Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Endemic–epidemic models to understand COVID-19 spatio-temporal evolution

Alessandro Celani\textsuperscript{a,}\textsuperscript{*}, Paolo Giudici\textsuperscript{b}

\textsuperscript{a} Dipartimento di Scienze Economiche e Sociali, Polytechnic University of Marche, Piazzale Raffaele Martelli 8, 60121 Ancona, Italy
\textsuperscript{b} Dipartimento di Scienze Economiche e Aziendali, University of Pavia, Via San Felice al Monastero 5, 27100 Pavia, Italy

\textbf{A R T I C L E   I N F O}

\textbf{Article history:}
Received 11 February 2021
Received in revised form 30 June 2021
Accepted 5 July 2021
Available online 12 July 2021

\textbf{Keywords:}
Contagion models
Multivariate statistics
COVID-19
Poisson processes
Spatio-temporal models

\textbf{A B S T R A C T}

We propose an endemic–epidemic model: a negative binomial space–time autoregression, which can be employed to monitor the contagion dynamics of the COVID-19 pandemic, both in time and in space. The model is exemplified through an empirical analysis of the provinces of northern Italy, heavily affected by the pandemic and characterized by similar non-pharmaceutical policy interventions.

© 2021 Elsevier B.V. All rights reserved.

1. Background and literature

As of 31 December 2020, over 84 million cases of Coronavirus disease (COVID-19) have been confirmed globally across 174 countries and regions. Continuous human-to-human transmission has been asserted in multiple countries and, in response, governments have enforced several non-pharmaceutical interventions (NPIs) to reduce transmission of SARS-CoV-2, in the absence of a COVID-19 vaccine. As NPIs imply a high socio-economic cost, most governments have differentiated policy measures at the regional level.

To understand the effectiveness of national policies, it is very important to understand the time dynamic of the contagion. To further understand the differential effectiveness of policies at the regional level, and the need to lift/apply restrictions only in some regions, it is also essential to understand the spatial diffusion of the contagion.

* Corresponding author.

E-mail addresses: aecelani94@gmail.com (A. Celani), paolo.giudici@unipv.it (P. Giudici).

https://doi.org/10.1016/j.spasta.2021.100528
2211-6753/© 2021 Elsevier B.V. All rights reserved.
Against this backdrop, this paper aims to contribute to the literature that links policy impacts to the epidemiological effects of the COVID-19 pandemic by means of a spatio-temporal approach. Most of the existing research literature, increased considerably in the last year, typically focuses on a country level analysis to assess the impact of timing and type of national policies undertaken to mitigate COVID-19 health outcomes, such as infected cases, severe hospitalized cases and deaths.

Among the studies that use statistical frameworks to model and predict the COVID-19 transmission, exponential growth models and time series based models are the most widespread. Within the first strand, Wu et al. (2020) apply the generalized version of the growth model, the logistic growth model, in 29 provinces of China and in 33 countries, to understand both the development of the epidemics and the impact of the very first drastic control measures implemented. A logistic model has also been used by Hermanowicz (2020) on the confirmed cases in mainland China to estimate and predict the maximum cases in a quasi-real time fashion. Within time series models, Ceylan (2020) uses different specifications of ARIMA models to predict the epidemiological trend of COVID-19 prevalence in Italy, Spain, and France. Perone (2020) analysed three different time series models (ARIMA, NNAR and ETS) and their hybrid combinations to forecast both the number of patients hospitalized with mild symptoms and the intensive care units in Italy in the second wave of the epidemics. A multivariate analysis on the Brazilian COVID-19 confirmed cases in some selected cities has been carried out by Silva et al. (2020); they estimate the transmission channels of every pair of cities using a regularized VAR approach.

Nevertheless, only a limited part of the previous studies allow for spatial dependence in their models, and even fewer are those that use spatial dependence models for count data, which characterize this type of epidemiological studies.

Among them, we follow the endemic–epidemic mixed-effects generalized linear model for areal disease counts originally developed by Held et al. (2005), and applied to the analysis of confirmed cases in several countries.

Among such applications, Giuliani et al. (2020) apply the model to Italy, with confirmed cases at the provincial level. The authors find strong evidence that homogeneous restrictions were necessary in the first phase, but they could have been more effective if the specificity of the local regions had been taken into account. The obtained results shed further light on the understanding the evolution of the COVID-19 pandemic in Italy, analysed recently in a series of papers, among which are Loro et al. (2021), Bartolucci and Farcomeni (2021), Farcomeni et al. (2021), Girand et al. (2020) and Agosto et al. (2021).

