A multi-group epidemic model to represent the COVID-19 spread among regions: analysis of the Italian case

Paolo Di Giamberardino[1], Daniela Iacoviello[1], Federico Papa[2], Carmela Sinisgalli[2]

***Abstract*— We formulate and implement a dynamical multi-group model to describe the diffusion of COVID-19 epidemic within homogeneous sub-populations, structurally or geographically separated but interconnected by a mobility network. Each sub-model provides a rather accurate description of infective sub-populations. Control actions, representing the intervention measures adopted to curb the disease dynamics, are also included. The multi-group structure of the model is specifically designed to investigate the effects of people mobility. The case of three Italian areas is considered, for the period March–October 2020 including the massive increment in people mobility occurred in summertime. The results of model simulation are in good agreement with the data and can provide important informations about the contribution of outcome individuals to the virus spread.**

I. Introduction

The current pandemic emergency caused by Coronavirus disease (COVID-19) has highlighted the vulnerability of all world countries in facing efficiently the unprecedented challenges thereby arisen, while evidencing the need of framing the determinants of the epidemic dynamics, especially to limit the impact of future possible outbreaks [1].

In addressing these questions, several literature studies have taken into account different types of possible interventions and two domains of variables: pathogen-associated variables and society-based variables. This latter variable domain is particularly relevant since individuals are actually the vectors of SARS-CoV-2 virus. Thus, social distancing and personal protective measures appear to be the primary mechanisms for controlling the COVID-19 spread, at least before vaccines become diffusely available. The effectiveness of social distancing is studied in many papers, which propose projections where the impact of introducing containment measures is shown along with the contribution in reducing the infection spread [2], [3], [4], [5].

In the early stage of pandemic, only qualitative data analysis was performed, as in [6] where data regarding how the human mobility changed in the United States at the beginning of the pandemic course are studied and the importance of quantifying the social distancing practices is emphasized. Moreover, apart from a number of examples for China, not many quantitative studies on the effect of limiting social distancing and human mobility exist. Such quantitative studies are generally based on mathematical modelling, they are data-driven, still they use rather different approaches, see e.g. [7], [8], [9].

In many recent works, both deterministic/stochastic and discrete/continuous models have been applied for the description, forecast and control of the COVID-19 epidemic spread. In the framework of compartmental models, the classes of Susceptible (S), Exposed (E), Infected (I) and Removed (R) subjects are generally introduced, yielding SEIR models. For the COVID-19 pandemic, like for other diseases with specific characteristics, additional categories are generally introduced, referring to the condition of infected patients, such as asymptomatic, or hospitalised, or in the quarantine condition [2], [10], [11].

Concerning the geographical aspects, the effect of interactions and travel restrictions on the pandemic evolution, multi-group epidemic compartmental models, suitable extensions of SIR/SEIR frameworks, are used to represent the COVID-19 spread among different (heterogeneous) populations, like e.g. in [12], [13], [14], [15]. The heterogeneity of the sub-populations is intended with respect to the epidemic properties and can depend on their different geographical allocation or other structural variables (e.g. age, population density, environment).

The model presented in this paper is based on a previous epidemic model reported in [10], representing the COVID-19 evolution by a SEIR-type model with two sub-populations of infected subjects, undiagnosed and diagnosed, and explicitly accounting for a fraction of asymptomatic infective subjects. The present model structure, depicted in Sect. II, incorporates N interconnected epidemic models of that kind, particularly with the aim of representing the effect of individual interactions and geographical exchanges among groups. In Sect. III the general model is specialized for $N = 3$ and it is applied to simulate the disease evolution in Italy: its territory is decomposed into three macro-regions, corresponding approximatively to North, Centre and South, in order to evidence some interesting aspects related to the increased human mobility following the first pandemic wave. The model proved to be apt in describing the summer period 2020 of the COVID-19 epidemic in Italy, by explicitly accounting for different scenarios characterizing geographically the distinct macro-regions of the Italian territory.

II. A multi-group epidemic model for the spread of COVID-19 among N groups

The model proposed here has a multi-group structure that incorporates different sub-units, each one describing

[1]Dept. of Computer, Control and Management Engineering, Sapienza University of Rome, Rome, Italy paolo.digiamberardino@uniroma1.it, daniela.iacoviello@uniroma1.it

[2]Institute for Systems Analysis and Computer Science "A. Ruberti", Rome, Italy federico.papa@iasi.cnr.it, carmela.sinisgalli@iasi.cnr.it

the dynamic evolution of the epidemic within a homogeneous population with specific evolution characteristics. For instance, the groups can represent different geographical regions or structurally different populations. The N groups (namely sub-units) are interconnected by a mobility network that accounts for the transfers of individuals who are allowed to travel from a group to another. Most typically, the model can describe a geographical system composed by N regions with people of each region moving for work, study or simply personal/holiday reasons. Some following numerical examples will be given referring to this situation.

