x,t)$ the density of the number of individuals that were contaminated x time units before time t, and that are not yet infectious at time t, i.e., the number of exposed individuals that began to be exposed at time $t - x$. Then, the size of the exposed population at time t is given by

$$E(t) = \int_0^\infty n_E(x,t)\,\mathrm{d}x \ . \tag{1}$$

Similarly, we denote by $n_I(x,t)$ the density of the number of individuals that became infectious x time units before time t, and that are not yet removed from the contamination chain at that time. Then, the size of the infectious population at time t is given by

$$I(t) = \int_0^\infty n_I(x,t)\,\mathrm{d}x \ . \tag{2}$$

Finally, we denote by $S(t)$ the number of susceptible individuals at time t, and by $R(t)$ the number of individuals that have been removed from the contamination chain before time t.

The total population at time t is given by

$$N(t) := S(t) + E(t) + I(t) + R(t) \ .$$

14

We consider the following system of PDE and ODE, with integral terms in the boundary conditions:

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -\frac{S(t)}{N(t)} \int_0^\infty K_{I \to E}(x,t) n_I(x,t)\,\mathrm{d}x \ , \tag{3a}$$

$$n_E(0,t) = \frac{S(t)}{N(t)} \int_0^\infty K_{I \to E}(x,t) n_I(x,t)\,\mathrm{d}x \ , \qquad \frac{\partial n_E}{\partial t}(x,t) + \frac{\partial n_E}{\partial x}(x,t) + K_{E \to I}(x,t) n_E(x,t) = 0 \ , \tag{3b}$$

$$n_I(0,t) = \int_0^\infty K_{E \to I}(x,t) n_E(x,t)\,\mathrm{d}x \ , \qquad \frac{\partial n_I}{\partial t}(x,t) + \frac{\partial n_I}{\partial x}(x,t) + K_{I \to R}(x,t) n_I(x,t) = 0 \ , \tag{3c}$$

$$\frac{\mathrm{d}R}{\mathrm{d}t} = \int_0^\infty K_{I \to R}(x,t) n_I(x,t)\,\mathrm{d}x \ . \tag{3d}$$

We assume that an initial condition at time 0, $S(0)$, $n_E(\cdot,0)$, $n_I(\cdot,0)$ and $R(0)$ is given.

This is inspired by the so called "age structured models" considered in population dynamics. Kermack and McKendrick developed the first model of this kind to analyze the Plague epidemy of Dec. 1905 – July 1906 in Mumbai [22]. Von Forster [37] studied a similar model. Nowadays, these models are used as a general tool in population dynamics, with applications to biology and ecology [31, 28],

In these models, "age" refers to the age elapsed in a compartment – each transition to a new compartment resets to zero the "age" of an individual. In contrast, in the classical SEIR literature based on ODE, the standard notion of age (time elapsed since birth) is taken into account, via a contact matrix tabulating age-dependent infectiosity rates [27]. These two notions of age should not be confused. In the sequel, we shall use quotes, as in "age", to denote the age in a compartment, and will omit quotes to denote the ordinary age (since birth).

We suppose that $K_{I \to E}$, $K_{E \to I}$ and $K_{I \to R}$ are given *nonnegative* functions. The value $K_{E \to I}(x,t)$ gives the departure rate from the compartment E to the compartment I, for individuals of "age" x in the compartment E, at time t. Similarly, $K_{I \to R}(x,t)$ gives the departure rate from the compartment I to the compartment R. As in the classical SEIR model, the departure term from the susceptible compartment, i.e., the right-hand-side of (3a) is bilinear in the number $S(t)$ of susceptible individuals and in the population of infectious individuals $n_I(\cdot,t)$, and we normalize by the size of the population $N(t)$. The term $K_{I \to E}(x,t)$ can interpreted as an infection rate.

Differentiating $N(t)$ with respect to time, using the system above, and assuming that for all $t \geqslant 0$, $n_E(x,t)$ and $n_I(x,t)$ vanish when x tends to infinity, we verify that the total population $N(t)$ is independent of time.

When the functions $K_{I \to E}, K_{E \to I}$ and $K_{I \to R}$ are constant, taking into account (1) and (2), we recover the classical SEIR model from the dynamics (3):

$$\dot{S} = -\frac{S}{N} K_{I \to E} I \ , \tag{4a}$$

$$\dot{E} = \frac{S}{N} K_{I \to E} I - K_{E \to I} E \ , \tag{4b}$$

$$\dot{I} = K_{E \to I} E - K_{I \to R} I \ , \tag{4c}$$

$$\dot{R} = K_{I \to R} I \ . \tag{4d}$$

In the sequel, we shall consider (3) instead of (4), and we shall assume that the rates $K_{E \to I}(x,t) = K_{E \to I}(x)$ and $K_{I \to R}(x,t) = K_{I \to R}(x)$ are functions of x, independent of time. The rate $K_{I \to E}$ will have the product form

$$K_{I \to E}(x,t) = \mu(t)\psi(x) \ .$$

The function $\psi(\cdot)$ is fixed, it is nonnegative and not *a.e.* zero. In this way, the infectiosity of an individual depends on his "age" in the infectious phase, whereas the term $\mu(t)$ represents the control of the epidemic by sanitary measures (social distancing, wearing masks, closing schools, lockdown, etc.). We shall assume that the infectiosity rate $K_{I \to E}(x,t)$ is the only parameter which can be controlled, hence, $\mu(\cdot)$ is a decision

variable. A variant of the ODE model (4), in which $K_{I \to E}$ depends on time, but not on x, is considered in [9].

For epidemics in their early stages, i.e., when the number of individuals in the exposed, infectious, or removed compartments is negligible with respect to the number of susceptible individuals, the classical SEIR model is well-approximated by a linear system (see e.g. [22, 3]) tracking only the populations in the (E) and (I) compartments. As noted in Section 3.2, the fraction of the French population exposed prior to May 11 is estimated of 5.7% (see [32]), which justifies reliance on this linear approximation in our context. The same approximation applies to the present PDE model. This is translated to the assumption $S(t)/N(t) \simeq 1$, and we are reduced to the following system:

$$n_E(0,t) = \int_0^\infty \mu(t)\psi(x)n_I(x,t)\,\mathrm{d}x \ , \qquad \frac{\partial n_E}{\partial t}(x,t) + \frac{\partial n_E}{\partial x}(x,t) + K_{E \to I}(x)n_E(x,t) = 0 \,, \quad \text{for } x > 0, \quad (5\text{a})$$

$$n_I(0,t) = \int_0^\infty K_{E \to I}(x)n_E(x,t)\,\mathrm{d}x \,, \qquad \frac{\partial n_I}{\partial t}(x,t) + \frac{\partial n_I}{\partial x}(x,t) + K_{I \to R}(x)n_I(x,t) = 0 \ , \quad \text{for } x > 0 \ . \quad (5\text{b})$$

This is a two-compartment generalization of the renewal equation, studied in Chapter 3 of [31].

In the sequel, we shall assume that there is a maximal "age" x_E^* of an individual in the exposed state. Similarly, we shall assume that is a maximal "age" x_I^* of an individual in the infectious state. These assumptions, which are consistent with epidemiological observations [25], will be incorporated in our model by forcing all remaining exposed individuals of "age" x_E^* to become infectious, with "age" 0. Similarly, all the remaining infectious individuals are removed when reaching "age" x_I^*. So, the function n_E is now only defined on the interval $[0, x_E^*]$, and similarly, n_I is only defined on $[0, x_I^*]$. This leads to the following system:

$$n_E(0,t) = \int_0^{x_I^*} \mu(t)\psi(x)n_I(x,t)\,\mathrm{d}x \,, \qquad \frac{\partial n_E}{\partial t}(x,t) + \frac{\partial n_E}{\partial x}(x,t) + K_{E \to I}(x)n_E(x,t) = 0 \,, \quad \text{for } 0 < x < x_E^* \,, \tag{6a}$$

$$n_I(0,t) = \int_0^{x_E^*} K_{E \to I}(x)n_E(x,t)\,\mathrm{d}x + n_E(x_E^*,t) \,, \tag{6b}$$

$$\frac{\partial n_I}{\partial t}(x,t) + \frac{\partial n_I}{\partial x}(x,t) + K_{I \to R}(x)n_I(x,t) = 0 \,, \quad \text{for } 0 < x < x_I^* \,, \tag{6c}$$

$$\frac{\mathrm{d}R}{\mathrm{d}t}(t) = \int_0^{x_I^*} K_{I \to R}(x)n_I(x,t)\mathrm{d}x + n_I(x_I^*) \ . \tag{6d}$$

This system may be obtained as a specialization of (5), in which $K_{E \to I}(x)$ is replaced by $K_{E \to I}(x)\mathbb{1}_{[0,x_E^*]}(x) + \delta_{x_E^*}(x)$, where $\mathbb{1}$ denotes the indicator function of a set, and δ denotes Dirac's delta function.

