Research article

Spatial spread of COVID-19 outbreak in Italy using multiscale kinetic transport equations with uncertainty

Giulia Bertaglia[1,2]**, Walter Boscheri** [1,2]**, Giacomo Dimarco** [1,2,*]**and Lorenzo Pareschi** [1,2]

[1] Department of Mathematics and Computer Science, University of Ferrara, Via Machiavelli 30, Ferrara 44121, Italy

[2] Center for Modeling, Computing and Statistic (CMCS), University of Ferrara, Via Muratori 9, Ferrara 44121, Italy

* **Correspondence:** Email: giacomo.dimarco@unife.it.

Abstract: In this paper we introduce a space-dependent multiscale model to describe the spatial spread of an infectious disease under uncertain data with particular interest in simulating the onset of the COVID-19 epidemic in Italy. While virus transmission is ruled by a SEIAR type compartmental model, within our approach the population is given by a sum of commuters moving on a extra-urban scale and non commuters interacting only on the smaller urban scale. A transport dynamics of the commuter population at large spatial scales, based on kinetic equations, is coupled with a diffusion model for non commuters at the urban scale. Thanks to a suitable scaling limit, the kinetic transport model used to describe the dynamics of commuters, within a given urban area coincides with the diffusion equations that characterize the movement of non-commuting individuals. Because of the high uncertainty in the data reported in the early phase of the epidemic, the presence of random inputs in both the initial data and the epidemic parameters is included in the model. A robust numerical method is designed to deal with the presence of multiple scales and the uncertainty quantification process. In our simulations, we considered a realistic geographical domain, describing the Lombardy region, in which the size of the cities, the number of infected individuals, the average number of daily commuters moving from one city to another, and the epidemic aspects are taken into account through a calibration of the model parameters based on the actual available data. The results show that the model is able to describe correctly the main features of the spatial expansion of the first wave of COVID-19 in northern Italy.

Keywords: kinetic transport equations; epidemic models; commuting flows; COVID-19; diffusion limit; asymptotic-preserving schemes; uncertainty quantification; unstructured grids

1. Introduction

The advent of the COVID-19 pandemic has caused a strong commitment of many researchers acting in different fields with the scope of trying to understand and give explanations to the global crisis we are experiencing. From the mathematical modeling point of view, several progresses have been done concerning the development of epidemic models capable of taking into account the different facets of this terrible disease. In particular, many recent researches have been addressed to the search of control strategies [1–4] to limit the spread and consequently hospitalizations and deaths, possibly reducing, at the same time, as much as possible the negative impact on the economy of the restrictive measures [5, 6].

Most of the recently proposed model represent improvements at various levels of the seminal works on compartmental epidemiological modeling proposed originally by Kermack and McKendrick [7]. These approaches [1, 3, 8–13] are typically focused on the epidemiological aspects of the virus spread under the hypothesis of global mixing of the population, hence, without taking into account the role of the spatial component in the evolution of an outbreak.

Despite in many situations the above description is sufficient to delineate the global trend of an epidemic, there are cases in which the spatial homogeneity assumption does not hold true. In these situations, one seeks for local interventions in order to reduce the spread without eventually affecting regions where the incidence of the infection does not require special care, as for instance in the recent case of the COVID-19 in Italy [14]. Consequently, from the modelling point of view, the inclusion of the spatial dependence represents a key challenge [15].

Indeed, with the increasing amount of information on population mobility and the computational resources available today, the design and simulation of epidemic models based on partial differential equations (PDEs) that include the details of spatial dynamics can be considered a realistic goal [16]. We recall that most of the existing epidemiological models taking into account spatial heterogeneities are based on reaction-diffusion equations [17–26]. Alternative modelling approaches are represented by the interaction of different homogeneous populations [4] or agent-based dynamics [27].

Most of these models do not allow for a clear distinction about the possible spatial behaviors of the population inside a given compartment. Consequently, although capable of originating realistic spatial patterns in situations where individuals move indiscriminately through the domain, such approaches are likely to be less effective in the case where one is interested in studying the spread of a virus in the human population. Under these circumstances, it is more realistic to consider only commuting individuals moving in major and preferred directions, not considering overall mass migration between distinct urban areas. Indeed, not all individuals move indiscriminately in the region of interest, as most of the population only interacts at the urban scale. Furthermore, in contrast to the use of diffusion models, the propagation rate of the infection along the spatial domain is obviously finite. Recently, trying to overcome some of the above criticisms, such as the paradox of the infinite propagation speed, alternative models based on hyperbolic PDEs have been proposed [28–32].

In this work, following the approach introduced in [32], we consider a realistic compartmental structure for the description of the epidemic dynamics of commuters and non commuters individuals in presence of uncertain data. The model consists of a system of kinetic transport equations describing a large-scale (extra-urban) commuting population by a continuous density [33–36] in a two-dimensional environment. This density can be interpreted as the probability for an individual to be at a given loc-

ation and move in a given direction at a given instant of time, in analogy to particle flows in rarefied gases [37–42]. The above system is coupled to a second system consisting of a set of diffusion equations that characterize the movement of the non-commuting population on a small (urban) scale. The epidemic spread is ruled by a Susceptible, Exposed, Infected, Asymptomatic and Removed (SEIAR) compartmental structure, in which both commuters and non commuters can interact. Using a suitable scaling process [32, 43], the model allows us to highlight, firstly, the relationship with well-known existing approaches based on reaction-diffusion equations and, secondly, to pass naturally from a hyperbolic description in peripheral areas to a diffusive regime when reaching an urban area, with regard to the commuting population.

An important aspect concerns overcoming the limitation caused by standard deterministic models that rely on the assumption that the initial conditions, boundary conditions, and all involved epidemic and mobility parameters are known. However, as observed in the case of the COVID-19 epidemic, this assumption, especially in the early stages of the epidemic, is not reliable. For example, the initial conditions in terms of the number of infected and asymptomatic persons are certainly affected by uncertainty because data are limited and population screening cannot be error-free. Epidemic parameters, although normally estimated or calibrated, are also often candidates for being random variables in a realistic approach. Therefore, in order to take into account these limitations underlying deterministic models, in this paper we resort to a stochastic approach based on the introduction of random terms into the initial modeling [1, 2, 30, 44].

Once the model is set up, its numerical solution on a computational domain describing a realistic geographic scenario poses several difficulties. In fact, the model consists of two coupled systems of PDEs, each characterized by five unknown functions living in a multidimensional space characterized by space, velocity, and stochastic variables. Additionally, we have to deal with the irregular shape of the spatial region and the multiscale nature of the dynamics (indeed, as previously stated, hyperbolic and parabolic behaviors coexist). Therefore, a particular care is devoted to the development of efficient and accurate numerical schemes. More precisely, a discretization of the system based on Gaussian quadrature points in velocity space [45, 46] and a finite volume approach on unstructured grids [47, 48] is considered. The adoption of asymptotic preserving time discretization techniques permits to avoid time step limitations introduced by the parabolic scaling without degradation of accuracy [49–51]. Finally, a non-intrusive stochastic collocation method, which guarantees spectral accuracy in the space of the uncertain parameters, is considered to deal with the uncertainty quantification process [30, 52].

The rest of the paper is organized as follows. In Section 2 the mathematical model is introduced. We first introduce the kinetic transport formulation for the epidemic compartments of commuters together with the corresponding diffusive dynamics of the non commuters in a deterministic setting. Subsequently, we link the two hyperbolic/parabolic dynamics through a formal passage to the limit for the system of commuters. A definition of the basic reproduction number of the epidemic for the resulting model is also reported. Next, we illustrate how to generalize the model in presence of uncertainty. The details of the numerical scheme used to approximate the resulting stochastic system are summarized in Appendix A.1. Section 3 is devoted to present an application of the current modelling to the first outbreak of COVID-19 in Italy and its spread in the Lombardy Region. The capability of the model to accurately represent the first wave of the COVID-19 epidemic in Italy is discussed in detail through comparisons with recorded data reported by official sources [53]. Conclusions and future perspectives are finally given in Section 4.

2. A compartmental kinetic transport model

Let $\Omega \subset \mathbb{R}^2$ characterizes a two-dimensional geographical area of interest and assume that individuals have been separated into two different groups: a commuting population, typically moving over long distances (extra-urban), and a non-commuting population, moving only in small-scale urban areas. In the first part of this section, for ease of presentation, the multiscale kinetic transport model is introduced in a deterministic setting. The relation between the current hyperbolic transport model and classical diffusion models is discussed in the second part. In the third part, details regarding the basic reproduction number associated with the model are given. Finally, in the last part, we discuss the generalization of the deterministic model to the case where uncertainty is taken into account.

2.1. Characterizing commuter and non commuter dynamics

We consider a population of commuters at position $x \in \Omega$ moving with velocity directions $v \in \mathbb{S}^1$ and denote the respective kinetic densities of susceptible (individuals who may be infected by the disease) by $f_S = f_S(x, v, t)$, exposed (individuals in the latent period, which are not yet infectious) by $f_E = f_E(x, v, t)$, severe symptomatic infected by $f_I = f_I(x, v, t)$, mildly symptomatic or asymptomatic infected by $f_A = f_A(x, v, t)$ and removed (individuals healed or died due to the disease) by $f_R = f_R(x, v, t)$. The kinetic distribution of commuters is then given by

$$f(x, v, t) = f_S(x, v, t) + f_E(x, v, t) + f_I(x, v, t) + f_A(x, v, t) + f_R(x, v, t),$$

and their total density is obtained by integration over the velocity space

$$\rho(x, t) = \frac{1}{2\pi} \int_{\mathbb{S}^1} f(x, v_*, t) \, dv_*.$$

As a consequence, one can recover the number of susceptible, exposed and recovered irrespective of their direction of displacement by integration over the velocity space. This gives

$$S(x, t) = \frac{1}{2\pi} \int_{\mathbb{S}^1} f_S(x, v, t) \, dv, \quad E(x, t) = \frac{1}{2\pi} \int_{\mathbb{S}^1} f_E(x, v, t) \, dv, \quad R(x, t) = \frac{1}{2\pi} \int_{\mathbb{S}^1} f_R(x, v, t) \, dv,$$

which we refer to as the density fractions of non-infectious individuals, whereas

$$I(x, t) = \frac{1}{2\pi} \int_{\mathbb{S}^1} f_I(x, v, t) \, dv, \quad A(x, t) = \frac{1}{2\pi} \int_{\mathbb{S}^1} f_A(x, v, t) \, dv,$$

are the density fractions of infectious individuals.

In this setting, the kinetic densities of the commuters satisfy the transport equations

$$\begin{aligned}
\frac{\partial f_S}{\partial t} + v_S \cdot \nabla_x f_S &= -F_I(f_S, I_T) - F_A(f_S, A_T) + \frac{1}{\tau_S}(S - f_S) \\
\frac{\partial f_E}{\partial t} + v_E \cdot \nabla_x f_E &= F_I(f_S, I_T) + F_A(f_S, A_T) - a f_E + \frac{1}{\tau_E}(E - f_E) \\
\frac{\partial f_I}{\partial t} + v_I \cdot \nabla_x f_I &= a\sigma f_E - \gamma_I f_I + \frac{1}{\tau_I}(I - f_I)
\end{aligned} \tag{2.1}$$

$$\frac{\partial f_A}{\partial t} + v_A \cdot \nabla_x f_A = a(1 - \sigma)f_E - \gamma_A f_A + \frac{1}{\tau_A}(A - f_A)$$

$$\frac{\partial f_R}{\partial t} + v_R \cdot \nabla_x f_R = \gamma_I f_I + \gamma_A f_A + \frac{1}{\tau_R}(R - f_R)$$

where the total densities of non infected individuals are defined by

$$S_T(x, t) = S(x, t) + S_u(x, t), \quad E_T(x, t) = E(x, t) + E_u(x, t), \quad R_T(x, t) = R(x, t) + R_u(x, t),$$

and similarly the total densities of infected by

$$I_T(x, t) = I(x, t) + I_u(x, t), \quad A_T(x, t) = A(x, t) + A_u(x, t).$$

In the above equations, $S_u(x, t)$, $E_u(x, t)$, $I_u(x, t)$, $A_u(x, t)$, $R_u(x, t)$ are the density fractions of non-commuters who, by assumption, move only on an urban scale. These densities satisfy a diffusion dynamics acting only at the same local scale

$$\frac{\partial S_u}{\partial t} = -F_I(S_u, I_T) - F_A(S_u, A_T) + \nabla_x \cdot (D_S^u \nabla_x S_u)$$

$$\frac{\partial E_u}{\partial t} = F_I(S_u, I_T) + F_A(S_u, A_T) - aE_u + \nabla_x \cdot (D_E^u \nabla_x E_u)$$

$$\frac{\partial I_u}{\partial t} = a\sigma E_u - \gamma_I I_u + \nabla_x \cdot (D_I^u \nabla_x I_u) \tag{2.2}$$

$$\frac{\partial A_u}{\partial t} = a(1 - \sigma)E_u - \gamma_A A_u + \nabla_x \cdot (D_A^u \nabla_x A_u)$$

$$\frac{\partial R_u}{\partial t} = \gamma_I I_u + \gamma_A A_u + \nabla_x \cdot (D_R^u \nabla_x R_u).$$

In the resulting Eqs (2.1) and (2.2) that couples the commuting and non-commuting dynamics, the velocities $v_i = \lambda_i(x)v$ in Eq (2.1), as well as the diffusion coefficients $D_i^u = D_i^u(x)$ in Eq (2.2), with $i \in \{S, E, I, A, R\}$, are designed to take into account the heterogeneity of geographical areas, and are thus chosen dependent on the spatial location. Similarly, also the relaxation times $\tau_i = \tau_i(x)$, $i \in \{S, E, I, A, R\}$ are space dependent. The quantities $\gamma_I = \gamma_I(x)$ and $\gamma_A = \gamma_A(x)$ are the recovery rates of symptomatic and asymptomatic infected (inverse of the infectious periods), respectively, while $a(x)$ represents the inverse of the latency period and $\sigma(x)$ is the probability rate of developing severe symptoms [3, 4, 9].

The transmission of the infection is governed by the incidence functions $F_I(\cdot, I_T)$ and $F_A(\cdot, A_T)$. We assume local interactions to characterize the nonlinear incidence functions [54, 55]

$$F_I(g, I_T) = \beta_I \frac{gI_T^p}{1 + \kappa_I I_T}, \qquad F_A(g, A_T) = \beta_A \frac{gA_T^p}{1 + \kappa_A A_T}, \tag{2.3}$$

where the classic bi-linear case corresponds to $p = 1$, $k_I = k_A = 0$. Parameters $\beta_I = \beta_I(x, t)$ and $\beta_A = \beta_A(x, t)$ characterize the contact rates of highly symptomatic and mildly symptomatic/asymptomatic infectious individuals, accounting for both the number of contacts and the probability of transmission. Hence, they may vary based on the effects of government control actions, such as wearing of masks, shutdown of specific activities or lockdowns [1, 7, 12]. On the other hand, parameters $\kappa_I = \kappa_I(x, t)$ and

$\kappa_A = \kappa_A(x,t)$ are the incidence damping coefficients based on the self-protective behavior assumed by the individuals due to the awareness of the epidemic risk [11, 26, 30].