Each of the N sub-models is a simplified version of the model previously proposed in [10] for the description of the first phase of the epidemic spread in our country (thereby modelled as a whole homogeneously-mixing group). In addition to a compartment of exposed individuals, which is proper of SEIR models, our model explicitly distinguishes between diagnosed and undiagnosed infective patients. As shown in [10], the proposed structure appears appropriate to mimic the Italian case, by incorporating also control actions reproducing government restrictions and emergency actions implemented to detect the infected cases, especially asymptomatic or mildly symptomatic cases. Precisely, each sub-model takes into account the following five state variables: $S(t)$ - number of susceptible individuals; $E(t)$ - number of exposed (infected but not yet infective) individuals; $I_u(t)$ - number of undiagnosed infective patients; it is specifically devised to represent subjects who will be asymptomatic or pauci-symptomatic during their whole infection period, in addition to subjects with recognisable symptoms, as long as they remain undiagnosed; $I_d(t)$ - number of diagnosed infective patients, possibly receiving medical treatments to cure the infection or its complications. We assume that this latter population cannot transmit the virus because of isolation (at home or at hospital); $R(t)$ - number of healed patients (treated or not).

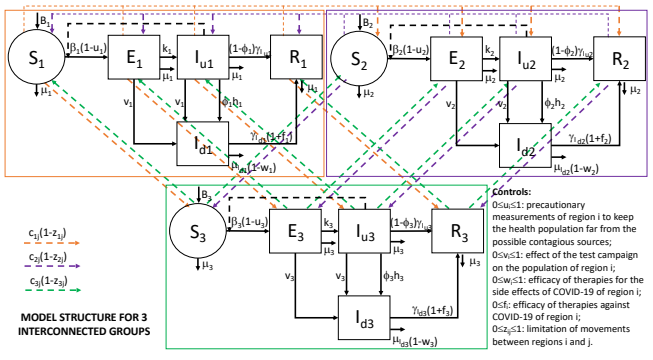

Fig. 1. Block diagram of model (1)-(5) applied to the mobility scheme among $N = 3$ interconnected regions of epidemic diffusion.

The complete model formulation includes N structurally identical groups or sub-systems, with each group i, $\{i = 1, 2, \ldots, N\}$, described by a SEIR-type model with undiagnosed and diagnosed infected subjects, as reported above. The N sub-systems are connected by means of a mobility network allowing people to move among groups. In the following, we refer to a specific epidemic group by its identifier i, also using the same subscript to denote the related state variables and parameters. In general, however, we assume that the only individuals allowed to move are the ones having no evidence and/or diagnosis of infection, i.e. the ones belonging to compartments S_i, E_i, I_{u_i}, R_i, $i = 1, 2, \ldots, N$. Fig. 1 shows, by means of a graphical example with $N = 3$, how the mobility network works, also illustrating the epidemic core model of each sub-system.

So, the spread of COVID-19 among N epidemiologically distinct groups can be formally described by means of N systems of time-varying ODE models of the following kind:

$$\dot{S}_i = B_i - \beta_i(1-u_i)S_i I_{u_i} - \mu_i S_i - \sum_{j=1,j\neq i}^{N} c_{i,j}(1-z_{i,j})S_i + \sum_{j=1,j\neq i}^{N} c_{j,i}(1-z_{j,i})S_j, \quad (1)$$

$$\dot{E}_i = \beta_i(1-u_i)S_i I_{u_i} - v_i E_i - k_i E_i - \mu_i E_i - \sum_{j=1,j\neq i}^{N} c_{i,j}(1-z_{i,j})E_i + \sum_{j=1,j\neq i}^{N} c_{j,i}(1-z_{j,i})E_j + In_E, \quad (2)$$

$$\dot{I}_{u_i} = k_i E_i - v_i I_{u_i} - h_i \phi_i I_{u_i} - \gamma_{I_{u_i}}(1-\phi_i)I_{u_i} - \mu_i I_{u_i} - \sum_{j=1,j\neq i}^{N} c_{i,j}(1-z_{i,j})I_{u_i} + \sum_{j=1,j\neq i}^{N} c_{j,i}(1-z_{j,i})I_{u_j} + In_{I_{u_i}}, \quad (3)$$

$$\dot{I}_{d_i} = h_i \phi_i I_{u_i} + v_i(E_i + I_{u_i}) - \gamma_{I_{d_i}}(1+f_i)I_{d_i} - \mu_{I_{d_i}}(1-w_i)I_{d_i}, \quad (4)$$

$$\dot{R}_i = \gamma_{I_{u_i}}(1-\phi_i)I_{u_i} + \gamma_{I_{d_i}}(1+f_i)I_{d_i} - \mu_i R_i - \sum_{j=1,j\neq i}^{N} c_{i,j}(1-z_{i,j})R_i + \sum_{j=1,j\neq i}^{N} c_{j,i}(1-z_{j,i})R_j, \quad (5)$$

