DOI: 10.5281/zenodo.6974199

Project	GUTTA
Work Package number	4
Work Package title	Developing tools for upgrading environmental sustainability and connectivity of regional mobility
Deliverable number	4.3.2.
Deliverable title	Understanding the COVID-19 dynamics from cross-border mobility data
Deliverable Responsible Partner	MMPI
Deliverable Lead authors	Corentin Cot (CNRS), Giacomo Cacciapaglia(CNRS) Dea Aksentijević (UNIRI), Alen Jugović (UNIRI), Tomislav Budić (MMPI), Gianandrea Mannarini (CMCC)
Deliverable Contributors	Marko Pripić, Davor Deželjin (MMPI)
Deliverable due date	2022-01-31
Deliverable latest review date	2022-08-29

Table of Contents

1. Executive Summary

The COVID-19 pandemic hit the European Countries starting from early 2020. After the first outbreak in February, a second wave developed with different timings during summer. It has been speculated about the role of Countries with a strong touristic appeal in triggering such a second wave. Therefore, a data-driven investigation was conducted to assess the possible role of people mobility to and within Croatia in the period between June and September 2020. To this end, various datasets were collected and processed through the modelling framework of the epidemic Renormalisation Group. The preliminary results ruled out a significant role of the maritime transportation in speeding-up the curve of contagion in the various regions of Croatia.

2. Introduction

With the growth of the human population and its impact on the environment, our societies are becoming increasingly vulnerable to new diseases, especially viral infectious diseases of zoonotic origin. At present, only 3% of the land ecosystems are untouched by human activities [1]. Furthermore, human-induced climate change is causing relocation of species and rapid migration of humans. Economic globalisation has also increased the mobility of both goods and people across countries and continents. All these factors play in favour of the transmission of viral pathogens from animal species to humans, and their rapid diffusion within the world population, with the COVID-19 pandemic being the latest example. The COVID-19 pandemic has also dramatically shown the unpreparedness of human society to face the threat of a pandemic and its inability to efficiently cope with it during the emergence of epidemiological waves.

Henceforth, it has become of paramount importance to introduce protocols and preparedness measures that help governments, private companies, and individual citizens to face a new viral pandemic. In this context, passenger mobility plays a crucial role as it allows for the spreading of infectious diseases across different regions of the world, especially at the beginning of a pandemic like COVID-19. In this study, within the GUTTA project, we aim at quantifying the impact of maritime transportation between Italy and Croatia on the diffusion of COVID-19. In fact, the 'second wave' of infections in Croatia, starting in July 2020, coincided with the restart of maritime connections after the first lockdowns. We performed a detailed analysis of mobility data in Croatia, both maritime and terrestrial. Also, we employed a novel approach to infectious disease spreading to determine the role played by maritime transportation. In our data-driven approach, a mathematical modelling of the infections is used to quantify the link between mobility data and data on the number of infections.

The main result of this study can provide guidance to local governments, decision-makers and healthcare stakeholders on the most effective actions to be taken at the beginning of a pandemic, in order to curb and limit the diffusion of the infections.

Note: for shortcuts, please refer to GUTTA Glossary available at: https://zenodo.org/record/5713897

3. Methodology

To study the diffusion of the COVID-19 infections in Croatia, we combined data describing the people's mobility via various transportation means with an epidemiological dataset. The latter describes the number of new infected people that were tested positive during a period of time. The two datasets were combined within the epidemic renormalisation group (eRG) framework, as detailed below. The main advantage of the mathematical model provided in the eRG is to allow us to characterise, in simple terms, the diffusion of an infectious disease within connected regions. Hence, the position of the peaks, i.e., the timings of the local maxima of new infections in different regions, can be predicted as a function of the mobility data. Comparing the predictions of the model with the actual data allows us to determine the role of various transportation means in enhancing the diffusion of COVID-19. This mechanism is expected to be the dominant one at the beginning of the epidemic. In Croatia, the first epidemiological wave, characterized by a temporary exponential increase of the cases, took place in March through April 2020. After a period of calm, a new increase was detected starting towards the end of June and lasting through the beginning of August. We identify the latter as the 'second wave' (see Fig.1).