We shall assume, in the sequel, that the following assumption holds.

Asssumption 1. *The functions $K_{E \to I}(\cdot)$, defined on $[0, x_E^*]$, and $\psi(\cdot)$ and $K_{I \to R}(\cdot)$, defined on $[0, x_I^*]$, are nonnegative, measurable and bounded. Moreover, the function ψ does not vanish a.e. and the point x_I^* is the maximum of the essential support of the function ψ.*

Indeed, considering the boundary condition in (6a), we see that a population of "age" $x > \max \operatorname{ess\,supp} \psi$ in the infected (I) compartment will not participate any more to the contamination chain. Hence, the last part of Assumption 1 is needed to interpret R has the number of *all* the removed individuals.

Systems of PDE of this nature have been studied in particular by Michel, Mischler and Perthame, see [28, 31], and also, with an abstract semigroup perspective, in the work by Mischler and Scher [29].

Then, using the boundedness of the coefficients (Assumption 1), and arguing as in the proof of Theorem 3.1 of [31] – which concerns the case of a single compartment – one can show that the system (6) admits a unique solution in the distribution sense $n := (n_E, n_I)$ with $n_E \in \mathcal{C}(\mathbb{R}_{\geqslant 0}, L^1([0, x_E^*]))$ and $n_I \in \mathcal{C}(\mathbb{R}_{\geqslant 0}, L^1([0, x_I^*]))$. Hence, we can associate to the PDE (5) a well defined family of time evolution linear operators $(T_{s,t})_{t \geqslant s \geqslant 0}$, acting on the space $L^1([0, x_E^*]) \times L^1([0, x_I^*])$. The operator $T_{s,t}$ maps an initial condition at time $s \geqslant 0$, that is a couple of functions $n(\cdot, s) := (n_E(\cdot, s), n_I(\cdot, s))$, to the couple of functions

16

$n(\cdot, t) := (n_E(\cdot, t), n_I(\cdot, t))$ at $t \geqslant s$. These operators are order preserving, meaning that, if $n^1(\cdot, s)$ and $n^2(\cdot, s)$ are two initial conditions such that $n_E^1(x, s) \leqslant n_E^2(x, s)$ and $n_I^1(x, s) \leqslant n_I^2(x, s)$ for all $x \geqslant 0$, then the inequalities $n_E^1(x, t) \leqslant n_E^2(x, t)$ and $n_I^1(x, t) \leqslant n_I^2(x, t)$ hold for all $x \geqslant 0$ and for all $t \geqslant s$.

An alternative modeling, more in the spirit of [22], would be to consider a single compartment, describing the evolution of the density $n(x, t)$ of individuals that were contaminated at time $t - x$ by the system:

$$n(0, t) = \int_0^\infty \mu(t) \psi(x) n(x, t) \, \mathrm{d}x \,, \quad \frac{\partial n}{\partial t}(x, t) + \frac{\partial n}{\partial x}(x, t) + K(x) n(x, t) = 0 \,, \quad \text{for } 0 < x < x^* \,, \qquad (7)$$

where $x^* > x_E^*$ is fixed, and $\psi(x) = 0$ for $x < x_E^*$. Then, $E(t) = \int_0^{x_E^*} n(x, t) \mathrm{d}x$ yields the size of the exposed compartment. However, we prefer the model (6) as it allows us to represent variable incubation times.

The system (6d) can be extended to represent infectiosity rates that depends on the ages (time elapsed since birth) of individuals, with infectiosity rates given by a contact matrix, as in [27]. It suffices to split each compartment in sub-compartments, corresponding to different age groups. This will be detailed in a further work.

7.2 A Perron-Frobenius Eigenproblem for Transport PDE

When the control $\mu(t)$ is constant and positive, the family of time evolution operators $(T_{s,t})_{t \geqslant s \geqslant 0}$ is determined by the semigroup $(S_t = T_{0,t})_{t \geqslant 0}$, and the long term evolution of the dynamical system (3) can be studied by means of the *Perron–Frobenius eigenproblem*

$$\bar{n}_E(0) = \int_0^{x_I^*} \mu \psi(x) \bar{n}_I(x) \, \mathrm{d}x \,, \quad \frac{\mathrm{d}\bar{n}_E}{\mathrm{d}x}(x) + (\lambda + K_{E \to I}(x)) \bar{n}_E(x) = 0 \quad \text{for } 0 < x < x_E^* \,, \qquad (8a)$$

$$\bar{n}_I(0) = \int_0^{x_E^*} K_{E \to I}(x) \bar{n}_E(x) \, \mathrm{d}x + n_E(x_E^*) \,, \quad \frac{\mathrm{d}\bar{n}_I}{\mathrm{d}x}(x) + (\lambda + K_{I \to R}(x)) \bar{n}_I(x) = 0 \quad \text{for } 0 < x < x_I^* \,, \quad (8b)$$

where $\bar{n} := (\bar{n}_I(\cdot), \bar{n}_E(\cdot))$ is a nonnegative eigenvector, and λ is the eigenvalue.

Since the functions $K_{E \to I}$ and $K_{I \to R}$ are independent of time, the existence of an eigenvector is an elementary result:

Proposition 1. *Suppose that Assumption 1 holds. Then, the eigenproblem (8) has a solution $(\bar{n}, \lambda)$, where $\bar{n} = (\bar{n}_E, \bar{n}_I)$, the functions $\bar{n}_E$ and $\bar{n}_I$ are continuous and positive, and $\lambda \in \mathbb{R}$. Moreover, the eigenvalue λ is unique, and the eigenvector $\bar{n}$ satisfying the latter conditions is unique up to a multiplicative constant.*

The proof of this proposition exploits a classical argument in renewal theory, see Lemma 3.1 p. 57 of [31]. We give the proof, leading to a semi-explicit representation of the eigenvector, which we shall need in Section 8.

We note first that the uniqueness of λ can be deduced from a general observation, of independent interest:

Lemma 1. *Let $w = (w_E, w_I)$, with $w_E \in L^1([0, x_E^*])$ and $w_I \in L^1([0, x_I^*])$, be such that*

$$\alpha \bar{n} \leqslant w \leqslant \beta \bar{n} \qquad (9)$$

for some $\alpha, \beta > 0$. Then,

$$\alpha \exp(\lambda t) \bar{n} \leqslant S_t w \leqslant \beta \exp(\lambda t) \bar{n}, \qquad \text{for all } t \geqslant 0 \,. \qquad (10)$$

Proof. This follows from the order preserving and linear character of the semigroup S_t, together with $S_t \bar{n} = \exp(\lambda t) \bar{n}$. $\qquad \square$

Therefore, Lemma 1 shows that λ is the growth rate of $n(\cdot, t)$ as $t \to \infty$, for all initial conditions $w = n(\cdot, 0)$ satisfying (9). In particular, λ is unique.