where $i = 1, 2, \ldots, N$ and the state variable dependence on t has been omitted for compactness' sake. A brief description of all the quantities included in the ODE system (1)-(5) is here given. B_i is the net input rate in compartment S_i, which accounts for both the newborn (susceptible) individuals and the balance between immigration and emigration; μ_i is the per capita death rate owing to causes not related to the infection (natural death of the population) and it represents the loss rate from any compartment of the model except for I_{d_i}; $\mu_{I_{d_i}}$ is the per capita death rate of diagnosed patients I_{d_i}; β_i is the relative contagiousness of individuals in compartment I_{u_i} and it accounts for two main factors, which are the contagion probability from one infected-susceptible contact (related to the aggressiveness of the virus) and the frequency of contacts; ϕ_i represents the fraction of the infective population I_{u_i} that will show recognisable symptoms and that will consequently be diagnosed and isolated (possibly receiving therapies); k_i describes the transition from E_i to I_{u_i}, and it is set to $k_i = 1/\tau_i$, where τ_i is the mean length of the incubation period; h_i refers to the transition from I_{u_i} to I_{d_i}, taken as $h_i = 1/\tau_{s_i}$, where τ_{s_i} is the mean time from infection until the occurrence of the first recognisable symptoms; $\gamma_{I_{u_i}}$ models the outflow from the infective compartment I_{u_i} associated to recovery from infection and, then, it is assumed $\gamma_{I_{u_i}} = 1/\tau_{r_i}$, with τ_{r_i} the mean recovery period

without any medical assistance; similarly $\gamma_{I_{d_i}}$ models the outflow from the infective compartments I_{d_i} due to recovery from the infection and, then, it is $\gamma_{I_{d_i}} = 1/\tau_{d_i}$, with τ_{d_i} denoting the mean recovery period of monitored patients; c_{ij} is a weight accounting for the transition probability of a subject moving from the *i*-th compartment S_i or E_i or I_{u_i} to the corresponding *j*-th compartment. Note that the coefficients c_{ij} can also be time-varying in order to represent the variation of transfer probabilities possibly occurring over time. This variability is especially required for long-term analysis when "ordinary" mobility regimens alternate with highly "intense" transfer periods, like summer or Christmas or Easter seasons. Note also that, for the sake of generality, the recovery rates from compartments I_{u_i} and I_{d_i} are taken into account separately by means of the rate constants $\gamma_{I_{u_i}}$ and $\gamma_{I_{d_i}}$, respectively, even though to a first approximation, and in the absence of experimental evidences, they can be assumed equal to each other.

As far as the control actions are concerned, the time-varying quantities $u_i(t)$, $v_i(t)$, $w_i(t)$, $z_{i,j}(t)$ (taking values in $[0, 1]$), and $f_i(t) \geq 0$ are introduced to represent the intervention measures adopted by the authorities to contain the disease outbreak. More precisely, $u_i(t)$ quantifies possible actions locally implemented by authorities to reduce the contact rate, and then the relative infectivity β_i, of population i. It accounts for all the government decrees introduced to limit the physical interactions among people, but also for the informative campaign for hygienic measures, TV/radio announcements, and so on. The quantity $v_i(t)$ represents the intensity of the swab test campaign performed on sub-population i, which changes daily depending on the number of swab tests actually performed. For the sake of simplicity, and in the absence of other indications, we assume that the amount of performed tests is equally distributed among people of compartments S_i, E_i and I_{u_i}, so that the same per capita test rate can be assumed for all these compartments. This implies that the exit fluxes of tested (positive) individuals leaving E_i and I_{u_i} are proportional to the number of individuals within the same compartment, i.e. $v_i(t)E_i(t)$ and $v_i(t)I_{u_i}(t)$, respectively. We notice that the flux $v_i(t)S_i(t)$ of (negative) test results exiting S_i does not explicitly appear in the model equation since it does not contribute to the dynamical evolution (actually an identical flux amount comes back to compartment S_i). The control actions $w_i(t)$ and $f_i(t)$ refer to the efficacy of the therapies adopted by the $i-$th health system, either to reduce side effects of COVID-19 and, respectively, to cure the infection. Furthermore, the time-varying controls $z_{i,j}(t)$, $i, j = 1, 2, \ldots, N$, represent the interventions and mobility restrictions implemented by the central government or local authorities to limit people transfers between groups i and j.

Finally, the pair of input fluxes $In_{E_i}(t)$ and $In_{I_{u_i}}(t)$ are introduced (see Eqs. (2), (3)) to model the cumulative entry of infected people coming from outer groups/regions whose epidemic dynamics is not incorporated in the N group model. Since diagnosed infective people are not allowed to travel, it is reasonable to account for such cumulative outer inputs only in the equations of E_i and I_{u_i}, $i = 1, \ldots, N$.