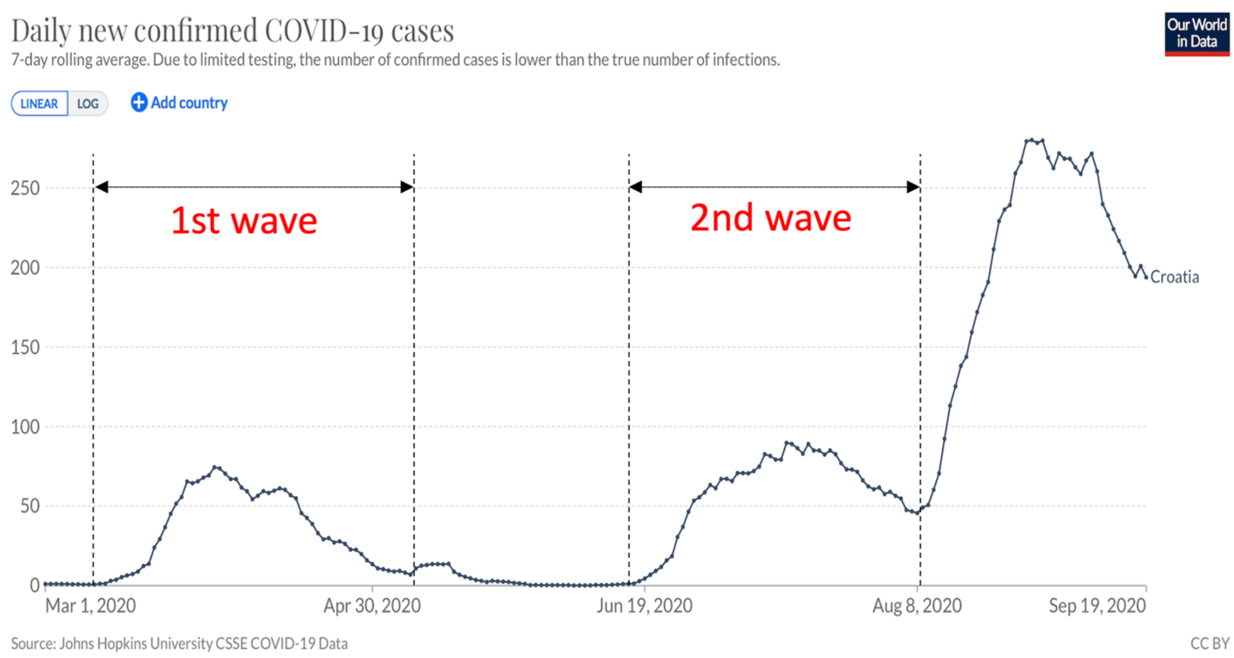

Figure 1 Number of new COVID-19 cases in Croatia in the period between March and September 2019. The first two waves are highlighted by vertical dashed lines.

3.1. Datasets

Following datasets for 2019 and 2020 (see also Tab.1) were used for the subsequent analyses:

1) Croatian integrated maritime information system (CIMIS). The dataset, provided by the GUTTA partner MMPI, is organized into seaports and include information regarding departures and arrival times of ferries along with the number of both embarking and disembarking passengers.

2) Car traffic. A dataset provided by UNIRI and Hrvatske ceste LLC containing the number of travellers per year (or per summer) crossing each Croatian border control checkpoint. It is therefore possible to assume the country of origin and the region of destination for each border.

We only included in the analysis the data regarding border crossing with the neighbouring countries (Slovenia, Hungary, Serbia, Bosnia-Herzegovina, Montenegro).

3) Railway traffic. IRG ("Independent Regulators' Group - Rail") provided by the Croatian government and from HŽ Infrastruktura that is taking care of the railway in Croatia: a report for 2020 which deals with the impact of the COVID crisis on the network. Data regarding cross-section of the railway network, how network looks like and under what conditions and in what way it is used (who are the stakeholders, infrastructure, etc.) was taken from the network report. Railway traffic was recorded on an annual basis.

4) Air traffic data. This information was procured by MMPI from the Eurocontrol Air traffic Directorate in Lyon and consisted in the number of travelers per month per airport in Croatia for 2020 and 2021.