We next provide a semi-explicit formula for the eigenvector. We set

$$F_{E \to I}^{\lambda}(x) := \int_0^x (\lambda + K_{E \to I}(z)) \mathrm{d}z \,, \text{ for } 0 \leqslant x \leqslant x_E^* \quad F_{I \to R}^{\lambda}(x) := \int_0^x (\lambda + K_{I \to R}(z)) \mathrm{d}z \,, \text{ for } 0 \leqslant x \leqslant x_I^* .$$

Integrating the differential equations (8), we see that a nonnegative eigenvector $\bar{n} = (\bar{n}_E, \bar{n}_I)$ necessarily satisfies:

$$\bar{n}_E(x) = \exp\left(-F_{E \to I}^{\lambda}(x)\right)\bar{n}_E(0) \,, \text{ for } 0 \leqslant x \leqslant x_E^* \quad \bar{n}_I(y) = \exp\left(-F_{I \to R}^{\lambda}(y)\right)\bar{n}_I(0) \,, \text{ for } 0 \leqslant y \leqslant x_I^* . \quad (11)$$

We deduce from the above relations that $\bar{n}_E(\cdot)$ is continuous on $[0, x_E^*]$, and that $\bar{n}_I(\cdot)$ is continuous on $[0, x_I^*]$. Using the boundary condition in (8b), we deduce that if $\bar{n}_E(0) = 0$, then $\bar{n}$ is identically 0, a contradiction. Hence, $\bar{n}$ is everywhere positive. Moreover, (11) and (8b) entail that the eigenvector $\bar{n}$ is unique, up to a scalar multiple.

Using the boundary conditions in (8), and specializing (11) to $x = x_E^*$ and $y = x_I^*$, we deduce that $\mu G^{\lambda} \bar{n}_E(0) = \bar{n}_E(0)$, where

$$G^{\lambda} = \left(\int_0^{x_I^*} \psi(x) \exp\left(-F_{I \to R}^{\lambda}(x)\right) \mathrm{d}x\right) \left(\int_0^{x_E^*} K_{E \to I}(y) \exp\left(-F_{E \to I}^{\lambda}(y)\right) \mathrm{d}y + \exp\left(-F_{E \to I}^{\lambda}(x_E^*)\right)\right).$$

Therefore, to find an eigenvector, we must solve the equation $\mu G^{\lambda} = 1$ (the so-called "characteristic equation" in renewal theory). Since the functions $K_{E \to I}$, $K_{I \to R}$ and ψ are integrable, and ψ is nonzero on a set of positive measure, we deduce that $\lim_{\lambda \to -\infty} G^{\lambda} = +\infty$. We also have $\lim_{\lambda \to +\infty} G^{\lambda} = 0$. Moreover, the map $\lambda \mapsto G^{\lambda}$ is continuous. Since $\mu > 0$, by the intermediate value theorem, we can find λ such that $\mu G^{\lambda} = 1$, and this λ is the eigenvalue. This concludes the proof of Proposition 1. $\qquad \square$

The asymptotic bound (10) can be reinforced, by showing that, for all positive initial conditions w,

$$S_t w = C_1(w)\bar{n}\exp(\lambda t) + O(\exp(\lambda_2 t)) \,, \qquad \text{as } t \to \infty \,, \quad (12)$$

for some positive constant $C_1(w)$, and $\lambda_2 < \lambda$. This result, with an explicit control of λ_2, can be obtained as follows. We make a diagonal scaling, using the positive eigenvector, and we normalize the semigroup to make the Perron eigenvalue λ equal to zero. This leads to the semigroup

$$(\tilde{S}_t w)(x) := \exp(-\lambda t)\bar{n}^{-1}(x)[S_t(w\bar{n})](x) \,.$$

In potential theory, a version of this scaling is known as *Doob's h-transform* (see e.g. [14]). The semigroup $\tilde{S}_t$ obtained in this way is associated with a Markov process, and, so, the spectral gap of this semigroup can be bounded in terms of Doeblin's ergodicity coefficient [16, 5], leading to (12). These aspects will be detailed elsewhere. Alternatively, the relative entropy inequality technique of [28] allows one to establish the convergence of $n(\cdot, t)$ to the eigenvector, modulo multiplicative constants, as t tends to infinity.

7.3 Universality of the log-rate of epidemic observables

Epidemic observables are obtained by applying a continuous linear form to the state variable. Supposing that $n_I(\cdot, t)$ is a continuous function, an epidemic observable will be of the form

$$Y_{\kappa}(t) = \varphi(n(\cdot, t)) := \int_0^{x_I^*} n_I(x, t) \, \mathrm{d}\kappa(x) \,, \quad (13)$$

where $\mathrm{d}\kappa(x)$ is a nonnegative nonzero Borel measure. Epidemic events anterior to the infectious phase, like contamination, are by nature hard to detect, so the observable depends only on n_I.

Proposition 2. *Suppose that Assumption 1 holds, let $(\lambda, \bar{n})$ denote the solution of the Perron-Frobenius eigenproblem (8), and suppose that for some $T > 0$, there exist positive constants α, β such that $\alpha \bar{n} \leqslant n(\cdot, T) \leqslant \beta \bar{n}$. Then, for all epidemic observables of the form (13), the map $t \mapsto \log Y_{\kappa}(t) - \lambda t$ is bounded. A fortiori,*

$$\lim_{t \to \infty} \frac{1}{t} \log Y_{\kappa}(t) = \lambda \,.$$

18

Proof. Taking $w = n(\cdot, T)$ in Lemma 1, we deduce that

$$\alpha \exp(\lambda t)\bar{n} \leqslant S_{T+t}n(\cdot, 0) = n(\cdot, T+t) \leqslant \beta \exp(\lambda t)\bar{n}, \text{ for } t > 0 \ .$$

It follows that $\log \alpha + \lambda t + \log \varphi(\bar{n}) \leqslant \log Y_\kappa(T+t) \leqslant \log \beta + \lambda t + \log \varphi(\bar{n})$. $\qquad\square$

A simple example of observable, discussed in Section 3.2, consists of *pure delays*. For instance, we assumed that the number of dispatches of MICU is given by $Y_{\mathrm{MICU}}(t) = \pi_{\mathrm{MICU}}C(t - \tau_{\mathrm{MICU}})$ where $C(t)$ is the number of contaminations at time t, π_{MICU} the proportion of contaminated patients who will need a MICU transport, and τ_{MICU} a fixed delay. This can be obtained as a special case of (13), taking $K_{E \to I} \equiv 0$, so that the transition tom E to I occurs always at time x_E^*, and $\mathrm{d}\kappa := \pi_{\mathrm{MICU}}\delta_{\tau_{\mathrm{MICU}}-x_E^*}$, where δ is the Dirac δ function.

Other events can be considered: medical advice, EMT dispatch, admission to ICU, or decease. These events corresponds to different values of the proportion π and of the delay τ. By Proposition 2, the rate $\lim_{t \to \infty} t^{-1} \log Y(t)$ will be the same for all the corresponding observables, although the convergence of the function $t^{-1} \log Y(t)$ to its limit will be observed in a delayed manner, for observables corresponding to the latest stages of the pathology.

7.4 Discrete versions of the epidemiological model

The reader interested in ODE model of epidemics might wish to note that the previous analysis applies to such finite dimensional models. Instead of the transport PDE (5), we may consider an ODE of the form

$$\dot{v} = Mv \tag{14}$$

where $v(t) \in \mathbb{R}^n$ and M is a $n \times n$ matrix with non-negative off-diagonal terms, a so-called *Metzler matrix*. In the original SEIR model [3], the matrix M, obtained by considering the (E, I)-block equations (4b), (4c), with $S/N \simeq 1$, is of dimension 2. In the generalizations of the SEIR model considered in [27, 13], the dimension n is increased to account for other compartments. One can also discretize the PDE system (5) using a monotone (upwind) finite difference scheme, and this leads to a system of the form (14).

In all these finite dimensional models, the matrix M is Metzler and irreducible. Then, the Perron–Frobenius theorem for linear, order-preserving semigroups (see [7]) implies that M admits a unique eigenvalue λ of maximal real part. Furthermore λ is algebraically simple and real, and its associated eigenvector u has strictly positive coordinates. Then, it follows from the spectral theorem that

$$v(t) = \exp(\lambda t)u + o(\exp(\lambda_2 t))$$

as $t \to \infty$, where λ_2 is the maximal real part of an eigenvalue of M distinct from λ. Again, in this discrete model, an epidemic observable $Y(t)$ is obtained by applying a nonnegative linear form to the vector $v(t)$, i.e, $Y(t) = \varphi^\top v(t)$, for some nonnegative column vector φ.

8 Tropicalization of the logarithm of nonnegative observables of switched Perron–Frobenius dynamics

8.1 Hilbert's geometry applied to piecewise linear approximation

We introduce an abstract setting, which captures epidemiological models in which most individuals are susceptible. This setting applies, in particular, to the transport PDE model of (5), when the transition functions are supported by compact intervals, and to the general finite dimensional Metzler model (14).