The endemic–epidemic model has also been applied by Fronterre et al. (2020) to investigate the dynamics of COVID-19 positive testing in England at Lower Tier Local Authorities level (LTA). They extend the formulation of Giuliani et al. (2020) using covariates such as data on inter LTA connectivity in terms of daily commuters and data on the reduction of car traffic due to restriction measures. The study shows that the COVID-19 epidemic exhibits high variability among the considered regions, which is discarded when looking only at national aggregate data.

A related paper is Berlemann and Haustein (2020) which relaxes the hypothesis of autoregressive dependence of order one in Held et al. (2005), following Bracher and Held (2020), but without using covariates. Their study focuses on selected German federal states, such as Bavaria and Saxony. The main evidence is that the second and third waves of containment measures were not essential to prevent problems in the German health system. Moreover, they find that although common containment policy turned out to be helpful to flatten the curve of new cases, this came at a too high price, unnecessarily increasing the total cost of containment.

Another recent study in this field is Ssentongo et al. (2020), in which a macro-analysis of the African continent has been carried out using data on most of the states. The formulation of Bracher and Held (2020) is extended introducing covariates such as country’s testing capacity, social policies, landlock status as well as metereological variables such as temperature and humidity, which revealed their importance in past epidemics. The authors find a good predictive accuracy of the model in terms of short term prediction of the infections, even if such accuracy heavily depends on the quality of testing and data reporting.

We extend the previous stream of literature from two main viewpoints. On one hand, we extend the literature that models the dependence of disease counts on policy measures by extending the
Poisson autoregressive model for disease counts proposed by Agosto and Giudici (2020) with a Negative Binomial distribution, that allows for over dispersion, as in Held et al. (2005) and the related papers. On the other hand, we extend the approach in Held et al. (2005) with a more general spatial dependence structure, that replaces neighbouring effects with a dependence based on the actual inter regional mobility. This allows to test for time-lagged local effects of policies on daily COVID-19 counts.

In terms of mobility estimation, our work is related to Cornelia et al. (2020), which develops a simple statistical model involving minimal assumption about disease dynamics (hence without explicit assuming an epidemiological mathematical framework) to measure the effectiveness of NPIs and forecast the spread of COVID-19 at local, state and national level. They find important gains in predictive accuracy, when public human mobility data (provided by Google, Facebook and Badoo, for example) is added to predict disease counts.

The rest of this paper is organized as follows. Section 2 presents the proposed methodology; Section 3 the considered data and the main empirical findings along with a forecast comparison with another count disease model. Section 4 summarizes the conclusions of the paper, and suggests future research directions.

2. Model and inference

2.1. Model

In the paper we assume, without loss of generality, that policies to mitigate COVID-19 effects are decided at the regional level, and that regions can be further subdivided into smaller (third level) subareas: provinces.

Let \( \{Y_{p,t}\}_{t \in \mathbb{Z}} \) denote the number of new confirmed cases in the provinces \( p = 1, \ldots, P \) of a given region, and on day \( t \). The endemic–epidemic multivariate time-series model for infectious disease proposed by Held et al. (2005) assumes that \( Y_{p,t} \mid F_{t-1} \sim NB(\mu_{p,t}, \psi_{p}) \), conditional on the history of the process up to time \( t - D \), is distributed as a negative binomial with mean:

\[
\mu_{p,t} = \lambda_{p,t} \sum_{d=1}^{D} u_d Y_{p,t-d} + \phi_{p,t} \sum_{d=1}^{D} \sum_{i \neq p} u_d \omega_{i,p} Y_{i,t-d} + \nu_t
\]

(1)

and over dispersion parameter \( \psi_p > 0 \) so that the conditional variance of \( Y_{p,t} \) is \( \mu_{p,t}(1 + \psi_p \mu_{p,t}) \).

The negative binomial model is used to allow for possible over dispersion due to under-reporting or unobserved covariates that affect the disease incidence. We remark that in the limiting case of \( \psi_p = 0 \) the negative binomial reduces to a Poisson distribution, for which the conditional mean and variance are identical. We assume that the over dispersion parameter is identical across provinces \( (\psi_p = \psi) \), which may be a realistic assumption within the same region.

The three terms on the right-hand side of Eq. (1) correspond to the three components of the model: the epidemic-within provinces, the epidemic-between provinces and the endemic.

The first term in (1) is the epidemic-within component. We propose to define it as an autoregression of the number of cases in province \( p \) on a weighted sum of the past number of cases up to day \( t - D \), in which the autoregressive parameter is province-specific and depends on an intercept and on four covariates: the population density \( (DE) \), the Stringency Index \( (SI) \),\(^1\) the Testing Policy \( (TEST) \) index\(^2\) (both indicators are available, at the time of writing, only at the national level, and hence are supposed to be constant across all the units), and an indicator for weekend days \( (W) \).

