Submitted to the Annals of Applied Statistics

USING POPULATION BASED KALMAN ESTIMATOR TO MODEL COVID-19 EPIDEMIC IN FRANCE: ESTIMATING THE EFFECTS OF NON-PHARMACEUTICAL INTERVENTIONS ON THE DYNAMICS OF EPIDEMIC.

BY ANNABELLE COLLIN[1,2,3], BORIS P. HEJBLUM[1,4,5], CAROLE VIGNALS[1,4,5,6], LAURENT LEHOT[1,4,5], RODOLPHE THIÉBAUT[1,4,5,6], PHILIPPE MOIREAU[7,8] AND MÉLANIE PRAGUE,[1,4,5,*]

[1] *Inria, Inria Bordeaux - Sud-Ouest, Talence, France*

[2] *Bordeaux INP, IMB UMR 5251, Talence, France*

[3] *IMB UMR 5251, Université Bordeaux, Talence, France*

[4] *Univ. Bordeaux, Inserm, Bordeaux Population Health Research Center, SISTM Team, UMR 1219, F-33000 Bordeaux, France*

[5] *Vaccine Research Institute, F-94000 Créteil, France*

[6] *CHU Pellegrin, F-33000 Bordeaux, France*

[7] *Inria, Inria Saclay-Ile-de-France, France*

[8] *LMS, CNRS UMR 7707, Ecole Polytechnique, Institut Polytechnique de Paris,* [*] *melanie.prague@inria.fr*

In response to the ongoing COVID-19 pandemic caused by SARS-CoV-2, governments are taking a wide range of non-pharmaceutical interventions (NPI). These measures include interventions as stringent as strict lockdown but also school closure, bar and restaurant closure, curfews and barrier gestures *i.e.* social distancing. Disentangling the effectiveness of each NPI is crucial to inform response to future outbreaks. To this end, we first develop a multi-level estimation of the French COVID-19 epidemic over a period of one year. We rely on a global extended Susceptible-Infectious-Recovered (SIR) mechanistic model of the infection including a dynamical (over time) transmission rate containing a Wiener process accounting for modeling error. Random effects are integrated following an innovative population approach based on a Kalman-type filter where the log-likelihood functional couples data across French regions. We then fit the estimated time-varying transmission rate using a regression model depending on NPI, while accounting for vaccination coverage, apparition of variants of concern (VoC) and seasonal weather conditions. We show that all NPI considered have an independent significant effect on the transmission rate. We additionally demonstrate a strong effect from weather conditions which decrease transmission during the summer period, and also estimate increased transmissibility of VoCs.

1. Introduction. The World Health Organization declared the COVID-19 pandemic on March 11, 2020. This disease is caused by an infection with the SARS-CoV-2 virus. As of April 30, 2021, more than 150 million cases have been confirmed worldwide, including 3.16 million deaths. While the majority of infected cases have a mild form (upper respiratory infection symptoms) without specific needs in terms of care Wu and McGoogan (2020), around 3% of cases, in particular the elderly, need hospitalization for treatment, such as oxygenation therapy Lapidus et al. (2021); Salje et al. (2020); Angulo, Finelli and Swerdlow (2021). Among those, about 17% are severe forms (severe acute respiratory syndrome) which

Keywords and phrases: Epidemics, Dynamical models, Kalman filters, Nonlinear mixed effect models, SARS-CoV-2, Non-pharmaceutical interventions.

will need to be admitted to intensive care units (ICU), with a potential need of mechanical ventilation Docherty et al. (2020).

The COVID-19 pandemic has pushed modern health care systems to a breaking point all over the world. Indeed, SARS-CoV-2 being a newly emerging pathogen, means the entire human population is susceptible to infection. Surges in hospitalizations and especially ICU needs wherever SARS-CoV-2 outbreaks arise have been observed. Due to an infectious phase starting before any symptoms are visible and a significant proportion of a- or pauci-symptomatic infections Jones et al. (2021), the spread of SARS-Cov-2 is extremely difficult to control Wu et al. (2021). In response, most governments have adapted drastic public health measures, also called Non-Pharmaceutical Interventions (NPI), in order to reduce the transmission of SARS-CoV-2 among their population, and consequently relax the pressure on their health-care system. In particular, the French government adopted the concept of a "graduated response" to the pandemic, deploying an arsenal of various NPI – some very stringent and others less so – in response to the COVID-19 national epidemic situation. Hale et al. Hale et al. (2021) built a stringency index that helps understand how strong the measures over time were. However, this indicator does not allow one to distinguish the effectiveness of each NPI, which is crucial to inform future preparedness response plans. Because NPI all have economic, psychological and social costs, it is paramount to evaluate their impact on the transmission of SARS-CoV-2, and on the dynamics of the COVID-19 epidemic.

Many studies relied on mechanistic models of epidemics in order to either predict its course Davies et al. (2020), evaluate vaccine prioritization strategies Bubar et al. (2021), or retrospectively measure NPIs impact. During the beginning of the epidemic, the focus was mostly on the timing of NPI initiation Pei, Kandula and Shaman (2020); Li et al. (2021) rather than the effect. However, disentangling the effect of each NPI is a complex problem as their allocation is not randomized and depends on the epidemic state. Many approaches aggregated data from multiple countries. Some worked on regression from time series based on incidence data Banholzer et al. (2021); Islam et al. (2020); Hsiang et al. (2020). Others used semi-mechanistic models and evaluated the percent reduction on the effective reproductive number Liu et al. (2021a); Haug et al. (2020); Flaxman et al. (2020). We preferably work at a country level so that the effect is not confounded by various behaviors and adherence levels of the population. Most of the work published at a country-level has focused on a on single aggregated NPI such as the Oxford COVID-19 Government Response Tracker Hale et al. (2021), very early epidemic Kraemer et al. (2020); Dehning et al. (2020), or a very limited set of interventions, see Brauner et al. (2021) for a review. Regarding France, Salje et al. (2020); Di Domenico et al. (2020) quantified the effect of various NPI. However, they based their results on the early epidemic and, they have a limited set of interventions. Our work rather focuses on the impact of each NPI on the transmission rate, a more valid indicator than direct epidemic curves or the reproductive number because it is independent of the infected proportion of the population. Finally, existing works do not account for vaccination roll out, introduction of variants, and the importance of weather while estimating the impact of NPI. The estimation of the independent effect of each of these factors leads to a challenging problem of estimation, including concerns about practical identifiability of each effect.

In this work, we propose a two-step approach. First, we estimate the transmission rate of SARS-CoV-2 and its variations in the 12 non-insular French regions over a period of more than a year - since March 2, 2020 until March 28, 2021. Second, we estimate, using linear regression, the effects of several NPI on the transmission rate while accounting for seasonal weather conditions throughout the pandemic, as well as the appearance of non-historical variants of concern (VoC) and an increasing proportion of vaccinated people. The first step consists of estimating the transmission rates in the 12 non-insular French regions.

This is extremely challenging due to the sparsity and noisiness of the available data, and also because the parametric shape of the transmission rates is unknown. Using data assimilation across multiple geographical regions, and coupling public data with a dynamic mechanistic model, smooth transmission rates can be estimated through a Kalman filtering approach Simon (2006); Bensoussan (2018) – as already used in epidemiology for COVID-19 spread Arroyo-Marioli et al. (2021); Islam, Hoque and Amin (2020) or for other epidemics with regional variability Hu et al. (2020). More precisely, we develop an ingenious methodology to tackle this difficult problem, based on two important methodological innovations: (1) in the model with the introduction of a time-varying dynamics for the transmission rate including a Wiener process accounting for modeling error, and (2) in the way the population is integrated as we follow a new method Kalman-type filter, compatible with population approaches. This method in which the log-likelihood functional – estimated using for example the Unscented Kalman Filter Julier and Uhlmann (2002) – elegantly couples data across multiple geographical regions is presented in Collin, Prague and Moireau (2020). These two innovations are coupled in a strategy allowing the estimation of smooth transmission rates without knowledge of their shapes. This second step allows us to provide estimations alongside associated uncertainties for a) the transmission rate of SARS-CoV-2 throughout the COVID-19 pandemic in France, b) the effect of the principal NPI implemented in France on this transmission rate, and c) the effect of seasonal weather conditions and new VoC circulation, from observed hospitalization data in French regions.

Section 2 presents the data and the SEIRAH model. In Sections 2.4 and 2.5, the two-step strategy to estimate the transmission rate, and the regression to estimate the effects of NPI are presented. Our results are highlighted in Section 3; and their limitations are discussed in Section 4.

2. Material and Methods. Open-data regarding the French COVID-19 epidemic, including hospitalization data, NPI implementations, VoC prevalence and the vaccination program are presented below. The dynamic model of the COVID epidemic, using an extended SIR type model, is then presented. Finally, we describe our strategies to estimate the transmission rate, using a population-based Kalman filter, to determine the impacts of NPI, seasonal weather conditions, and VoCs.

2.1. *Available Data.*

Hospitalization Data. Hospitalization data are extracted from the SI-VIC database (Système d'Information pour le suivi des VICtimes), a governmental system created in 2016 in order to identify and follow victims in exceptional circumstances (e.g. terror attacks). Since March 18th 2020, the SI-VIC database provides, to Santé Publique France, the daily number of hospitalized COVID-19 patients, at multiple geographical scales. In this work, we focus on the 12 French non-insular regions.

An entry in the SI-VIC database corresponds to a patient hospitalized in connection with COVID-19. Specifically, it requires the presence of at least one of two criteria: i) a biologically confirmed diagnosis of COVID-19 (e.g. RT-PCR positive test result), or ii) a chest CT scan suggestive of COVID-19. In this analysis, we rely on the daily incident number of hospitalizations ($Y^{H_{in}}$), and the total number of individuals hospitalized daily (Y^H). The two data series are displayed in Figure 1 over a period of 391 days (from March 2, 2020 to March 28, 2021). Of note, in order to be able to compare the magnitude of epidemic in each region, we standardized the data by the size of the population in each region: direct interpretation would be the number of incident or prevalent hospitalization for 100,000 inhabitants.