We consider $(V, \leqslant)$, a partially ordered Banach space, with topological dual V'. We denote by $V_{\geqslant 0} := \{v \in V \mid v \geqslant 0\}$ the set of nonnegative elements of V, which is a convex cone. This cone must be pointed (i.e., $V_{\geqslant 0} \cap (-V_{\geqslant 0}) = \{0\}$), since the relation $\leqslant$ is a partial order. We require this cone to be closed.

We consider a sequence of m semigroups $S^i = (S^i_t)_{t \geqslant 0}$, for $i \in [m]$, where $[m] := \{1, \ldots, m\}$. We assume that for all $i \in [m]$, and for all $t \geqslant 0$, S^i_t is a bounded linear operator from V to itself, and that the semigroup property holds, i.e., $S^i_{t+s} = S^i_t \circ S^i_s$. We shall say that the semigroup S^i is *order preserving* if, for all $v \in V_{\geqslant 0}$, and for all $t \geqslant 0$, $S^i_t v \in V_{\geqslant 0}$.

We shall consider commutation instants, $t_0 := 0 < t_1 < \cdots < t_{m-1}$, These instants will correspond to significant epidemiological dates, for instance, dates at which sanitary measures are taken. We set $t_m := +\infty$.

We select an initial condition $v_0 \in V_{\geqslant 0}$, and consider the abstract dynamical system obtained by switching between the evolutions determined by the semigroups $S^1, \ldots, S^m$, at the successive times $t_1, \ldots, t_{m-1}$. The state of this dynamical system, at time $t \in [t_j, t_{j+1})$, is given by

$$v_t := S^{j+1}_{t-t_j} \circ S^j_{t_j - t_{j-1}} \circ \cdots \circ S^1_{t_1 - t_0}(v_0) \ . \tag{15}$$

Recall that a *part* of the closed convex cone $V_{\geqslant 0}$ is an equivalence class for the relation $\sim$ such that, for $v, w \in V_{\geqslant 0}$, we have $v \sim w$ if and only if there exists two positive constants α and β such that $\alpha v \leqslant w \leqslant \beta v$. A part is *trivial* if it is reduced to the equivalence class of the zero vector. *Hilbert's projective metric d_H* is defined on every non-trivial part of $V_{\geqslant 0}$ by the following formula

$$d_H(v, w) = \log \inf \left\{ \frac{\beta}{\alpha} : \alpha, \beta > 0, \ \alpha v \leqslant w \leqslant \beta v \right\} \ .$$

The infimum is achieved, since $V_{\geqslant 0}$ is closed. The map d_H is nonnegative, it satisfies the triangular inequality, and $d_H(v, w)$ vanishes if, and only if, v and w are proportional – this justifies the term "projective metric". This metric plays a fundamental role in Perron–Frobenius theory and in metric geometry, and also in tropical geometry, see [26, 30, 10] for background.

When $V_{\geqslant 0} = (\mathbb{R}_{\geqslant 0})^n$ is the standard orthant, and when all the entries of the vectors v and w are positive, we have

$$d_H(v, w) = \max_{k \in [n]} (\log v_k - \log w_k) - \min_{k \in [n]} (\log v_k - \log w_k) \ .$$

Denoting by e the unit vector of $\mathbb{R}^n$, we observe that

$$d_H(v, w) = \| \log v - \log w \|_H$$

where the notation $\log v$ is understood entrywise, and

$$\|z\|_H = 2 \min_{c \in \mathbb{R}} \|x - ce\|_\infty \ .$$

In other words, up to a logarithmic change of variables, d_H arises by modding out the normed space $(\mathbb{R}^n, \| \cdot \|_\infty)$ by the one-dimensional space $\mathbb{R}e$.

We shall suppose that every semigroup S^i has an eigenvector $u^i \geqslant 0$, with eigenvalue λ^i, meaning that

$$S^i_t u^i = \exp(\lambda^i t) u^i, \quad \forall t \geqslant 0 \ .$$

Since S^i_t preserves $V_{\geqslant 0}$, this entails that λ^i is real.

We choose a linear form $\varphi \in V'$ which we require to take nonnegative values on $V_{\geqslant 0}$. We shall think of V has the *state space* and φ as an *observable*. We consider the following scalar observation of the dynamics

$$Y_t := \varphi(v_t) \ .$$

We shall assume, in addition, that φ does not vanish on v_t, for all $t \geqslant 0$. Then, we can define the image of the observation by the logarithmic map

$$y_t := \log Y_t, \qquad \forall t \geqslant 0 \ .$$

The following result shows that the logarithm of the observation stays at finite distance from a piecewise linear map.

Theorem 1. *Suppose that the semigroups $S^1, \ldots, S^m$ are order preserving. Suppose in addition that the initial condition v_0 and the eigenvectors $u^1, \ldots, u^m$ all lie in the same non-trivial part of $V_{\geqslant 0}$, and that the linear form φ takes positive values on this part. Then, there exists a constant C such that the piecewise linear map $t \mapsto y_t^{\mathrm{trop}}$ defined, for $t \in [t_j, t_{j+1})$, by*

$$y_t^{\mathrm{trop}} := \lambda_{j+1}(t - t_j) + \lambda_j(t_j - t_{j-1}) + \cdots + \lambda_1(t_1 - t_0) + C \ ,$$

satisfies

$$|y_t - y_t^{\mathrm{trop}}| \leqslant \frac{\Delta}{2}, \qquad \forall t \geqslant 0 \ ,$$

where

$$\Delta = d_H(v_0, u^1) + d_H(u^1, u^2) + \cdots + d_H(u^{m-1}, u^m) \ .$$

Proof. By definition of Hilbert's projective metric, we can find positive constants α_0, β_0, such that $\alpha_0 u^1 \leqslant v_0 \leqslant \beta_0 u^1$ and $d_H(v_0, u^1) = \log(\beta_0/\alpha_0)$. Similarly, for all $i \in [m-1]$, we can find positive constants α_i, β_i, such that $\alpha_i u^{i+1} \leqslant u^i \leqslant \beta_i u^{i+1}$, and $d_H(u^i, u^{i+1}) = \log(\beta_i/\alpha_i)$. For all $j \geqslant 0$, with $j \leqslant m-1$, and for all $t \in [t_j, t_{j+1})$, we set

$$z_t := \lambda_{j+1}(t - t_j) + \lambda_j(t_j - t_{j-1}) + \cdots + \lambda_1(t_1 - t_0) \ .$$

Since the semigroups S^i are linear and order preserving, we prove by induction

$$\exp(z_t)\alpha_j \ldots \alpha_0 u^{j+1} \leqslant v_t \leqslant \exp(z_t)\beta_j \ldots \beta_0 u^{j+1} \ .$$

We observe that

$$\alpha_{m-1} \ldots \alpha_{j+1} u^m \leqslant u^{j+1} \leqslant \beta_{m-1} \ldots \beta_{j+1} u^m$$

and so

$$\exp(z_t)\alpha_{m-1} \ldots \alpha_0 u^m \leqslant v_t \leqslant \exp(z_t)\beta_{m-1} \ldots \beta_0 u^m \ .$$

Applying the linear form φ to latter inequalities, taking the image by the log map, and setting

$$C := \log \varphi(u^m) + \frac{1}{2} \sum_{j=0}^{m-1} \log(\beta_j \alpha_j) \ ,$$

we arrive at the bound of the theorem. $\qquad\qquad\square$

A general principle from tropical geometry states that using "logarithmic glasses" reveals a piecewise linear structure [36, 20]. Theorem 1 is inspired by this principle. This motivates the notation y^{trop}, for the "tropicalization" of the logarithm of the observable Y.

Theorem 1 carries over to discrete time systems in a straightforward manner.