Alternative incidence functions are given by

$$F_I(g, I_T) = \beta_I \frac{g I_T^p}{1 + \kappa_I \int_{\bar{\Omega}} I_T \, dx}, \qquad F_A(g, A_T) = \beta_A \frac{g A_T^p}{1 + \kappa_A \int_{\bar{\Omega}} A_T \, dx}, \qquad (2.4)$$

where $\bar{\Omega}$ is a chosen portion of the domain which permits to take into account the fact that social distancing may depend on the average level of infection of the region rather than only on the local situation. Since the proposed model separates the contribution of the spread of the disease between the two infectious compartments, through the definition of two different incidence functions, F_I and F_A, we chose to keep the same distinction also for the factors multiplying the damping coefficients κ_I and κ_A in each function respectively. Clearly, the choice could have been different, accounting for the sum $I_T + A_T$ in both the incidence functions, or even considering only the contribution of I_T for both. The resulting Eqs (2.1) and (2.2) will be referred to as multiscale kinetic SEIAR (MK-SEIAR) model. Note that, because of the presence of two populations acting at different scales, the model allows a more realistic description of the typical commuting dynamics involving only a fraction of the population and distinguishes it from the epidemic process affecting the entire population.

2.2. Commuters behavior in urban areas

The hyperbolic transport model for the commuters deserves some remarks. In fact, while it is clear that a hyperbolic description permits to describe correctly the daily extra-urban commuting part, the same individuals when moving inside the urban area are better described by a traditional diffusion model. A remarkable feature of the transport Eq (2.1), is that it permits to recover a classical diffusion behavior under the hypothesis that the relaxation times $\tau_{S,I,R}$ tend to zero while keeping finite the diffusion coefficients

$$D_S = \frac{1}{2}\lambda_S^2 \tau_S, \quad D_E = \frac{1}{2}\lambda_E^2 \tau_E, \quad D_I = \frac{1}{2}\lambda_I^2 \tau_I, \quad D_A = \frac{1}{2}\lambda_A^2 \tau_A, \quad D_R = \frac{1}{2}\lambda_R^2 \tau_R. \qquad (2.5)$$

More precisely, let us introduce the flux functions

$$J_S = \frac{\lambda_S}{2\pi} \int_{\mathbb{S}^1} v f_S(x,v,t) \, dv, \quad J_E = \frac{\lambda_E}{2\pi} \int_{\mathbb{S}^1} v f_E(x,v,t) \, dv, \quad J_I = \frac{\lambda_I}{2\pi} \int_{\mathbb{S}^1} v f_I(x,v,t) \, dv$$

$$J_A = \frac{\lambda_A}{2\pi} \int_{\mathbb{S}^1} v f_A(x,v,t) \, dv, \quad J_R = \frac{\lambda_R}{2\pi} \int_{\mathbb{S}^1} v f_R(x,v,t) \, dv.$$

Then, integrating system Eq (2.1) in v, we get the following set of equations for the macroscopic densities of commuters

$$\frac{\partial S}{\partial t} + \nabla_x \cdot J_S = -F_I(S, I_T) - F_A(S, A_T)$$

$$\frac{\partial E}{\partial t} + \nabla_x \cdot J_E = F_I(S, I_T) + F_A(S, A_T) - aE$$

$$\frac{\partial I}{\partial t} + \nabla_x \cdot J_I = a\sigma E - \gamma_I I \qquad (2.6)$$

$$\frac{\partial A}{\partial t} + \nabla_x \cdot J_A = a(1-\sigma)E - \gamma_A A$$

$$\frac{\partial R}{\partial t} + \nabla_x \cdot J_R = \gamma_I I + \gamma_A A$$

whereas the flux functions satisfy

$$\frac{\partial J_S}{\partial t} + \frac{\lambda_S^2}{2\pi} \int_{\mathbb{S}^1} (v \cdot \nabla_x f_S) v \, dv = -F_I(J_S, I_T) - F_A(J_S, A_T) - \frac{1}{\tau_S} J_S$$

$$\frac{\partial J_E}{\partial t} + \frac{\lambda_E^2}{2\pi} \int_{\mathbb{S}^1} (v \cdot \nabla_x f_E) v \, dv = \frac{\lambda_E}{\lambda_S} (F_I(J_S, I_T) + F_A(J_S, A_T)) - aJ_E - \frac{1}{\tau_E} J_E$$

$$\frac{\partial J_I}{\partial t} + \frac{\lambda_I^2}{2\pi} \int_{\mathbb{S}^1} (v \cdot \nabla_x f_I) v \, dv = \frac{\lambda_I}{\lambda_E} a\sigma J_E - \gamma_I J_I - \frac{1}{\tau_I} J_I \qquad (2.7)$$

$$\frac{\partial J_A}{\partial t} + \frac{\lambda_A^2}{2\pi} \int_{\mathbb{S}^1} (v \cdot \nabla_x f_A) v \, dv = \frac{\lambda_A}{\lambda_E} a(1-\sigma) J_E - \gamma_A J_A - \frac{1}{\tau_A} J_A$$

$$\frac{\partial J_R}{\partial t} + \frac{\lambda_R^2}{2\pi} \int_{\mathbb{S}^1} (v \cdot \nabla_x f_R) v \, dv = \frac{\lambda_R}{\lambda_I} \gamma_I J_I + \frac{\lambda_R}{\lambda_A} \gamma_A J_A - \frac{1}{\tau_R} J_R.$$

Clearly, the above system is not closed because the evolution of the fluxes in Eq (2.7) involves higher order moments of the kinetic densities. The diffusion limit can be formally recovered, by introducing the space dependent diffusion coefficients Eq (2.5) and letting $\tau_{S,I,R} \to 0$. We get from the r.h.s. in Eq (2.1)

$$f_S = S, \quad f_E = E, \quad f_I = I, \quad f_A = A, \quad f_R = R,$$

and, consequently, from Eq (2.7) we recover Fick's law

$$J_S = -D_S \nabla_x S, \quad J_E = -D_E \nabla_x E, \quad J_I = -D_I \nabla_x I, \quad J_A = -D_A \nabla_x A, \quad J_R = -D_R \nabla_x R, \qquad (2.8)$$

since

$$\int_{\mathbb{S}^1} (v \cdot \nabla_x S) v \, dv = \int_{\mathbb{S}^1} (v \otimes v) \, dv \nabla_x S = \pi \nabla_x S,$$

and similarly for the other densities. Thus, substituting Eqs (2.8) into (2.6) we get the diffusion system for the population of commuters [21, 23, 56]

$$\frac{\partial S}{\partial t} = -F_I(S, I_T) - F_A(S, A_T) + \nabla_x \cdot (D_S \nabla_x S)$$

$$\frac{\partial E}{\partial t} = F_I(S, I_T) + F_A(S, A_T) - aE + \nabla_x \cdot (D_E \nabla_x E)$$

$$\frac{\partial I}{\partial t} = a\sigma E - \gamma_I I + \nabla_x \cdot (D_I \nabla_x I) \qquad (2.9)$$

$$\frac{\partial A}{\partial t} = a(1-\sigma)E - \gamma_A A + \nabla_x \cdot (D_A \nabla_x A)$$

$$\frac{\partial R}{\partial t} = \gamma_I I + \gamma_A A + \nabla_x \cdot (D_R \nabla_x R)$$

which is coupled with Eq (2.2) for the non-commuting counterpart. The capability of the model to account for different regimes, hyperbolic or parabolic, accordingly to the space dependent relaxation times τ_i, $i \in \{S, E, I, A, R\}$, makes it suitable for describing the dynamics of human beings. Indeed, it is clear that the daily routine is a complex mixing of individuals moving at the scale of a city and individuals moving among different urban centers. In this situation, due to the lack of microscopic information and the high complexity of the dynamics, it is reasonable to avoid describing the details of movements within an urban area and model this through a diffusion operator. On the other hand, commuters when moving from one city to another follow well-established connections for which a description via transport operators is more appropriate.

2.3. Basic reproduction number

The standard threshold of epidemic models is the well-known reproduction number R_0, which defines the average number of secondary infections produced when one infected individual is introduced into a host population in which everyone is susceptible [7] during its entire period of infectiousness. This number determines when an infection can invade and persist in a new host population. For many deterministic epidemic models, an infection begins in a fully susceptible population if and only if $R_0 > 1$. Its definition in the case of spatially dependent dynamics is not straightforward, particularly when considering its spatial dependence. In the following, assuming no inflow/outflow boundary conditions in Ω, integrating over velocity and space, we derive the following definition for the average reproduction number value on the domain Ω

$$
\begin{aligned}
R_0(t) = &\frac{\int_\Omega F_I(S_T, I_T)\,dx}{\int_\Omega \gamma_I(x)I_T(x,t)\,dx} \cdot \frac{\int_\Omega a(x)\sigma(x)E_T(x,t)\,dx}{\int_\Omega a(x)E_T(x,t)\,dx} \\
&+ \frac{\int_\Omega F_A(S_T, A_T)\,dx}{\int_\Omega \gamma_A(x)A_T(x,t)\,dx} \cdot \frac{\int_\Omega a(x)(1-\sigma(x))E_T(x,t)\,dx}{\int_\Omega a(x)E_T(x,t)\,dx} \ .
\end{aligned}
\tag{2.10}
$$

The derivation of the above expression for $R_0(t)$, computed following the *next-generation matrix* approach [57], is presented in detail in [30] using a suitable linearization of the corresponding nonlinear process for the space averaged quantities.

It is worth to underline that from definition Eq (2.10) it can be deduced that it is a combination of the growth of E_T, I_T and A_T that determines the persistence of the epidemic, not solely the growth of E_T in time, neither the growth of the simple sum $E_T + I_T + A_T$. If, additionally, compartments I and A are considered homogeneously mixed in a unique compartment, allowing $\beta_I = \beta_A = \beta$, $\kappa_A = \kappa_I = \kappa$ and $\gamma_I = \gamma_A = \gamma$, we recover a SEIR-type compartmental model and the reproduction number results as in [32]:

$$
R_0(t) = \frac{\int_\Omega F(S_T, I_T)\,dx}{\int_\Omega \gamma(x)I_T(x,t)\,dx}.
\tag{2.11}
$$

Let us finally observe that, under the same no inflow/outflow boundary conditions, integrating in Ω Eqs (2.1) and (2.2) yields respectively the conservation of the total populations of commuters and

non-commuters

$$\frac{\partial}{\partial t} \int_{\Omega} (S(x,t) + E(x,t) + I(x,t) + A(x,t) + R(x,t)) \, dx = 0,$$

$$\frac{\partial}{\partial t} \int_{\Omega} (S_u(x,t) + E_u(x,t) + I_u(x,t) + A_u(x,t) + R_u(x,t)) \, dx = 0.$$

2.4. Including data uncertainty

To extend the Eqs (2.1) and (2.2) to the case in which uncertainties are taken into account, let us suppose that the population of commuters depend on an additional random vector $z = (z_1, \ldots, z_d)^T \in \Omega_z \subseteq \mathbb{R}^d$, where $z_1, \ldots, z_d$ are independent random variables. This vector is used to characterize possible sources of uncertainty in the physical system due to lack of information on the actual number of infected or specific epidemic characteristics of the infectious disease.

Thus, in the system we have the following high-dimensional unknowns

$$f_S = f_S(x, v, t, z), \ f_E = f_E(x, v, t, z), \ f_I = f_I(x, v, t, z), \ f_A = f_A(x, v, t, z), f_R = f_R(x, v, t, z).$$

The same considerations apply to the non-commuter population, yielding

$$S_u = S_u(x, t, z), \ E_u = E_u(x, t, z), \ I_u = I_u(x, t, z), A_u = A_u(x, t, z), R_u = R_u(x, t, z).$$

Notice that besides the introduction of a new vector of variables, the structure of the Eqs (2.1) and (2.2) does not change, i.e., there is no direct variation of the unknowns with respect to z, which, instead, have to be intended as parameters into the equations. We will further assume that also some epidemic parameters are affected by uncertainty. Therefore, for instance, parameters acting inside the incidence function may have an additional dependence of the form

$$\beta_I = \beta_I(x, t, z), \ \beta_A = \beta_A(x, t, z).$$

$$k_I = k_I(x, t, z), \ k_A = k_A(x, t, z),$$

In the next section, we discuss several numerical examples based on the Eq (2.1) and (2.2) with uncertainty. We will consider that the initial number of detected infected I, derived from the data at disposal [53, 58], represents only a lower bound while the true values are not known but affected by uncertainty. Consequently, also initial conditions of exposed E, asymptomatic A and susceptible S will contain a stochastic dependence.