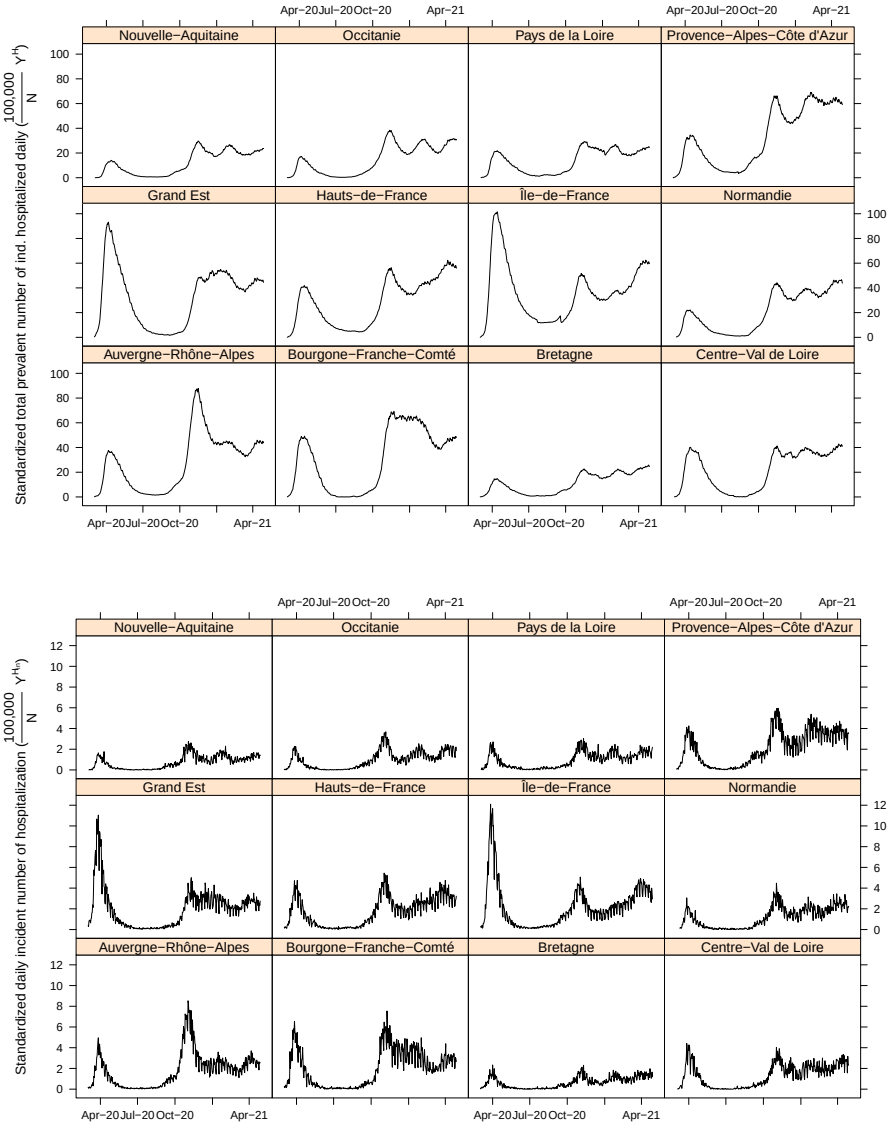

FIG 1. *Top: standardized total prevalent number of individuals hospitalized daily for 100,000 inhabitants* ($\frac{100,000}{N} Y^H$). *Bottom: standardized daily incident number of hospitalizations for 100,000 inhabitants*($\frac{100,000}{N} Y^{H_{in}}$).

Non-Pharmaceutical interventions. Timing and modalities of the various NPI implemented in France over the course of the epidemic have been gathered from the French government's action summary website . In France, public health interventions have been highly multi-pronged. In our analysis, we considered the following summarized NPI occurring during the first year of the epidemic in France: i) first lockdown (with two phases of relaxation/reopening as described bellow), ii) second lockdown (with one phase of relaxation/reopening as described bellow), iii) 8PM curfew, iv) 6PM curfew, v) school closure, vi) bar and restaurant closure, vii) barrier gestures (including all mandatory sanitary protocols: physical distancing, hand washing, part-time remote work and, mask wearing in public spaces). Of note, NPI such as travel bans, enhanced testing, contact-trace-isolate, were ignored to ensure identifia-

bility, as they were either i) with a magnitude difficult to quantify, or ii) enforced in complete overlap with other NPI. In addition, partial interventions at a sub-regional level were not considered as an implementation of a measure. This resulted in a rather similar profile of interventions across regions as most of them were applied simultaneously in the 12 regions of interest. Figure 2 summarizes all 10 considered NPI over time.

The first and the second lockdowns have been disconnected due to their different modalities resulting in different behaviors, and thus possibly different impact on transmission. For example, during the first lockdown (from March 17, 2020 to May 11, 2020), the entire population had to work from home – the only exceptions were for workers in vital sectors, such as medical, security, or food sectors – and outings could not exceed one hour in a one kilometer perimeter. Whereas during the second lockdown (from October 29, 2020 to December 15, 2020), working on-site was authorized when working from home was not feasible, and outings were limited to only 3 hours in a 20 kilometers perimeter. In addition, the end of the first lockdown was gradual and separated by the government in three official phases (phase 1: May 11, 2020 to June 2, 2020, phase 2: June 2, 2020 to June 22, 2020 and phase 3: after June 22, 2020) with many evolving measures as the authorized distance of travels, the reopening of cultural places (*i.e.*, museums), the reopening of non-essential stores, etc. A government campaign to raise awareness about barrier gestures started at the end of the first lockdown. Masks and hand washing were mandatory in many places such as public transports, schools and companies. We assume that *barrier gestures* start on May 11th 2020, assuming that most of mandatory measures with a potential strong impact started to be implemented at this date. Concerning the second lockdown, a reopening of the non-essential stores happened 2 weeks prior to the end of the lockdown prior to Christmas. To account for this, we separate the second lockdown in two phases: a full lockdown until November 28th 2020 and a reduced lockdown afterwards.

School closures were documented as of the holiday schedule, which can vary between regions. In France, schools were also closed during the first lockdown. Furthermore, the reopening of schools from May 11, 2020 (end of first lockdown) to July 4, 2020 (end of term) was very progressive with phased school and level-reopening and attendance rising slowly back to normal. We averaged the estimated school attendance to be 30% of school capacities during that transition. Complete closure of all of bars and restaurants happened twice over the study period: first a few days before the implementation of the first lockdown in March 2020 (as for schools closure), and along the beginning of the second lockdown in October 2020. The measure was lifted progressively in all regions after the end of the first lockdown in June 2020. At the end of the second lockdown, the measure was not lifted considering the epidemic situation was not good enough. It was still in place at the end of the study period.

Curfew measures started to be implemented October 17, 2020 in the 12 regions of interest in several major cities and in *Île-de-France*, with a curfew from 9PM to 6AM. It was extended on October 22nd to 54 departments (sub-regional administrative unit). It was suspended during the second lockdown from October 30th to December 15th 2020, when a new one was implemented at 8PM at a national level. From January 2-12, 2021, the starting hour was changed to 6PM progressively at the departmental level until January 16th when it switched to 6PM for all 12 regions of interest. Finally on March 20, 2021, the starting hour was changed again nationally to 7PM. To model the measures we considered only two variables, grouping curfews starting at 9PM or 8PM and those starting at 6PM or 7PM, and we set the value to 1 only when a whole region was under the curfew. Assuming that curfews and lockdowns induce different behaviors, curfews are not considered to be included in lockdowns, even if it is forbidden to go out in the evening during lockdowns. Similarly, 6PM curfew and 8PM curfew were considered as different interventions instead of nested ones as they are likely to induce different behaviors.

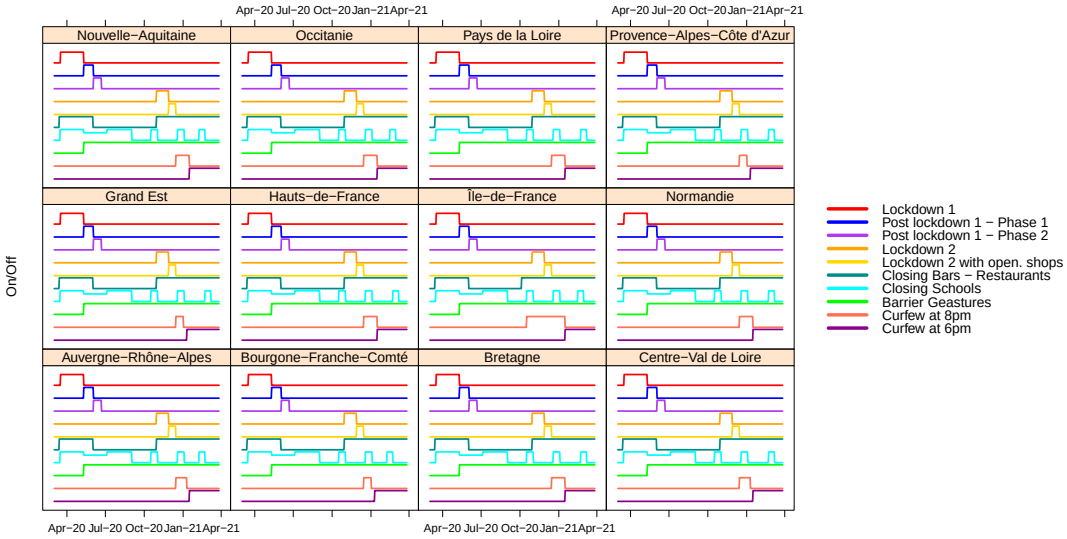

FIG 2. *Major NPI at the regional level*

2.2. *Other exogenous variables: weather, VoC, vaccination coverage.*

Weather conditions. The role of weather conditions on the transmission of the SARS-CoV-2 remains disputed, and early publications were criticized for their inconsistency in results Zaitchik et al. (2020). Nonetheless, the potential impact of both temperature and humidity on aerosolized and fomite transmission routes based on comparisons with other respiratory infections is based on sound mechanistic arguments Rodó et al. (2021). Additionally, with the Northern Hemisphere going through a second winter season during the pandemic, evidence of the association of transmission with seasonal trends of temperature and humidity appears more robust in recent publications Rodó et al. (2021); Liu et al. (2021b). Daily weather data – namely temperature in Celsius degrees (T), relative humidity in percentage (RH) and absolute humidity in g.m^{-3} (AH) - measured by meteorological stations were extracted from the National Oceanic and Atmospheric Administration database using `R` package *worldmet*. Regional daily means were estimated with all stations located in the region or at less than 10 km from the region border. To account for variations of population density across a region, a weighting was used based the on population living in a 10km buffer around each station, giving more importance to weather conditions around densely populated areas. We use the PREDICT index of weather transmissibility of COVID-19 (IPTCC), as defined by Bukhari and Jameel (2020) and Roumagnac et al. (2021):

$$IPTCC = 100\, e^{-\frac{1}{2}\left[\frac{(T-7.5)^2}{196}+\frac{(RH-75)^2}{625}+\frac{(AH-6)^2}{2.89}\right]}.$$

As for interpretation, this rough index ranges from 0% to 100%, the smaller the less favorable are the conditions for COVID-19 transmissions. For France, we observe a seasonality of IPTCC being small during summer months and higher otherwise, with a north-east/south-west gradient.

From this index, we created a weather variable by normalizing (min-max range equal to 1), subtracting the global average value and by reversing it, see black curves in Figure 3. Finally, to focus on the seasonal variation, a *loess* smoothing with a span of 0.2 was applied before use, leading to a smooth weather variable, denoted by W in what follows. Resulting seasonal

variations of this variable for each of the 12 regions of interest over the study period are shown on Figure 3 (red curves). The lower the value, the closer temperature and humidity conditions were close to optimal transmission conditions defined by Bukhari and Jameel (2020). Again, two clear period appear, summer with higher values of this weather variable and winter with lower values. This variable is denoted $W_i(t)$ for the ith region. When taking $W = 0$, one considers the global average value over all French regions over the study period.