8.2   Application to the transport PDE model

Theorem 1 applies in particular to the transport model (6). Then, as noted above, the evolution operator of the system (5) preserves the space $V = L^1([0, x_E^*]) \times L^1([0, x_I^*])$. Moreover, when the epidemiological control term $\mu(t)$ is constant, Proposition 1 shows that the eigenproblem (8) has a positive and continuous solution $\bar{n}$, with a real eigenvalue λ. Different stages of sanitary policies correspond to successive values $\mu^1, \ldots, \mu^m$ of $\mu(t)$, leading to different semigroups S^i, $i \in [m]$. Then, the solution $v_t := n(\cdot, t)$ of (5) is determined as in (15). Each semigroup S^i yields a continuous and positive eigenvector $u^i := \bar{n}^i$ satisfying (8) associated with a real eigenvalue λ^i of S^i. Two continuous and positive functions defined on a compact interval are always in the same part of the cone of nonnegative functions of V, so Theorem 1 applies to this model.

We next give an explicit estimate for the Hilbert projective distances between eigenvectors, arising in Theorem 1.

Proposition 3. *Suppose that Assumption 1 holds, and that for $i = 1, 2$, $(\lambda^i, \bar{n}^i)$ is the solution $(\lambda, \bar{n})$ of the Perron-Frobenius eigenproblem* (8) *when $\mu = \mu^i$. Then, we have*

$$d_H(\bar{n}^1, \bar{n}^2) \leqslant |\lambda_1 - \lambda_2|(x_E^* + x_I^*) \ .$$

The term $x_E^* + x_I^*$ is the maximal time elapsed between contamination and the end of infectiosity.

Proof of Proposition 3. Suppose, without loss of generality, that $\bar{n}_E^i(0) = 1$ for $i = 1, 2$. Let $\phi_{E \to I}(x) := -\int_0^x K_{E \to I}(z)\mathrm{d}z$ and $\phi_{I \to R}(x) := -\int_0^x K_{I \to R}(z)\mathrm{d}z$. Then, (11) yields

$$\bar{n}_E^i(x) = \exp(-\lambda^i x)\phi_{E \to I}(x) \,, \tag{16}$$

$$\bar{n}_I^i(0) = \int_0^{x_E^*} K_{E \to I}(x)\phi_{E \to I}(x)\exp(-\lambda^i x)\mathrm{d}x + \phi_{E \to I}(x_E^*)\exp(-\lambda^i x_E^*) \,, \tag{17}$$

$$1 = \bar{n}_E^i(0) = \mu^i\Big(\int_0^{x_I^*}\psi(x)\exp(-\lambda^i x)\mathrm{d}x\Big)\bar{n}_I^i(0) \ . \tag{18}$$

Let $j \in \{1, 2\}$ be distinct from i. Then, setting $t^+ := \max(t, 0)$, bounding $\exp(-\lambda^i x)$ by $\exp(-\lambda^j x)\exp\big((\lambda^j - \lambda^i)^+ x_E^*\big)$ in (17), and applying a similar bound in (18), we obtain:

$$\bar{n}_E^i(0) \leqslant \frac{\mu^i}{\mu^j}\exp\big((\lambda^j - \lambda^i)^+(x_I^* + x_E^*)\big)\bar{n}_E^j(0) \,.$$

which yields $\mu^j/\mu^i \leqslant \exp\big((\lambda^j - \lambda^i)^+\big)(x_I^* + x_E^*)$. We deduce that $\bar{n}_I^i(x) \leqslant \bar{n}_I^j(x)\exp\big((\lambda^j - \lambda^i)^+\big)(x_I^* + x_E^*)$. So,

$$d_H((\bar{n}_E^i, \bar{n}_I^i), (\bar{n}_E^j, \bar{n}_I^j)) \leqslant (\lambda^j - \lambda^i)^+(x_I^* + x_E^*) + (\lambda^i - \lambda^j)^+(x_I^* + x_E^*) = |\lambda^i - \lambda^j|(x_I^* + x_E^*) \,. \qquad \square$$

The bound of Theorem 1 may be refined. This is left for further work.

9 Short term predictions

We now describe the basic methodology we propose to build confidence intervals for future occurrences of medical events related to epidemic progression, and raise alarms about its potential resurgence. We first consider a single time series of numbers of event occurrences. We then describe how to consolidate several time series corresponding to distinct medical events in order to construct improved alarm criteria. The simpler case of least squares fitting is considered first, the more robust ℓ_1 alternative is described next.

Time series for a single type of events: Let $X(1), \ldots, X(n)$ be indices of days, and we aim to do a forecast based on observations made on these days. Typically, on day d_0, we may select $n = 7$ and let $X(1) = d_0 - n, \ldots, X(n) = d_0 - 1$ to perform a forecast on the basis of the last seven days. Let $Y(t)$ denote the count of medical events (for instance, dispatches of ambulances) on day $X(t)$, and let $Z(t) = \log Y(t)$. Based on the previous discussion (epidemiological modeling) we assume that for all $t = 1, \ldots, n$,

$$Z(t) = \alpha + \beta X(t) + \epsilon_t$$

for constants α, β, where ϵ_t denotes some random noise. For simplicity we assume here i.i.d. noise sequence $\epsilon_1, \ldots, \epsilon_n$, and that each ϵ_t admits a Gaussian distribution $\mathcal{N}(0, \sigma^2)$ with zero mean and variance σ^2.

Least-square estimates for the parameters α, β are then provided by

$$\hat{\beta} = \frac{\sum_{t=1}^n (X(t) - \bar{X})(Z(t) - \bar{Z})}{\sum_{t=1}^n (X(t) - \bar{X})^2}, \quad \hat{\alpha} = \bar{Z} - \hat{\beta}\bar{X}, \tag{19}$$

where

$$\bar{X} = \frac{1}{n}\sum_{t=1}^{n} X(t), \quad \bar{Z} = \frac{1}{n}\sum_{t=1}^{n} Z(t). \tag{20}$$

The variance σ^2 can be estimated as

$$\hat{\sigma}^2 = \frac{1}{n-2}\sum_{t=1}^{n}(Z(t) - \hat{Z}(t))^2, \tag{21}$$

where

$$\hat{Z}(t) := \hat{\alpha} + \hat{\beta}X(t). \tag{22}$$

Under the assumptions of i.i.d. Gaussian errors ϵ_t, we have that, for each t corresponding to a future day day $X(t)$ (in particular, $t \notin \{1, \ldots, n\}$), the three following variables:

$$\frac{\hat{\alpha} - \alpha}{\hat{\sigma}\sqrt{\frac{1}{n} + \frac{\bar{X}^2}{\sum_{t'=1}^{n}(X(t')-\bar{X})^2}}}, \quad \frac{\hat{\beta} - \beta}{\hat{\sigma}\sqrt{\frac{1}{\sum_{t'=1}^{n}(X(t')-\bar{X})^2}}}, \quad \frac{Z(t) - \hat{Z}(t)}{\hat{\sigma}\sqrt{1 + \frac{1}{n} + \frac{(X(t)-\bar{X})^2}{\sum_{t'=1}^{n}(X(t')-\bar{X})^2}}},$$

all admit a bilateral Student distribution with $n-2$ degrees of freedom (see [21, Ch. 28] or [11]). Denote by t_γ^{n-2} the γ-th quantile of this distribution. For $\epsilon \in [0,1]$, this provides us with the following confidence intervals with confidence $1 - \epsilon$:

$$\begin{aligned}
\alpha &\in \left[\hat{\alpha} - \hat{\sigma}\sqrt{\frac{1}{n} + \frac{\bar{X}^2}{\sum_{t'=1}^{n}(X(t')-\bar{X})^2}}\, t_{1-\epsilon/2}^{n-2}, \hat{\alpha} + \hat{\sigma}\sqrt{\frac{1}{n} + \frac{\bar{X}^2}{\sum_{t'=1}^{n}(X(t')-\bar{X})^2}}\, t_{1-\epsilon/2}^{n-2}\right], \\
\beta &\in \left[\hat{\beta} - \hat{\sigma}\sqrt{\frac{1}{\sum_{t'=1}^{n}(X(t')-\bar{X})^2}}\, t_{1-\epsilon/2}^{n-2}, \hat{\beta} + \hat{\sigma}\sqrt{\frac{1}{\sum_{t'=1}^{n}(X(t')-\bar{X})^2}}\, t_{1-\epsilon/2}^{n-2}\right] \\
Z(t) &\in \left[\hat{Z}(t) - \hat{\sigma}\sqrt{1 + \frac{1}{n} + \frac{(X(t)-\bar{X})^2}{\sum_{t'=1}^{n}(X(t')-\bar{X})^2}}\, t_{1-\epsilon/2}^{n-2}, \hat{Z}(t) + \hat{\sigma}\sqrt{1 + \frac{1}{n} + \frac{(X(t)-\bar{X})^2}{\sum_{t'=1}^{n}(X(t')-\bar{X})^2}}\, t_{1-\epsilon/2}^{n-2}\right]
\end{aligned} \tag{23}$$