More precisely, we assume that the parameter \( \lambda_p \) is determined by the following linear equation:

\[
\log(\lambda_{p,t}) = \alpha_p \lambda_{p,t} + \beta_1^{(DE)}DE_p + \beta_2^{(SI)}SI_{t-\tau} + \beta_3^{(TEST)}TEST_{t-\tau} + \beta_4^{(W)}W_t
\]

(2)

\(^1\) The Stringency Index is a real valued variable that ranges from 0 to 100 representing the level of restriction imposed at a national level. It is calculated as a linear combination of eight indices representing containment measures and an indicator which records public information campaigns.

\(^2\) The Testing Policy Index is an ordinal variable indicating who was access to testing. It ranges from 0 (no tests) to 3 (open public screening).
where $\alpha_p^{(i)}$ is a random effect parameter, allowed to vary across provinces; $\beta_1^{(i)}, \ldots, \beta_4^{(i)}$ are the four regression parameters associated with the abovementioned covariates. As the effects of containment and testing policies have at least two weeks of delay on new infections we assume, without loss of generality, $\tau = 14$.

To complete the specification of the epidemic-within component, we assume that the weights $u_d$ are chosen according to a Poisson distribution, with a mean parameter $(\gamma)$ that will be chosen minimizing the Akaike Information Criterion (AIC). In this way the larger weight is not necessarily the first, and the mean parameter corresponds to the average waiting time between the infection appearing in an individual and in a next one, as described in Nishiura et al. (2020) and Bracher and Held (2020). In formuale:

$$u_d = \frac{\gamma^{d-1}}{(d-1)!} \exp(-\gamma), \quad \gamma > 0, \quad (3)$$

for $d = 1, \ldots, D$. We will choose the best value of $D$ by scoring rules minimization (Czado et al., 2009).

The second term in Eq. (1) is the epidemic-between component, suggested by Held et al. (2005) to explain the spread of a disease across different locations. We propose to define it as a regression of the number of cases in province $p$ on a weighted sum of the past number of cases up to time $t-D$ in other provinces ($i \neq p$), in which the autoregressive parameter is province-specific and depends on four covariates: the province population $p$ (POP) and, as in the epidemic-within component: the Stringency Index (SI), the Testing Policy (TEST) index and the indicator for weekend days (W).

More precisely, we assume that the parameter $\phi_p$ is specified by the following linear equation:

$$\log(\phi_{p,t}) = \alpha_p^{(0)} + \beta_1^{(0)} \text{POP}_p + \beta_2^{(0)} \text{SI}_{t-\tau} + \beta_3^{(0)} \text{TEST}_{t-\tau} + \beta_4^{(0)} W_t \quad (4)$$

where $\alpha_p^{(0)}$ is a random effect intercept and $\beta_1^{(0)}, \ldots, \beta_4^{(0)}$ are the four regression parameters.

The specification of the between-epidemic component is completed by the system of weights, which has two factors. While $u_d$ are the same as for the epidemic-within component, $\omega_{i,p}$ are chosen to account for the spatial distance between provinces. Precisely, we assume that $\omega_{i,p}$ is specified as a function of a variable $o_{i,p}$, obtained by discretizing the commuting flows between pairs of provinces. In more detail, we consider the inbound and outbound number of commuters for both study and work reasons, and categorize the resulting value into a variable with six increasing levels: the larger the number of commuters, the lower the level.\(^3\)

We believe that the resulting variable is able to explain the concept of “closeness” between two provinces better than a simple neighbouring degree, as the virus transmission depends more on the movement of people rather than on the actual physical distance between two provinces. We also believe that the spatial weights could represent a structural component with short-term variations but stable over time, so that we can safely use official census data, updated every decade. In addition, the short-term variations are captured by the covariates present in the model, such as the stringency index, which reflects all the various restrictions imposed by central and local governments.

Mathematically, the dependence of $\omega_{i,p}$ on $o_{i,p}$ can be specified by a power law function (Meyer and Held, 2014):

$$\omega_{i,p} = o_{i,p}^{-d},$$

where $d \geq 0$ is a decay parameter. The greater the $d$, the faster the decaying of the power law function and, therefore, the less important is the difference between different levels of $o_{i,p}$.