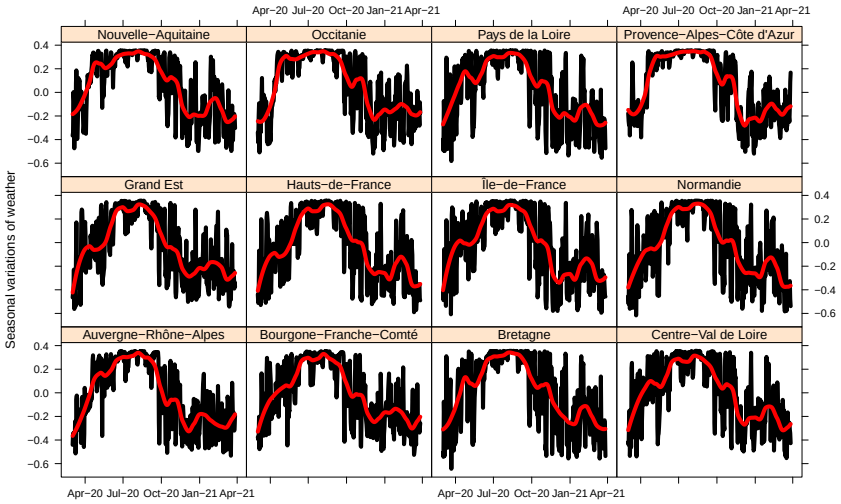

FIG 3. *Weather variable modeling the seasonal weather conditions of the 12 regions of interest in black and after smoothing in red (denoted by W in this paper). The higher the value, the lower the favorable conditions (temperature and humidity) for COVID-19 transmission.*

VoC of SARS-CoV-2. Some variants of the SARS-CoV-2 virus have been classified by national and international health authorities as VoCs as they impact transmissibility or virulence, or decrease the effectiveness of measures World Health Organization. Since January 2021, French health authorities have conducted surveys to estimate the prevalence of three VoCs: 20I/501Y.V1 (Alpha), 20H/501Y.V2 (Beta) and 20J/501Y.V3 (Gamma). The survey of VoC Delta started after this study. We therefore included the cumulative proportion of cases infected by one of these VoCs as a potential covariate explaining the transmission. We used data from two cross-sectional "flash" surveys performed on January 7-8, 2021 and January 27, 2021, see Santé Publique France; Santé Publique France and the weekly estimation of VoC circulation provided by the SI-DEP database at the regional level from February 12, 2021 to March 28,2021 .

Between January 8, 2021 and February 12, 2021, the estimated proportion of the sum of the three VoC have increased from a national mean of 3.3% to a national mean of 46%. To complete the missing data, we assume that the proportion before January 8, 2021 equals to 0% and that the evolution is linear between January 8 and January 27, 2021, and then between January 27 and February 12, 2021. Of note, a logistic and an exponential growth were also tested with no significant changes in conclusions (results not shown). With no data reported for *Bourgogne-Franche-Comté* region on January 27, 2021, only one slope was estimated on the time window. This variable is denoted $VoC_i(t)$ in what follows. Representations of the VoC proportion in each region over time are given in Section 1 of the Web Supplementary Material.

Vaccination process. Vaccination started in France on December 27, 2020. Three COVID vaccines were at that time authorized: BNT162b2 mRNA (Pfizer), ChAdOx1 nCoV-19 (AztraZeneca) and mRNA-1273 (Moderna). To take into account the vaccination process, we used the VAC-SI database which provides the cumulative percentage of the population who have been vaccinated with at least one dose of vaccine over time. Vaccination was first targeted to the elderly 75+ years of age. The proportion of vaccinated individuals, denoted here within as V, ramped up to about 12% by the end of study period. Representations of the vaccination coverage in the population over time are given in Section 1 of the Web Supplementary Material.

2.3. *Model of the epidemic.*

The mechanistic model:. We model the evolution of the COVID-19 epidemic using an extended SEIR type model Wang et al. (2020), called a SEIRAH model, adapted from Wang et al. (2020); Prague et al. (2020) where the population of size N is divided into 5 compartments: susceptible S, latent exposed E, symptomatic infectious I, asymptomatic/paucisymptomatic infectious A, hospitalized H, removed R (i.e. both recovered and deceased), see Figure 4. The number of vaccinated people denoted by V is assumed to be known, see Section 2.2. The dynamics of such model is given by

$$
(1) \qquad \begin{cases}
\dot{S} = -\underline{b}\left(1 - \dfrac{V}{N}\right)\dfrac{S(I+\alpha A)}{N} \\
\dot{E} = \underline{b}\left(1 - \dfrac{V}{N}\right)\dfrac{S(I+\alpha A)}{N} - \dfrac{E}{D_E} \\
\dot{I} = \dfrac{r_E}{D_E}E - \dfrac{1-r_I}{D_Q}I - \dfrac{r_I}{D_I}I, \\
\dot{R} = \dfrac{r_I I + A}{D_I} + \dfrac{H}{D_H} \\
\dot{A} = \dfrac{1-r_E}{D_E}E - \dfrac{A}{D_I} \\
\dot{H} = \dfrac{1-r_I}{D_Q}I - \dfrac{H}{D_H}
\end{cases}
$$

where $\alpha, r_E, D_E, r_I, D_I, D_Q, D_H$ are time-independent parameters described in Table 1 while b is a function of time modeling the disease transmission rate.

Parameter	Interpretation	Value
$\underline{b}$	Transmission rate of infectious cases	Region Specific - **Estimated**
r_E	Ascertainment rate	0.844 He et al. (2021)
r_I	Non hospitalized rate	0.966 Angulo, Finelli and Swerdlow (2021)
α	Ratio of transmission between A and I	0.55 Li et al. (2020)
D_E	Latent (incubation) period (days)	5.1 Lauer et al. (2020)
D_I	Infectious period (days)	5 Cevik et al. (2020)
D_Q	Duration from I onset to H (days)	11-D_E = 5.9 Delfraissy et al. (2020)
D_H	Hospitalization period (days)	18.3*
N	Population size	Region Specific

TABLE 1

*Model parameters for the SEIRAH model, and associated values. *Computed using the correlation between the data $Y^{H_{in}}$ and Y^H when considering region data, see Section 2 of the Web Supplementary Material.*

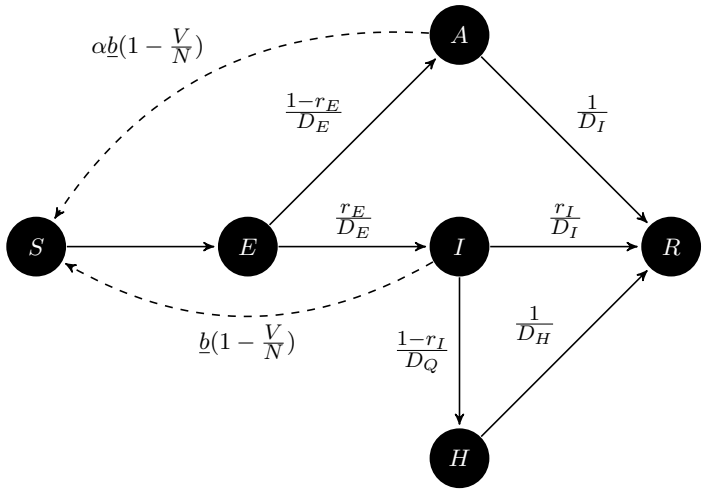

FIG 4. *SEIRAH model representation – adapted from Wang et al. (2020); Prague et al. (2020)*

The observation model:. The two quantities Y^H and $Y^{H_{in}}$ relate to the solutions of System (1) respectively as for all regions $i = 1, \ldots, 12$, for all observation time in days $j = 1, \ldots, 391$: $Y^H_{ij} = H_i(j) + \epsilon^H_{ij}$ and $Y^{H_{in}}_{ij} = \frac{(1-r_I)}{D_Q} I_i(j) + \epsilon^{H_{in}}_{ij}$, in which ϵ^H_{ij} and $\epsilon^{H_{in}}_{ij}$ represent normally distributed constant measurement errors.

Effective reproductive number and attack rates:. When individuals are homogeneous and mix uniformly, the effective reproductive ratio $R_{\text{eff}}(t)$ is defined as the mean number of infections generated during the infectious period of a single infectious case at time t. In this model, the effective reproductive ratio can be written as a function of model parameters (see Section 3 of the Web Supplementary Material for details), namely

$$R_{\text{eff}}(t) = \underline{b}(t)\Big(1 - \frac{V}{N}\Big)\frac{S(t)}{N}\left(D_I\alpha(1-r_E) + \frac{D_I D_Q r_E}{(1-r_I)D_I + r_I D_Q}\right). \tag{2}$$

When neglecting the deaths, the proportion of infected individuals assuming no waning immunity – also called attack rates – among the population in each region at a given date is given by:

$$\frac{E+I+R+A+H}{N} = 1 - \frac{S}{N}.$$

2.4. *A population-based Kalman filter to estimate the transmission rate.*

Framing the problem.. In order to sustain the modeling choices behind the disease transmission rate $\underline{b}$, and more generally the selected inference procedure, we used Kalman-based estimation strategies Simon (2006); Bensoussan (2018) applied to the epidemic model (1). We define the global transmission rate $t \mapsto b(t)$, which accounts for the proportion of susceptible removed from the system thanks to vaccination:

$$b(t) \stackrel{\text{def}}{=} \underline{b}(t)\left(1 - \frac{V(t)}{N}\right).$$

We propose to estimate b and ultimately to retrieve $\underline{b}$ using an estimation of vaccinated people V, see Section 2.2. In each region, we then introduce a dynamic equation for b of the form

$$\mathrm{d}b^i(t) = g^i(t)\mathrm{d}t + \mathrm{d}\nu^i(t), \tag{3}$$

where ν^i consists in a Wiener process such as for all $t, s \geq 0$, $\nu^i(t) - \nu^i(s) \sim \mathcal{N}(0, (t-s)\sigma_\nu)$ with σ_ν^2 is known and constant in all regions and g^i is a function describing the evolution of the global transmission rates b^i in each region.
After discretization using forward Euler time-scheme with small-enough time-step δt, we end up with a discrete-time dynamical system applied to the variable $x = (E, I, R, A, H)^\intercal \in \mathbb{R}^5$, for each region $i = 1, \ldots, 12$:

$$z_{n+1}^i \stackrel{\text{def}}{=} \begin{pmatrix} x_{n+1}^i \\ b_{n+1}^i \\ \theta_{n+1}^i \end{pmatrix} = \begin{pmatrix} x_n^i + \delta t\, f(x_n^i, b_n^i, \theta_n^i) \\ b_n^i + \delta t\, g^i(t_n, \theta_n^i) \\ \theta_n^i \end{pmatrix} + \begin{pmatrix} 0_5 \\ 1 \\ 0_{N_p} \end{pmatrix} \nu_n^i, \tag{4}$$

where f accounts for the dynamics of (E, I, R, A, H) in (1) while S is reconstructed afterwards using $S = N - (E + I + R + A + H)$ in each region. In the time discrete system, $(\nu_n^i)_{n \geq 0}$ now represent independent random variables, normally distributed with 0 mean and a variance equal to $\sigma_\nu^2 \delta t$. Moreover, if constant parameters have to be estimated, the vector $\theta \in \mathbb{R}^{N_p}$ gathers all of them. For the estimation in the following, we transform the variable z to account for biological constraints. Remarking that all state variables are positive and bounded by N, the total population size of the region, we transform the state variable using $x \mapsto \text{logit}(x/N)$. We also apply a similar transformation to $b \mapsto \text{logit}(b/max_b)$. Of note, state and transmission rate variables are more likely to have Gaussian distributions in the transformation space. After calibration, we fixed $max_b = 1.5$ and checked that other values did not substantially changed the results (result not shown).