As an illustration, for $n = 7$ and $\epsilon = 5\%$, we can plug in $t_{0.975}^5 = 2.571$ in the last interval, and thus obtain a 95%-confidence interval centered around $\hat{Z}(t)$ for $Z(t) = \log Y(t)$, the logarithm of the count $Y(t)$ on a future day $X(t)$, that is:

$$Z(t) \in \left[\hat{Y}(t) - 2.571 \times \hat{\sigma}\sqrt{1 + \frac{1}{n} + \frac{(X(t)-\bar{X})^2}{\sum_{t'=1}^{n}(X(t')-\bar{X})^2}}, \hat{Z}(t) + 2.571 \times \hat{\sigma}\sqrt{1 + \frac{1}{n} + \frac{(X(t)-\bar{X})^2}{\sum_{t'=1}^{n}(X(t')-\bar{X})^2}}\right] \tag{24}$$

Although we could extend this definition of the confidence interval for the short terms predictions of the value of $Z(t)$, we propose a more conservative confidence domain, in the shape of a trapezoid. It is obtained by extending the upper-bound $Z(t)^+$ (resp. the lower bound $Z(t)^-$ of the 95% confidence interval on $Z(t)$ by a line with slope equal to the upper-bound β^+ (resp. lower-bound β^-) of 95% confidence interval on β. For a given day t, the upper and lower envelopes of the trapezoid have ordinates

$$\left(\hat{\beta}(t - t_n) + \hat{Z}_n\right) \pm \left(\sqrt{\mathrm{Var}(\hat{\beta})}(t - t_n) + \sqrt{\hat{\sigma}^2 + \mathrm{Var}(\hat{Z}_n)}\right) t_{1-\epsilon/2}^{n-2}.$$

If instead of the count $Y(t)$ on a particular day $X(t)$, we are interested in the trend of the epidemic, whether exploding or contracting, we should then consider the confidence interval for parameter β. Again for $n = 7$ and $\epsilon = 5\%$ this gives

$$\beta \in \left[\hat{\beta} - 2.571 \times \hat{\sigma}\sqrt{\frac{1}{\sum_{t'=1}^{n}(X(t')-\bar{X})^2}}, \hat{\beta} + 2.571 \times \hat{\sigma}\sqrt{\frac{1}{\sum_{t'=1}^{n}(X(t')-\bar{X})^2}}\right] \tag{25}$$

One-sided confidence intervals may also be provided, and are in fact more natural for the definition of alarm indicators.

For concreteness, assume we want to raise an alarm when the doubling time, $\delta = (\log 2)/\beta$, is δ^* days or less, where δ^* could be 10 for instance. This is equivalent to β exceeding $(\log 2)/\delta^*$. Thus δ is less than δ^* days with confidence $1 - \epsilon$ when

$$\frac{\log 2}{\delta^*} < \hat{\beta} - t_{1-\epsilon}^{n-2}\sqrt{V},$$

where

$$V = \hat{\sigma}\sqrt{\frac{1}{\sum_{t'=1}^{n}(X(t') - \bar{X})^2}}.$$

Raising an alarm under this condition then amounts to calibrating the false positive probability at ϵ. For instance, for $\epsilon = 5\%$, and $n = 7$, we would plug in $t_{0.95}^7 = 2.015$ in the above expression.

Alternatively, raising an alarm under the condition

$$\frac{\log 2}{\delta^*} < \hat{\beta} + t_{1-\epsilon}^{n-2}\sqrt{V},$$

corresponds to calibrating the false negative probability (probability of not raising an alarm while $\delta \leqslant 10$) at ϵ.

Our alarm indicators correspond to the first choice, i.e. calibration of a false positive rate, with δ^* set to $+\infty$.

Alarm indicators based on multiple types of events: Assume that several types j of events are available, and let J denote the corresponding set of events. For instance, we could distinguish between dispatches of ambulances bringing patients to Intensive Care Units as opposed to Non-intensive Care Units, thereby producing two distinct time series. Let $X_j(t)$, $t = 1, \ldots, n_j$ denote the days on which counts $Y_j(t)$ of type j event occurrences are to be used. Let $Z_j(t) = \log Y_j(t)$. We assume as before the linear regression model

$$Z_j(t) = \alpha_j + \beta_j X_j(t) + \epsilon_j(t), \ t = 1, \ldots, n^j \ .$$

Now for each of these times series, we can produce, based on the previous discussion, the estimator

$$\hat{\beta}_j := \frac{\sum_{t=1}^{n_j}(X_j(t) - \bar{X}_j)(Z_j(t) - \bar{Z}_j)}{\sum_{t=1}^{n_j}(X_j(t) - \bar{X}_j)^2},$$

where

$$\bar{X}_j = \frac{1}{n_j}\sum_{t=1}^{n_j} X_j(t), \quad \bar{Z}_j = \frac{1}{n_j}\sum_{t=1}^{n_j} Z_j(t).$$

Suppose in addition that the noise terms $\epsilon_j(t)$ are mutually independent, Gaussian, with zero mean and variance σ_j^2 for errors $\epsilon_j(t)$. Suppose finally that the exponents β_j all coincide with β, the exponent that is characteristic of the epidemic's progression. Denote by

$$V_j := \hat{\sigma}_j^2 \sqrt{\frac{1}{\sum_{t=1}^{n_j}(X_j(t) - \bar{X}_j)^2}}, \tag{26}$$

where, reproducing the computations for a single time series, we let

$$\hat{\sigma}_j^2 := \frac{1}{n_j - 2}\sum_{t=1}^{n_j}(Z_j(t) - \hat{Z}_j(t))^2,$$

and

$$\hat{Z}_j(t) := \hat{\alpha}_j + \hat{\beta}_j X_j(t).$$

24

As previously, V_j is our estimate of the variance of estimate $\hat{\beta}_j$. We finally propose to combine the individual estimators $\hat{\beta}_j$ into

$$\hat{\beta} := \frac{\sum_{j \in J} \frac{1}{V_j} \hat{\beta}_j}{\sum_{j \in J} \frac{1}{V_j}}. \tag{27}$$

For the sake of simplicity, let us approximate the bilateral Student distribution with $n-2$ degrees of freedom by the standard distribution $\mathcal{N}(0,1)$. We then have the approximate distributions $\hat{\beta}_j \approx \mathcal{N}(\beta, V_j)$, and hence the approximate distribution $\hat{\beta} - \beta \approx \mathcal{N}(0, V)$,

$$V := \frac{1}{\sum_{j \in J} \frac{1}{V_j}}. \tag{28}$$

Weighing the individual estimators $\hat{\beta}_j$ by the reciprocal of their variances as just done minimizes the variance of the resulting estimator. The same approach as previously considered then leads to the following conditions for alarm raising:

To raise an alarm when the doubling time $\delta = (\log 2)/\beta$ exceeds δ^* days (e.g., $\delta^* = 10$), if we target a false alarm probability of ϵ, we are led to raise an alarm when Condition

$$\frac{\log 2}{\delta^*} < \hat{\beta} - g_{1-\epsilon}\sqrt{V}, \tag{29}$$

where $g_{1-\epsilon}$ is the $1 - \epsilon$-quantile of the standard Gaussian distribution.

If instead we target a false negative probability (probability of not raising an alarm) at ϵ, we would then raise an alarm when

$$\frac{\log 2}{\delta^*} < \hat{\beta} + g_{1-\epsilon}\sqrt{V}, \tag{30}$$

More robust ℓ_1-based approach: The previous estimators and derived alarm conditions have the appeal of simplicity, but can be advantageously replaced by more robust versions, that are less sensitive to the presence of outliers.