We notice that the proposed model does not incorporate the possibility of re-infection. Indeed in our model, once recovered, a patient (R) is no longer susceptible. However, this simplifying hypothesis, which is actually the object of clinical studies and debates on the persistence and actual length of the immunity period, can be a valid assumption if a short-term analysis, like the one presented in the next section, is performed. Moreover, our model is designed to describe the infection transmission using a nonlinear S-I dynamics that is the main difference with meta-population models which focus on pathogen transferring instead of transmission.

The following section is devoted to the application of the model for the numerical simulation of the COVID-19 evolution in Italy, considering $N = 3$ groups (regions) and using real demographic and epidemic data for parameter estimation. Setting $N = 3$ coupled to a summer observation period is motivated in order to highlight the influence of human movements on the epidemic spread. In the period of interest, following the first epidemic wave, a considerable number of movements could be counted across Italy, with non uniform flow, but with predominant flow orientation from North towards Centre-South. N is chosen as a trade-off between the expected level of detail and of fragmentation.

III. Parameter estimation and analysis of the mobility in Italy up to 21-10-2020

We aim to evaluate, by numerical simulation, the impact of human mobility and people transfer on the diffusion of COVID-19 in our country, considering three geographical areas of the Italian territory. The areas are North, Centre, and South, each one gathering different Italian districts, as reported by Table I. The analysis aims to evaluate the effects of restoring the human mobility after the first strong lock-down in Italy (implemented by the central government as of March 9, 2020) on the trigger of the second wave, with a specific focus on the role played by the exodus occurring during the past summer holidays 2020.

TABLE I

Macro-regions vs. Italian districts correspondences.

Identifier	Region	Italian districts
1	North	Piemonte, Trentino Alto Adige, Lombardia Valle d'Aosta, Liguria, Emilia Romagna, Veneto, Friuli Venezia Giulia
2	Centre	Toscana, Umbria, Marche, Lazio
3	South	Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sardegna, Sicilia

A. Parameter estimation from data in the interval March 9 - June 3, 2020

Concerning the model parameters B_i, μ_i, k_i, h_i, $i = 1, 2, 3$, they have been inferred from official data sources, following the approach described in [10]. Table II reports the values of these parameters for the three macro-regions. Note that the net input rates B_i, $i = 1, 2, 3$, have been computed as $B_i \approx P_i \cdot \mu_i$, where P_i are the total numbers of people within the i-th region (almost constant along the "short" period of interest), i.e. $P_1 = 27,746,113$; $P_2 = 12,016,009$; $P_3 = 20,597,424$.

TABLE II

FIXED MODEL PARAMETERS.

Parameter	Value	Source
B_1	779.67 persons $\cdot$ day^{-1}	[16]
B_2	337.65 persons $\cdot$ day^{-1}	[16]
B_3	578.79 persons $\cdot$ day^{-1}	[16]
μ_1, μ_2, μ_3	$2.81 \cdot 10^{-5}$ day^{-1}	[16]
k_1, k_2, k_3	0.167 day^{-1}	[17], [18]
h_1, h_2, h_3	0.20 day^{-1}	[17], [18]

The remaining parameters β_i, ϕ_i, $\mu_{I_{d_i}}$, $\gamma_{I_{u_i}}$, $\gamma_{I_{d_i}}$, $i = 1, 2, 3$, have been estimated by means of a least square fitting procedure, which, for the sake of simplicity, assumed $\gamma_{I_{u_i}} = \gamma_{I_{d_i}}$ for any i. The epidemiological data exploited for the fitting are: i) daily number of diagnosed individuals that are currently positive, ii) total number of recoveries among all diagnosed positives, iii) total number of notified deaths. For each region i, such data are respectively reproduced by the computational quantities: a) $I_{d_i}(\delta_j)$, b) $\int_{\sigma=\delta_0}^{\sigma=\delta_j} \gamma_{I_{d_i}}(1 + f_i(\sigma))I_{d_i}(\sigma)d\sigma$, c) $\int_{\sigma=\delta_0}^{\sigma=\delta_j} \mu_{I_{d_i}}(1 - w_i(\sigma))I_{d_i}(\sigma)d\sigma$, where δ_0 is March 9, that is the beginning of the national lock-down in Italy, while the j-th notification day δ_j can run until June 3, i.e. until the starting day of the restored free mobility among regions (about one month after the end of the lock-down on May 4). The choice of exploiting epidemiological data for the model training in the selected time interval is made to simplify the identification procedure. Indeed, since in the chosen period the human mobility was practically forbidden (except for necessity reasons), we can set $z_{i,j}(t) = 1$, for any pair (i, j) and any t in the identification interval, performing three independent procedures for the identification of the parameter sets $\{\beta_i,\phi_i,\mu_{I_{d_i}},\gamma_{I_{u_i}},\gamma_{I_{d_i}}\}$, for $i = 1, 2, 3$. In view of the strong restrictions carried out during the lock-down on the arrivals to Italy from abroad, we also assume null infective inputs external to the system, i.e. $In_{E_i}(t) = In_{I_{u_i}}(t) = 0$, $i = 1, 2, 3$, from March 9 to June 3, 2020.