3. Application to COVID-19 spread in Italy

To validate the proposed methodology in a realistic geographical and epidemiological scenario, a numerical test reproducing the epidemic outbreak of COVID-19 in the Lombardy Region of Italy, from February 27, 2020 to March 22, 2020, is designed, taking into account the uncertainty underlying initial conditions of infected individuals. We underline that the discretization of the multiscale system of PDEs Eqs (2.1) and (2.2) presented in this work is not trivial and requires the construction of a specific numerical method able to correctly describe the transition from a convective to a diffusive regime in realistic geometries. Additionally the method should be capable to deal efficiently with the uncertainty characterized by the stochastic nature of the model. Details on these numerical aspects are given in Appendix A.1 together with the tables of the data used for population mobility in Appendix A.2.
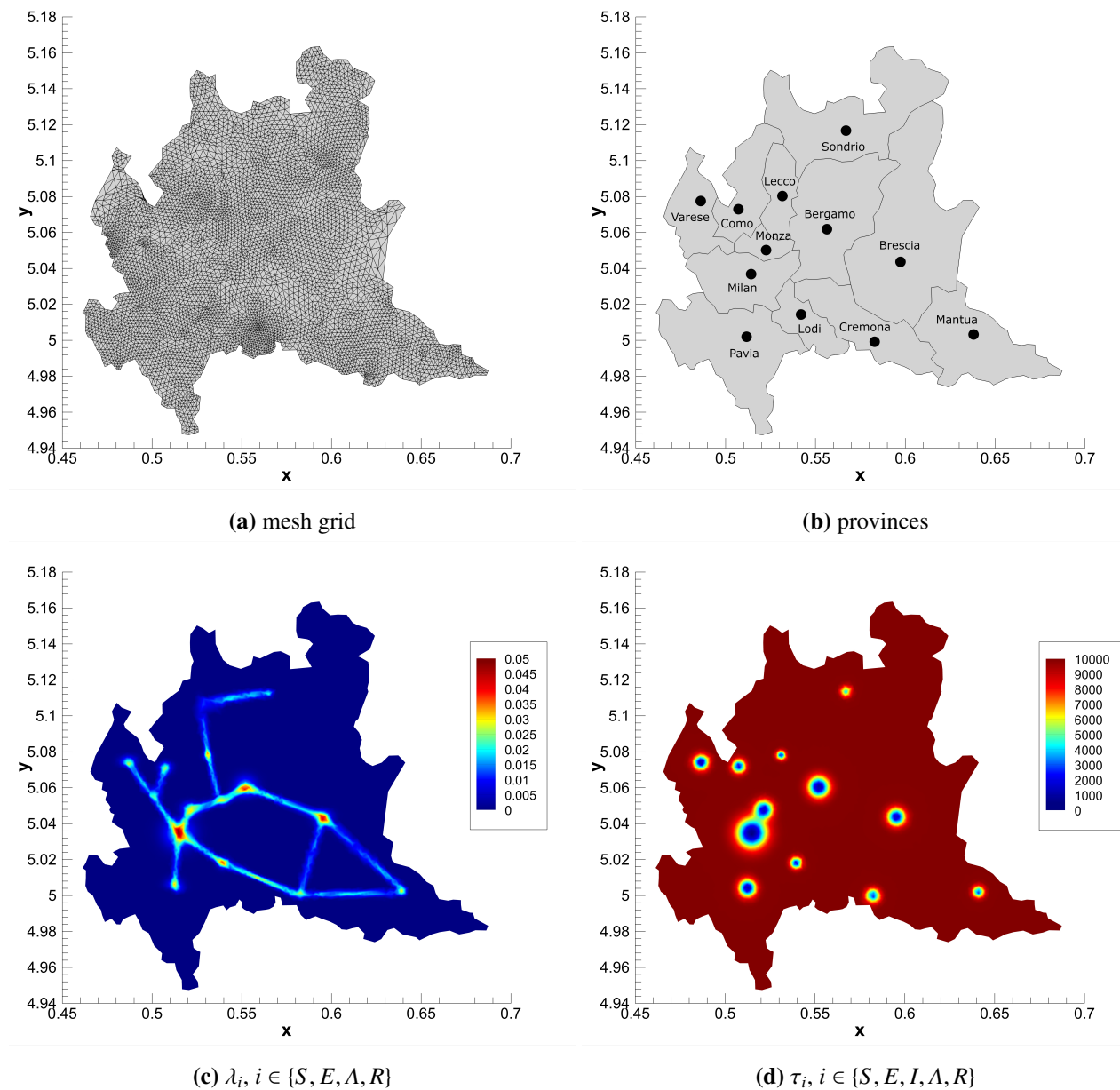
(a) mesh grid

(b) provinces

(c) λ_i, $i \in \{S, E, A, R\}$

(d) τ_i, $i \in \{S, E, I, A, R\}$

Figure 1. Top: unstructured computational mesh used to discretize the Lombardy Region (a) and identification of the provinces (b). Bottom: initial condition imposed for characteristic speeds λ (c) and relaxation times τ (d).

3.1. Computational setup

The computational domain is defined in terms of the boundary that circumscribes the Lombardy Region, which is available in [58] as a list of georeferenced points in the ED50/UTM Zone 32N reference coordinate system. To avoid ill-conditioned reconstruction matrices and other related problems that arise when dealing with large numbers in finite arithmetic, all coordinates are re-scaled by a factor of 10^6. The resulting computational grid is composed of $N_E = 10,792$ triangular control volumes. The mesh grid is presented in Figure 1a. No-flux boundary conditions are imposed in the whole boundary

of the domain, assuming that the population is not moving from/to the adjacent Regions. The domain is then subdivided in the $\mathcal{N}_c = 12$ provinces of Lombardy: Pavia (PV), Lodi (LO), Cremona (CR), Mantua (MN), Milan (MI), Bergamo (BG), Brescia (BS), Varese (VA), Monza-Brianza (MB), Como (CO), Lecco (LC) and Sondrio (SO). The identification of these cities is shown in Figure 1b.

The units of measure chosen for this numerical test can be summarized as follows:

$$1\,\text{km} = 10^{-3}\,\text{L}, \quad 1\,\text{person} \approx 10^7\,\text{P}, \quad 1\,\text{day} = 2\,\text{T},$$

with [L], [P] and [T] being the length, person and time units used in the simulation, respectively. Notice that the normalization of the population is made with respect to the total number of individuals of the Region (taken from [59]), which is $M = 10.027.602$, to properly work in a context in which the total population is equal to the unit.

To avoid the mobility of the population in the entire territory and to simulate a more realistic geographical scenario in which individuals travel along the main traffic paths of the Region, different values of propagation speeds are assigned in the domain which reflect, as close as possible, the real characteristics of the territory. Along the main connections of the Region, a mean value of $\lambda = 0.04$ is prescribed for compartments S, E, A and R, which ensures a maximum travel distance of 80 km within a day; while, for the same compartments, $\lambda = 0.02$ is ulterior fixed in the urbanized circles. A spatial width of $h = 0.5$ km is assigned to the traveling paths. On the other hand, assuming that highly infectious subjects are mostly detected in the most optimistic scenario, being subsequently quarantined or hospitalized, the speed assigned to compartment I is set null. However, the infected people, even if limited by quarantines and social distancing, can still contribute to the spread the disease via the diffusion process at the urban scale (mimicking for instance the still possible infections happening at the family level). A null value $\lambda = 0$ is set in the rest of the computational domain for all the epidemiological compartments. The resulting distribution of the characteristic speeds is visible from Figure 1c.

The relaxation time is set $\tau_i = 10^4$ for $i \in \{S, E, I, A, R\}$, so that the model recovers a hyperbolic regime in the entire region, apart from the main cities, where a parabolic setting is prescribed to correctly capture the diffusive behavior of the disease spreading which typically occurs in highly urbanized zones. Hence, to obtain a smooth change of the relaxation time between extra-urban and urban scale, in the urban area of radius r_c of each city $c = 1, \ldots, \mathcal{N}_c$ the following relaxation time $\tau_{c,i}$, with Gaussian-like shape centered in the city center (x_c, y_c), is prescribed:

$$\tau_{c,i}(x, y) = \tau_i + (\tau_0 - \tau_i) \sum_{c=1}^{\mathcal{N}_c} e^{-\frac{(x-x_c)^2 + (y-y_c)^2}{2r_c^2}}, \qquad (x, y) \in \Omega,$$

with a diffusive relaxation time chosen to be $\tau_0 = 10^{-4}$. The resulting distribution of the relaxation times is presented in Figure 1d.

3.2. Uncertain initial data and epidemic parameters

Considering p_c the number of citizens of a generic city (province) denoted with subscript c, the initial spatial distribution of the generic population $f(x, y)$ is assigned, for each province and each epidemiological compartment, as a multivariate Gaussian function with the variance being the radius
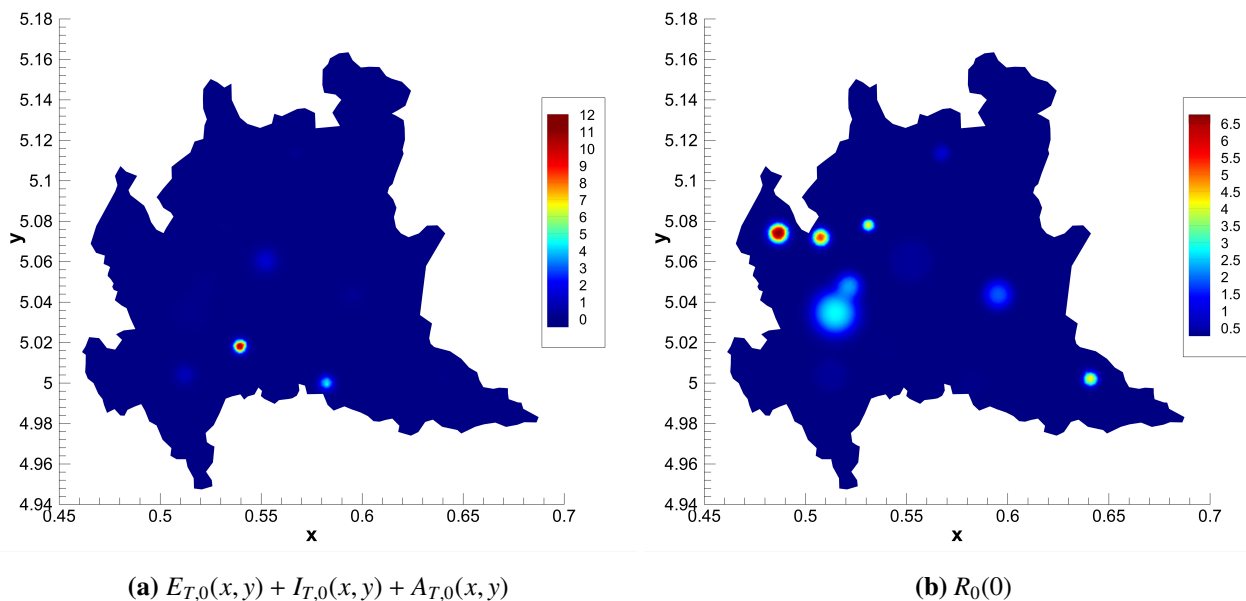
(a) $E_{T,0}(x,y) + I_{T,0}(x,y) + A_{T,0}(x,y)$

(b) $R_0(0)$

Figure 2. Initial distribution (on February 27, 2020) of the infected population $E_{T,0}+I_{T,0}+A_{T,0}$ (a) and of the reproduction number $R_0(0)$ (b) in the Lombardy Region.

of the urban area r_c:

$$f(x,y) = \frac{1}{2\pi r_c} e^{-\frac{(x-x_c)^2+(y-y_c)^2}{2r_c^2}} p_c,$$

with (x_c, y_c) representing the coordinates of a generic city center. The initial population setting, for each province of the Lombardy Region, is taken from [59] and reported in Table A.1, Appendix A.2. Note that, the radius r_c associated to each city, defined in Table A.1 (first column), permits to exactly re-obtain the population p_c when integrating over the computational domain the initial spatially distributed population.

Since at the beginning of the pandemic, tracking of positive individuals in Italy was very scarce, in this numerical test we consider that the initial amount of infected people is the leading quantity affected by uncertainty. To this aim, we introduce a single source of uncertainty z having uniform distribution, $z \sim \mathcal{U}(0, 1)$ so that the initial conditions for compartment I, in each control volume, are prescribed as

$$I_T(0, z) = I_{T,0}(1 + \mu z),$$

with $I_{T,0}$ initial amount of highly infectious corresponding to the values reported by February 27, 2020 in the GitHub repository [53] daily updated by the Civil Protection Department of Italy for each city, listed in the last column of Table A.1. This choice effectively associates all infected individuals detected with the I compartment, as a result of the screening policy adopted during February–March 2020 in Italy. In fact, tests to assess the presence of SARS-CoV-2 were performed almost exclusively on patients with consistent symptoms and fever at the beginning of this pandemic. Regarding the uncertainty of these data, we impose $\mu = 1$, assuming that at least half of the actual highly symptomatic infected were detected at the beginning of the pandemic outbreak. For all the cities with zero infected detected by February 27, 2020 (e.g., Mantua, Varese, Como and Lecco), we choose to fix $I_{T,0} = 1$ in order to assign an effective uncertainty.

Based on the estimations reported in [30], the expected initial amount of exposed $E_{T,0}$ and asymptomatic/mildly symptomatic individuals $A_{T,0}$ is imposed so that $E_{T,0} = 10\,I_{T,0}$ and $A_{T,0} = 9\,I_{T,0}$ in each location. Therefore, also initial conditions for compartments E, A and S become stochastic, depending on the initial amount of severe infectious at each location:

$$E_T(0, z) = 10\,I_T(0, z), \qquad A_T(0, z) = 9\,I_T(0, z),$$

$$S_T(0, z) = N - E_T(0, z) - I_T(0, z) - A_T(0, z).$$

Finally, removed individuals are initially set null everywhere in the network, $R_T(0, z) = 0.0$.

To properly subdivide the population in commuters and non-commuters, regional mobility data are considered. In particular, the matrix of commuters presented in Table A.2, Appendix A.2, reflects mobility data provided by the Lombardy Region for the regional fluxes of year 2020 [60], which is in agreement with the one derived from ISTAT data released in October, 2011, as also confirmed in [4]. Therefore, to each control volume, the total percentage of commuters referred to the province where it is located is assigned, and non-commuting individuals are computed as a result of conservation principles. From Table A.2 it can be noticed that some connections are not taken into account simply because the amount of commuters along these routes is negligible if compared to the amount of individuals traveling in the other paths and with respect to the dimension of the populations.

Concerning epidemiological parameters, accordingly with values reported in [4, 9, 30], we set $\gamma_I = 1/14$, $\gamma_A = 2\gamma_I = 1/7$, $a = 1/3$, considering these parameters as clinical ones and therefore deterministic. We also assume the probability rate of developing severe symptoms $\sigma = 1/12.5$, as in [9, 30].

From the first day simulated in this test, the population was aware of the risk associated with COVID-19 and recommendations such as washing hands often, not touching their faces, avoiding handshakes and keeping distances had already been disseminated by the government, hence, we initially fix coefficients $k_I = k_A = 50$.

The initial value of the transmission rate of asymptomatic/mildly infectious people is calibrated as the result of a least square problem with respect to the observed cumulative number of infected $I_T(t)$, through a deterministic SEIAR ODE model set up for the whole Lombardy Region, which provide the estimate $\beta_A = 0.58 \times 10^{-3}$. As previously mentioned, since we are assuming that highly infectious subjects are mostly detected in the most optimistic scenario, being subsequently quarantined or hospitalized, the transmission rate of I is set $\beta_I = 0.03\,\beta_A$, as in [4, 9, 30]. Finally, in the incidence function we fix $p = 1$.

With this parametric setup, we obtain an initial expected value of the basic reproduction number for the Region, evaluated as from definition Eq (2.10), $R_0(0) = 3.2$, which is in agreement with estimations reported in [4, 9, 30, 61]. Nevertheless, with the proposed methodology it is possible to present the heterogeneity underlying the basic reproduction number at the local scale, as shown in Figure 2 for $\mu = 0$, together with the initial global amount of infected people $E_{T,0}(x, y) + I_{T,0}(x, y) + A_{T,0}(x, y)$ present in the domain.