The last term in (1) is the endemic component: it includes terms that describe the seasonality of the difference between regions. We follow Held and Paul (2012) who proposed to model seasonal variations of infectious disease through an harmonic wave, assuming that:

$$\log(v_{p,t}) = \alpha_p^{(v)} + \gamma \sin \left(\frac{2\pi}{365} t\right) + \theta \cos \left(\frac{2\pi}{365} t\right) + \log(\text{POP}_p) \quad (5)$$

where $\alpha_p^{(v)}$ is a province varying intercept and $2\pi/365$ are Fourier frequencies for daily data.

\(^3\) To obtain the range of each class we equally split the range of the series of daily commuters for pairs of provinces (in logarithm) into six levels and then we allocate each pair to the class to which it belongs.
To simplify the notation, note that the seasonal terms can also be expressed as
\[ y \sin \left( \frac{2\pi}{365} t \right) + \theta \cos \left( \frac{2\pi}{365} t \right) = A \sin \left( \frac{2\pi}{365} t + \rho \right) \]  
with \( A = \sqrt{\gamma^2 + \theta^2} \) describing the amplitude of the seasonal component and \( \rho = \arctan(\delta/\theta) \) describing the phase shift of each sinusoidal wave.

What remains is the specification of the random effects appearing in the epidemic components. Paul and Held (2011) suggest to allow for a diversified time origin between provinces. This can be implemented assuming that the two intercepts in Eqs. (2)–(4) are random effects which follow a bidimensional Gaussian distribution:
\[ \xi = \left[ \alpha^{(\lambda)}_p, \alpha^{(\phi)}_p \right]^\top \sim N_2(\xi_0, \Sigma), \]  
where \( \xi_0 = \left[ \alpha^{(\lambda)}_0, \alpha^{(\phi)}_0 \right]^\top \) and \( \Sigma = \text{diag}(\sigma^2_\lambda, \sigma^2_\phi) \) is a diagonal variance covariance matrix. Indeed a diversification between provinces, especially in the first stage of the epidemics is not trivial; allowing for province-specific random effects is expected to improve the model fit.

2.2. Inference

Our assumed model is characterized by three structural components, plus two random ones, which depend on a number of parameters that are to be estimated from the data. For the sake of clarity, we summarize in Fig. 1 all model parameters and their functional dependences.

Without random effects, the maximum likelihood (ML) estimates of the model (1) can be obtained by directly maximizing the negative binomial log-likelihood using numerical optimization routines (see Paul et al., 2008 for details). By stacking together all the model parameters in 1 with the exception for the overdispersion one in the vector \( \Theta \) we can express the global log-likelihood as
\[ \log L(\Theta) = \sum_{p,t} \log L_{p,t}(\hat{\Theta}, \psi) \]  

where $\Theta = [\bar{\Theta}, \psi]^\top$ and

$$
\log L_{p,t}(\bar{\Theta}, \psi) \propto \log \Gamma\left(Y_{p,t} + \frac{1}{\psi}\right) - \log \Gamma\left(\frac{1}{\psi}\right) + \frac{1}{\psi} \log \frac{1}{1 + \psi} \mu_{p,t}(\bar{\Theta}) + Y_{p,t} \log \left(\frac{\psi \mu_{p,t}(\bar{\Theta})}{1 + \psi \mu_{p,t}(\bar{\Theta})}\right)
$$

(9)

is the negative binomial log-likelihood contribution of the observation $Y_{p,t}$ and $\Gamma(\cdot)$ is the Gamma function.

If there are random effects, penalized likelihood approaches can be carried out to obtain parameter estimates (see Paul and Held, 2011). That is, variance components embedded in $\Sigma$ are treated as fixed when estimating the fixed effects. After integration with respect to the parameters in $\Theta$ of the penalized log-likelihood we obtain the marginal likelihood that can be maximized to obtain variance parameters.

Inference for the regression parameters $\Theta$, given known variance components, is based on the penalized log-likelihood

$$
\log L_{\text{pen}}(\Theta | \Sigma) = \log L(\Theta) + \log p(\xi | \Sigma)
$$

(10)

where $\log L(\Theta)$ has been defined in (8)–(9) and the penalty term $\log p(\xi | \Sigma)$ is the log distribution of the vector of random effects, a bivariate Gaussian, that can be expressed as

$$
\log p(\xi | \Sigma) = \propto \frac{1}{2} \xi^\top \Sigma^{-1} \xi
$$

(11)

As mentioned before, variance parameters are obtained by maximizing the following marginal likelihood:

$$
L_{\text{marg}}(\Sigma) = \int \exp[\log L_{\text{pen}}(\Theta | \Sigma)] d\Theta.
$$

(12)

In practice, the following algorithm is used to estimate all the parameters

1. Update regression coefficient given the current variance ones.
2. Update variance components given current regression coefficients.
3. Repeat the first two steps until convergence is reached.