For the identification of each parameter set $\{\beta_i,\phi_i,\mu_{I_{d_i}},\gamma_{I_{u_i}},\gamma_{I_{d_i}}\}$ from data, we exploited suitable time-varying controls $u_i(t)$, $v_i(t)$, $w_i(t)$, $f_i(t)$ taking into account the variations of the social behaviour (owing to both government restrictions and increasing health risk awareness in the population), of the swab test campaign, and of the health system efficiency. The time course of the control actions assumed over the interval March 9 - June 3 is reported in Fig. 2, lower panels, together with the optimal fitting curves a)-c) of data i)-iii), upper panels, for the three macro-regions specified by Table I. Note that in each region: $u_i(t)$ is assumed rapidly increasing from March 9 as a consequence of the decree ratifying the beginning of the national lock-down (see also [10]); $v_i(t)$, is derived directly from data on the number of swab tests performed in region i. We observe that rapid changes of $w_i(t)$ and $f_i(t)$ are required to adequately fit the data; they can be explained in view of two main factors: progressive increase of the medical experience and initial lack in the notification process of the daily recoveries. The estimated model parameters for each region are reported in Table III.

An *ad hoc* best fitting procedure was developed in MATLAB accounting for positivity constraints on the parameters by a logarithmic transformation. The procedure incorporates native MATLAB routines for integration of the ODE system (1)-(5) (`ode45`) and for OLS unconstrained minimization (`fminsearch`).

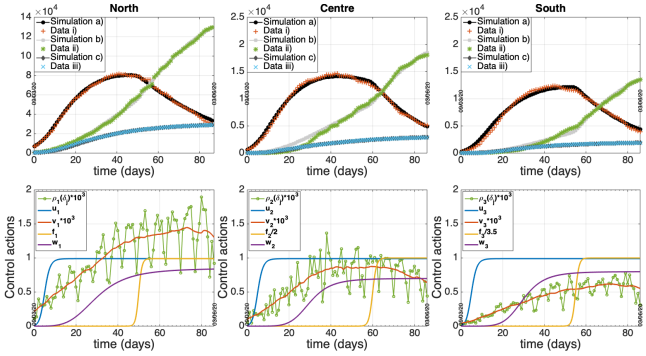

Fig. 2. Fitting curves (upper panels) and corresponding control actions (lower panels) for the three macro-regions of Table I. Data retrieved from [19]. $v_i(t) = M_{10}(\rho_i(\delta_j))$, $i = 1, 2, 3$, for $t \in [\delta_j, \delta_{j+1})$, where $\rho_i(\delta_j)$ is the ratio between the number of swab tests in region i at δ_j and P_i; $M_{10}(\rho_i(\delta_j))$ returns the average value of $\rho_i(\cdot)$ on a moving window of 21 days length centered on δ_j. Initial conditions for $i = 1, 2, 3$: $I_{d_i}(\delta_0)$ is the measured number of current positives of region i at δ_0 (March 9); $E_i(\delta_0)$, $I_{u_i}(\delta_0)$, $R_i(\delta_0)$ are computed from $E_i(\delta_0)/I_{d_i}(\delta_0) = 11$, $I_{u_i}(\delta_0)/I_{d_i}(\delta_0) = 9$ and $R_i(\delta_0)/I_{d_i}(\delta_0) = 1$ (simulations in [10]), and $S_i(\delta_0)$ is computed by the relation $S_i + E_i + I_{u_i} + I_{d_i} + R_i = P_i$.

TABLE III

ESTIMATED MODEL PARAMETERS (BASED ON EPIDEMIOLOGICAL DATA FROM MARCH 9 TO JUNE 3).

i	Parameters			
	β_i (persons^{-1} $\cdot$ day^{-1})	ϕ_i (-)	$\mu_{I_{d_1}}$ (day^{-1})	$\gamma_{I_{u_i}}$, $\gamma_{I_{d_i}}$ (day^{-1})
1	$1.289 \cdot 10^{-8}$	0.0992	0.0144	0.0192
2	$3.290 \cdot 10^{-8}$	0.1947	0.0069	0.0145
3	$2.900 \cdot 10^{-8}$	0.2173	0.0070	0.0090

B. Increasing the mobility among regions after June 3

We want now to investigate the effect of increasing the human mobility, officially restored among regions as of June 3, on the time-course of the epidemic in each area. Estimating the related parameters by best fitting of data from June 3 onwards is out of the scope of this section. We rather propose a quantitative analysis of the whole system aimed to evaluate the impact of some policies and social behaviours, with particular emphasis on the mobility aspects after the lock-down end.