To model the effects of the lockdown imposed by the government from March 9, 2020 in north of Italy [14], from that day the transmission rate β_A is reduced by 50%, also increasing the coefficients $k_I = k_A = 80$ as a result of the population becoming increasingly aware of the epidemic risks. In addition, the percentage of commuting individuals is reduced by 60% for each compartment according

to mobility data tracked through GPS systems of mobile phones and made temporarily available by Google [61, 62].

3.3. Results and discussion

Numerical results of the test are reported in Figures 3–7. In Figure 3, the expected evolution in time of the infected individuals, together with 95% confidence intervals, is shown for exposed E, highly symptomatic subjects I and asymptomatic or weakly symptomatic people A, for each city of the Lombardy Region. Here it is already appreciable the heterogeneity of the diffusion of the virus. Indeed, from the different y-axis scales adopted for the plot of the provinces, it can be noticed that Milan, Bergamo and Brescia present a consistently higher contagion.

From Figure 4 it can be observed that the lower bound of the confidence interval of the cumulative amount in time of highly symptomatic individuals is in line with data reported by the Civil Protection Department of Italy [53]. As expected, due to the uncertainty taken into account, the mean value of the numerical result in each city is higher than the registered one.

The comparison between the expected evolution in time of the cumulative amount of severe infectious with respect to the effective cumulative amount of total infectious people, including asymptomatic and mildly symptomatic individuals, is shown in Figure 5. From this figure it is clear that the number of infections recorded during the first outbreak of COVID-19 in Lombardy represents a clear underestimation of the actual trend of infection suffered by the Region and by Italy as a whole, and how the presence of asymptomatic subjects, not detected, has affected the pandemic evolution.

Numerical results presented in terms of integrated variables for the whole Lombardy are shown in Figure 6, together with the expected temporal evolution of the reproduction number $R_0(t)$, again with 95% confidence bands. It is here highlighted that the drop of $R_0(t)$ on March 9 reflects the imposition of lockdown restrictions, as presented in the previous Section. Moreover, in this plot, results are reported starting from February 28, instead of February 27, to permit to the system to achieve a correct initialization of the commuters (who are totally placed in the origin location at the beginning of the simulation) during the first day simulated.

In Figure 7, final expectation and variance of the cumulative amount of infected people, namely $E_T + A_T + I_T$, are reported in the 2D framework (first row). If comparing Figure 7a with 2a, it can be noticed that, at the end of March, the virus is no longer majorly affecting the province of Lodi and Cremona, but has been spread arriving to hit most of all Brescia, Milan and Bergamo. Finally, in the second row of Figure 7, the expectation of the susceptible population S_T on the initial day simulated (February 27, 2020) is compared with the one obtained at the end of the simulation (March 22, 2020). Here it can be verified that the majority of the population does not leave their home city, as per real behavior, but there is only a small percentage of commuters who move within the domain, along the prescribed routes. Similar results are obtained for compartments E and A, whose commuting part, even though small, strongly contributes to the spatial spread of the epidemic.

4. Conclusions

In this paper we introduced a realistic model for the spatial spread of a virus with a focus on the case of COVID-19. Unlike models currently in the literature, which typically ignore spatial details or alternatively introduce them as simple diffusive dynamics, in our approach we have tried to capture
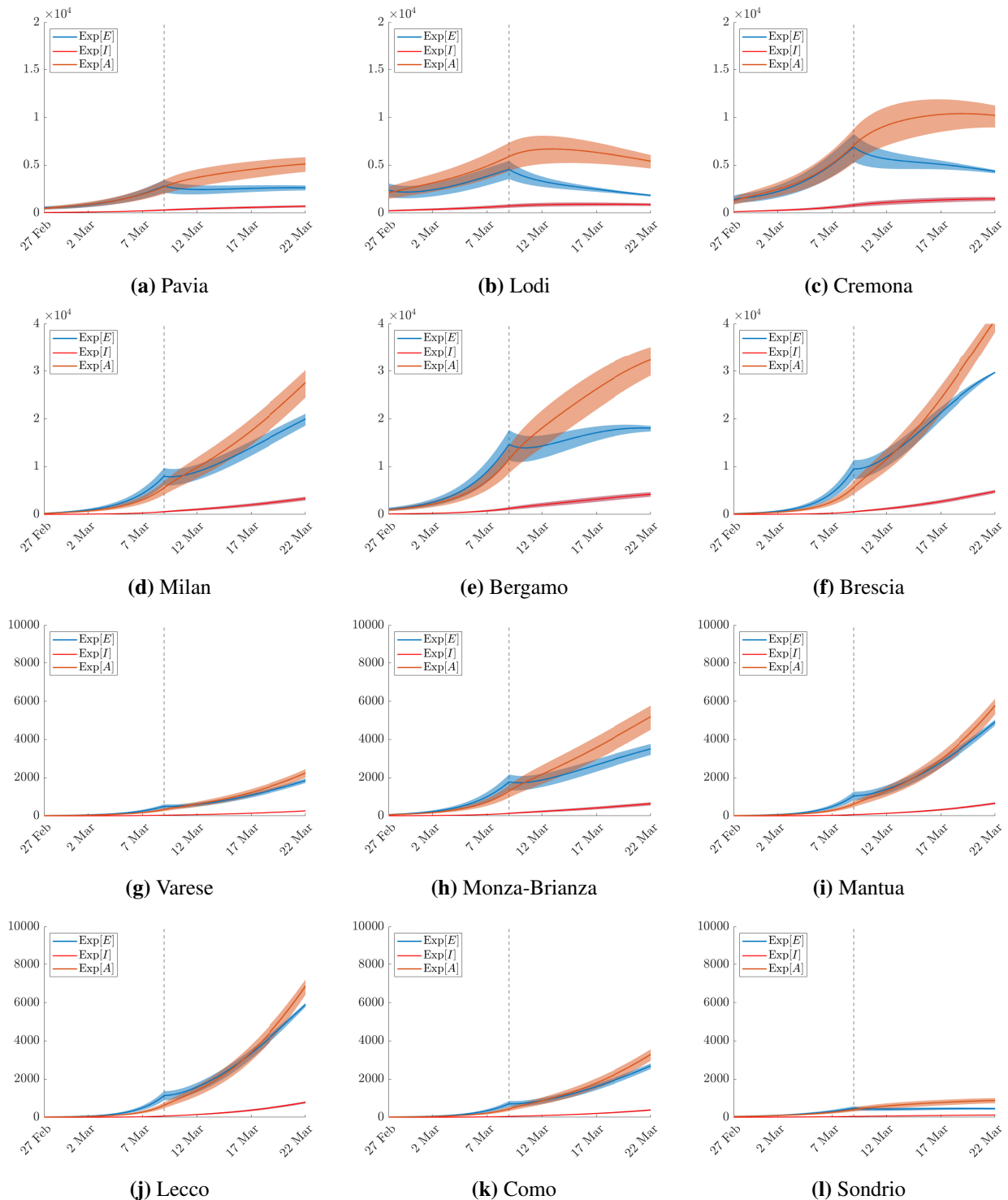
Figure 3. Numerical results, with 95% confidence intervals, of the simulation of the first outbreak of COVID-19 in Lombardy, Italy. Expected evolution Exp[·] in time of compartments E, A, I. Vertical dashed lines identify the onset of governmental lockdown restrictions.
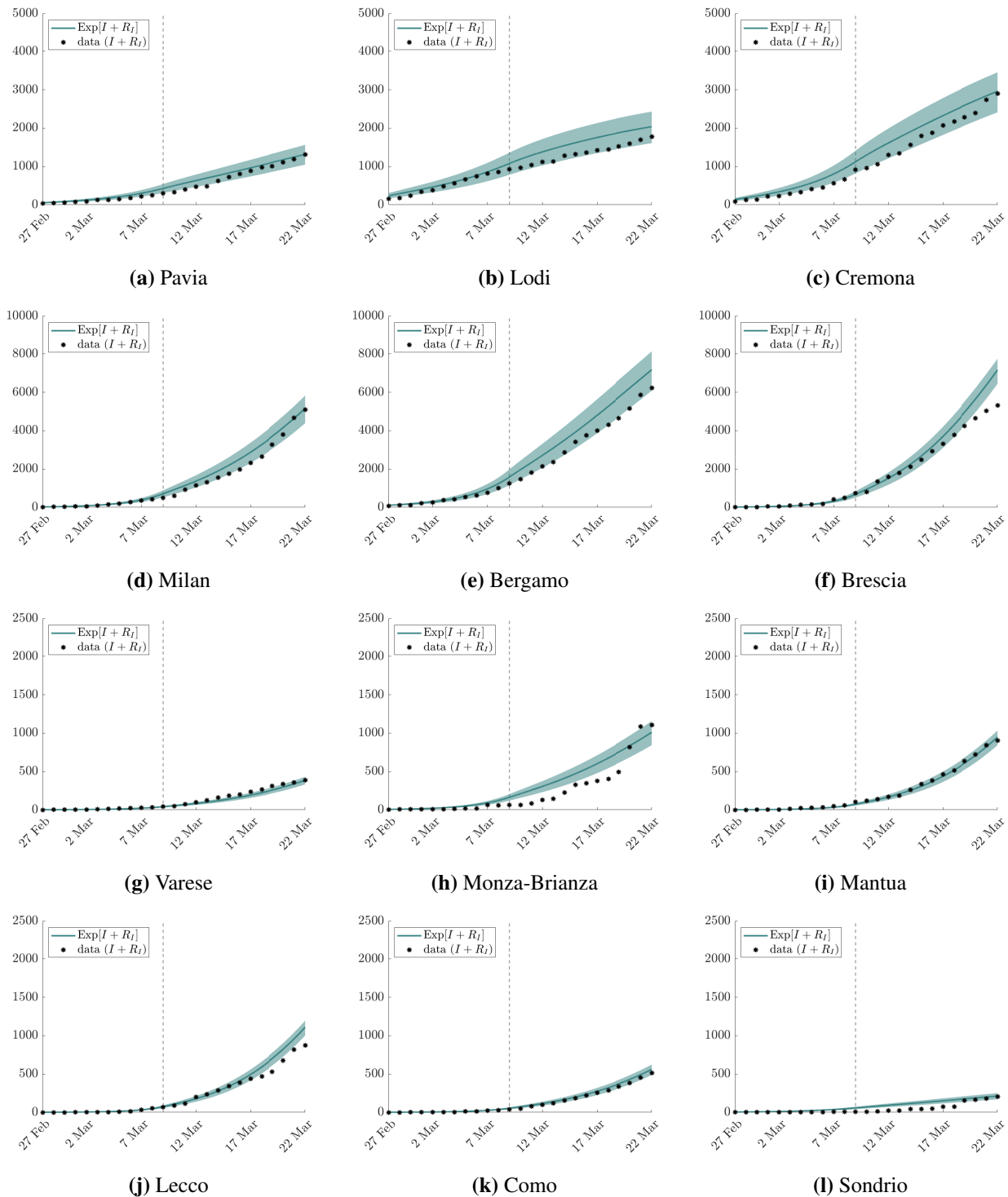
Figure 4. Numerical results, with 95% confidence intervals, of the simulation of the first outbreak of COVID-19 in Lombardy, Italy. Expected evolution Exp[·] in time of the cumulative amount of severe infectious $(I + R_I)$ compared with data of cumulative infectious taken from the COVID-19 repository of the Civil Protection Department of Italy [53]. Vertical dashed lines identify the onset of governmental lockdown restrictions.