We remark that the above estimation procedure is implemented in the R packages surveillan ce (Meyer et al., 2017) and hhh4addon,4 which we have used to fit our model and to produce the results of this paper.

3. Data and results

The data used for this study are the confirmed cases of COVID-19 in the forty provinces of five regions of northern Italy: Piedmont, Lombardy, Liguria, Veneto and Emilia-Romagna. All the above mentioned regions belong to the same country and, therefore, have a similar cultural, institutional and legal environment. However, in Italy, health policies are decided at the regional level and, therefore, the actions taken to tackle the COVID-19 pandemic, and especially the operational ones, may be different in different regions.

The decision to take into consideration only the above mentioned regions comes from the fact that northern Italy was the first area affected by the virus, and the most hit overall since then. As of 21/12/2020, the last data of the sample, the cumulative confirmed cases of the five regions considered (1 144 718) were almost 60% of the total cases of the whole country (1 964 000) (Presidenza del Consiglio dei Ministri, 2020; Dong et al., 2020).

Data is available from the first day of recording (25/02/2020) until 21/12/2020, for a total of 301 days. It is downloaded from the official GitHub repository of the Italian Civil Protection

---

4 Available at: https://github.com/jbracher/hhh4addon.
Table 1
COVID-19 daily incidence (DI) (total number of infections per 1000 inhabitants) and total population in millions (Pop) for the forty provinces studied.

|     | DI   | Pop |     | DI   | Pop |     | DI   | Pop |
|-----|------|-----|-----|------|-----|-----|------|-----|
| TO  | 46.51| 2.25| SO  | 47.46| 0.18| IM  | 32.73| 0.21|
| VC  | 47.21| 0.17| MI  | 53.05| 3.28| SV  | 28.92| 0.27|
| NO  | 42.10| 0.37| BG  | 25.31| 1.12| GE  | 42.62| 0.84|
| CN  | 46.84| 0.59| BS  | 33.08| 1.27| SP  | 37.95| 0.22|
| AT  | 46.96| 0.21| PV  | 45.07| 0.55| VR  | 53.40| 0.93|
| AL  | 42.12| 0.42| CR  | 40.40| 0.36| VI  | 50.00| 0.86|
| BI  | 39.72| 0.17| MN  | 39.52| 0.41| BL  | 72.48| 1.02|
| VB  | 45.97| 0.16| LC  | 39.41| 0.34| TV  | 55.90| 0.89|
| VA  | 58.02| 0.89| LO  | 46.12| 0.23| VE  | 45.64| 0.85|
| CO  | 56.14| 0.60| MB  | 54.93| 0.88| PD  | 49.36| 0.94|

Department (Presidenza del Consiglio dei Ministri, 2020). Data about the population size and density of provinces and on the number of commuters between pairs of provinces are taken from the Italian Statistical Institute.\(^5\) Data about the stringency index and the testing policy indicator are obtained from The Oxford Coronavirus Government Tracker (OxCGRT).\(^6\)

Table 1 presents, for all the provinces, two statistics: the mean daily incidence of the pandemic and the population.

Panel (a) of Fig. 2 describes the evolution of the new infections in the regions considered for the whole sample. Panel (b) of Fig. 2 instead represents the weighted undirected graphs obtained from the adjacency matrix with entries \(o_{i,p}\). In each graph, the length of each edge is inversely proportional to the number of commuters.

Panel (a) of Fig. 2 clearly shows the two phases of the epidemics. Panel (b) emphasizes the centrality role of some provinces (those are located in the centre of the graph), such as Milan (MI), Bergamo (BG) and Cremona (CR), and the more peripheral role of others (those located on the border) such as Sondrio (SO), Ravenna (RN) and Belluno (BL).

The best autoregressive specification is found by performing a lag selection for different values of the autoregressive parameter \(D\) (from the first up to the tenth lag) and for the model with and without random intercepts on both the epidemic components. For each specification, the optimal value of the parameters is obtained using the training set; then an out of sample forecast is performed on the test set. The best model is the one with the lowest Ranked probability (RP) score, a proper scoring rules for count data (Czado et al., 2009), defined as:

\[
RP(P, Y_{p,t}) = \sum_{k=0}^{\infty} \{P_k - 1(Y_{p,t} \leq k)\}
\]

where \(1(\cdot)\) is the indicator function. As can be seen in Panel (a) of Table 2, the RP score is minimized by the AR(7) specification with random intercepts.

The estimated coefficients of the selected AR(7) model are stored in Panel (b) of Table 2 whereas the estimated autoregressive weights \(u_d\) are shown in Panel a of Fig. 3.