Population approach.. To perform the estimation, we rely on an extension of the classical Unscented Kalman filter (UKF) Julier and Uhlmann (1997, 2002); Simon (2006). The particularity in this application, as described in Section 2.4, is that multiple series of data are observed jointly in multiple regions as we observe multiple realizations of the same epidemic. In order to take into account in our Kalman estimation that parameters in different regions are correlated as in a population approach, we follow a recently proposed population-based Kalman formulation Collin, Prague and Moireau (2020). As in mixed-effect models Lavielle (2014), each initial uncertainty variable z_0^i is assumed to be randomly distributed around a common population intercept z_0^{pop} with a Gaussian distribution of unknown covariance $\boldsymbol{Q}_0$, namely:

$$z_0^i \sim_{\text{i.i.d.}} \mathcal{N}(z_0^{\text{pop}}, \boldsymbol{Q}_0).$$

When we treat the population intercept as the empirical mean over the population members in the construction of the objective function, we recover a classical filtering problem Simon (2006) on the aggregated variable $\mathbf{z} = (z^1 \cdots z^{N_r})^\intercal$. The only difference is the formulation of the initial covariance prior $\hat{\boldsymbol{P}}_0$ which couples the observations across regions and can be written as:

$$\hat{\boldsymbol{P}}_0^{-1} = \frac{1}{N_r^2} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} (1 \cdots 1) \otimes \boldsymbol{M} + \left[\mathbb{1}_{N_r} - \frac{1}{N_r} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} (1 \cdots 1) \right] \otimes \hat{\boldsymbol{Q}}_0^{-1},$$

where $\otimes$ indicates a Kronecker product, $\hat{\boldsymbol{Q}}_0$ is a prior of $\boldsymbol{Q}_0$ and $\boldsymbol{M}$ a small penalization matrix that guarantees the overall invertibility of $\hat{\boldsymbol{P}}_0$. As a consequence, the matrix $\hat{\boldsymbol{P}}_0$ is not block diagonal with respect to the region i and thus all the region dynamics are coupled. The resulting time-discrete Kalman estimator couples all the regions together to produce a population-based estimation. Note that in such strategy, forcing a variable to be constant in the population is possible by simply choosing $\hat{\boldsymbol{Q}}_0$ such that $\text{Tr}(\hat{\boldsymbol{Q}}_0^{-1})$ is very small with

respect to $\mathrm{Tr}(\boldsymbol{M})$. Conversely, $\mathrm{Tr}(\boldsymbol{M})$ small with respect to $\mathrm{Tr}(\hat{\boldsymbol{Q}}_0^{-1})$ in a large population $N_r \gg 1$ will encourage each region to remain decoupled from each others. Starting from a given prior knowledge, our Kalman implementation uses the available measurements $(Y_{ij}^H, Y_{ij}^{H_{in}})_{1\leq i\leq 12, 0\leq j\leq 391}$, to compute recursively in time the following estimates:

$$\hat{\boldsymbol{z}}_n \simeq \mathbb{E}(\boldsymbol{z}_n | (Y_{ij}^H, Y_{ij}^{H_{in}}), 1 \leq i \leq 12, 1 \leq j \leq n), \quad 0 \leq n \leq 391,$$

and

$$\hat{\boldsymbol{P}}_n \simeq \mathbb{C}\mathrm{ov}(\boldsymbol{z}_n - \hat{\boldsymbol{z}}_n | (Y_{ij}^H, Y_{ij}^{H_{in}}), 1 \leq i \leq 12, 1 \leq j \leq n), \quad 0 \leq n \leq 391.$$

Then exploiting that $\hat{\boldsymbol{z}}_n$ gather the augmented state $(x_{n+1}^i, b_{n+1}^i, \theta_{n+1}^i)$, $1 \leq i \leq 12$ of the 12 regions, a simple post-processing over the regions provides estimations of b and the state variable. We refer to Collin, Prague and Moireau (2020) for further details about this Kalman-based population approach.

Estimation strategy.. As we want to inject as few information as possible on the shape of the transmission rate b, we will first consider that the Wiener process ν defined in Equation (3) is a time-dependent function which thus encompass the complete dynamics of b. This corresponds to choosing $g \equiv 0$, meaning we have no prior knowledge on the evolution of transmission rates, *i.e.* on the NPI effects. However, to avoid overfitting, our objective is to disentangle the latent trajectory of b from possible noise in the data which can be observed in Figure 1. We adopt a 3-step approach described below consisting in smoothing the trajectory of b:

1. Estimate a reasonable prior for the initial transmission rate b before the start of any NPI. We use data before the first lockdown (10 days available) and assume the transmission rate $b^i(t) = b_{init}^i$ for $t = 1 \cdots 10$ days to be a constant. In other words, we apply the population Kalman filter estimation described above with θ from Equation (4) being reduced to b_{init}.
2. Estimate the shape of b with prior on initial value but without any prior on the dynamics. We set the initial value of b^i for times $t = 1 \cdots 10$ days to b_{init}^i and take $g(t) \equiv 0$ such that Equation (3) rewrites $\mathrm{d}b(t) = \mathrm{d}\nu(t)$. We apply the population Kalman filter estimation described above. Of note, the parameter vector θ from Equation (4) is now empty. As there is no information on b, the model error is very important and could lead to an over-fitting of the data. We then build a prior for the dynamics of b by fitting a parametric shape based on sum of logistic functions on the weighted average trajectories of b over all regions with least-square method. Then, we build a prior for the dynamics of b by fitting a parametric shape based on sum of logistic functions on the weighted average trajectories of b over all regions with least-square method. We choose logistic functions as they seem well adapted to model variations due to for example stiff lockdowns or smooth unlocks. We set the number of logistic functions from the number of main changes of variations of the weighted average trajectories of b. This function will represent a prior g on the dynamics.
3. Estimate the shape of b with prior on initial value and informative prior on the dynamics. We set the initial value of b^i for times $t = 1 \cdots 10$ days to b_{init}^i and take $\mathrm{d}b(t) = g(t)\mathrm{d}t + \mathrm{d}\nu(t)$ such as in Equation (3). We apply the population Kalman filter estimation described above. As the dynamics of b still includes a modeling noise, the shape of b could be different from the prior transmission rate defined from Step 2. This is the result of this final analysis that is used for further description of the effect of NPI.

2.5. *Explanatory model for the transmission rate.*

Mixed effects model. Using the function b obtained with population Kalman filter as described in Section 2.4, one can determine the impacts of NPIs, seasonal weather conditions, and VoCs on the transmission rate $\underline{b} = \frac{b}{(1-V/N)}$. We followed the example of Flaxman et al. Flaxman et al. (2020) - in considering interventions effects to be multiplicative. Therefore, we fitted a linear mixed effect model on the logarithm transformation of the transmission. The model equations are given in Section 4.2 of the Web Supplementary material. It consists of an intercept, which represents the average transmission of COVID-19 across all regions without NPI without VoC circulation and for an average French weather condition (W=0), the sum of effect of the 10 NPIs described in Section 2.1, the effect of weather and the effect of the VoC percentage. It is known that the transmission of SARS-CoV-2 is very different indoors and outdoors Bulfone et al. (2021). Thus, we also added an interaction effect between bar and restaurant closure and the weather, accounting for the opening of outdoor sitting areas. Of note, this is the only interaction tested to avoid overfitting. We finally added random effects to account for heterogeneity between regions. We added a random intercept and random slopes for the effect of first, second lockdown and 6PM curfew as they may be highly variable between regions. We assume a full covariance matrix for the random effects, allowing effects to be correlated together in each region. In particular, we believe that they may be impacted by various factors not accounted for in the epidemic model, such as the density, the age distribution or the urbanization of the region.

Note that we added a 7-day delay to the lockdowns, which can be interpreted as a necessary time to allow people to organize themselves for adaptation (implementation of working from home, childcare etc ...). This choice has been motivated by the observed 7-day delay in the transmission rates obtained by Kalman filters and will be discussed. To ease interpretation of the estimated effects, parameters were transformed back and expressed as a decrease or an increase in transmission in percentage by applying the function $x \mapsto 100(e^x - 1)$. Classical 95% confidence intervals were obtained using $100(e^{x+/-1.96SE(x)} - 1)$, were SE is the standard error obtained from the regression.

Interpretation of the impact of seasonal weather conditions. To facilitate understanding, we calculated some NPI effects during the summer and winter periods. To do so, we replace in the results W by its summer (resp. winter) average from June 21st, 2020 to October 21st, 2020 (resp. before June 21st, 2020 and after October 21st, 2020) over all regions.

Basic reproductive number. The intercept of the regression presented above represents the mean transmission rate over all regions when there is no NPI in place, no VoC and the weather condition is taken to be the average weather condition over a year in France. Thus, inserting it in Equation (2) directly provides the basic reproductive number and 95% confidence intervals can be computed using the standard error of this parameter.