Figure 5. Numerical results, with 95% confidence intervals, of the simulation of the first outbreak of COVID-19 in Lombardy, Italy. Expected evolution Exp[·] in time of the cumulative amount of severe infectious $(I + R_I)$ with respect to the effective cumulative amount of total infectious people, including asymptomatic and mildly symptomatic individuals $(I + A + R)$. Data of cumulative infectious is taken from the COVID-19 repository of the Civil Protection Department of Italy [53]. Vertical dashed lines identify the onset of governmental lockdown restrictions.
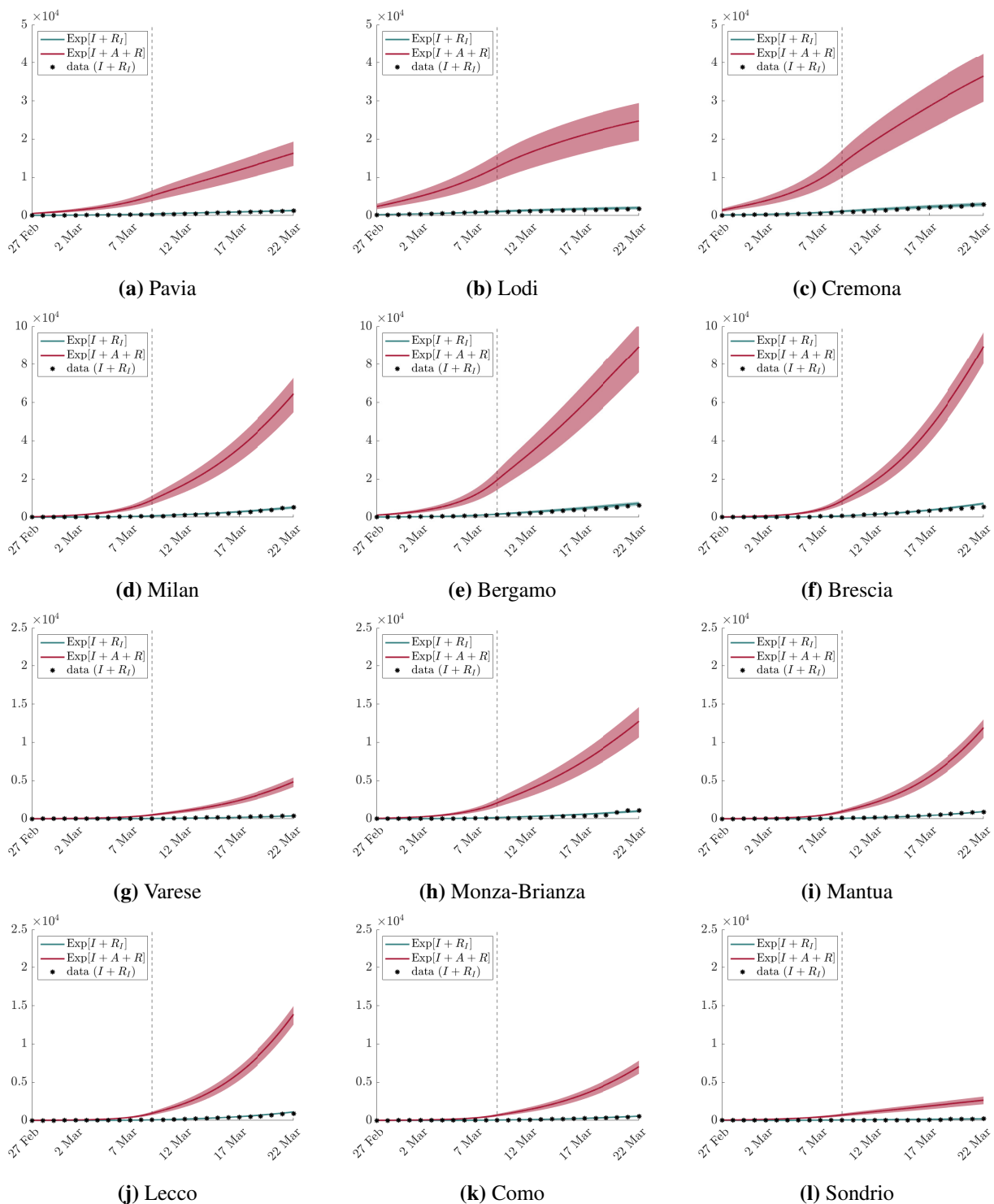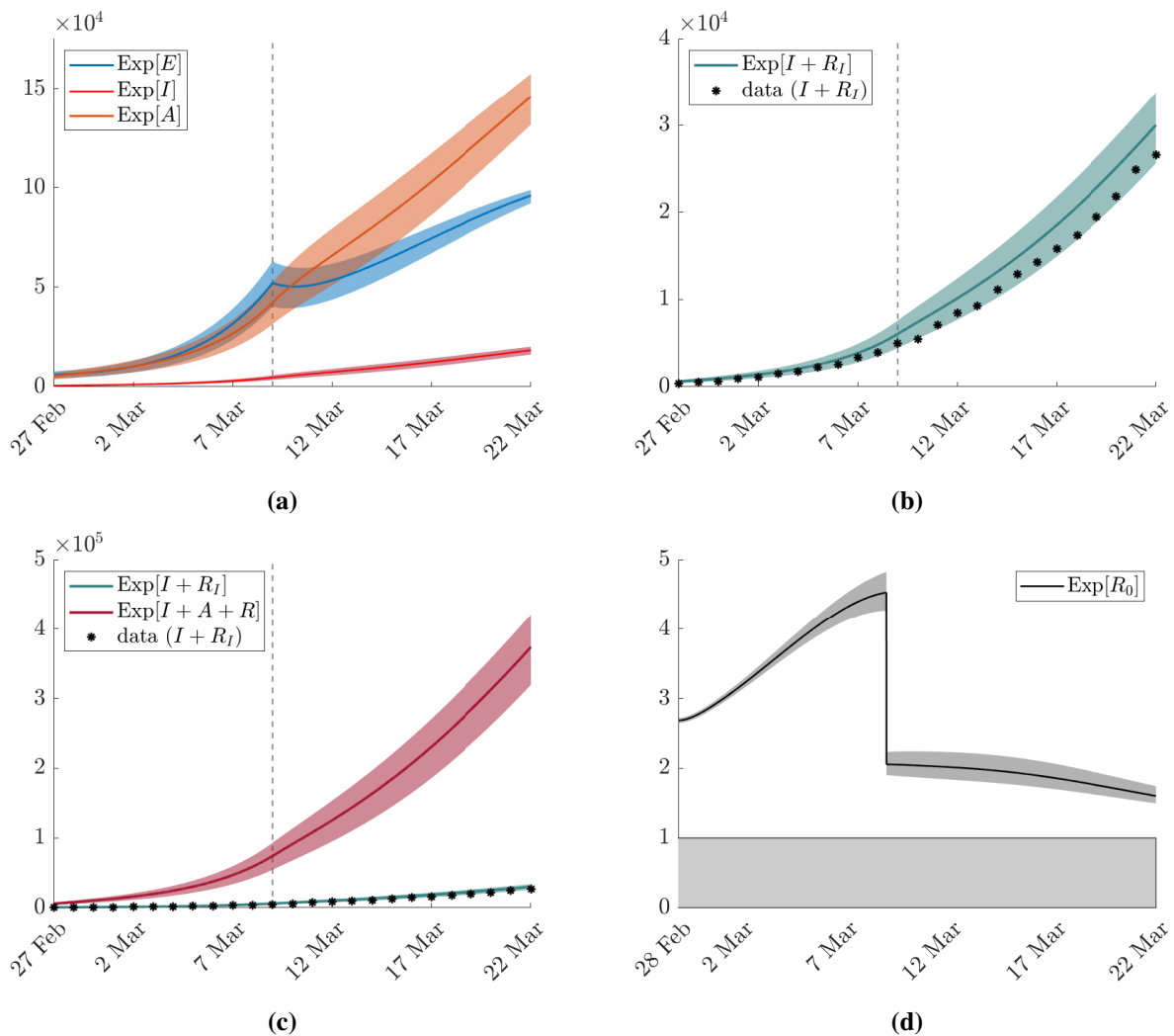
Figure 6. Numerical results, with 95% confidence intervals, of the simulation of the first outbreak of COVID-19 in Lombardy, Italy. Expected evolution Exp[·] in time, for the whole Region, of: compartments E, A, I (a); cumulative amount of severe infectious $(I + R_I)$ compared with data of cumulative infectious (b); cumulative amount of severe infectious $(I + R_I)$ with respect to the effective cumulative amount of total infectious people, including asymptomatic and mildly symptomatic individuals $(I + A + R)$ (c); reproduction number $R_0(t)$ (d). Vertical dashed lines identify the onset of governmental lockdown restrictions.

the essential characteristics of the movements of individuals, which are very different if we consider commuting individuals who for work reasons move over long distances, from one city to another, to individuals who instead carry out their activities on an urban scale. The separation of individuals into these two classes, and the use of different spatial dynamics characterized by appropriate systems of transport and diffusive equations, allows in particular to avoid mass migration phenomena typical of models based on a single population and the instantaneous propagation of infectious disease over long distances. From the epidemiological point of view, modeling is developed in a compartmental context described by a SEIAR-type framework capable of describing the effect of exposed and asymptomatic
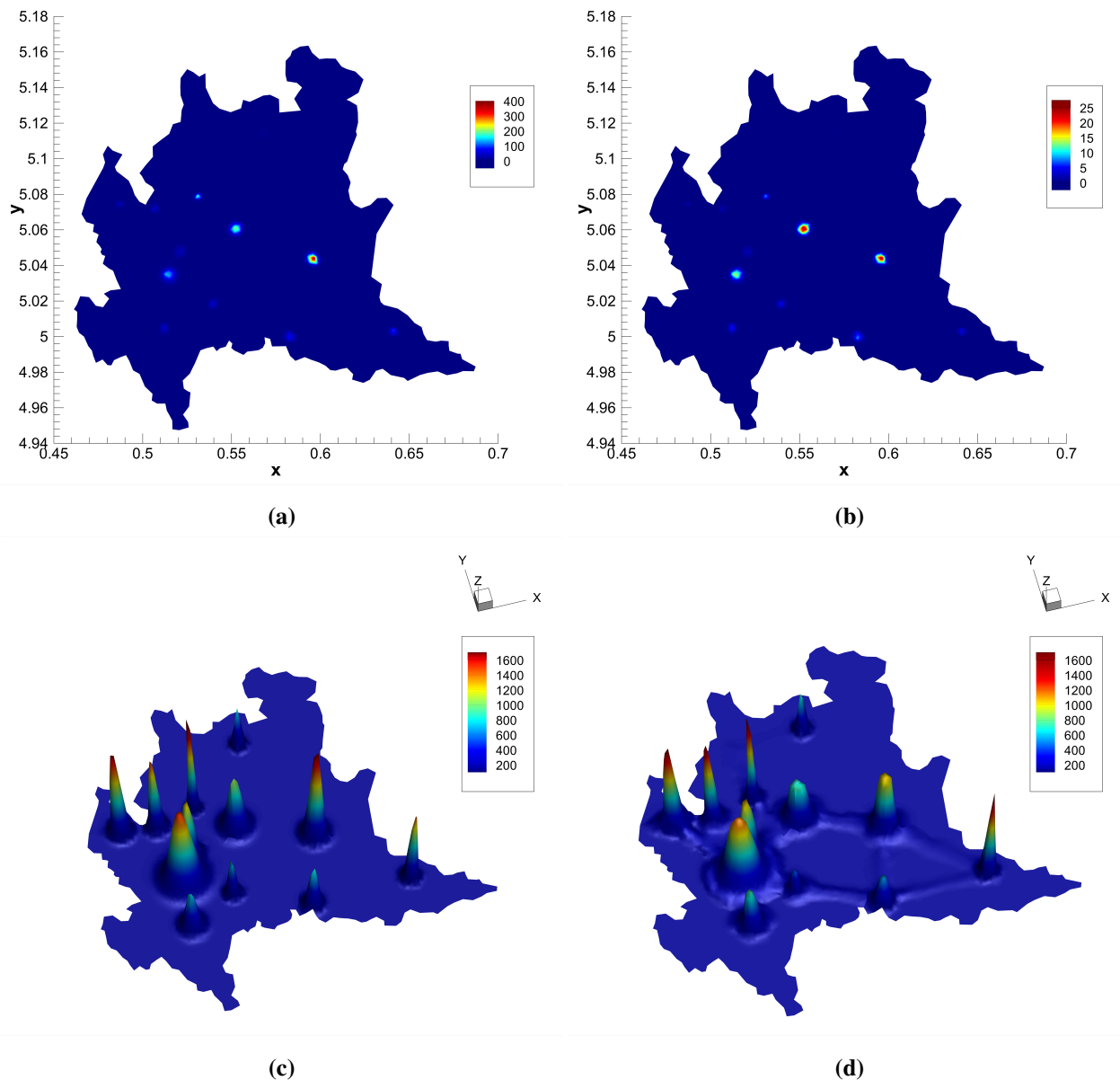
Figure 7. Numerical results of the simulation of the first outbreak of COVID-19 in Lombardy, Italy. Top: expectation (a) and variance (b) of the cumulative amount of infected people $E_T + A_T + I_T$ at the end of the simulation (March 22, 2020). Bottom: expectation of the susceptible population S_T on the initial day simulated (February 27, 2020) (c) and at the end of the simulation (March 22, 2020) (d).

individuals within the spatial spread of the disease. In addition, given the high uncertainty on the actual amount of individuals in the various compartments able to propagate the infection, the model was developed taking into account the presence of stochastic variables that therefore require an appropriate process of quantification. The resulting multiscale system of partial differential equations was then solved on realistic spatial geometries by a numerical method combining finite volume IMEX techniques for the deterministic part, with a non-intrusive collocation approach for the stochastic com-

ponent. After a careful calibration of the model parameters based on the available data, an in-depth analysis of the results is reported in the case of the initial phase of the spread of COVID-19 in Italy occurred in the Lombardy region. The results show the ability of the model to correctly describe the epidemic dynamics and the importance of a heterogeneous spatial description and of the inclusion of stochastic parameters.

Acknowledgments

Conflict of interest

The Authors declare that they have no conflict of interest.

A. Supplementary material

In this appendix, we report the details of the numerical scheme adopted to approximate the MK-SEIAR Eqs (2.1) and (2.2) together with the data tables concerning the details of the population and of the commuting flows.

A.1. Numerical method

We first give the details of the method in the case in which uncertainty is not present and successively we explain how the Eqs (2.1) and (2.2) is solved in the case of stochasticity. For the commuters, the numerical scheme for the deterministic case is based on a discrete ordinate method in velocity in which the even and odd parity formulation is employed [46, 51, 63]. The details of such approach are given in A.1.1. Then a finite volume method working on two-dimensional unstructured meshes [48, 64] for the discrete ordinate approximation of the commuters is introduced in A.1.2. The full discretization of the Eq (2.1) is obtained through the use of suitable IMEX Runge-Kutta schemes [49, 50]. In particular, the above choices permit to obtain a scheme which consistently captures the diffusion limit from the kinetic system when the scaling parameters $\tau_{S,I,R}$ tends toward zero. This part of the method is discussed in A.1.3. Finally, the discretization of the stochastic part for the Eq (2.1) and (2.2) when uncertainty is present is explained in A.1.4.

A.1.1. Even and odd parities formulation

We rewrite Eq (2.1) by using the so-called even and odd parities formulation. In other words, we denote $v = (\eta, \xi) \in \mathbb{S}^1$ and then we obtain four equations with non-negative $\xi, \eta \geq 0$ for each compartment of the commuters. The change of variables reads, omitting the time and space dependence for simplicity, as [46]

$$r_i^{(1)}(\xi, \eta) = \frac{1}{2}(f_i(\xi, -\eta) + f_i(-\xi, \eta)), \quad r_i^{(2)}(\xi, \eta) = \frac{1}{2}(f_i(\xi, \eta) + f_i(-\xi, -\eta))$$

while for the scalar fluxes one has

$$j_i^{(1)}(\xi,\eta) = \frac{\lambda_i}{2}(f_i(\xi,-\eta) + f_i(-\xi,\eta)), \quad j_i^{(2)}(\xi,\eta) = \frac{\lambda_i}{2}(f_i(\xi,\eta) + f_i(-\xi,-\eta))$$

with $i = S, E, I, A, R$. An equivalent formulation with respect to Eq (2.1) can then be obtained thanks to this change of variables and reads as

$$\frac{\partial r_S^{(1)}}{\partial t} + \xi\frac{\partial j_S^{(1)}}{\partial x} - \eta\frac{\partial j_S^{(1)}}{\partial y} = -F_I(r_S^{(1)}, I_T) - F_A(r_S^{(1)}, A_T) + \frac{1}{\tau_S}\left(S - r_S^{(1)}\right)$$

$$\frac{\partial r_S^{(2)}}{\partial t} + \xi\frac{\partial j_S^{(2)}}{\partial x} + \eta\frac{\partial j_S^{(2)}}{\partial y} = -F_I(r_S^{(2)}, I_T) - F_A(r_S^{(2)}, A_T) + \frac{1}{\tau_S}\left(S - r_S^{(2)}\right)$$

$$\frac{\partial r_E^{(1)}}{\partial t} + \xi\frac{\partial j_E^{(1)}}{\partial x} - \eta\frac{\partial j_E^{(1)}}{\partial y} = F_I(r_S^{(1)}, I_T) + F_A(r_S^{(1)}, A_T) - ar_E^{(1)} + \frac{1}{\tau_E}\left(E - r_E^{(1)}\right)$$

$$\frac{\partial r_E^{(2)}}{\partial t} + \xi\frac{\partial j_E^{(2)}}{\partial x} + \eta\frac{\partial j_E^{(2)}}{\partial y} = F_I(r_S^{(2)}, I_T) + F_A(r_S^{(2)}, A_T) - ar_E^{(2)} + \frac{1}{\tau_E}\left(E - r_E^{(2)}\right)$$

$$\frac{\partial r_I^{(1)}}{\partial t} + \xi\frac{\partial j_I^{(1)}}{\partial x} - \eta\frac{\partial j_I^{(1)}}{\partial y} = a\sigma r_E^{(1)} - \gamma_I r_I^{(1)} + \frac{1}{\tau_I}\left(I - r_I^{(1)}\right)$$