From Table 2 note first that the hypothesis that a Poisson model is suited for the data is strongly rejected, given the statistical significance of the parameter \(\psi\): the more general and dispersed negative binomial better fits the observed data.

Overall, from Table 2 it emerges that the endemic portion of the process is strong and statistically robust in all its parts, as shown by the parameters \(A\), \(\rho\) and the intercept \(\alpha^{(\nu)}\). This denotes that the seasonal component is very strong, reflecting the intuition that climatic conditions are very important. COVID-19 like many other viruses (such as flu) tends to manifest its effect especially in the colder months and therefore with the arrival of winter, also because in these periods people

\(^5\) Available at: https://www4.istat.it/it/archivio/139381.
\(^6\) Available at: https://github.com/OxCGRT.
aggregation takes place especially in closed places, facilitating the contagion spread. This is the main reason why we did not observe a significant increase in infections in summer, but only since October, and in general after the reopening of the schools.

The epidemic potential of the process can be analysed by looking at the components embedded in $\lambda$ and $\phi$. Among the demographic variables population density and population, the latter has a
significative positive effect. The larger the province, in terms of residents, the more likely it attracts people, according to a gravity effect, and therefore the more likely it has a high infection spread.

The effect of the population size may be mitigated by the effect of policy containment measures that limit social mobility, measured by the stringency index parameter \( \beta_2 \). Table 2 shows that this effect is statistically strong, revealing that the NPIs introduced since the beginning of the epidemic have had a positive effect in mitigating the COVID-19 infections.

On the other hand, the effect of the testing policy variable is statistically significant, even if only for the epidemic-within component. This evidence confirms that the lower number of infections observed during the first wave of the epidemic, compared to those of the second one, was influenced by the fact that number of available tests were not sufficient to cover number of vaccines that it was able to guarantee from the summer and the second wave in general.

The effect of the dummy weekend is also significant: the number of daily tests carried out on weekends is, on average.

A province specific analysis can be done graphically by looking at the Panel (b) of Fig. 3, where the fitted means are split into the three components for the twelve provinces most affected by the epidemics and at both Figs. 4–5, where the random intercepts of the epidemic components are plotted along with their 95% confidence interval.

Among the twelve most affected provinces there are the five regional capitals: Milan, Turin, Genoa, Venice and Bologna. Without considering the summer period, in which the number of cases has been relatively small compared to the rest of the sample, it is evident that the epidemic-between component assumed a different role in the first and second phases of pandemic. In the first one its impact was limited, mainly due to the general lockdown imposed by the Italian government. Since the reopening of the schools, and more generally with the arrival of the winter, the impact
of the epidemic-between started to rise: this is due to the fact that the restrictions imposed have been milder compared to those of the first phase and have affected the various units differently. The adoption of region varying policies is fundamental to manage well the COVID-19 emergency heterogeneity, given that the health emergency started to assume different characteristics in the various regions.

Note that the remaining share of the fitted means is explained by the epidemic-within component. This is consistent especially if we consider provinces with a high population and high population density, such as the capitals of the regions. Consider as example Milan (MI), which is the most affected province, mainly driven by the metropolitan city of Milan, the second-largest Italian city by population. With a dense transport network and a large number of internal daily commuters (over 1.5 millions), the evolution of contagion in this province is determined in a large portion by internal matters.
3.1. Model accuracy

To evaluate the goodness of fit of the model, we propose two measures to compare the estimates with the observed counts in the training set. The first one is the count data Pearson $R^2$, defined by Cameron and Windmeijer (1996), whereas the second one is the Spearman’s rank correlation coefficient (SRCC), similar to the Pearson’s correlation coefficient, but based on the ranked variables. For the sake of simplicity, both of them are plotted for the twelve most affected provinces in Fig. 6.

From Fig. 3 note that the variability of the data has been well captured by our model. Indeed, by averaging over all provinces, it can be shown that $R^2_{PE}$ is about (80%). In addition, SRCC is equal to 0.885 and this shows that, on average, the predicted and observed counts have a very similar rank.