Possible realistic values of the control actions have been inferred based on epidemiological data of the total number of cases and of the daily number of new cases during the period June 3 - October 21 with a particular focus on the role played by the human mobility increase in triggering the second pandemic wave. Such epidemiological data are reproduced by the model based quantities $C_i(\delta_j) = \int_{\sigma=\delta_0}^{\sigma=\delta_j} [h_i\phi_i I_{u_i}(\sigma) + v_i(\sigma)(E_i(\sigma) + I_{u_i}(\sigma))]d\sigma$, for the total cases of region i at day δ_j, and by the related increment $C_i(\delta_j) - C_i(\delta_{j-1})$, for the daily new cases. In order to match such data, particular attention has been devoted to the calibration of the quantities $u_i(t)$, $z_{i,j}(t)$, and $c_{i,j}(t)$, as well as of the inputs $In_{E_i}(t)$, $In_{I_{u_i}}(t)$, $i = 1, 2, 3$. Since there was not any differentiation

of the mobility restrictions across Italy until the middle of October, we assume in this analysis $z_{i,j}(t) = z(t)$, for any pair i, j (and for any time). Moreover, for the sake of simplicity, we consider a single cumulative flux entering the infected communities of region i from outside, $In_i(t)$, which is assumed equally distributed between E_i and I_{u_i}, that is $In_{E_i}(t) = In_{I_{u_i}}(t) = In_i(t)/2$. The time behaviour of the coefficients $c_{i,j}(t)$ is also chosen *ad hoc* so as to represent both the "ordinary mobility", before and after summer, and the unbalanced transfer of people characterizing the holiday exodus, which can be observed in the available and official source data. Concerning the other control actions, $v_i(t)$ is driven by the data on the number of swab tests, as in Section III-A, while $w_i(t)$ and $f_i(t)$ are kept fixed to the value reached at the end of the identification period. Note that the controls $w_i(t)$ and $f_i(t)$ do not play a crucial role on the total cases (as it can be deduced from the model equations), so an accurate tuning of these functions is not required for the following analysis.

Fig. 3 depicts the time course of the time-varying quantities adopted to reproduce the epidemiological data (right panel): $u_i(t)$, $In_i(t)$ (left panels) and $z(t)$, $c_{i,j}(t)$ (central panels). The time behaviour before June 3 is the same

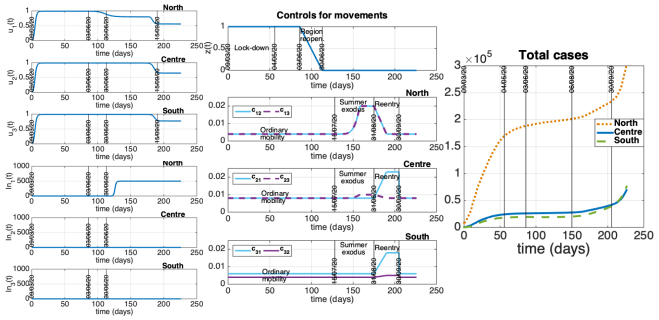

Fig. 3. Time-course of the controls (left and central panels) chosen to reproduce the total number of cases (left panel) until October 21.

of Section III-A. Concerning the evolution after June 3, we highlight the following aspects related to $u_i(t)$, $In_i(t)$, $i = 1, 2, 3$: (A) there is no increase of cases in Centre and South Italy from the end of the lock-down to mid-August (see Fig. 3, right panel), suggesting that neither a remarkable change in the social behaviour nor a sensible influx of new infections from outside is present in regions $i = 2, 3$ during the first phase of summer. So we keep $u_i(t) \approx 0.99$, $i = 2, 3$, until mid September, and $In_i(t) = 0$, $i = 2, 3$; (B) the northern cases keep on increasing after the restored mobility (although the increasing rate is lighter than before May 4). So a first decrease of $u_1(t)$ (before mid September) and a concomitant increase of $In_1(t)$ are assumed after the reopening; (C) the inputs $In_i(t)$, $i = 1, 2, 3$, are not changed anymore (i.e. until October 21), while a sensible change for the controls $u_i(t)$, $i = 1, 2, 3$ is assumed almost at the middle of September when the restart of production activities and schools may have increased the contacts among people.

As far as the mobility controls after June 3 are concerned, we make the following assumptions on $z(t)$, $c_{i,j}(t)$, $i \neq j$, $i, j = 1, 2, 3$: (D) the mobility is slowly restored throughout one month after the government decree allowing free people transfers as of June 3, i.e. $z(t)$ decreases from 1 (no mobility) on June 3 to 0 (completely restored mobility) on June 30; (E) we assume a "regular mobility" (people transfers for study/work or for visiting relatives/friends under "ordinary" conditions) before and after the holidays (that is before July 20 and after September 30) and, for any t in these periods, we fix $c_{i,j}(t) = \bar{c}_{i,j}$, where the constant coefficients $\bar{c}_{i,j}$ are set in order to provide an order of magnitude of 100K persons travelling each day between each pair of regions. In particular we choose $\bar{c}_{1,2} = 4 \cdot 10^{-3}$, $\bar{c}_{1,3} = 4 \cdot 10^{-3}$, $\bar{c}_{2,1} = 8 \cdot 10^{-3}$, $\bar{c}_{2,3} = 8 \cdot 10^{-3}$, $\bar{c}_{3,1} = 6 \cdot 10^{-3}$, $\bar{c}_{3,2} = 4 \cdot 10^{-3}$; (F) we change rapidly the coefficients $c_{i,j}(t)$ from July 20 to September 30 (see Fig. 3) to simulate a summer exodus from North towards Centre and South that produces a demographic unbalance of a few million people in each region. A statistical analysis reported by ENIT (the Italian National Agency for Tourism) on September 15 shows that almost 24 millions of Italians went to Centre and South for their holidays [20].