5) COVID-19 epidemiology. The data were extracted from a Croatian public resource[1]. They included the total number and the number of daily new infected for each Croatian county. This number corresponds to the individuals that reported a positive test. We extracted the data from the 21st of March 2020 to the 18th October 2021, corresponding to 575 days in total. The raw data were pre-processed to smoothen daily fluctuations. For that, we used a moving 7-day average.

Tab. 1 Description of mobility datasets used in this work. MMPI is the Croatian Minister for Sea, Transport and Infrastructure (a GUTTA project Partner) and UNIRI is the University of Rijeka (a subcontractor of MMPI).

ID		**Vehicles**	**Name**	**Provider**	**Resolution**		**#data-points**
					Time	**Space**	
D1	Maritime	Ferries	CIMIS	MMPI	weekly	by port	1360
D2	Terrestrial	Cars	Highway data	UNIRI	summer	borders	1
D3		Trains	Railway data	UNIRI	annual	borders	1

[1] https://www.koronavirus.hr/podaci/otvoreni-strojno-citljivi-podaci/526

D4	Airborne	Planes	Air traffic data	MMPI	monthly	airports	21
D5	Epidemiology	—	New cases	MMPI	daily	county	575

Some features of the datasets used are provided in Tab.1. The largest number of people was moved on the terrestrial mode via cars; hence the D2 dataset is supposed to play a major role in determining the diffusion of the infections.

3.2. The eRG framework

The epidemiological Renormalisation Group (eRG) framework has been proposed in [2] to describe the exponential increase in the number of new infections, followed by a reduction back to approximately zero. We will refer to this phenomenon as a "wave". As we have seen in Figure 1, Croatia experienced three waves of COVID-19 infections between March and September 2020, and additional ones in the following years.

Without entering into the mathematical details, the eRG framework consists in a differential equation that describes a single epidemiological wave propagating in an isolated region [2,3].

A strength of the method lays in the fact that two parameters are sufficient to characterise the wave: an effective infection rate, γ, related to the duration of the wave, and the total number of individuals that are infected during a single wave, A. The latter can be normalised to the total population of the region, so that in the following we will express it in number of infections per million inhabitants. The values of γ and A can be obtained by fitting the epidemiological data in a specific region. Note that γ does not depend on the number of infected. As such, it does not suffer from biases coming from the number of available test kits and the policy of testing adopted during various phases of the pandemic, nor on any possible regional differences. Hence, γ offers a faithful characterisation of the severity of each wave in different regions and at different times.

The eRG has also been extended to include the mobility of people between different regions, as long as the flow only involves a small fraction of the region inhabitants [3,4,5]. This extension allows to study the relation between the emergence of epidemiological waves in different regions, and the mobility of people in between regions. In particular, the timing of the wave peaks can be directly related to the mobility flow, providing a handle to quantify the impact of various transportation means on the COVID-19 diffusion.

Application of the eRG framework to the datasets of Tab.1 available for Croatia implies the following steps:

- We subdivide Croatia into homogeneous regions (see next section).

- For each wave, we fit the eRG parameters on the available epidemiological data. For this study, we only focus on the first two waves, occurring between March and August 2020.

- We use the eRG equations, together with the fitted parameters, to numerically calculate the diffusion of the waves to the different regions.

- We create scenarios, where various mobility datasets are included with a weight. The latter parametrises the effective impact of the actual transportation means on the virus diffusion. In practice, this weight corresponds to the probability of finding infected individuals among the passengers of that specific transportation mean.

- We compare the result of the numerical equations to the observed data, to establish which configuration best fits the timing of each wave in the different regions.

The numerical computations have been performed on a personal computer, using the software Wolfram's Mathematica (licensed via the CNRS).

3.3. NUTS-2 regions and the issue of Zagreb

Croatia is a diverse country in terms of geography, population density, and demography. It consists of a long coastal region of great touristic interest, and an internal region where the capital Zagreb is located. Furthermore, it has numerous neighbouring countries: Slovenia and Hungary in the North, Serbia in the East, Bosnia-Hercegovina and Montenegro in the South. Finally, it shares a maritime border with Italy in the west (along the Adriatic Sea). The main connections with abroad, therefore, are realised via road and railway, via maritime transport (with Italy mainly) and flights.