3. Results.

Estimation of the transmission rate using a population-based Kalman filter. Step 1 of estimation provided the initial values for the transmission rate and exhibited quite similar values between regions (average 0.78 sd 0.012, see Section 4.1 of the Web Supplementary Material for the values at the regional level). Although higher values are found for the regions with the higher first wave, the small variability between regions may indicate that the magnitude of the first wave was not fully driven by a higher transmission rate, but also by different initial epidemic states (i.e. number of exposed and infectious cases resulting in differences in virus introduction timelines or the occurrence of super-spreading events Roux, Massonnaud and Crépey (2020)). In Step 2, the transmission rates obtained without knowledge of their shapes are given in Figure 5 (top, left) for all the regions. As there is no information on b,

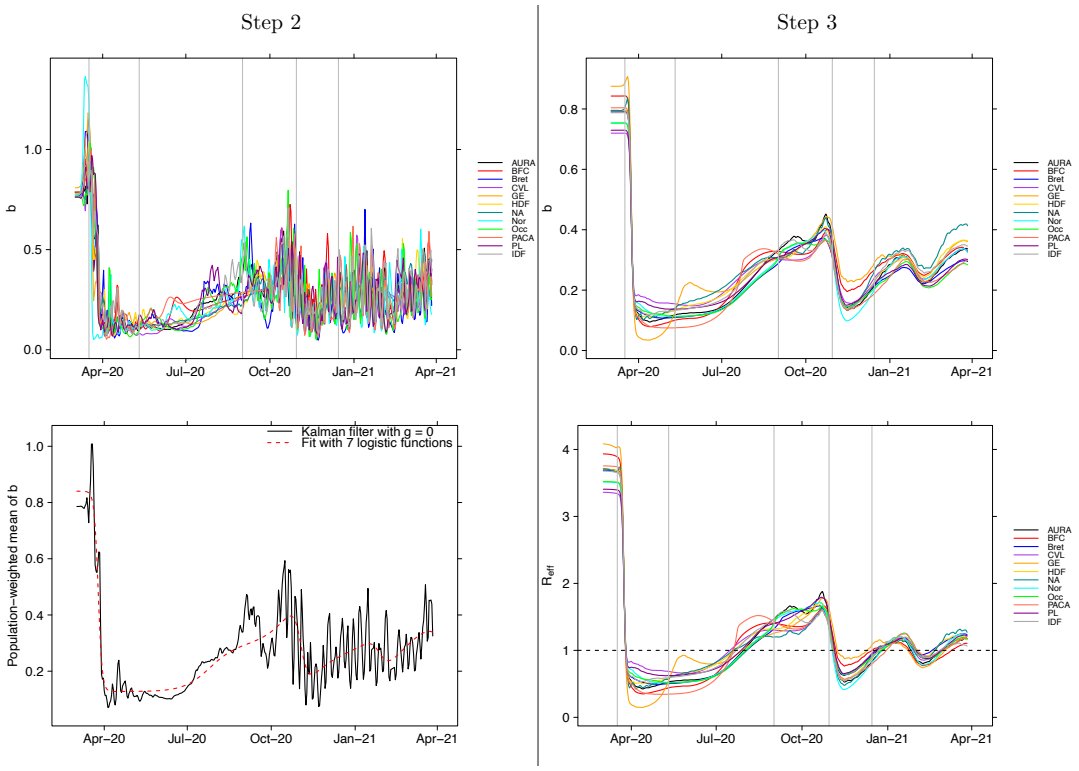

FIG 5. *First column (Step 2): Top - Time evolution of b in the 12 French regions with $g = 0$ (no* apriori*). Bottom - Mean value of b over time (black line) obtained at Step 2 fitted using 7 logistic functions (dashed red line). Second column (Step 3): Top - Time evolution of b in the 12 French regions when g corresponds to 7 logistic functions obtained at Step 2 (with $\nu \neq 0$). Bottom - Time evolution of R_{eff} in the 12 French regions when g corresponds to 7 logistic functions (with $\nu \neq 0$). If R_{eff} is superior to 1 (see the horizontal dashed black line), the infection will spread. The first (resp. second, third, fourth, fifth) vertical gray line corresponds to the first day of the first lockdown (resp. the last day of the first lockdown, the start of the academic year, the first day of the second lockdown, the last day of the second lockdown).*

the model error is very important and leads to an over-fitting of the data. In particular, oscillations with zero mean and with period close to 7 days are visible for many regions. This is related to a known under-reporting during week-ends. The weighted average trajectory of b is displayed in Figure 5 (bottom, left). We approximated the function g by a sum of 7 logistic functions. Results of the least-square fitting are displayed in dashed line and constitute the prior for the following. Finally, in Step 3, the transmission rates b are obtained and displayed in Figure 5 (top, right). As the dynamics of b still includes a modeling noise, the shape of b is different from the prior transmission rate defined from Step 2 but still smoother.

Effective reproductive number. The resulting effective reproductive ratio R_{eff} is given at the regional level in Figure 5 (bottom, right). It starts from a value ranging between 3.5 and 4 in all regions. This is partially driven by winter-like conditions during the first days of March 2020. The basic reproductive number will be later computed adjusting for weather. The time variation of R_{eff} shows that it rapidly gets under the critical value of one after the first lockdown initiation and then oscillates around this value.

Attack rates. The attack rate represents one intrinsic feature of interest in any epidemic model (as described in Section 2.3). On top of adding an insight on the epidemic dynamics by adding extra knowledge on the number of possible hidden / unmeasured cases, attack rates are a good way to validate how realistic the model fitting is. Figure 6 represents the attack

rate at various important times: at the end of first lockdown (May 11, 2020), October 5, 2020 and at the end of our study period (March 28, 2021). The national French attack rates are respectively estimated to be 5.7%, 8.8%, and 25.3% at these dates.

Basic reproductive number. After adjusting for average weather over a period of one year in France, we estimate the national average basic reproductive number at 3.10 [2.95 ; 3.26].

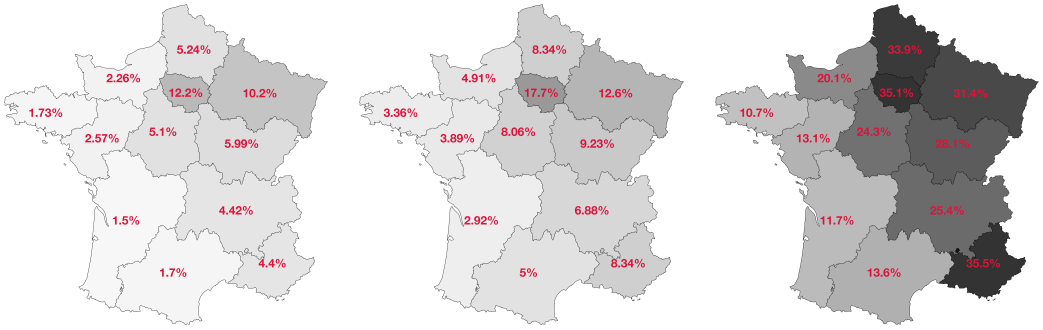

FIG 6. *Model estimation for the proportion of naturally immunized individuals in the population (deaths and vaccinated people not taken into account) on May 11th, 2020 (left), on Oct. 5th, 2020 (middle) and on March 28th, 2021 (right).*

	Modification of the transmission rate	
Covariate	**scale**	**increase/decrease**
NPI		
Lockdown 1 - delay of 7 days	-78% [-82% ; -74%]	↘
Post lockdown 1 - Phase 1	-54% [-56% ; -52%]	↘
Post lockdown 1 - Phase 2	-48% [-50% ; -47%]	↘
Lockdown 2 - delay of 7 days	-54% [-57% ; -49%]	↘
Lockdown 2 with opened shops	-51% [-53% ; -49%]	↘
Closing schools	- 7% [- 8% ; - 5%]	↘
Barrier gestures	-46% [-48% ; -44%]	↘
Curfew at 6PM	-30% [-33% ; -26%]	↘
Curfew at 8PM	-28% [-31% ; -25%]	↘
Bar and restaurant closure (ref. $W = 0$)	-10% [-13% ; - 8%]	↘
Bar and restaurant closure, summer (vs. ref.)	- 8% [-11% ; - 4%]	↘
winter (vs. ref.)	-11% [-14% ; - 8%]	↘
Other factors		
100% of VoC circulating	22% [15% ; 28%]	↗
Weather effect during summer (ref. $W = 0$)	-22% [-24% ; -21%]	↘
during winter (ref. $W = 0$)	10% [9% ; 11%]	↗

TABLE 2

Estimation and 95% confidence intervals of the effects of seasonal weather conditions, VoC proportion, and NPI on the transmission rates. Model AIC = -1,388.

Effects of NPI on transmission rate. Table 2 summarizes the estimation of the effect of NPI on the transmission rates derived from the fixed effects of the model provided in Section 2.5. Regression residuals, fixed and random effects are given in Section 4.2 of the Web Supplementary Materials. Figure 7 shows the corresponding individual fits. The very restrictive first lockdown (respectively the less restrictive second lockdown) decreases the transmission rate by $\sim 80\%$ (respectively $\sim 55\%$). The impact of school closure is around 7% and the bar and restaurant closure around 10%. The barrier gestures decrease the transmission by $\sim 46\%$.

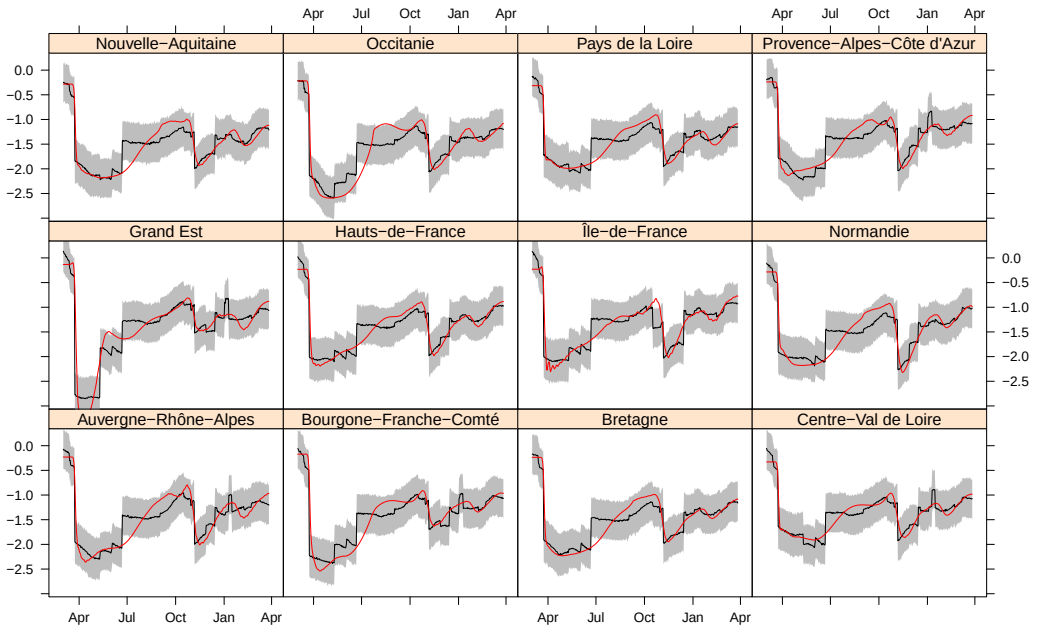

FIG 7. *Transmission rates. Estimated using the population-based Kalman filter. Individual fits for each region. Random effects on* $\log(b_{init})$, b_{l1}, b_{l2} *and* $b_{curf6PM}$.

One can see that the values between the curfew at 6PM and the curfew at 8PM (around 30%) are very close and not statistically different (3% [-2% ; 8%]). We demonstrate that all NPI in this analysis decrease the transmission and have an effect significantly different from zero.

Effect of weather. The seasonal weather conditions variable is unitless, so the interpretation of its estimated effect should be made by comparing a value to a reference, which is set to be the average weather in France. We find that transmission is on average significantly increased by 10% during the winter period, and significantly decreased by 22% during the summer period, compared to average weather. Figure 8 (top) presents the estimated effect of seasonal weather conditions over the period of study, by region.

The estimated interaction between the closure of bars and restaurants and the seasonal weather conditions is statistically significant (p=0.037) but is also complex to interpret. We show that although closing bars and restaurants always has a significant effect, decreasing the transmission, it is slightly more effective in winter (11% [8%;14%] decrease) rather than in summer (8% [4%;11%] decrease). Figure 8-Middle presents the estimated effect of bar and restaurant closure for the 12 regions of interest over the study period and its interaction with weather conditions. The weather conditions impacts the estimation, with a stronger effect on transmission reduction in Northern regions compared to Southern ones, and also a stronger effect is observed in the Spring 2021 due to more favorable weather.