Overall, as shown in Fig. 4, the assumed control actions allow to fit well the data on the total number of cases (middle panels) and on the daily number of new cases (lower panels) from June 3 until October 21 The first comment that we

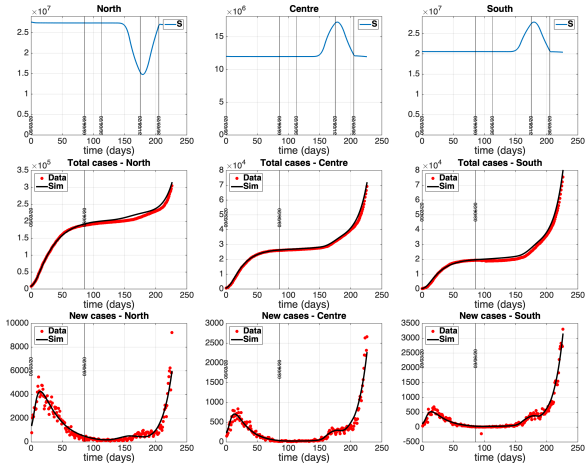

Fig. 4. Model assessment from epidemiological data between March 9 and October 21, 2020. Upper panels: size of the northern, central and southern susceptible population. Middle panels: total number of cases. Lower panels: new daily cases. Red dots: ISS data [19]. Black lines: model reconstruction.

can give is that there is a slight increment of cases in the interval 150-200 days occurred in all the regions, especially of Centre-South, (note the slope change in total cases, middle panels of Fig. 4, or the bell-shaped course of new cases, lower panels of Fig. 4). This increment can be explained by the combination of the summer exodus from North to Centre-South together with the enhanced virus circulation in the North ($u_1(t)$ lowered and $In_1(t)$ increased after June 3). However, such an increment is rather contained, proving that the social behaviour and the movements during summer were not enough to immediately trigger a second wave. However, they probably contributed to the second wave since the fast increase of cases after September 15 can be actually reproduced only combining the sensible reduction after this date of $u_i(t)$, $i = 1, 2, 3$, with the mobility and

the social behaviour assumptions made for the holiday time. This combination of effects has been demonstrated also by predictions obtained removing one control action at a time (simulations not shown).

IV. Concluding remarks

In containing the spread of COVID-19, robust mathematical models are important quantitative and predictive tools that can help understanding the disease and forecasting its future dynamics. A dynamical multi-group model designed to describe the epidemic evolution of COVID-19 within N sub-populations is formulated and implemented in the present paper. Each sub-population is homogeneous from an epidemic point of view, and it is geographically or structurally distinct from the others, but interconnected by a mobility network which allows (inwards and outwards) fluxes of people among groups. The epidemic model for each sub-population suitably extends a SEIR-type model accounting for both diagnosed and undiagnosed infectives, as well as for asymptomatic subjects, in the sub-groups. Moreover, control actions representing intervention measures adopted to contain the disease outbreak are modeled.

After giving a general formulation of the multi-group model, we have presented an application with each group representing an Italian macro-region. The multi-group model can be used to highlight social and economic aspects influencing the epidemic spread among groups of people geographically separated. The number of sub-groups within a whole reference population, and then the extension of the areas where the groups are located, depends on the level of detail required by the performed analysis: a lower number of groups belonging to larger areas allows high-level, coarse-grained analyses which focus on "macroscopic" aspects related to the group interconnection.

The present work aims to evaluate how the variations of human mobility, following the application of the government decrees, contributed and influenced the diffusion of COVID-19 in Italy. In particular, we intended to present a quantitative analysis of the Italian situation, focusing mainly on mobility related aspects. For this reason, the parametric identification of the model is performed on demographic and epidemiological data in a period centred around the summer 2020, which was characterized by a huge number of people movements across the Country. Also, the choice of the model sub-units (northern, central, southern macro-regions) is representative for the simulation period. Overall, the model simulation produced fitting results that show a very good agreement with the data, which is confirmed by the model predictions of Fig.4. The simulation evidenced how the combination of the summer exodus (mostly from North to Centre-South) along with the enhanced virus circulation in Northern Italy produces a rather moderate increment of cases during the late summer. Although the mobility aspects alone appear to be insufficient to immediately trigger a second epidemic wave, their combination with the re-opening of production activities and schools in mid September was crucial in determining the observed sharp increase of cases at the beginning of October.