A popular alternative is the following ℓ_1 criterion. We again consider $Z_j(t) := \log Y_j(t)$, where $Y_j(t)$ is the number of type j events on day $X_j(t)$. We then let $\hat{\alpha}_j$, $\hat{\beta}_j$ achieve the minimum of the criterion $\sum_{t=1}^{n^j} |\alpha + \beta X_t^j - Y_t^j|$. They are obtained by solving a linear program. Here we assume that observations $Z_j(t)$ are mutually independent and distributed according to density $f_{j,t}(z) = \frac{1}{2\lambda^j} \exp(-|z - \alpha_j - \beta_j X_j(t)|/\lambda_j)$. In other words this corresponds to adding a Laplace observation noise with density $\frac{1}{2\lambda^j} \exp(-|z|/\lambda_j)$ to the signal of interest $\alpha_j + X_j(t)\beta_j$. The above ℓ_1 minimization criterion corresponds to maximum likelihood estimation of α_j, β_j in this observational noise model, as its log-likelihood is given by

$$-|T| \log(2\lambda_j) - \sum_{t=1}^{n_j} \frac{|Z_j(t) - \alpha_j - \beta_j X_j(t)|}{\lambda_j}.$$

A rich theory for the performance of the resulting estimators is available, see for instance [23]. The latter work treats general i.i.d. errors, and do not restrict itself to e.g. Laplacian distribution of errors; recent work like [33] experiments techniques to obtain confidence intervals when distribution of errors is unknown. Here we make the choice of Laplace-distributed errors for sake of simplicity. In particular, the asymptotic theory in [23] suggests the approximation

$$\hat{\beta}_j \sim \mathcal{N}(\beta_j, V_j)$$

where

$$\hat{\lambda}_j := \frac{1}{n_j} \sum_{t=1}^{n_j} |Z_j(t) - \hat{\alpha}_j - \hat{\beta}_j X_j(t)|, \tag{31}$$

and

$$V_j := (\hat{\lambda}_j)^2 \frac{1}{\sum_{t=1}^{n_j} X_j(t)^2 - \frac{1}{n_j}(\sum_{t=1}^{n_j} X_j(t))^2}. \tag{32}$$

We again consider that multiple types $j \in J$ of time series are conjointly available, and that each β^j coincides with β, the parameter to be estimated. Assuming the $\hat{\beta}_j$ to be independent with $\hat{\beta} \sim \mathcal{N}(\beta_j, V_j)$, leads us to define the estimator

$$\hat{\beta} := \frac{\sum_{j \in J} \frac{\hat{\beta}_j}{(\hat{\lambda}_j)^2}}{\sum_{j \in J} \frac{1}{(\hat{\lambda}_j)^2}}, \tag{33}$$

whose distribution is then given by $\hat{\beta} \sim \mathcal{N}(\beta, V)$ where

$$V = \frac{1}{\sum_{j \in J} \frac{\sum_{t=1}^{n_j} X_j(t)^2 - (\sum_{t=1}^{n_j} X_j(t))^2/n_j}{(\hat{\lambda}_j)^2}}. \tag{34}$$

A symmetric $(1 - \epsilon)$-confidence interval for β is then provided by

$$\beta \in I := \left[\hat{\beta} - g_{1-\epsilon/2}\sqrt{V}, \hat{\beta} + g_{1-\epsilon/2}\sqrt{V}\right]. \tag{35}$$

Similarly, $(1 - \epsilon)$-confidence one-sided intervals for β are obtained by letting

$$\beta \in I' := [\hat{\beta} - g_{1-\epsilon}\sqrt{V}, +\infty), \ \beta \in I'' := (-\infty, \hat{\beta} + g_{1-\epsilon}\sqrt{V}]. \tag{36}$$

The doubling time δ is given by $(\log 2)/\beta$ if $\beta > 0$, and $+\infty$ otherwise. This gives the $1 - \epsilon$-confidence conditions for δ:

$$\text{if } \hat{\beta} - g_{1-\epsilon}\sqrt{V} > 0, \ \delta \in I_1 = \left[0, \frac{\log 2}{\hat{\beta} - g_{1-\epsilon}\sqrt{V}}\right], \tag{37}$$

and

$$\delta \in I_2 = \left[\frac{\log 2}{\max(0, \hat{\beta} + g_{1-\epsilon}\sqrt{V})}, +\infty\right). \tag{38}$$

For concreteness assume we want to raise an alarm when δ is δ^* days or less, where δ^* could be 10. From the above consideration, δ is below δ^* days with confidence $1 - \epsilon$ when

$$\frac{\log 2}{\delta^*} < \hat{\beta} - g_{1-\epsilon}\sqrt{V}.$$

Raising an alarm under this condition then amounts to calibrating the false positive probability at ϵ.

Alternatively, we may consider to raise an alarm under the condition

$$\frac{\log 2}{\delta^*} < \hat{\beta} + g_{1-\epsilon}\sqrt{V}.$$

This would correspond to calibrating the false negative probability (probability of not raising an alarm while $\delta \leqslant \delta^*$) at ϵ.

10   Conclusion

We have shown that monitoring of emergency calls to EMS allows to anticipate the evolution of an epidemic by providing several *early signals*, each with specific characteristics in terms of time lag and reliability.

Our study illustrates the spatially differentiated nature of the epidemic kinetics, with significant doubling time differences between neighboring departments.

Such spatial differentiation, if present at a granularity finer than that of departments considered here, could be exploited using the methods described in the present work in order to detect potential epidemic resurgences at the corresponding spatial granularity. This shows great promise in enabling detection of so-called epidemic clusters.

There is thus huge potential in the extension of this work and its application to finer spatial resolution.

Notwithstanding such extensions, monitoring epidemic kinetics through EMS calls at regional levels can already be exploited to define region-specific sanitary measures, such as lifting of travel bans, proportionate to the regional situation, and to allow early detection of epidemic resurgence. Importantly, we expect this finding to be applicable in full generality to EMS organizations worldwide. Thus the methods introduced here may be of wide applicability to combat Covid-19. Beyond Covid-19, EMS organizations have a unique role to play in early detection of sanitary crises.

11   Acknowledgments

References

[1] D. M. Anderson, N. A. Ciletti, H. Lee-Lewis, D. Elli, J. Segal, K. L. DeBord, K. A. Overheim, M. Tretiakova, R. R. Brubaker, and O. Schneewind. Pneumonic plague pathogenesis and immunity in brown norway rats. *Am. J. Pathol.*, 174(3):910–921, 2009.

[2] Andrew S. Azman, Kara E. Rudolph, Derek A. T. Cummings, and Justin Lessler. The incubation period of cholera: A systematic review. *J. Infect.*, 66(5):432–438, 2013.

[3] N. Bacaer. Un modèle mathématique des débuts de l'épidémie de coronavirus en France. hal-02509142, 2020.

[4] D. J Baker, C. Télion, and P. Carli. Multiple casualty incidents: the prehospital role of the anesthesiologist in europe. *Anesthesiology clinics*, 25(1):179–188, 2007.

[5] V. Bansaye, B. Cloez, and P. Gabriel. Ergodic behavior of non-conservative semigroups via generalized doeblin's conditions. *Acta Applicandae Mathematicae*, 166:29–72, 2020.

[6] R. Bellman and R. Roth. Curve fitting by segmented straight lines. *J. Am. Stat. Assoc.*, 64(327):1079–1084, 1969.

[7] A. Berman and R.J. Plemmons. *Nonnegative matrices in the mathematical sciences*. Academic Press, 1979.

[8] G. C. Calafiore, S. Gaubert, and C. Possieri. Log-sum-exp neural networks and posynomial models for convex and log-log-convex data. *IEEE Transactions on Neural Networks and Learning Systems*, 31(3):827–838, 2020.

[9] Yi-Cheng Chen, Ping-En Lu, Cheng-Shang Chang, and Tzu-Hsuan Liu. A time-dependent SIR model for COVID-19 with undetectable infected persons. `http://gibbs1.ee.nthu.edu.tw/A_TIME_DEPENDENT_SIR_MODEL_FOR_COVID_19.PDF`, 2020.

[10] G. Cohen, S. Gaubert, and J.-P. Quadrat. Duality and separation theorems in idempotent semimodules. *Linear Algebra and Appl.*, 379:395–422, 2004.

[11] H. Cramér. *Mathematical methods of statistics*, volume 43. Princeton university press, 1999.