Effect of VoC. Concerning the effect of the three VoCs, our results show that an increase from 0 to 100% of the proportion would increase the transmission rate by an estimated 22% [15% ; 28%]. Figure 8 (bottom) presents how the increased transmissibility varies over time due to VoC proportion.

4. Discussion. In this paper, we use an innovative method to infer the transmission rate over time from hospitalization data, and estimate the effect of multiple NPI, weather and

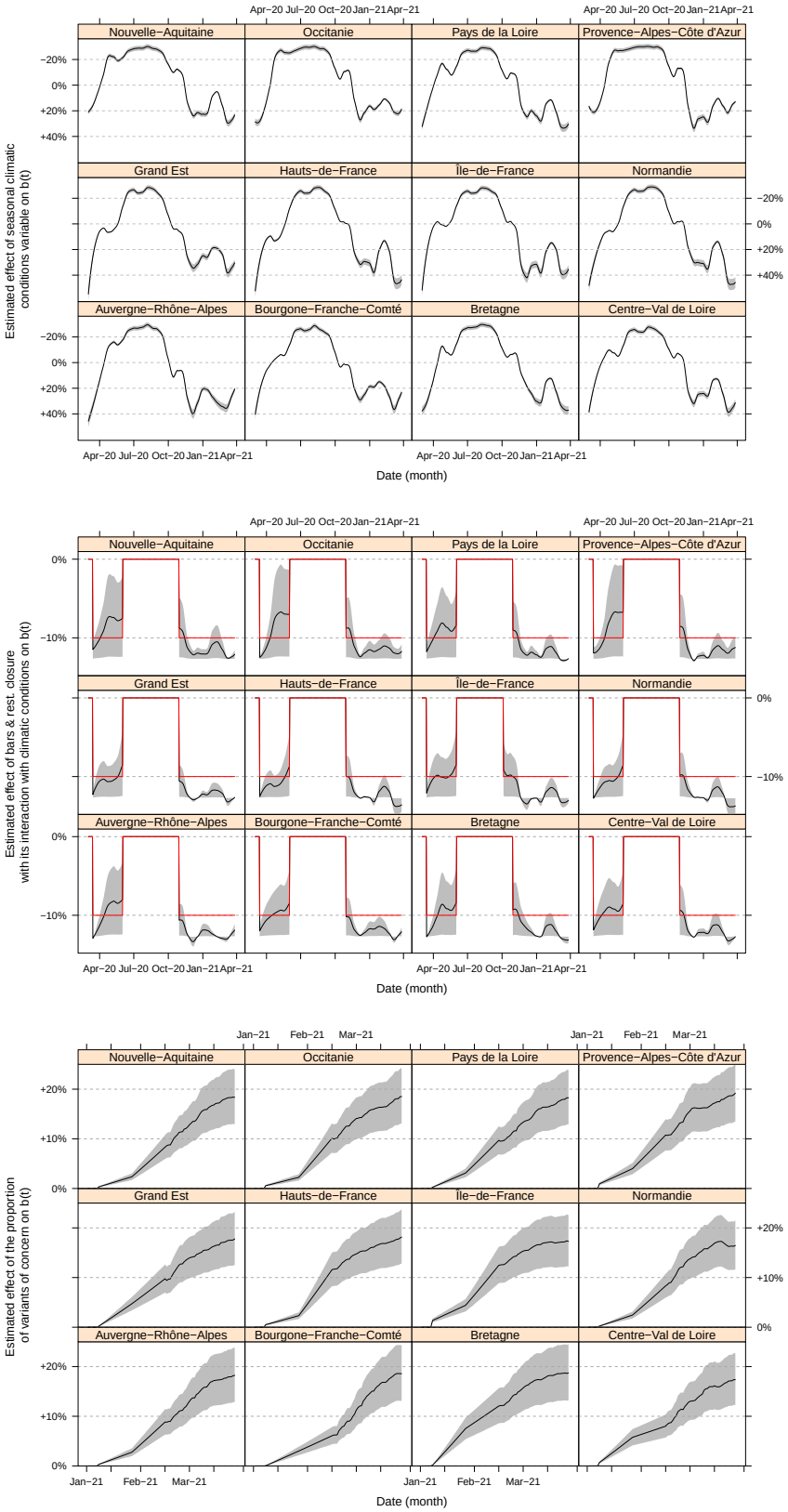

FIG 8. *Top: Estimated effect of seasonal weather conditions and its 95% confidence band for the 12 regions of interest over the study period using the global average value over the period as reference for comparison. Middle: Estimated effect of bar and restaurant closure for the 12 regions of interest over the study period. In red, the main effect of -10%. In black and grey, the effect with the interaction with weather conditions and its 95% confidence band. Bottom: Estimated effect of VoC appearance and its 95% confidence band from January 1, 2021.*

VoC on it. We showed that all NPI considered have a significant and independent effect on the transmission rate. We additionally demonstrated a strong effect of weather conditions, decreasing the transmission in the summer period, and increasing it in the winter period, and we observed an increased transmissibility due to VoCs.

Concerning the transmission rate, interpretation of our results is conditional on the mechanistic model illustrated in Figure 4, and careful attention must be given to the parameters values set from the scientific literature which are detailed in Table 1. This model does not take into account the age-structure of the population on the contrary to the works of Salje et al. (2020); Di Domenico et al. (2020) using French data. However, we believe that using a population approach accounts for different intrinsic characteristics such as population density, age structure, transportation habits, etc. that influence disease spread through each region-specific transmission rate. Another shortcoming of the SEIRAH model is that it does not take into account waning immunity or travel between regions. The model could be slightly modified to take these aspects into account but would necessitate further fixing of more parameters while no other data is available, especially on inter-regional travel. For this reason we did not proceed with this development. Concerning the available data, the drawback to use hospitalization data is its lack of exhaustiveness given that hospital services declaring cases may vary over time. However, one may think that this variation is less important than in other sources of data such as the daily number of new confirmed cases for which the testing policy has greatly changed during the course of the epidemic.

Interestingly, although the models were different and the data not fully identical, our results were comparable in terms of attack rates with existing modeling works Hozé et al. (2021) and sero-prevalence studies Warszawski et al. (2020); Carrat et al. (2020); Santé Publique France. A comparison is available in Web Supplementary material, Section 4.1. Overall, our estimates tend to be slightly higher by less than 5% than other estimation and sero-prevalence studies. This is presumably due to the strong assumption that there is no waning immunity in our model. Indeed, we assume that all individuals, once infected, have a high enough titer of antibody response to systematically test positive in sero-prevalence studies. In terms of reproductive numbers, we found 3.10 [2.95 ; 3.26] (after removing the weather impact using the regression model) compared to 3.18 [3.09 ; 3.24] in Di Domenico et al. (2020) and 2.90 [2.81 ; 3.01] in Salje et al. (2020). These comparisons allow us to validate the estimation strategy, which has the great advantage of estimating transmission rates without assumptions on the shape of the transmission functions. Furthermore, the computational times are not excessive (a few minutes for the full estimation on a classical work station without code optimization).

We restricted our analysis to the estimation of ten NPI effects. More variables could have been added, e.g. partial interventions (applied, for example, only in large cities in regions highly affected by COVID-19) – but those would be inconsistent with our model defined at a regional level. Other potential variables suffer from inconsistent definitions. Working from home is a good example. First, the adherence to this intervention (which can vary due to employee fatigue, organization difficulties and the absence of legal constraint) is difficult to evaluate. Second, regarding quantitative indicators, the DARES carried out surveys for companies with 10 or more employees and showed that the mean percentage of remote work is around 29.6% [18.8%, 40.4%] with the max value during the two lockdowns . However, this indicator accounts for working from home and paid vacation which may be very different in terms of behaviors. Hence by omitting the working from home in the model, its effect is captured by the effect of lockdown and barrier gestures. Identifying the proper effect of working from home is very difficult, because injunctions and implementations of "working from home" fluctuate greatly over time and over geographical areas without a valid way of measuring adherence.

We also had to make a certain number of modeling assumptions. For some interventions, such as lockdowns, we considered a 7-day delay of implementation (which is visible on the transmission rates obtained with the Kalman filter). This choice is validated by other studies such as Dehning et al. (2020), in which the delay can be up to 15 days. However, we investigated its impact on our regression adjustment. Not considering the delay of 7 days greatly deteriorates the fits, see Section 4.2 of the Web Supplementary Materials for more details. Considering our choice to model the effect of the partial opening of schools during the May 11th to July 4th 2020 period as equal to 70% of the effect of a full closure could be considered as an oversimplification. Over this period following the first lockdown, schools reopened very gradually according to three different phases and school attendance growth was even more progressive. Up to June 2nd, strong disparities in openings and attendance existed among regions, with an average of only 30% of pupils under 12 years old attending. Other levels of secondary schools ("collèges" and "lycées") reopened in early June and progressively received more students until vacation, starting in early July. Not to risk facing identifiability issues, we chose not to differentiate regions or phases of reopening and a ratio of 0.7 of closure effect was applied for all regions. Regarding the effect of curfews, we failed at demonstrating a statistically significant difference between 6PM and 8PM. This is may be due to an identifiability issue as the 8PM curfew was in place for less than 3 weeks in many regions. An alternative explanation could be that both curfews were impacting globally social gathering the same way. As for example, both are preventing most private dinners and parties (or at least reducing drastically the number of guests).

In this article, we decided to take into account the weather condition using past works of Bukhari and Jameel (2020) and Bukhari and Jameel (2020). Using separately temperature, absolute, and relative humidity was not explored in this work. All in all, we saw a clear point in including the weather condition as a variable modifying the transmission. Section 4.2 of the Web Supplementary Materials explored simpler models. All deteriorated the fits. However, we remain extremely cautious about the interpretation of the estimated effects of weather conditions and, beyond, of associated mechanisms. Finally, we consider an interaction between seasonal weather conditions and bar and restaurant closure justified by the use of terraces (which have been expanded in many places since the beginning of the pandemic). Of course other interactions can be considered and more complex models can be written but it may lead to over-fitting.

Looking back at the original data, the variant 20I/501Y.V1 (Alpha) appears to have always been predominant (over 90% at any time) compared to the two other VoCs (Beta and Gamma) considered in the French non-insular territory . Widely variable estimations of transmissibility increase for the variant 20I/501Y.V1 (Alpha) have been proposed in the scientific literature, ranging from 29% to 90% Graham et al. (2021); Campbell et al. (2021); Washington et al. (2021); Volz et al. (2021); Davies et al. (2021) the lower values being in line with our own findings. The higher estimates could be partially explained because new VoCs appeared at the beginning of winter in England and in the USA. This may lead to confusion between weather conditions and VoCs, which both increase weather conditions, as mentioned by Campbell et al. (2021). We tested this assumption by removing weather conditions from our model and found an increased transmissibility of 43%, see Section 4.2 of the Web Supplementary Material. We retained the weather condition in our model as it greatly improved the fit.