To better assess the validity of the proposed model, we consider how it performs in predictions. Specifically, we use the data of the last 21 days of the sample considered (test set) to carry out a one day ahead forecasting exercise. For each day of the test set we estimate the model parameters using the data up to time $t$ and we use them to predict the number of new cases at $t + 1$ for each province. In order to evaluate the forecast accuracy of our model not only in absolute terms, but also in relation to a benchmark model belonging to the same class, we repeat the same exercise with
We recall that an integer valued random variable \( \{Y_t\}_{t \in \mathbb{Z}} \) follows an INGARCH(p,q) process if

\[
\begin{align*}
Y_t | \mathcal{F}_{t-1} &\sim \mathcal{P}(\lambda_t) \quad \forall t \in \mathbb{Z} \\
\lambda_t &= \gamma_0 + \sum_{i=1}^{q} \gamma_i Y_{t-i} + \sum_{j=1}^{p} \delta_j \lambda_{t-j}
\end{align*}
\]

where \( \gamma_i > 0 \forall i = 0, \ldots, q \) and \( \delta_j > 0 \forall j = 1, \ldots, p \) are parameters to be estimated, \( \mathcal{P}(\cdot) \) denotes the Poisson distribution and \( \mathcal{F}_t \) is the set containing the history of the process up to time \( t \). Unlike the HHH, the INGARCH is a univariate model, hence we must estimate the parameters for each series separately; the \( p \) and \( q \) lag orders are decided performing a lag selection and choosing the specification that minimizes the AIC among all the combinations up to lag \( p = q = 3 \).
Table 3

|       | $R^2_{PE}$ | SRCC = $r_s$ |
|-------|------------|--------------|
| HHH   | 0.868(0.196) | 0.244(0.3) |
| INGARCH | 0.699(0.34)  | 0.132(0.25) |

The predictive accuracy comparison between the two models is described in Table 3, where both the Pearson $R^2$ and the SRCC based on the out of sample predictions of the two models are presented (as an average over all the provinces).

From Table 3 note that both indicators agree on the superiority of the HHH forecasts with respect to the INGARCH. The $R^2$ of the HHH model is about 17 percentage points higher. Similarly, the SRCC is almost twice higher.

The evidence so far allows to conclude that the HHH model is more accurate than the INGARCH model.

To further compare the performance of the two models we use a Monte Carlo permutation test for paired individual scores, following Paul and Held (2011). As a test statistics, we can calculate the difference between the mean of the RP scores of the two models over the 21 days of the test period. Under the null hypothesis of no difference, the actually observed difference between mean scores should not be so different from the distribution of the test statistic under permutation. In Table 4, the results of the test are summarized.
The test results in 4 further confirm the superiority of HHH forecasts, as the INGARCH RP score is much higher. The difference also turns out to be statistically different from 0.

4. Conclusions

We have proposed a novel implementation of the endemic–epidemic model proposed by Held et al. (2005), to understand the time and space contagion dynamics of COVID-19. The model, applied to data from the provinces of northern Italy, quite satisfactorily identifies the most important factors that affect contagion spread: within province effects, between provinces effects, and seasonality.

The model could be extended, from an applied viewpoint, to other countries and, from a methodological viewpoint, within a Bayesian framework, following the approaches of Giudici et al. (2000) and Giudici et al. (2003).

Acknowledgement

This research has received funding from the European Union’s Horizon 2020 research and innovation programme “PERISCOPE: Pan European Response to the ImpactS of COVID-19 and future Pandemics and Epidemics”, under the Grant Agreement Nos. 101016233, (H2020-SC1-PHE-CORONAVIRUS-2020-2-RTD).