An important development of the present study may consist in extending the data analysis and the estimation procedure to predict and then prevent the disease diffusion.

References

[1] World Heath Organization (WHO), "Coronavirus Disease (COVID-19) Dashboard," last accessed January 2020.

[2] P.Di Giamberardino and D. Iacoviello, "Evaluation of the effect of different policies in the containment of epidemic spreads for the COVID-19 case," *Biomed Signal Process Control*, vol. 65, no. 102325, pp. 1–15, 2021.

[3] M. Casares and H. Khan, "The timing and intensity of social distancing to flatten the COVID-19 curve: The case of Spain," *International Journal of Environmental Research and Public Health*, vol. 1, no. 7283, pp. 1–14, 2020.

[4] H. Sjödin, A. Wilder-Smith, S. Osman, Z. Farooq, and J. Rocklöv, "Only strict quarantine measures can curb the coronavirus disease (COVID-19) outbreak in italy," *Euro Surveill.*, vol. 25, pp. 1–6, 2020.

[5] S. M. Kissler, C. Tedijanto, E. Goldstein, Y. H. Grad, and M. Lipsitch, "Projecting the transmission dynamics of SARS–CoV–2 through the postpandemic period," *Science*, vol. 368, no. 6493, pp. 860–868, 2020.

[6] M. Lee, J. Zhao, Q.Sun, Y.Pan, W. Zhou, C. Xiong, and L. Zhang, "Human mobility trends during the early stage of the COVID-19 pandemic in the United States," *PLOS ONE*, vol. 15, pp. 1–15, 2020.

[7] H. S. Badr, H. Du, M. Marshall, E. Dong, M. M. Squire, and L. M. Gardner, "Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study," *Lancet Infect Dis*, vol. 20, no. 11, pp. 1247–1254, 2020.

[8] J.De Vos, "The effect of COVID-19 and subsequent social distancing on travel behavior," *Transportation Research Interdisciplinary Perspectives*, vol. 5, no. 100121, pp. 1–3, 2020.

[9] H. Gibbs and et al , "Changing travel patterns in China during the early stages of the COVID-19 pandemic," *Nature Communications*, vol. 11, pp. 1–5, 2020.

[10] P. Di Giamberardino, D. Iacoviello, F. Papa, and C. Sinisgalli, "Dynamical evolution of COVID-19 in Italy with an evaluation of the size of the asymptomatic infective population," *IEEE Journal of Biomedical and Health Informatics*, 2020.

[11] G. Giordano, F. Blanchini, R. Bruno, P. Colaneri, A. Filippo, A. Matteo, and M. Colaneri, "Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy," *Nature Medicine*, vol. 26, pp. 1–6, 2020.

[12] S. Contreras, H. A. Villavicencio, D. M.-O. abd J. P. Biron-Lattes, and A. Olivera-Nappa, "A multi-group SEIRA model for the spread of COVID-19 among heterogeneous populations," *Chaos Solitons Fractals*, vol. 136, p. 109925, 2020.

[13] K. L. K, M. Peirlinck, F. S. Costabal, and E. Kuhl, "Outbreak dynamics of COVID-19 in Europe and the effect of travel restrictions," *Comput Methods Biomech Biomed Engin*, vol. 23, no. 11, pp. 710–717, 2020.

[14] A. Scala, A. Flori, A. Spelta, E. Brugnoli, M. Cinelli, W. Quattrociocchi, and F. Pammolli, "Time, space and social interactions: exit mechanisms for the Covid-19 epidemics," *Nature Sci Rep*, vol. 10, p. 13764, 2020.

[15] E.Bertuzzo, L. Mari, D.Pasetto, and et al, "The geography of COVID-19 spread in Italy and implications for the relaxation of confinement measures," *Nature Communication*, vol. 11, pp. 855–860, 2020.

[16] ISTAT, *Intercensal population estimates. Demographic balance*. Italian National Institute of Statistics, 2018, http://demo.istat.it/index.html.

[17] J. T. Wu, K. Leung, and G. M. Leung, "Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study," *The Lancet*, vol. 395, pp. 689–697, 2020.

[18] B. Tang, X. Wang, Q. Li, N. L. Bragazzi, S. Tang, Y. Xiao, and J. Wu, "Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions," *Journal of Clinical Medicine*, vol. 9, p. 462, 2020.

[19] Dipartimento della Protezione Civile, *COVID-19 Italia - Monitoraggio della situazione*. GitHub, 2020, https://github.com/pcm-dpc/COVID-19/tree/master/dati-andamento-nazionale.

[20] ENIT - Agenzia Nazionale Italiana del Turismo, *Bollettino N.8-Ufficio Studi ENIT*, Ministero del Turismo, https://www.enit.it/wwwenit/images/multimedia/Bollettino_Ufficio_Studi/Bollettino_8/BOLLETTINO-ENIT-N8.pdf.