Altogether, this work is one of the first attempts to evaluate retrospectively the effect of multiple NPIs over a period of one-year of the COVID-19 epidemic. On top of applying a novel methodology to current and important application, this work could be extended to generate “what-if” scenarios and help determine appropriate NPI implementations for future waves of infection.

Acknowledgements. The authors thank the `opencovid-19` initiative for their contribution in opening the data used in this article. This work is supported in part by Inria Mission COVID19, project GESTEPID. The authors warmly thank Jane Heffernan for scientific discussions and a thorough proof-reading of the article. We also thank Linda Wittkop, Jane Heffernan, Quentin Clairon, Thomas Ferté and Maria Pietro for constructive discussions about this work.

Author contributions. AC, BPH, PM, MP and RT designed the study. LL and CV analyzed the data. AC, PM and MP implemented the software code. AC, BPH and MP interpreted the results. AC, BPH, LL, PM and MP wrote the manuscript.

REFERENCES

ANGULO, F. J., FINELLI, L. and SWERDLOW, D. L. (2021). Estimation of US SARS-CoV-2 Infections, Symptomatic Infections, Hospitalizations, and Deaths Using Seroprevalence Surveys. *Journal of the American Medical Association Network Open* **4** e2033706.

ARROYO-MARIOLI, F., BULLANO, F., KUCINSKAS, S. and RONDÓN-MORENO, C. (2021). Tracking R of COVID-19: A new real-time estimation using the Kalman filter. *PLOS One* **16** e0244474.

BANHOLZER, N., VAN WEENEN, E., LISON, A., CENEDESE, A., SEELIGER, A., KRATZWALD, B., TSCHERNUTTER, D., SALLES, J. P., BOTTRIGHI, P., LEHTINEN, S., FEUERRIEGEL, S. and VACH, W. (2021). Estimating the effects of non-pharmaceutical interventions on the number of new infections with COVID-19 during the first epidemic wave. *PLOS One* **16** e0252827.

BENSOUSSAN, A. (2018). *Estimation and Control of Dynamical Systems*. *Interdisciplinary Applied Mathematics*. Springer.

BRAUNER, J. M., MINDERMANN, S., SHARMA, M., JOHNSTON, D., SALVATIER, J., GAVENČIAK, T., STEPHENSON, A. B., LEECH, G., ALTMAN, G., MIKULIK, V., NORMAN, A. J., MONRAD, J. T., BESIROGLU, T., GE, H., HARTWICK, M. A., TEH, Y. W., CHINDELEVITCH, L., GAL, Y. and KULVEIT, J. (2021). Inferring the effectiveness of government interventions against COVID-19. *Science* **371**.

BUBAR, K. M., REINHOLT, K., KISSLER, S. M., LIPSITCH, M., COBEY, S., GRAD, Y. H. and LARREMORE, D. B. (2021). Model-informed COVID-19 vaccine prioritization strategies by age and serostatus. *Science* **371** 916–921.

BUKHARI, Q. and JAMEEL, Y. (2020). Will coronavirus pandemic diminish by summer? *Available at SSRN 3556998*.

BULFONE, T. C., MALEKINEJAD, M., RUTHERFORD, G. W. and RAZANI, N. (2021). Outdoor transmission of SARS-CoV-2 and other respiratory viruses: a systematic review. *The Journal of Infectious Diseases* **223** 550–561.

CAMPBELL, F., ARCHER, B., LAURENSON-SCHAFER, H., JINNAI, Y., KONINGS, F., BATRA, N., PAVLIN, B., VANDEMAELE, K., VAN KERKHOVE, M. D., JOMBART, T., MORGAN, O. and LE POLAIN DE WAROUX, O. (2021). Increased transmissibility and global spread of SARS-CoV-2 variants of concern as at June 2021. *Eurosurveillance* **26** 2100509.

CARRAT, F., DE LAMBALLERIE, X., RAHIB, D., BLANCHÉ, H., LAPIDUS, N., ARTAUD, F., KAB, S., RENUY, A., DE EDELENYI, F. S., MEYER, L., LYDIÉ, N., CHARLES, M.-A., ANCEL, P.-Y., JUSOT, F., ROUQUETTE, A., PRIET, S., SABA VILLARROEL, P. M., FOURIÉ, T., LUSIVIKA-NZINGA, C., NICOL, J., LEGOT, S., DRUESNE-PECOLLO, N., ESSEDDIK, Y., LAI, C., GAGLIOLO, J.-M., DELEUZE, J.-F., BAJOS, N., SEVERI, G., TOUVIER, M., ZINS, M. and THE SAPRIS AND SAPRIS-SERO STUDY GROUPS (2020). Seroprevalence of SARS-CoV-2 among adults in three regions of France following the lockdown and associated risk factors: a multicohort study. *medRxiv*.

CEVIK, M., TATE, M., LLOYD, O., MARAOLO, A. E., SCHAFERS, J. and HO, A. (2020). SARS-CoV-2, SARS-CoV, and MERS-CoV viral load dynamics, duration of viral shedding, and infectiousness: a systematic review and meta-analysis. *The Lancet Microbe*.

COLLIN, A., PRAGUE, M. and MOIREAU, P. (2020). Estimation for dynamical systems using a population-based Kalman filter – Applications to pharmacokinetics models. *Archive ouverte HAL*.

DAVIES, N. G., KUCHARSKI, A. J., EGGO, R. M., GIMMA, A., EDMUNDS, W. J. and CENTRE FOR THE MATHEMATICAL MODELLING OF INFECTIOUS DISEASES COVID-19 WORKING GROUP (2020). Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study. *The Lancet Public Health* **5** e375–e385.

DAVIES, N. G., ABBOTT, S., BARNARD, R. C., JARVIS, C. I., KUCHARSKI, A. J., MUNDAY, J. D., PEARSON, C. A. B., RUSSELL, T. W., TULLY, D. C., WASHBURNE, A. D., WENSELEERS, T., GIMMA, A., WAITES, W., WONG, K. L. M., VAN ZANDVOORT, K., SILVERMAN, J. D., CMMID COVID-19 WORKING GROUP1, COVID-19 GENOMICS UK (COG-UK) CONSORTIUM, DIAZ-ORDAZ, K., KEOGH, R., EGGO, R. M., FUNK, S., JIT, M., ATKINS, K. E. and EDMUNDS, W. J. (2021). Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science* **372** eabg3055.

DEHNING, J., ZIERENBERG, J., SPITZNER, F. P., WIBRAL, M., NETO, J. P., WILCZEK, M. and PRIESEMANN, V. (2020). Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science* **369**.

DELFRAISSY, J. F., ATLANI DUAULT, L., BENAMOUZIG, D., BOUADMA, L., CAUCHEMEZ, S., CHAUVIN, F., FONTANET, A., HOANG, A., MALVY, D. and YAZDANPANAH, Y. (2020). Une deuxième vague entrainant une situation sanitaire critique. In *Note du Conseil scientifique COVID-19*.

DI DOMENICO, L., PULLANO, G., SABBATINI, C. E., BOËLLE, P.-Y. and COLIZZA, V. (2020). Impact of lockdown on COVID-19 epidemic in Île-de-France and possible exit strategies. *BioMed Central Medicine* **18** 1–13.

DOCHERTY, A. B., HARRISON, E. M., GREEN, C. A., HARDWICK, H. E., PIUS, R., NORMAN, L., HOLDEN, K. A., READ, J. M., DONDELINGER, F., CARSON, G., MERSON, L., LEE, J., PLOTKIN, D., SIGFRID, L., HALPIN, S., JACKSON, C., GAMBLE, C., HORBY, P. W., NGUYEN-VAN-TAM, J. S., HO, A., RUSSELL, C. D., DUNNING, J., OPENSHAW, P. J., BAILLIE, J. K. and SEMPLE, M. G. (2020). Features of 20 133 UK patients in hospital with COVID-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. *British Medical Journal* **369** m1985.

FLAXMAN, S., MISHRA, S., GANDY, A., UNWIN, H. J. T., COUPLAND, H., MELLAN, T. A., WHITTAKER, C., ZHU, H., BERAH, T., EATON, J. W., MONOD, M., IMPERIAL COLLEGE COVID-19 RESPONSE TEAM, GHANI, A. C., DONNELLY, C. A., RILEY, S., VOLLMER, M. A. C., FERGUSON, N. M., OKELL, L. C. and BHATT, S. (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* **584** 257–261.

SANTÉ PUBLIQUE FRANCE COVID-19 – Point épidémiologique hebdomadaire du 28 Janvier 2021.

SANTÉ PUBLIQUE FRANCE COVID-19 – Point épidémiologique hebdomadaire du 11 Mars 2021.

SANTÉ PUBLIQUE FRANCE COVID-19 – Point épidémiologique hebdomadaire du 32 Décembre 2020.

GRAHAM, M. S., SUDRE, C. H., MAY, A., ANTONELLI, M., MURRAY, B., VARSAVSKY, T., KLÄSER, K., CANAS, L. S., MOLTENI, E., MODAT, M., DREW, D. A., NGUYEN, L. H., POLIDORI, L., SELVACHANDRAN, S., HU, C., CAPDEVILA, J. C. H., COVID-19 GENOMICS UK (COG-UK) CONSORTIUM, HAMMERS, A., CHAN, A. T., WOLF, J., SPECTOR, T. D., STEVES, C. J. and OURSELIN, S. (2021). Changes in Symptomatology, Reinfection, and Transmissibility Associated with the SARS-CoV-2 Variant B.1.1.7: An Ecological Study. *The Lancet Public Health* **6** e335-e345.

HALE, T., ANGRIST, N., GOLDSZMIDT, R., KIRA, B., PETHERICK, A., PHILLIPS, T., WEBSTER, S., CAMERON-BLAKE, E., HALLAS, L., MAJUMDAR, S. and TATLOW, H. (2021). A Global Panel Database of Pandemic Policies (Oxford COVID-19 Government Response Tracker). *Nature Human Behaviour* **5** 529–538.

HAUG, N., GEYRHOFER, L., LONDEI, A., DERVIC, E., DESVARS-LARRIVE, A., LORETO, V., PINIOR, B., THURNER, S. and KLIMEK, P. (2020). Ranking the effectiveness of worldwide COVID-19 government interventions. *Nature Human Behaviour* **4** 1303–1312.

HE, J., GUO, Y., MAO, R. and ZHANG, J. (2021). Proportion of asymptomatic coronavirus disease 2019: A systematic review and meta-analysis. *Journal of Medical Virology* **93** 820–830.

HOZÉ, N., PAIREAU, J., LAPIDUS, N., KIEM, C. T., SALJE, H., SEVERI, G., TOUVIER, M., ZINS, M., DE LAMBALLERIE, X., LÉVY-BRUHL, D., CARRAT, F. and CAUCHEMEZ, S. (2021). Monitoring the proportion of the population infected by SARS-CoV-2 using age-stratified hospitalisation and serological data: a modelling study. *The Lancet Public Health* **6** e408–e415.