$$\frac{\partial r_I^{(2)}}{\partial t} + \xi\frac{\partial j_S^{(2)}}{\partial x} + \eta\frac{\partial j_I^{(2)}}{\partial y} = a\sigma r_E^{(2)} - \gamma_I r_I^{(2)} + \frac{1}{\tau_I}\left(I - r_I^{(2)}\right) \tag{A.1}$$

$$\frac{\partial r_A^{(1)}}{\partial t} + \xi\frac{\partial j_A^{(1)}}{\partial x} - \eta\frac{\partial j_A^{(1)}}{\partial y} = a(1-\sigma)r_E^{(1)} - \gamma_A r_A^{(1)} + \frac{1}{\tau_A}\left(A - r_A^{(1)}\right)$$

$$\frac{\partial r_A^{(2)}}{\partial t} + \xi\frac{\partial j_A^{(2)}}{\partial x} + \eta\frac{\partial j_A^{(2)}}{\partial y} = a(1-\sigma)r_E^{(2)} - \gamma_A r_A^{(2)} + \frac{1}{\tau_A}\left(A - r_A^{(2)}\right)$$

$$\frac{\partial r_R^{(1)}}{\partial t} + \xi\frac{\partial j_R^{(1)}}{\partial x} - \eta\frac{\partial j_R^{(1)}}{\partial y} = \gamma_I r_I^{(1)} + \gamma_A r_A^{(1)} + \frac{1}{\tau_R}\left(R - r_R^{(1)}\right)$$

$$\frac{\partial r_R^{(2)}}{\partial t} + \xi\frac{\partial j_R^{(2)}}{\partial x} + \eta\frac{\partial j_R^{(2)}}{\partial y} = \gamma_I r_I^{(2)} + \gamma_A r_A^{(2)} + \frac{1}{\tau_R}\left(R - r_R^{(2)}\right)$$

and

$$\frac{\partial j_S^{(1)}}{\partial t} + \lambda_S^2\xi\frac{\partial r_S^{(1)}}{\partial x} - \lambda_S^2\eta\frac{\partial r_S^{(1)}}{\partial y} = -F_I(j_S^{(1)}, I_T) - F_A(j_S^{(1)}, A_T) - \frac{1}{\tau_S}j_S^{(1)}$$

$$\frac{\partial j_S^{(2)}}{\partial t} + \lambda_S^2\xi\frac{\partial r_S^{(2)}}{\partial x} + \lambda_S^2\eta\frac{\partial r_S^{(2)}}{\partial y} = -F_I(j_S^{(2)}, I_T) - F_A(j_S^{(1)}, A_T) - \frac{1}{\tau_S}j_S^{(2)}$$

$$\frac{\partial j_E^{(1)}}{\partial t} + \lambda_E^2\xi\frac{\partial r_E^{(1)}}{\partial x} - \lambda_E^2\eta\frac{\partial r_E^{(1)}}{\partial y} = \frac{\lambda_E}{\lambda_S}\left(F_I(j_S^{(1)}, I_T) + F_A(j_S^{(1)}, A_T)\right) - aj_E^{(1)} - \frac{1}{\tau_E}j_E^{(1)}$$

$$\frac{\partial j_E^{(2)}}{\partial t} + \lambda_E^2\xi\frac{\partial r_E^{(2)}}{\partial x} + \lambda_E^2\eta\frac{\partial r_E^{(2)}}{\partial y} = \frac{\lambda_E}{\lambda_S}\left(F_I(j_S^{(2)}, I_T) + F_A(j_S^{(2)}, A_T)\right) - aj_E^{(2)} - \frac{1}{\tau_E}j_E^{(2)}$$

$$\frac{\partial j_I^{(1)}}{\partial t} + \lambda_I^2\xi\frac{\partial r_I^{(1)}}{\partial x} - \lambda_I^2\eta\frac{\partial r_I^{(1)}}{\partial y} = \frac{\lambda_I}{\lambda_E}a\sigma j_E^{(1)} - \gamma_I j_I^{(1)} - \frac{1}{\tau_I}j_I^{(1)}$$

$$\frac{\partial j_I^{(2)}}{\partial t} + \lambda_I^2 \xi \frac{\partial r_I^{(2)}}{\partial x} + \lambda_I^2 \eta \frac{\partial r_I^{(2)}}{\partial y} = \frac{\lambda_I}{\lambda_E} a\sigma j_E^{(2)} - \gamma_I j_I^{(2)} - \frac{1}{\tau_I} j_I^{(2)}. \tag{A.2}$$

$$\frac{\partial j_A^{(1)}}{\partial t} + \lambda_A^2 \xi \frac{\partial r_A^{(1)}}{\partial x} - \lambda_A^2 \eta \frac{\partial r_A^{(1)}}{\partial y} = \frac{\lambda_A}{\lambda_E} a(1-\sigma) j_E^{(1)} - \gamma_A j_A^{(1)} - \frac{1}{\tau_A} j_A^{(1)}$$

$$\frac{\partial j_A^{(2)}}{\partial t} + \lambda_A^2 \xi \frac{\partial r_A^{(2)}}{\partial x} + \lambda_A^2 \eta \frac{\partial r_A^{(2)}}{\partial y} = \frac{\lambda_A}{\lambda_E} a(1-\sigma) j_E^{(2)} - \gamma_A j_A^{(2)} - \frac{1}{\tau_A} j_A^{(2)}$$

$$\frac{\partial j_R^{(1)}}{\partial t} + \lambda_R^2 \xi \frac{\partial r_R^{(1)}}{\partial x} - \lambda_R^2 \eta \frac{\partial r_R^{(1)}}{\partial y} = \frac{\lambda_R}{\lambda_I} \gamma_I j_I^{(1)} + \frac{\lambda_R}{\lambda_A} \gamma_A j_A^{(1)} - \frac{1}{\tau_R} j_R^{(1)}$$

$$\frac{\partial j_R^{(2)}}{\partial t} + \lambda_R^2 \xi \frac{\partial r_R^{(2)}}{\partial x} + \lambda_R^2 \eta \frac{\partial r_R^{(2)}}{\partial y} = \frac{\lambda_R}{\lambda_I} \gamma_I j_I^{(2)} + \frac{\lambda_R}{\lambda_A} \gamma_A j_A^{(2)} - \frac{1}{\tau_R} j_R^{(2)}.$$

One can observe that due to symmetry, we need to solve these equations for ξ, η in the positive quadrant only. Thus the number of unknowns in Eq (2.1) and Eqs (A.1) and (A.2) is effectively the same.

A.1.2. Space discretization on unstructured grids

We consider now a spatial two-dimensional computational domain Ω which is discretized by a set of non overlapping polygons $P_i, i = 1, \ldots N_p$. Each element P_i exhibits an arbitrary number N_{S_i} of edges $e_{j,i}$ where the subscripts indicates that this is the edge shared by elements P_i and P_j. The boundary of the cell is consequently given by $\partial P_i = \bigcup_{j=1}^{N_{S_i}} e_{ji}$. The governing equations for the commuters rewritten in the odd and even formulation are then discretized on the unstructured mesh by means of a finite volume scheme which is conveniently rewritten in condensed form as

$$\frac{\partial \mathbf{Q}}{\partial t} + \nabla_x \cdot \mathbf{F}(\mathbf{Q}) = \mathbf{S}(\mathbf{Q}), \qquad (x, y) \in \Omega \subset \mathbb{R}^2, \quad t \in \mathbb{R}_0^+, \tag{A.3}$$

where $\mathbf{Q}$ is the vector of conserved variables

$$\mathbf{Q} = \left(r_i^{(1)}, r_i^{(2)}, j_i^{(1)}, j_i^{(2)} \right)^\top, \quad i = S, E, I, A, R$$

while $\mathbf{F}(\mathbf{Q})$ is the linear flux tensor and $\mathbf{S}(\mathbf{Q})$ represents the stiff source term defined in Eqs (A.1) and (A.2). As usual for finite volume schemes, data are represented by spatial cell averages, which are defined at time t^n as

$$\mathbf{Q}_i^n = \frac{1}{|P_i|} \int_{P_i} \mathbf{Q}(\mathbf{x}, t^n) \, d\mathbf{x}, \tag{A.4}$$

where $|P_i|$ denotes the surface of element P_i.

A first order in time finite volume method is then obtained by integration of the governing Eq (A.3) over the space control volume $|P_i|$

$$\mathbf{Q}_i^{n+1} = \mathbf{Q}_i^n - \frac{\Delta t}{|P_i|} \sum_{P_j \in \mathcal{N}_{S_i}} \int_{e_{ij}} \mathbf{F}_{ij}^n \cdot \mathbf{n}_{ij} \, dl + \int_{P_i} \mathbf{S}_i^n \, d\mathbf{x}. \tag{A.5}$$

Higher order in space is then achieved by substituting the cell averages by piecewise high order polynomials. We refer to these polynomial reconstructions to as $\mathbf{w}_i(\mathbf{x})$ which are obtained from the given

cell averages Eq (A.4). In particular, our choice is to rely on a second order Central WENO (CWENO) reconstruction procedure along the lines of [64, 65]. We omit the details for brevity. The numerical flux function $\mathbf{F}_{ij} \cdot \mathbf{n}_{ij}$ is given by a simple and robust local Lax-Friedrichs flux yielding

$$\mathbf{F}_{ij} \cdot \mathbf{n}_{ij} = \frac{1}{2} \left(\mathbf{F}(\mathbf{w}_{i,j}^+) + \mathbf{F}(\mathbf{w}_{i,j}^-) \right) \cdot \mathbf{n}_{ij} - \frac{1}{2} s_{\max} \left(\mathbf{w}_{i,j}^+ - \mathbf{w}_{i,j}^- \right), \tag{A.6}$$

where $\mathbf{w}_{i,j}^+, \mathbf{w}_{i,j}^-$ are the high order boundary extrapolated data evaluated through the CWENO reconstruction procedure. The numerical dissipation is given by $s_{\max}$ which is the maximum eigenvalue of the Jacobian matrix in spatial normal direction,

$$\mathbf{A_n} = \frac{\partial \mathbf{F}}{\partial \mathbf{Q}}. \tag{A.7}$$

Let us notice now that in the diffusion limit, i.e., as $(\tau_S, \tau_I, \tau_R) \to 0$, the source term $\mathbf{S}(\mathbf{Q})$ becomes stiff, therefore in order to avoid prohibitive time steps we need to discretize the commuters system implicitly. To this aim, a second order IMEX method which preserves the asymptotic limit given by the diffusion Eq (2.9) is proposed and briefly described hereafter.

A.1.3. Time integration and numerical diffusion limit

We consider again Eq (2.1) formulated using the parities Eqs (A.1) and (A.2). We also assume $\tau_{S,I,R} = \tau$ and rewrite Esq (A.1) and (A.2) in partitioned form as

$$\begin{aligned}
\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{f}(\mathbf{v})}{\partial x} + \frac{\partial \mathbf{g}(\mathbf{v})}{\partial y} &= \mathbf{E}(\mathbf{u}) + \frac{1}{\tau} (\mathbf{U} - \mathbf{u}) \\
\frac{\partial \mathbf{v}}{\partial t} + \mathbf{\Lambda}^2 \frac{\partial \mathbf{f}(\mathbf{u})}{\partial x} + \mathbf{\Lambda}^2 \frac{\partial \mathbf{g}(\mathbf{u})}{\partial y} &= \mathbf{E}(\mathbf{v}) - \frac{1}{\tau} \mathbf{v},
\end{aligned} \tag{A.8}$$

in which

$$\begin{aligned}
\mathbf{u} &= \left(r_S^{(1)}, r_S^{(2)}, r_E^{(1)}, r_E^{(2)}, r_I^{(1)}, r_I^{(2)}, r_A^{(1)}, r_A^{(2)}, r_R^{(1)}, r_R^{(2)} \right)^T, \\
\mathbf{v} &= \left(j_S^{(1)}, j_S^{(2)}, j_E^{(1)}, j_E^{(2)}, j_I^{(1)}, j_I^{(2)}, j_A^{(1)}, j_A^{(2)}, j_R^{(1)}, j_R^{(2)} \right)^T, \\
\mathbf{f}(\mathbf{v}) &= \xi \mathbf{v}, \quad \mathbf{g}(\mathbf{v}) = \eta \mathbf{J} \mathbf{v}, \quad \mathbf{J} = \mathrm{diag}\{-1, 1, -1, 1, -1, 1, -1, 1, -1, 1\}, \\
\mathbf{E}(\mathbf{u}) &= \left(-F_I(r_S^{(i)}, I_T) - F_A(r_S^{(i)}, A_T), F_I(r_S^{(i)}, I_T) + F_A(r_S^{(i)}, A_T) - a r_E^{(i)}, a\sigma r_E^{(i)} - \gamma_I r_I^{(i)}, \right. \\
&\quad \left. a(1 - \sigma) r_E^{(i)} - \gamma_A r_A^{(i)}, \gamma_I r_I^{(i)} \gamma_A r_A^{(i)} \right)^T, \quad i = 1, 2 \\
\mathbf{U} &= (S, S, E, E, I, I, A, A, R, R)^T, \quad \mathbf{\Lambda} = \mathrm{diag}\{\lambda_S, \lambda_S, \lambda_E, \lambda_E, \lambda_I, \lambda_I, \lambda_A, \lambda_A, \lambda_R, \lambda_R\},
\end{aligned} \tag{A.9}$$

and $\mathbf{f}(\mathbf{u})$, $\mathbf{g}(\mathbf{u})$, $\mathbf{E}(\mathbf{v})$ are defined similarly. Now, following [50], the Implicit-Explict Runge-Kutta (IMEX-RK) approach appid to system (A.8) reads as

$$\begin{aligned}
\mathbf{u}^{(k)} &= \mathbf{u}^n - \Delta t \sum_{j=1}^{k} a_{kj} \left(\frac{\partial \mathbf{f}(\mathbf{v}^{(j)})}{\partial x} + \frac{\partial \mathbf{g}(\mathbf{v}^{(j)})}{\partial y} - \frac{1}{\tau} \left(\mathbf{U}^{(j)} - \mathbf{u}^{(j)} \right) \right) + \Delta t \sum_{j=1}^{k-1} \tilde{a}_{kj} \mathbf{E} \left(\mathbf{u}^{(j)} \right) \\
\mathbf{v}^{(k)} &= \mathbf{v}^n - \Delta t \sum_{j=1}^{k-1} \tilde{a}_{kj} \left(\mathbf{\Lambda}^2 \frac{\partial \mathbf{f}(\mathbf{u}^{(j)})}{\partial x} + \mathbf{\Lambda}^2 \frac{\partial \mathbf{g}(\mathbf{u}^{(j)})}{\partial y} - \mathbf{E}(\mathbf{v}^{(j)}) \right) + \Delta t \sum_{j=1}^{k} a_{kj} \frac{1}{\tau} \mathbf{v}^{(j)},
\end{aligned} \tag{A.10}$$

where $\mathbf{u}^{(k)}$, $\mathbf{v}^{(k)}$ are the so-called internal stages. The numerical solution reads

$$
\mathbf{u}^{n+1} = \mathbf{u}^n - \Delta t \sum_{k=1}^{s} b_k \left(\frac{\partial \mathbf{f}(\mathbf{v}^{(k)})}{\partial x} + \frac{\partial \mathbf{g}(\mathbf{v}^{(k)})}{\partial y} - \frac{1}{\tau} \left(\mathbf{U}^{(k)} - \mathbf{u}^{(k)} \right) \right) + \Delta t \sum_{k=1}^{s} \tilde{b}_k \mathbf{E} \left(\mathbf{u}^{(k)} \right)
$$

$$
\mathbf{v}^{n+1} = \mathbf{v}^n - \Delta t \sum_{k=1}^{s} \tilde{b}_k \left(\mathbf{\Lambda}^2 \frac{\partial \mathbf{f}(\mathbf{u}^{(k)})}{\partial x} + \mathbf{\Lambda}^2 \frac{\partial \mathbf{g}(\mathbf{u}^{(k)})}{\partial y} - \mathbf{E}(\mathbf{v}^{(k)}) \right) + \Delta t \sum_{k=1}^{s} b_k \frac{1}{\tau} \mathbf{v}^{(k)}.
$$