To use the data to characterise the diffusion of the virus within Croatia, we need to first establish regions within Croatia to be associated to the eRG equations. One possibility is to study Croatian counties, taking into account the subtlety of the mobility within the counties and with neighbouring countries. This corresponds, in the Eurostat nomenclature, to the "NUTS-3" level. However, this subdivision level is problematic for multiple reasons. Firstly, the population of different counties is highly unequal, going from 50.000 inhabitants in Lika-Senj to 800.000 in Zagreb city. This unequal distribution of the population would amplify statistical fluctuations in the epidemiological data for counties with low population. Furthermore, the unequal weight of different counties can bias the numerical output of the eRG computations. Thirdly, it is hard to quantify the mobility flows between counties due to the smallness of these territorial units. We had access to the flow along main highways as provided by the Croatian ministry of transport (MMPI), however this data would be insufficient to provide a reliable estimate at county level. In fact, drivers passing through the checkpoints may drive across many counties, and we did not have any information on their departure point and destination. Henceforth, we deemed this level of geographical granularity to be inappropriate for our purposes.

We opted instead for the 2021 NUTS-2 level, where Croatia is subdivided into four statistical regions as in Fig.2:

- Pannonia (HR02, Panonska Hrvatska), 1,054,000 inhabitants, stretching to the East and touching Hungary, Serbia and Bosnia-Hercegovina.
- Adriatic (HR03, Jadranska Hrvatska), 1,372,000 inhabitants, comprising of the coastal region and touching Slovenia, Bosnia-Hercegovina and Montenegro.
- Zagreb (HR05, Grad Zagreb), 800,000 inhabitants, surrounded by the Northern region.
- Northern (HR06, Sjeverna Hrvatska), 813,000 inhabitants, in the north and touching Slovenia and Hungary.

This division is much more uniform in terms of population and minimises the statistical uncertainties on the data. It is also more suitable for studying the diffusion of the infectious disease, as each region has some specificities that makes their role unique. For instance, Adriatic is the only region that is connected by maritime transportation. Furthermore, the main international airports are in Adriatic (Split) and Zagreb. All regions except Zagreb are connected to neighbouring Countries via border road and train connections.

The above facts highlight an issue with the mobility data related to Zagreb: the absence of a direct border with other countries implies that the available road and train data are absent in the national dataset available for this study. Hence, we miss direct information about the mobility directed from abroad to the capital city of Zagreb, which is expected to be of major importance. To consider this missing information, in the numerical results, we reroute part of the road and train flow across boundaries to the Zagreb region. This is justified by the fact that highways and railway lines connect the boundaries directly to Zagreb, while crossing any of the other three NUTS-2 regions of Croatia. We will investigate a few scenarios where a fraction of this traffic is attributed to Zagreb instead of the other regions.

Figure 2 NUTS-2 regions of Croatia with a schematic indication of people's mobility flows. Stars indicate the major airport, while the blue lines show an example of Italy-Croatia maritime routes.

4. Results

We here present results from different eRG simulations that consider the different mobility datasets of Tab.1 as an input. In these simulations, we will investigate each transportation mean by considering them individually, then by proposing multiple scenarios in which we mix all of them. By use of statistical indicators, we will show that some scenarios completely misrepresent the observed epidemiological data, while others match them more closely.

We focus on reproducing the epidemiological data of the second wave, that hit Croatia in the summer months of 2020 between June and August. The main reason for this choice is that this wave may have been influenced by the holiday season, where all types of mobility were intensely used for touristic reasons.