HSIANG, S., ALLEN, D., ANNAN-PHAN, S., BELL, K., BOLLIGER, I., CHONG, T., DRUCKENMILLER, H., HUANG, L. Y., HULTGREN, A., KRASOVICH, E. et al. (2020). The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature* **584** 262–267.

HU, B., QIU, W., XU, C. and WANG, J. (2020). Integration of a Kalman filter in the geographically weighted regression for modeling the transmission of hand, foot and mouth disease. *BioMed Central Public Health* **20** 479.

ISLAM, S., HOQUE, E. and AMIN, M. R. (2020). Integration of Kalman filter in the epidemiological model: a robust approach to predict COVID-19 outbreak in Bangladesh. *medRxiv*.

ISLAM, N., SHARP, S. J., CHOWELL, G., SHABNAM, S., KAWACHI, I., LACEY, B., MASSARO, J. M., D'AGOSTINO, R. B. and WHITE, M. (2020). Physical distancing interventions and incidence of coronavirus disease 2019: natural experiment in 149 countries. *British Medical Journal* **370** m2743.

JONES, T. C., BIELE, G., MÜHLEMANN, B., VEITH, T., SCHNEIDER, J., BEHEIM-SCHWARZBACH, J., BLEICKER, T., TESCH, J., SCHMIDT, M. L., SANDER, L. E., KURTH, F., MENZEL, P., SCHWARZER, R., ZUCHOWSKI, M., HOFMANN, J., KRUMBHOLZ, A., STEIN, A., EDELMANN, A., CORMAN, V. M. and DROSTEN, C. (2021). Estimating infectiousness throughout SARS-CoV-2 infection course. *Science*.

JULIER, S. J. and UHLMANN, J. K. (1997). A new extension of the Kalman filter to nonlinear systems. In *Signal Processing, Sensor Fusion, and Target Recognition VI* **3068** 182–193.

JULIER, S. J. and UHLMANN, J. K. (2002). Reduced Sigma Point Filters for the Propagation of Means and Covariances Through Nonlinear Transformations. In *In Proceedings of the 2002 American Control Conference* **2** 887–892.

KRAEMER, M. U. G., YANG, C.-H., GUTIERREZ, B., WU, C.-H., KLEIN, B., PIGOTT, D. M., DU PLESSIS, L., FARIA, N. R., LI, R., HANAGE, W. P., BROWNSTEIN, J. S., LAYAN, M., VESPIGNANI, A., TIAN, H., DYE, C., PYBUS, O. G. and SCARPINO, S. V. (2020). The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* **368** 493–497.

LAPIDUS, N., PAIREAU, J., LEVY-BRUHL, D., DE LAMBALLERIE, X., SEVERI, G., TOUVIER, M., ZINS, M., CAUCHEMEZ, S., CARRAT, F. and SAPRIS-SERO STUDY GROUP (2021). Do not neglect SARS-CoV-2 hospitalization and fatality risks in the middle-aged adult population. *Infectious Diseases Now* **51** 380–382.

LAUER, S. A., GRANTZ, K. H., BI, Q., JONES, F. K., ZHENG, Q., MEREDITH, H. R., AZMAN, A. S., REICH, N. G. and LESSLER, J. (2020). The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Annals of Internal Medicine* **172** 577–582.

LAVIELLE, M. (2014). *Mixed effects models for the population approach: models, tasks, methods and tools*. CRC press.

LI, R., PEI, S., CHEN, B., SONG, Y., ZHANG, T., YANG, W. and SHAMAN, J. (2020). Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). *Science* **368** 489–493.

LI, Y., CAMPBELL, H., KULKARNI, D., HARPUR, A., NUNDY, M., WANG, X., NAIR, H. and THE USHER NETWORK FOR COVID-19 EVIDENCE REVIEWS (UNCOVER) GROUP (2021). The temporal association of introducing and lifting non-pharmaceutical interventions with the time-varying reproduction number (R) of SARS-CoV-2: a modelling study across 131 countries. *The Lancet Infectious Diseases* **21** 193–202.

LIU, Y., MORGENSTERN, C., KELLY, J., LOWE, R., CMMID COVID-19 WORKING GROUP and JIT, M. (2021a). The impact of non-pharmaceutical interventions on SARS-CoV-2 transmission across 130 countries and territories. *BioMed Central Medicine* **19** 1–12.

LIU, X., HUANG, J., LI, C., ZHAO, Y., WANG, D., HUANG, Z. and YANG, K. (2021b). The role of seasonality in the spread of COVID-19 pandemic. *Environmental Research* **195** 110874.

WORLD HEALTH ORGANIZATION COVID-19 Weekly Epidemiological Update – 25 February 2021.

PEI, S., KANDULA, S. and SHAMAN, J. (2020). Differential effects of intervention timing on COVID-19 spread in the United States. *Science advances* **6** eabd6370.

PRAGUE, M., WITTKOP, L., COLLIN, A., CLAIRON, Q., DUTARTRE, D., MOIREAU, P., THIEBAUT, R. and HEJBLUM, B. P. (2020). Population modeling of early COVID-19 epidemic dynamics in French regions and estimation of the lockdown impact on infection rate. *medRxiv*.

RODÓ, X., SAN-JOSÉ, A., KIRCHGATTER, K. and LÓPEZ, L. (2021). Changing climate and the COVID-19 pandemic: more than just heads or tails. *Nature Medicine* **27** 576–579.

ROUMAGNAC, A., DE CARVALHO FILHO, E., BERTRAND, R., BANCHEREAU, A.-K. and LAHACHE, G. (2021). Étude de l'influence potentielle de l'humidité et de la température dans la propagation de la pandémie COVID-19. *Médecine de Catastrophe-Urgences Collectives* **5** 87–102.

ROUX, J., MASSONNAUD, C. and CRÉPEY, P. (2020). COVID-19: One-month impact of the French lockdown on the epidemic burden. *medRxiv*.

SALJE, H., TRAN KIEM, C., LEFRANCQ, N., COURTEJOIE, N., BOSETTI, P., PAIREAU, J., ANDRONICO, A., HOZÉ, N., RICHET, J., DUBOST, C.-L., LE STRAT, Y., LESSLER, J., LEVY-BRUHL, D., FONTANET, A., OPATOWSKI, L., BOËLLE, P.-Y. and CAUCHEMEZ, S. (2020). Estimating the burden of SARS-CoV-2 in France. *Science* **369** 208–211.

SIMON, D. (2006). *Optimal State Estimation: Kalman, H_∞, and Nonlinear Approaches*. Wiley–Blackwell.

VOLZ, E., MISHRA, S., CHAND, M., BARRETT, J. C., JOHNSON, R., GEIDELBERG, L., HINSLEY, W. R., LAYDON, D. J., DABRERA, G., O'TOOLE, Á., AMATO, R., RAGONNET-CRONIN, M., HARRISON, I., JACKSON, B., ARIANI, C. V., BOYD, O., LOMAN, N. J., MCCRONE, J. T., GONÇALVES, S., JORGENSEN, D., MYERS, R., HILL, V., JACKSON, D. K., GAYTHORPE, K., GROVES, N., SILLITOE, J., KWIATKOWSKI, D. P., FLAXMAN, S., RATMANN, O., BHATT, S., HOPKINS, S., GANDY, A., RAMBAUT, A. and FERGUSON, N. M. (2021). Transmission of SARS-CoV-2 Lineage B. 1.1. 7 in England: Insights from linking epidemiological and genetic data. *medRxiv* 1–12.

WANG, C., LIU, L., HAO, X., GUO, H., WANG, Q., HUANG, J., HE, N., YU, H., LIN, X., PAN, A., WEI, S. and WU, T. (2020). Evolving Epidemiology and Impact of Non-pharmaceutical Interventions on the Outbreak of Coronavirus Disease 2019 in Wuhan, China. *medRxiv*.

WARSZAWSKI, J., BAJOS, N., MEYER, L., DE LAMBALLERIE, X. and ET AL. (2020). ER 1167: In May 2020, 4.5% of the population 2020 of metropolitan France had developed antibodies against SARS-CoV-2 – the first results of the EpiCov national survey Technical Report, DREES - INSERM.

WASHINGTON, N. L., GANGAVARAPU, K., ZELLER, M., BOLZE, A., CIRULLI, E. T., BARRETT, K. M. S., LARSEN, B. B., ANDERSON, C., WHITE, S., CASSENS, T., JACOBS, S., LEVAN, G., NGUYEN, J., RAMIREZ, J. M., RIVERA-GARCIA, C., SANDOVAL, E., WANG, X., WONG, D., SPENCER, E., ROBLES-SIKISAKA, R., KURZBAN, E., HUGHES, L. D., DENG, X., WANG, C., SERVELLITA, V., VALENTINE, H., DE HOFF, P., SEAVER, P., SATHE, S., GIETZEN, K., SICKLER, B., ANTICO, J., HOON, K., LIU, J., HARDING, A., BAKHTAR, O., BASLER, T., AUSTIN, B., MACCANNELL, D., ISAKSSON, M., FEBBO, P. G., BECKER, D., LAURENT, M., MCDONALD, E., YEO, G. W., KNIGHT, R., LAURENT, L. C., DE FEO, E., WOROBEY, M., CHIU, C. Y., SUCHARD, M. A., LU, J. T., LEE, W. and ANDERSEN, K. G. (2021). Emergence and rapid transmission of SARS-CoV-2 B. 1.1. 7 in the United States. *Cell* **184** 2587-2594.e7.

WU, Z. and MCGOOGAN, J. M. (2020). Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72,314 Cases From the Chinese Center for Disease Control and Prevention. *Journal of the American Medical Association* **323** 1239–1242.

WU, P., LIU, F., CHANG, Z., LIN, Y., REN, M., ZHENG, C., LI, Y., PENG, Z., QIN, Y., YU, J., GENG, M., YANG, X., ZHAO, H., LI, Z., ZHOU, S., RAN, L., COWLING, B. J., LAI, S., CHEN, Q., WANG, L., TSANG, T. K. and LI, Z. (2021). Assessing asymptomatic, pre-symptomatic and symptomatic transmission risk of SARS-CoV-2. *Clinical Infectious Diseases*.

French government's action summary website. https://www.gouvernement.fr/info-coronavirus/les-actions-du-gouvernement.

Données de laboratoires pour le dépistage : Indicateurs sur les variants (SI-DEP). https://www.data.gouv.fr/en/datasets/.

Données relatives aux personnes vaccinées contre la COVID-19. https://www.data.gouv.fr/fr/datasets/donnees-relatives-aux-personnes-vaccinees-contre-la-covid-19-1/.

Déchiffrer le monde du travail pour éclaire le débat public (DARES) - Principaux résultats des derniers mois - avril 2021. https://dares.travail-emploi.gouv.fr/publication/activite-et-conditions-demploi-de-la-main-doeuvre-pendant-la-crise-sanitaire-covid-19-1.

ZAITCHIK, B. F., SWEIJD, N., SHUMAKE-GUILLEMOT, J., MORSE, A., GORDON, C., MARTY, A., TRTANJ, J., LUTERBACHER, J., BOTAI, J., BEHERA, S. et al. (2020). A framework for research linking weather, climate and COVID-19. *Nature Communications* **11** 1–3.