References

Agosto, A., Campmas, A., Giudici, P., Renda, A., 2021. A statistical model to monitor COVID-19 contagion growth. Stat. Med. URL https://onlinelibrary.wiley.com/doi/10.1002/sim.9020.
Agosto, A., Giudici, P., 2020. A Poisson autoregressive model to understand Covid-19 contagion dynamics. Risks 8, 77.
Bartolucci, F., Farcomeni, A., 2021. A spatio-temporal model based on discrete latent variables for the analysis of COVID-19 incidence. Spatial Stat. 100504, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.9004.
Berlemann, M., Haustein, E., 2020. Right and Yet Wrong: A Spatio-Temporal Evaluation of Germany’s Covid-19 Containment Policy. CESifo Working Paper Series 8446, CESifo.
Bracher, J., Held, L., 2020. Endemic-epidemic models with discrete-time serial interval distributions for infectious disease prediction. Int. J. Forecast. 69.
Cameron, A.C., Windmeijer, F.A.G., 1996. R-squared measures for count data regression models with applications to health-care utilization. J. Bus. Econom. Statist. 14 (2), 209–220.
Ceylan, Z., 2020. Estimation of Covid-19 prevelance in Italy, Spain and France. Scie. Total Environ. 729.
Cornelia, I., Annan-Phan, S.E., Thai, X.H., Mehra, S., Hsiang, S.M., Blumenstock, J.E., 2020. Public Mobility Data Enables COVID-19 Forecasting and Management at Local and Global Scales. Working Paper 28120, National Bureau of Economic Research.
Czado, C., Gneiting, T., Held, L., 2009. Predictive model assessment for count data. Biometrics 65, 1254–1261.
Dong, E., Du, H., Gardner, L., 2020. An interactive web-base dashboard to track COVID-19 in real time. Lancet Inf Dis. 20 (5), 533–534.
Farcomeni, A., Divino, F., Jona-Lasinio, G., Lloison, G., 2021. An ensemble approach to short-term forecast of COVID-19 intensive care occupancy in Italian regions. Biom. J. 63, 503–513.
Ferland, R., Lautor, A., Oraichi, D., 2006. Integer-valued GARCH processes. J. Time Series Anal. 27 (6), 923–942.
Fronterre, C., et al., 2020. COVID-19 in England: spatial patterns and regional outbreaks. MedRxiv.
Girandi, P., Greco, L., Mameli, V., Musio, M., Racugno, W., Ruli, E., Ventura, L., 2020. Robust inference for nonlinear regression models from the tsallis score: application to Covid-19 contagion in Italy. arXiv:2004.03187.
Giudici, P., Held, L.K., Rasser, G., 2000. Modelling categorical covariates in bayesian disease mapping by partition structures. Stat Med 17-18, 2579–2593.
Giudici, P., Mezzetti, M., Muliere, P., 2003. Mixtures of products of dirichlet processes for variable selection in survival analysis. Journ. Stat. Plan. Inf. 111, 101–115.
Giuliani, D., Dickson, M.M., Espa, G., Santi, F., 2020. Modelling and predicting the spatio-temporal spread of Covid-19 in Italy. BMC Infect. Dis. 20 (700).
Held, L., Hohle, M., Hoffman, M., 2005. A statistical framework for the analysis of multivariate infectious disease surveillance counts. Stat. Model. 5, 187–199.
Held, L., Paul, M., 2012. Modeling seasonality in space-time infectious disease surveillance data. Biom. J. 54 (6), 824–843.
Hermanowicz, S., 2020. Forecasting the wuhan coronavirus (2019-ncov) epidemics using a simple (simplicistic) model. MedRxiv.
Loro, P.A.D., Divino, F., Farcomeni, A., Jona-Lasinio, G., Lovison, G., Maruotti, A., Mingione, M., 2021. Nowcasting COVID-19 incidence indicators during the italian first outbreak. Stat. Med. 1–22, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.9004.
Meyer, S., Held, L., 2014. Power-law models for infectious disease spread. Ann. Appl. Stat. 8 (23), 1612–1639.
Meyer, S., Held, L., Hohle, M., 2017. Spatio-temporal analysis od epidemic phenomena using the R package surveillance. J. Stat. Softw. 77 (11), 1–55.
Nishiura, H., Linton, N.M., Akhmetzhanov, A.R., 2020. Serial interval of novel coronavirus (COVID-19) infections. Int. J. Infect. Dis. 93, 284–286.
Paul, M., Held, L., 2011. Predictive assessment of a non-linear effects for multivariate time series of infectious disease counts. Stat. Med. 30 (10), 1118–1136.
Paul, M., Held, L., Toschke, A., 2008. Multivariate modelling of infectious disease surveillance data. Stat. Med. 27 (29), 6250–6267.
Perone, G., 2020. Comparison of ARIMA, ETS, NNAR and hybrid models to forecast the second wave of COVID-19 hospitalizations in Italy. Health, Econometrics and Data Group (HEDG) Working Papers 20/18, HEDG, c/o Department of Economics, University of York.
2020. Presidenza del consiglio dei ministri, dipartimento della protezione civile. Dati COVID-19 Italia. https://github.com/pcm-dpc/COVID-19.
Silva, T., Anghinoni, L., Zhao, L., 2020. Quantitative analysis of the effectiveness of public health measures on Covid-19 transmission. MedRxiv.
Ssentongo, P., Fronterre, C., Geronimo, A., Greybush, S.J., Mbabazi, P.K., Muvawala, J., Nahalama, S.B., Omadi, P.O., Opar, B.T., Sinnar, S.A., Wang, Y., Whalen, A.J., Held, L., Jewell, C., Muwanguzi, A.J.B., Greatrex, H., Norton, M.M., Diggle, P., Schiff, S.J., 2020. Tracking and predicting the african COVID-19 pandemic. MedRxiv.
Wu, K., Darce, D., Wang, Q., Sornette, D., 2020. Generalized logistic growth modeling of the Covid-19 outbreak in 29 provinces in China and in the rest of the world. pp. 1561–1581. ArXiv Preprint, 101.