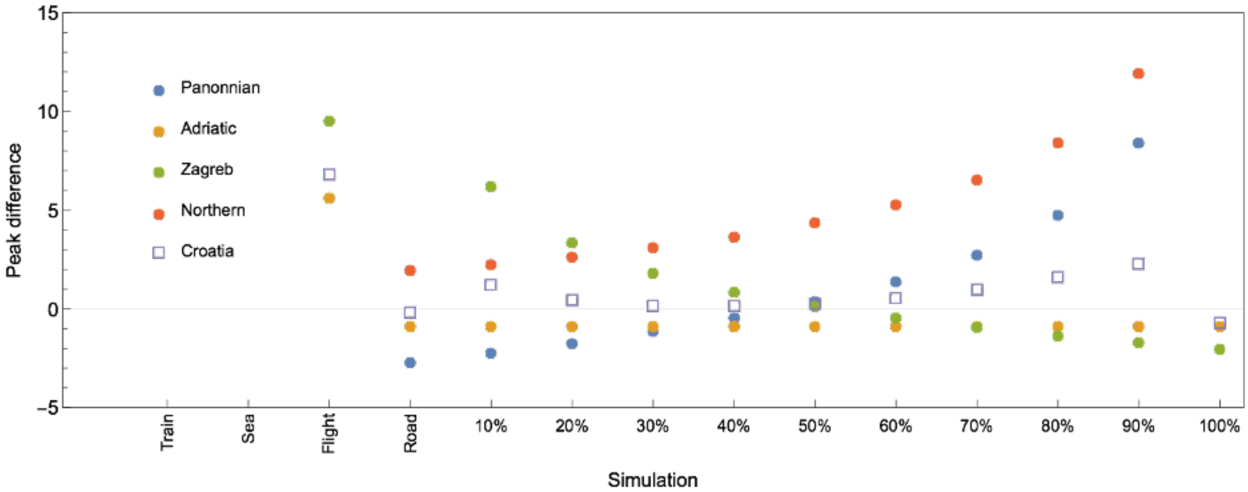

Fig. 3 Difference in the peak timing (in days) between the result of the eRG calculation and the fitted data. Here, we only include one type of mobility at the time. The points to the right of "Road" indicate the percentage of road mobility data rerouted to Zagreb, where only road mobility is included.

4.2 The role of maritime transportation

We first want to check what are the main effects of each type of mobility on the simulation. To do this, we generated a simulation for each transportation type (maritime, train, flights and road) to see how the data can be reproduced by only using one of them at the time, as shown in Fig.3. To quantify the fitness of the eRG calculation, we show in the figure the time difference between the peak obtained by the calculation and that of the fits. The first four entries ("Train", "Sea", "Flight", "Road") show the impact of the four vectors we consider here, where coloured points correspond to the NUTS-2 regions, while the square corresponds to the cumulative data for the whole country. We can directly see that train and sea are not enough to explain the data as all related peak differences are even out of scale in the plot. Instead, flight and road data seem to improve the simulation while not being able to explain the whole trend. In particular, road data cannot cover the Zagreb region, as it is enclosed as explained above.

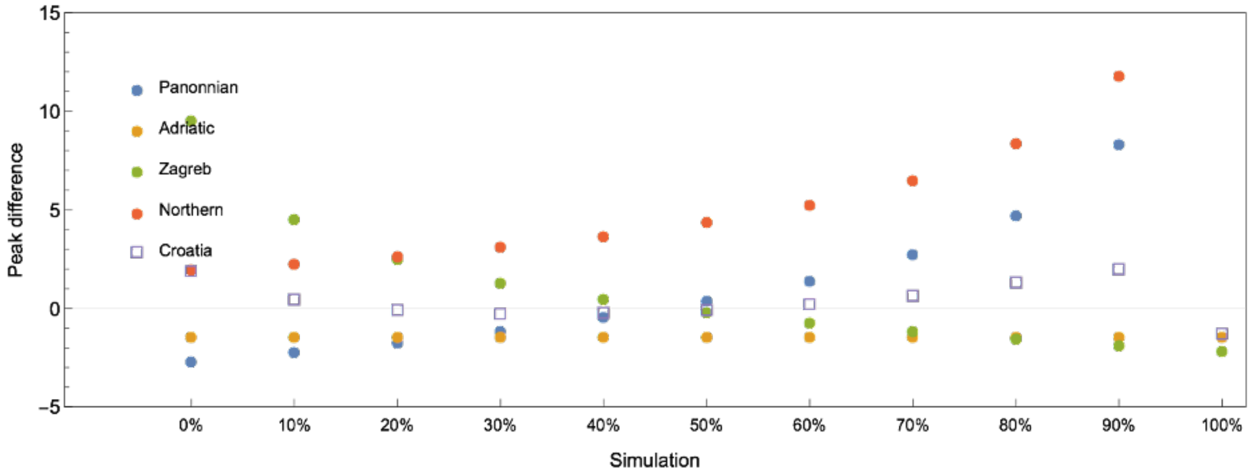

Figure 4 Peak differences (in days) once all transportation vectors are included. The percentage indicates the rerouting percentage to Zagreb from the other regions. We use the same rerouting rate for all regions of Croatia percentage to Zagreb from the other regions.