[12] Agence Régionale de Santé Grand Est. Coronavirus Covid 19 en Grand Est: Point de situation. Press release of March 8th, available from `https://www.grand-est.ars.sante.fr/system/files/2020-03/Covid19_point_GrandEst080320.pdf`, 2020.

[13] L. Di Domenico, G. Pullano, Ch. E. Sabbatini, P.-Y. Boëlle, and V. Colizza. Expected impact of lockdown in Île-de-France and possible exit strategies. Report #9 , www.epicx-lab.com/covid-19.html, 2020.

[14] E.B. Dynkin. Boundary theory of Markov processes (the discrete case). *Russian Math. Surveys*, 24(7):1–42, 1969.

[15] Santé Publique France. Données hospitalières relatives à l'épidémie de Covid-19, 2020. `https://www.data.gouv.fr/fr/datasets/donnees-hospitalieres-relatives-a-lepidemie-de-covid-19/`, Retrieved on May 10th, 2020.

[16] S. Gaubert and Z. Qu. Dobrushin's ergodicity coefficient for Markov operators on cones. *Integral Equations and Operator Theory*, 81(1):127–150, 2015.

[17] N. Hansen and A. Ostermeier. Completely derandomized self-adaptation in evolution strategies. *Evolutionary Computation*, 9(2):159–195, 2001.

[18] M. Hirsch, P. Carli, R. Nizard, B. Riou, B. Baroudjian, Th. Baubet, V. Chhor, Ch. Chollet-Xemard, N. Dantchev, N. Fleury, J.-P. Fontaine, Y. Yordanov, M. Raphael, C. Paugam-Burtz, L. Lafont, and health professionals of AP-HP. The medical response to multisite terrorist attacks in paris. *The Lancet*, 386(10012):2535–2538, 2015.

[19] Insee. Ménages selon la taille en 2016. Comparaisons régionales et départementales, 2019. `https://www.insee.fr/fr/statistiques/2012714`.

[20] I. Itenberg, G. Mikhalkin, and E. Shustin. *Tropical algebraic geometry*. Oberwolfach seminars. Birkhäuser, 2007.

[21] N. L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous univariate distributions*. Wiley, New York, 1994.

[22] W. O Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. A*, 115:700–721, 1927.

[23] R. W Koenker and G. Bassett. Regression quantiles. *Econometrica*, 46(1):33–50, 1978.

[24] J. C. Lagarias, J. A. Reeds, M. H. Wright, , and P. E. Wright. Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM Journal of Optimization*, 9(1):112–147, 1998.

[25] S. A. Lauer, K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, and J. Lessler. The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: Estimation and application. *Annals of Internal Medicine*, 172(9):577–582, 2020. PMID: 32150748.

[26] B. Lemmens and R. Nussbaum. *Nonlinear Perron-Frobenius Theory*, volume 189 of *Cambridge Tracts in Mathematics*. Cambridge University Press, May 2012.

[27] C. Massonnaud, J. Roux, and P. Crépey. Covid-19: Forecasting short term hospital needs in France. Report available from Sfar.org, 2020.

[28] P. Michel, S. Mischler, and B. Perthame. General relative entropy inequality: an illustration on growth models. *J. Math. Pures et Appl.*, 84(9):1235–1260, May 11 2005.

[29] S. Mischler and J. Scher. Spectral analysis of semigroups and growth-fragmentation equations. *Annales de l'Institut Henri Poincare (C) Non Linear Analysis*, 33(3):849 – 898, 2016.

[30] A. Papadopoulos and M. Troyanov. Weak Finsler structures and the Funk weak metric. *Math. Proc. Cambridge Philos. Soc.*, 147(2):419–437, 2009.

[31] B. Perthame. *Transport equations in biology*. Birkhäuser, 2007.

[32] H. Salje, C. Tran Kiem, N. Lefrancq, N. Courtejoie, P. Bosetti, J. Paireau, A. Andronico, N. Hoze, J. Richet, C.-L. Dubost, Y. Le Strat, J. Lessler, D. Bruhl, A. Fontanet, L. Opatowski, P.-Y. Boëlle, and S. Cauchemez. Estimating the burden of SARS-CoV-2 in France. pasteur-02548181, 2020.

[33] G. Stangenhaus, S. C. Narula, and F. F. Pedro. Bootstrap confidence intervals for the minimum sum of absolute errors regression. *Journal of statistical computation and simulation*, 48(3-4):127–133, 1993.

[34] Virlogeux V., Fang V. J., Park M., Wu J. T., and Cowling B. J. Comparison of incubation period distribution of human infections with mers-cov in south korea and saudi arabia. *Sci. Rep.*, 6(35839), 2016.

[35] M. Varia, S. Wilson, Sh. Sarwal, A. McGeer, E. Gournis, Eleni Galanis, B. Henry, and Hospital Outbreak Investigation Team. Investigation of a nosocomial outbreak of severe acute respiratory syndrome (sars) in toronto, canada. *CMAJ*, 169(4):285–292, 2003.

[36] O. Viro. Dequantization of real algebraic geometry on logarithmic paper. In *European Congress of Mathematics, Vol. I (Barcelona, 2000)*, volume 201 of *Progr. Math.*, pages 135–146. Birkhäuser, Basel, 2001.

[37] H. von Forster. Some remarks on changing populations. In Jr. F. Stohlman, editor, *The Kinetics of Cellular Proliferation*, pages 382–407. Grune & Stratton, New York, 1959.

A   Appendix: algorithms to compute a best approximation of the logarithm of the number of events by a piecewise linear map

Given an epidemiologic observable $Y(t)$, we need to approximate $\log Y(t)$ by a function

$$\mathcal{L}(t) := \min_{1 \leqslant j \leqslant \nu} \left(\lambda_j t + c_j \right),$$

where ν is the number of phases with constant sanitary policy during the considered time period. The parameters λ_j, c_j are assumed without loss of generality to satisfy $\lambda_1 \leqslant \lambda_2 \cdots \leqslant \lambda_\nu$. The concavity constraint imposed on the approximating function $\mathcal{L}(t)$ makes the problem different from standard function approximation problems, and contributes to the robustness of the fitting procedure by reducing the amount of overfitting.

The two most natural criteria for fitting function $\mathcal{L}(t)$ to observations $\log Y(t)$ are to minimize either a least squares, or ℓ_2 loss function $\sum_{t \in \mathcal{T}} |\mathcal{L}(t) - \log Y(t)|^2$, or an ℓ_1 loss function $\sum_{t \in \mathcal{T}} |\mathcal{L}(t) - \log Y(t)|$, where $\mathcal{T}$ is a finite set of time instants at which observations have been made. As discussed in Section 9, the ℓ_1 formulation is more robust in being less sensitive to outliers, and is the one used on Figure 4.

The corresponding optimization problem over parameters λ_i, c_i is non-convex as soon as $\nu \geqslant 2$. A straightforward option is to use a derivative free procedure, like the Nelder-Mead [24] algorithm. Depending on the initial point, this algorithm may converge to a local minimum, which may not be epidemiologically significant. So, a possibility is to guide the algorithm by providing it a initial guess of the optimal solution. To do, we start by an a priori selection of the time periods over which function $\mathcal{L}(t)$ is linear (which could be obtained by prior knowledge of delay parameters τ and times of policy changes, or found by brute force search). We then determine a minimum cost linear fit of target function $\log Y(t)$ over each such period, and use the concave envelope of the resulting function as our initial condition for local search. This is how we initially obtained the best ℓ_1 approximation shown on Figure 4. We also used CMA-ES for comparison [17]. Both Nelder-Mead and CMA-ES algorithms appear to be sensitive to the initial conditions. Notice in this respect that the objective function is linear on the cells of a polyhedral complex and that it can be constant on certain unbounded cells of this complex, so a local search algorithm may be trapped in a cell in which the function is constant. Another perspective is to observe that this best approximation problem is equivalent to a learning problem, looking for the parameters of a neural networks with a single hidden layer and min-type activation functions, see [8]. This allows one to apply (nonsmooth) optimization algorithms used in learning, still leading in general to a local optimum. An approach leading to the global optimum is dynamic programming, originating from Bellman [6]. Ayoub Foussoul (École polytechnique) provided us with a dynamic programming solver, implementing several refinements, and allowing us to certify the global optimality of the approximation shown in Figure 4, up to a fixed precision.