(A.11)

In the above equations, the matrices $\tilde{A} = (\tilde{a}_{kj})$, with $\tilde{a}_{kj} = 0$ for $j \geq k$, and $A = (a_{kj})$, with $a_{kj} = 0$ for $j > k$ are $s \times s$ matrices, with s number of Runge-Kutta stages, defining respectively the explicit and the implicit part of the scheme, and vectors $\tilde{b} = (\tilde{b}_1, ..., \tilde{b}_s)^T$ and $b = (b_1, ..., b_s)^T$ are the quadrature weights. Furthermore, we choose the Runge-Kutta scheme in such a way that the following relations hold true

$$
a_{kj} = b_j, \qquad j = 1, \dots, s, \qquad \tilde{a}_{kj} = \tilde{b}_j, \qquad j = 1, \dots, s - 1.
$$

(A.12)

The scheme Eqs (A.10) and (A.11) treats implicitly the stiff terms and explicitly all the rest. Moreover, one can prove that the above scheme is a consistent discretization of the limit system in the diffusive regime. In fact, assuming for simplicity $D_{S,I,R}$ independent from space, the second equation in Eq (A.10) can be rewritten as

$$
\tau \mathbf{v}^{(k)} = \tau \mathbf{v}^n - \Delta t \sum_{j=1}^{k-1} \tilde{a}_{kj} \left(\tau \mathbf{\Lambda}^2 \frac{\partial \mathbf{f}(\mathbf{u}^{(j)})}{\partial x} + \tau \mathbf{\Lambda}^2 \frac{\partial \mathbf{g}(\mathbf{u}^{(j)})}{\partial y} - \tau \mathbf{E}(\mathbf{v}^{(j)}) \right) + \Delta t \sum_{j=1}^{k} a_{kj} \mathbf{v}^{(j)},
$$

therefore, assuming Eq (2.5), the limit $\tau \to 0$ yields

$$
\sum_{j=1}^{k} a_{kj} \mathbf{v}^{(j)} = \sum_{j=1}^{k-1} \tilde{a}_{kj} \left(2\mathbf{D} \frac{\partial \mathbf{f}(\mathbf{U}^{(j)})}{\partial x} + 2\mathbf{D} \frac{\partial \mathbf{g}(\mathbf{U}^{(j)})}{\partial y} \right),
$$

(A.13)

where $\mathbf{D} = \text{diag} \{D_S, D_S, D_E, D_E, D_I, D_I, D_A, D_A, D_R, D_R\}$ and where we used the fact that from the first equation in Eq (A.10) as $\tau \to 0$ we have $\mathbf{u}^{(j)} = \mathbf{U}^{(j)}$. Note also that Eq (A.13) implies that $j^{(1)}_{S,E,I,A,R} = j^{(2)}_{S,E,I,A,R}$ in $\mathbf{v}^{(j)}$, i.e., we restore perfect symmetry in direction of propagation of the information. Using now the identity $\mathbf{u}^{(j)} = \mathbf{U}^{(j)}$ into the first equation in Eq (A.10) we get

$$
\mathbf{U}^{(k)} = \mathbf{U}^n - \Delta t \sum_{j=1}^{k} a_{kj} \left(\frac{\partial \mathbf{f}(\mathbf{v}^{(j)})}{\partial x} + \frac{\partial \mathbf{g}(\mathbf{v}^{(j)})}{\partial y} \right) + \Delta t \sum_{j=1}^{k-1} \tilde{a}_{kj} \mathbf{E} \left(\mathbf{U}^{(j)} \right),
$$

(A.14)

and using Eqs (A.13) into (A.14) thanks to the definitions of $\mathbf{f}$ and $\mathbf{g}$ gives

$$
\mathbf{U}^{(k)} = \mathbf{U}^n - 2\Delta t \mathbf{D} \sum_{j=1}^{k-1} \tilde{a}_{kj} \left(\xi^2 \frac{\partial^2 \mathbf{U}^{(j)}}{\partial x^2} + 2\xi \eta \mathbf{J} \frac{\partial^2 \mathbf{U}^{(j)}}{\partial x \partial y} + \eta^2 \frac{\partial^2 \mathbf{U}^{(j)}}{\partial y^2} \right)
$$

$$
+ \Delta t \sum_{j=1}^{k-1} \tilde{a}_{kj} \mathbf{E} \left(\mathbf{U}^{(j)} \right).
$$

(A.15)

Finally, integrating over the velocity field one has

$$
S^{(k)} = S^n - \Delta t D_S \sum_{j=1}^{k-1} \tilde{a}_{kj} \left(\frac{\partial^2 S^{(j)}}{\partial x^2} + \frac{\partial^2 S^{(j)}}{\partial y^2} \right) - \Delta t \sum_{j=1}^{k-1} \tilde{a}_{kj} F(S^{(j)}, I_T^{(j)}),
$$

$$
I^{(k)} = I^n - \Delta t D_I \sum_{j=1}^{k-1} \tilde{a}_{kj} \left(\frac{\partial^2 I^{(j)}}{\partial x^2} + \frac{\partial^2 I^{(j)}}{\partial y^2} \right) + \Delta t \sum_{j=1}^{k-1} \tilde{a}_{kj} \left(F(S^{(j)}, I_T^{(j)}) - \gamma I^{(j)} \right), \qquad \text{(A.16)}
$$

$$
R^{(k)} = R^n - \Delta t D_R \sum_{j=1}^{k-1} \tilde{a}_{kj} \left(\frac{\partial^2 R^{(j)}}{\partial x^2} + \frac{\partial^2 R^{(j)}}{\partial y^2} \right) + \Delta t \sum_{j=1}^{k-1} \tilde{a}_{kj} \gamma I^{(j)}
$$

and thus, the internal stages correspond to the stages of the explicit scheme applied to the reaction-diffusion Eq (2.9). To conclude the proof one has to notice that thanks to the choice Eq (A.12), the last stage is equivalent to the numerical solution. Thus, this is enough to guarantee that the scheme is a consistent discretization of the limit equation.

Note that the limit system is consistent with the discretization of the non commuters diffusive Eq (2.2). In this latter case, we adopt the same finite volume setting for the unknowns

$$
\mathbf{Q}^u = (S_u, E_u, I_u, A_u, R_u)^\top,
$$

This simply reads

$$
\frac{\partial \mathbf{Q}^u}{\partial t} + \nabla_x \cdot \mathbf{F}^u(\mathbf{Q}^u) = \mathbf{S}^u(\mathbf{Q}^u), \qquad (x, y) \in \Omega \subset \mathbb{R}^2, \quad t \in \mathbb{R}_0^+, \qquad \text{(A.17)}
$$

with

$$
\mathbf{F}^u = \begin{pmatrix} -D_S^u (S_u)_x & -D_S^u (S_u)_y \\ -D_E^u (E_u)_x & -D_E^u (E_u)_y \\ -D_I^u (I_u)_x & -D_I^u (I_u)_y \\ -D_A^u (A_u)_x & -D_A^u (A_u)_y \\ -D_R^u (R_u)_x & -D_R^u (R_u)_y \end{pmatrix}, \qquad
\mathbf{S}^u = \begin{pmatrix} -F_I(S_u, I_T) - F_A(S_u, A_T) \\ F_I(S_u, I_T) + F_A(S_u, A_T) - aE_u \\ a\sigma E_u - \gamma_I I_u \\ a(1 - \sigma)E_u - \gamma_A A_u \\ \gamma_I I_u + \gamma_A A_u \end{pmatrix}.
$$

Then, the same CWENO reconstruction and the same numerical local Lax-Friedrichs flux is employed where however, we account for a dissipation proportional to the diffusive terms. In other words, the numerical viscosity is given by the maximum eigenvalue of the viscous operator $s_{max}^V = \max\left(D_S^u, D_E^u, D_I^u, D_A^u, D_R^u\right)$. Concerning the time discretization, the explicit part of the Runge-Kutta scheme introduced in the previous paragraph is employed.

A.1.4. Stochastic collocation method

In the case in which we deal with the stochastic Eqs (2.1) and (2.2), we employ a generalized Polynomial Chaos (gPC) expansion technique [52, 66]. We restrict to the case in which there is only one stochastic variable z in the system. The extension to the case of a vector of random variables is straightforward. The probability density function of the single random input is supposed known: $\rho_z : \Gamma \to \mathbb{R}^+$. In this case, the approximate solution for the commuters $\mathbf{Q}_M(x, v, t, z)$ and the non commuters $\mathbf{Q}_M^u(x, t, z)$ are represented as truncations of the series of the orthonormal polynomials describing the

Table A.1. Setting of the Lombardy provinces: urban radius r_c, total inhabitants M and initial amount of highly infectious individuals $I_{T,0}$, detected by February 27, 2020 (initial day of the simulation). The total population is given by ISTAT data of December 31, 2019 [59], while data of highly infectious correspond to those reported in the GitHub repository of the Civil Protection Department of Italy [53]. Null values, listed with *, in the simulation are substituted with 1 to permit an effective assignation of uncertain initial condition.

Province	Urban radius r_c [km]	Total population M	Infectious $I_{T,0}$
Pavia (PV)	3.24	540376	36
Lodi (LO)	2.04	227412	159
Cremona (CR)	2.40	355908	91
Mantua (MN)	1.92	406919	0^*
Milan (MI)	5.76	3265327	15
Bergamo (BG)	3.96	1108126	72
Brescia (BS)	3.24	1255437	10
Varese (VA)	2.76	884876	0^*
Monza-Brianza (MB)	3.24	870193	5
Como (CO)	2.40	597642	0^*
Lecco (LC)	3.24	334961	0^*
Sondrio (SO)	1.56	180425	3

random space, i.e..

$$\mathbf{Q}_M(x, v, t, z) = \sum_{j=1}^{M} \hat{\mathbf{Q}}_j(x, v, t)\phi_j(z), \quad \mathbf{Q}_M^u(x, t, z) = \sum_{j=1}^{M} \hat{\mathbf{Q}}_j^u(x, t)\phi_j(z) \tag{A.18}$$

where M is the number of terms of the truncated series and $\phi_j(z)$ are orthonormal polynomials, with respect to the measure $\rho_z(z)\,dz$. The expansion coefficients are obtained by

$$\hat{\mathbf{Q}}_j(x, v, t) = \int_\Gamma \mathbf{Q}(x, v, t, z)\,\phi_j(z)\,\rho_z(z)\,dz, \quad \hat{\mathbf{Q}}_j^u(x, t) = \int_\Gamma \mathbf{Q}^u(x, t, z)\,\phi_j(z)\,\rho_z(z)\,dz, \quad j = 1, \dots, M. \tag{A.19}$$

Then, the exact integrals for the expansion coefficients in Eq (A.19) are replaced by a suitable quadrature formula characterized by the set $\{z_n, w_n\}_{n=1}^{N_p}$, where z_n is the n-th collocation point, w_n is the corresponding weight and N_p represents the number of quadrature points. We then have

$$\hat{\mathbf{Q}}_j(x, v, t) \approx \sum_{n=1}^{N_p} \mathbf{Q}(x, v, t, z_n)\,\phi_j(z_n)\,w_n, \quad \hat{\mathbf{Q}}_j^u(x, t) \approx \sum_{n=1}^{N_p} \mathbf{Q}^u(x, t, z_n)\,\phi_j(z_n)\,w_n, \quad j = 1, \dots, M \tag{A.20}$$

where $\mathbf{Q}(x, v, t, z_n)$ and $\mathbf{Q}^u(x, t, z_n)$ with $n = 1, \dots, N_p$ are the solutions of the problem evaluated at the n-th collocation point for the commuters and non commuters. Thanks to the computation of the above coefficients than it is possible to compute all quantities of interest concerning the random variable. For example, the expectations are approximated as

$$\mathbb{E}[\mathbf{Q}] \approx \mathbb{E}[\mathbf{Q}_M] = \int_\Gamma \mathbf{Q}_M(x, v, t, z)\,\rho_z(z)\,dz \approx \sum_{n=1}^{N_p} \mathbf{Q}(x, v, t, z_n)\,w_n, \tag{A.21}$$

Table A.2. Matrix of commuters of the Lombardy Region (Italy). Departure provinces are listed on the first left column, while arrival provinces are reported in the following columns. Each entry is given as number of commuters of the origin province and the corresponding percentage with respect to the total population of the city. This matrix is extracted from the origin-destination matrix provided by the Lombardy Region for the regional fluxes of year 2020 [60].