Hence, we tried to reroute a fixed percentage of the road data from the other regions to Zagreb, obtaining the results on the right (labelled by the percentage of rerouting). This shows that the diffusion of the virus during the second wave in Croatia can be well modelled by just road data. In order to minimise the mismatch for all Croatian NUTS-2 regions, a rerouting of 30% of the traffic to Zagreb is needed. This result allows us to conclude that, during the summer 2020's wave, both maritime and railway transportation played just a minor role in the propagation of COVID-19 in Croatia.

4.3 Optimal traffic diversion to Zagreb and flights

Now that we investigated the possibility of having a dominant mobility mean for the infection in Croatia, we can clearly see that road traffic is the main factor of propagation. As argued previously, in order to advocate for a road traffic going to Zagreb we had to consider that a certain percentage of the total car traffic going to Croatia should be rerouted to Zagreb instead of other regions. We simulated, therefore, a scenario where all mobility data are included, while a certain percentage of the road traffic is rerouted to Zagreb from the other regions. The results are shown in Fig.4, where we can clearly see that a rerouting of 30—40% can well reproduce the epidemiological data in all regions of Croatia. Quantitatively, this means that the peaks are reproduced in the eRG numerical analysis with a mismatch of maximum five days from the observations.

Conclusions

We have quantitatively analysed mobility data with the epidemiological data during the period of March-August of 2020 in Croatia, within the eRG framework. The main goal was to establish the impact of various mobility vectors on the diffusion of the virus, at the origin of the second wave in Croatia (June-August 2020). Our results show that maritime transportation (and train) did not play any significant role in the onset of the second wave at the end of June 2020, even though the timing coincides with the restart of the maritime routes to Italy after the first 2020 lockdown.

Instead, we demonstrated that road and flight mobility were the main players, as they can successfully reproduce the timing of the waves in all NUTS-2 regions of Croatia when integrated to the eRG framework. However, since the Zagreb city region is enclosed within the Northern region, we had to assume that a rerouting of road traffic to Zagreb took place from the other regions of Croatia. We find that a rerouting rate of 30% is optimal in reproducing the epidemiological data.

This information can provide guidance to local governments, decision-makers and healthcare stakeholders on the most effective actions to be taken at the beginning of a pandemic, in order to curb

and limit the diffusion of the infections.

References

[1] Plumptre, A.J., Baisero, D., Belote, R.T., Vázquez-Domínguez, E., Faurby, S., Jędrzejewski, W., Kiara, H., Kühl, H., Benítez-López, A., Luna-Aranguré, C. and Voigt, M., 2021. Where might we find ecologically intact communities? Frontiers in Forests and Global Change, p.26., https://doi.org/10.3389/ffgc.2021.626635

[2] Della Morte, M., Orlando, D. & Sannino, F. Renormalization Group Approach to Pandemics: The COVID-19 Case. Front. Phys. 8, 144 (2020). https://doi.org/10.3389/fphy.2020.00144.

[3] Cacciapaglia, G. & Sannino, F. Interplay of social distancing and border restrictions for pandemics (COVID-19) via the epidemic Renormalisation Group framework. Sci Rep 10, 15828 (2020). https://doi.org/10.1038/s41598-020-72175-4. 2005.04956.

[4] Cacciapaglia, G., Cot, C. & Sannino, F. Second wave COVID-19 pandemic in Europe: A temporal playbook. Sci Rep 10, 15514 (2020). DOI https://doi.org/10.1038/s41598-020-72611-5. 2007.13100.

[5] Cacciapaglia, G., Cot, C., Islind, A. S., Óskarsdóttir, M. & Sannino, F. Impact of US vaccination strategy on COVID-19 wave dynamics. Sci. Reports 11, 10960 (2021). DOI https://doi.org/10.1038/s41598-021-90539-2. 2012.12004.