From/To	PV	LO	CR	MN	MI	BG	BS	VA	MB	CO	LC	SO	Total
PV	–	9601	–	–	83825	–	–	–	–	–	–	–	93426 (17.3%)
LO	9169	–	13712	–	56717	–	–	–	–	–	–	–	79598 (35.0%)
CR	–	13264	–	11654	23142	12025	17681	–	–	–	–	–	77766 (21.8%)
MN	–	–	1157	–	–	2267	–	22142	–	–	–	–	35980 (8.8%)
MI	82617	55397	21622	1946	–	74168	26709	144681	234682	41575	20801	1000	705198 (21.6%)
BG	–	–	12016	–	76337	–	78348	–	14826	–	17611	–	199138 (18.0%)
BS	–	–	16967	21643	26594	70879	–	–	–	–	–	–	136083 (10.8%)
VA	–	–	–	–	143152	–	–	–	–	33529	–	–	176681 (20.0%)
MB	–	–	–	–	247183	14938	–	–	–	–	37761	–	299882 (34.5%)
CO	–	–	–	–	44412	–	–	36249	–	–	–	–	80661 (13.5%)
LC	–	–	–	–	23621	19392	–	–	40317	–	–	4851	88181 (26.3%)
SO	–	–	–	–	1227	–	–	–	–	–	4545	–	5772 (3.2%)

and

$$\mathbb{E}\left[\mathbf{Q}^u\right] \approx \mathbb{E}\left[\mathbf{Q}^u_M\right] = \int_\Gamma \mathbf{Q}^u_M(x, t, z)\,\rho_z(z)\,dz \approx \sum_{n=1}^{N_p} \mathbf{Q}^u(x, t, z_n)\,w_n. \quad \text{(A.22)}$$

In the same way, all other quantities of interest such as the variance of $\mathbf{Q}$ and $\mathbf{Q}^u$ can be computed.

A.2. Population and mobility data

In this appendix we report tables containing data on the initial distribution of populations in the various urban areas considered in the Lombardy region, see Table A.1, and data on the corresponding flows of commuters between cities, see Table A.2.

References

1. G. Albi, L. Pareschi, M. Zanella, Control with uncertain data of socially structured compartmental epidemic models, *J. Math. Bio.*, **82** (2021), 63.

2. G. Albi, L. Pareschi, M. Zanella, Modelling lockdown measures in epidemic outbreaks using selective socio-economic containment with uncertainty, preprint medRxiv doi: 10.1101/2020.05.12.20099721, 2020.

3. B. Tang, X. Wang, A. Li, N. L. Bragazzi, S. Tang, Y. Xiao, et al., Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions, *J. Clinical Med.*, **9** (2020), 462.

4. M. Gatto, E. Bertuzzo, L. Mari, S. Miccoli, L. Carraro, R. Casagrandi et al., Spread and dynamicss of the COVID-19 epidemic in Italy: Effects of emergency containment measures, *Proceed. Nat. Acad. Sci.*, **117** (2020), 10484–10491.

5. B. N. Ashraf, Economic impact of government interventions during the COVID-19 pandemic: International evidence from financial markets, *J. Behav. Exp. Finance*, **27** (2020), 100371.

6. G. Dimarco, L. Pareschi, G. Toscani, M. Zanella, Wealth distribution under the spread of infectious diseases, *Phys. Rev. E*, **102** (2020), 022303.

7. H. W. Hethcote, The Mathematics of Infectious Diseases, *SIAM Rev.*, **42** (2000), 599–653.

8. D. Balcan, B. Gonçalves, H. Hu, J. J. Ramasco, V. Colizza, A. Vespignani, Modeling the spatial spread of infectious diseases: the GLobal Epidemic and Mobility computational model, *J. Comput. Sci.*, **1** (2010), 132–145.

9. B. Buonomo, R. Della Marca, Effects of information-induced behavioural changes during the COVID-19 lockdowns: The case of Italy: COVID-19 lockdowns and behavioral change, *R. Soc. Open Sci.*, **7** (2020), 201635.

10. V. Colizza, A. Vespignani, A. Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations, *J. Theor. Biol.*, **251** (2008), 450–467.

11. E. Franco, A feedback SIR (fSIR) model highlights advantages and limitations of infection-based social distancing, preprint, arXiv:2004.13216, 2020.

12. G. Giordano, F. Blanchini, R. Bruno, P. Colaneri, A. Di Filippo, A. Di Matteo et al., Modelling the COVID-19 epidemic and implementation of populationwide interventions in Italy, *Nat. Med.*, **26** (2020), 855-860.

13. E. L. Piccolomini, F. Zama, Monitoring Italian COVID-19 spread by a forced SEIRD model, *PloS One*, **15** (2020), e0237417.

14. Chronology of main steps and legal acts taken by the Italian Government for the containment of the COVID-19 epidemiological emergency, (`http://www.protezionecivile.gov.it/documents/20182/1227694/Summary+of+\measures+taken+against+the+spread+of+C-19/c16459ad-4e52-4e90-90f3-\c6a2b30c17eb`)

15. S. Riley, K. Eames, V. Isham, D. Mollison, P. Trapman, Five challenges for spatial epidemic models, *Epidemics*, **10** (2015), 68–71.

16. R. Dutta, S. Gomes, D. Kalise, L. Pacchiardi, Using mobility data in the design of optimal lockdown strategies for the COVID-19 pandemic, *PLoS Comput. Biol.*, **17** (2021), e1009236.

17. L. J. S. Allen, B. M. Bolker, Y. Lou, A. L. Nevai, Asymptotic profiles of the steady states for an SIS epidemic reaction-diffusion model, *Discrete Contin. Dyn. Syst.*, **21** (2008), 1–20.

18. V. Capasso, Global solution for a diffusive nonlinear deterministic epidemic model, *SIAM J. Appl. Math.*, **35** (1978), 274–284.

19. W. E. Fitzgibbon, J. J. Morgan, G. F. Webb, An outbreak vector-host epidemic model with spatial structure: the 2015-2016 Zika outbreak in Rio De Janeiro, *Theor. Biol. Med. Model.*, **14** (2017), 7.

20. Q. X. Liu, Z. Jin, Formation of spatial patterns in an epidemic model with constant removal rate of the infectives, *J. Stat. Mech. Theory Exp.*, (2007), P05002.

21. P. Magal, G. F. Webb, X. Wu, Spatial spread of epidemic diseases in geographical settings: Seasonal influenza epidemics in Puerto Rico, *Discrete Cont. Dyn. Sys. B*, **25** (2019), 2185–2202.

22. J. P. Keller, L. Gerardo-Giorda, A. Veneziani, Numerical simulation of a susceptible-exposed-infectious space-continuous model for the spread of rabies in raccoons across a realistic landscape, *J. Biol. Dyn.*, **7** (2014), 31–46.

23. G. Sun, Pattern formation of an epidemic model with diffusion, *Nonlinear Dyn.*, **69** (2012), 1097–1104.

24. A. Viguerie, G. Lorenzo, F. Auricchio, D. Baroli, T. J. R. Hughes, A. Patton et al., Simulating the spread of COVID-19 via a spatially-resolved susceptible-exposed-infected-recovered-deceased (SEIRD) model with heterogeneous diffusion, *Appl. Math. Lett.*, **101** (2021), 106617.

25. A. Viguerie, A. Veneziani, G. Lorenzo, D. Baroli, N. Aretz-Nellesen, A. Patton et al., Diffusion–reaction compartmental models formulated in a continuum mechanics framework: application to COVID-19, mathematical analysis, and numerical study, *Comput. Mech.*, **66** (2020), 1131–1152.

26. J. Wang, F. Xie, T. Kuniya, Analysis of a reaction-diffusion cholera epidemic model in a spatially heterogeneous environment, *Comm. Nonlin. Sci. Num. Simul.*, **80** (2020), 104951.

27. E. Frias-Martinez, G. Williamson, V. Frias-Martinez, An agent-based model of epidemic spread using human mobility and social network information, in *Proceedings of the 3rd International Conference on Social Computing*, Boston, MA, USA, (2011), 49–56.

28. E. Barbera, G. Consolo, G. Valenti, Spread of infectious diseases in a hyperbolic reaction-diffusion susceptible-infected-recovered model, *Phys. Rev. E*, **88** (2013), 052719.

29. G. Bertaglia, L. Pareschi, Hyperbolic models for the spread of epidemics on networks: kinetic description and numerical methods, *ESAIM Math. Model, Numer. Anal.*, **55** (2020), 381–407.

30. G. Bertaglia, L. Pareschi, Hyperbolic compartmental models for epidemic spread on networks with uncertain data: application to the emergence of Covid-19 in Italy, *Math. Mod. Meth. Appl. Sci.*, 2021.

31. R. M. Colombo, M. Garavello, F. Marcellini, E. Rossi, An age and space structured SIR model describing the COVID-19 pandemic, *J. Math. Ind.*, **10** (2020), 22.

32. W. Boscheri, G. Dimarco, L. Pareschi, Modeling and simulating the spatial spread of an epidemic through multiscale kinetic transport equations, *Math. Mod. Meth. App. Math.*, **31** (2021), 1059–1097.

33. K. M. Case, P. F. Zweifel, Existence and uniqueness theorems for the neutron transport equation, *J. Math. Phys.*, **4** (1963), 1376–1385.

34. F. A. C. C. Chalub, P. A. Markovich, B. Perthame, C. Schmeiser, Kinetic models for chemotaxis and their drift-diffusion limits, *Monatsh. Math.*, **142** (2004), 123–141.

35. T. Hillen, A. Swan, The diffusion limit of transport equations in biology, in *Mathematical Models and Methods for Living Systems*, Springer, **2167** (2016).

36. B. Perthame, *Transport Equations in Biology*, Birkhäuser, Boston, 2007.

37. L. Pareschi, G. Toscani, *Interacting multiagent systems: kinetic equations and Monte Carlo methods*, Oxford University Press, Oxford, UK, 2014.

38. N. Bellomo, R. Bingham, M. A. J. Chaplain, G. Dosi, G. Forni, D. A. Knopoff et al., A multi-scale model of virus pandemic: Heterogeneous interactive entities in a globally connected world, *Math. Mod. Meth. Appl. Sci.*, **30** (2020), 1591–1651.

39. C. Cercignani, R. Illner, M. Pulvirenti, *The Mathematical Theory of Diluted Gases*, Springer, New York, 1994.

40. M. Delitala, Generalized kinetic theory approach to modeling spread and evolution of epidemics, *Math. Compet. Mod.*, **39** (2004), 1–12.

41. M. Pulvirenti, S. Simonella, A kinetic model for epidemic spread, *Math. Mech. Complex Syst.*, **8** (2020), 249–260.

42. R. Yano, Kinetic modeling of local epidemic spread and its simulation, *J. Sci. Comput.*, **73** (2017), 122–156.

43. E. W. Larsen, J. B. Keller, Asymptotic solution of neutron transport problems for small free mean paths, *J. Math. Phys.*, **15** (1974), 75–81.

44. M. Peirlinck, K. Linka, F. Sahli Costabal, J. Bhattacharya, E. Bendavid, J. P. Ioannidis et al., Visualizing the invisible: The effect of asymptomatic transmission on the outbreak dynamicss of COVID-19, *Comp. Meth. Appl. Mech. Eng.*, **372** (2020), 113410.

45. F. Golse, S. Jin, C. Levermore, The convergence of numerical transfer schemes in diffusive regimes I: Discrete-ordinate method, *SIAM J. Num. Anal.*, **36** (1999), 1333–1369.

46. S. Jin, L. Pareschi, G. Toscani, Uniformly accurate diffusive relaxation schemes for multiscale transport equations, *SIAM J. Num. Anal.*, **38** (2000), 913–936.

47. M. Dumbser, M. Kaeser, Arbitrary high order non-oscillatory finite volume schemes on unstructured meshes for linear hyperbolic systems, *J. Comput. Phys.*, **221** (2007), 693–723.

48. E. Gaburro, W. Boscheri, S. Chiocchetti, C. Klingenberg, V. Springel, M. Dumbser, High order direct Arbitrary-Lagrangian-Eulerian schemes on moving Voronoi meshes with topology changes, *J. Comput. Phys.*, **407** (2020), 109167.

49. S. Boscarino, L. Pareschi, G. Russo, Implicit-explicit Runge-Kutta schemes for hyperbolic systems and kinetic equations in the diffusion limit, *SIAM J. Sci. Comput.*, **35** (2013), 22–51.

50. S. Boscarino, L. Pareschi, G. Russo, A unified IMEX Runge-Kutta approach for hyperbolic systems with multiscale relaxation, *SIAM J. Numer. Anal.*, **55** (2017), 2085–2109.

51. G. Dimarco, L. Pareschi, Numerical methods for kinetic equations, *Acta Numer.*, **23** (2014), 369–520.

52. D. Xiu, *Numerical Methods for Stochastic Computations: A Spectral Method Approach*, Princeton University Press, Princeton, NY, (2010).

53. Presidenza del Consiglio dei Ministri, Dipartimento della Protezione Civile, Italia, COVID-19 epidemiological data in Italy, (`https://github.com/pcm-dpc/COVID-19`).

54. A. Korobeinikov, P. K. Maini, Non-linear incidence and stability of infectious disease models, *Math. Med. Bio.: J. IMA*, **22** (2005), 113–128.

55. V. Capasso, G. Serio, A generalization of the Kermack-McKendrick deterministic epidemic model, *Math. Biosci.*, **42** (1978), 43.

56. G. F. Webb, A reaction-diffusion model for a deterministic diffusion epidemic, *J. Math. Anal. Appl.*, **84** (1981), 150–161.

57. O. Diekmann, J. Heesterbeek, M. Roberts, The construction of next-generation matrices for compartmental epidemic models, *J. Roy. Soc. Interface*, **7** (2010), 873–885.

58. Istituto Nazionale di Statistica, Italia. Dati Geografici, (`https://www4.istat.it/it/archivio/209722`)

59. Istituto Nazionale di Statistica, Italia, Dati Demografici, (`http://demo.istat.it/`)

60. Regione Lombardia, Italia. Open Data, (`https://www.dati.lombardia.it/Mobilit-e-trasporti/Matrice-OD2020-Passeggeri/hyqr-mpe2`)

61. M. A. C. Vollmer, S. Mishra, H. J. T. Unwin, A. Gandy, T. A. Mellan, H. Zhu et al., Using mobility to estimate the transmission intensity of COVID-19 in Italy: a subnational analysis with future scenarios, Technical Report May, Imperial College London, 2020.

62. A. Aktay, S. Bavadekar, G. Cossoul, J. Davis, D. Desfontaines, A. Fabrikant et al., Google COVID-19 community mobility reports: anonymization process description (version 1.1), pre-print, *arXiv:2004.04145*, (2020).

63. A. Klar, An asymptotic-induced scheme for nonstationary transport equations in the diffusive limit, *SIAM J. Numer. Anal.*, **35** (1998), 1073–1094.

64. W. Boscheri, G. Dimarco, High order central WENO-Implicit-Explicit Runge Kutta schemes for the BGK model on general polygonal meshes, *J. Comput. Phys.*, **422** (2020), 109766.

65. M. Dumbser, W. Boscheri, M. Semplice, G. Russo, Central weighted ENO schemes for hyperbolic conservation laws on fixed and moving unstructured meshes, *SIAM J. Sci. Comp.*, **39** (2017), A2564–A2591.

66. S. Jin, H. Lu, L. Pareschi, Efficient stochastic asymptotic-preserving implicit-explicit methods for transport equations with diffusive scalings and random inputs, *SIAM J. Sci. Comput.*, **40** (2018), A671–A696.

AIMS Press