sustainability

MDPI

Article

Using Transport Activity-Based Model to Simulate the Pandemic

Moez Kilani [1,*], Ousmane Diop [1] and Ngagne Diop [2]

[1] Lille Economie et Management, UMR 9221, Département Economie Gestion, Université du Littoral Côte d'Opale, 59140 Dunkerque, France
[2] Territoires, Ville, Environnement et Socialété, ULR 4477, Département Economie Gestion, Université du Littoral Côte d'Opale, 59140 Dunkerque, France
* Correspondence: moez.kilani@univ-littoral.fr

Abstract: We use an activity-based transport model to simulate the progression of a virus at the regional scale. We analyse several scenarios corresponding to distinct situations and describing how small initial clusters of infected agents expand and reach a pandemic level. We evaluate the effectiveness of some public restrictions and compare the number of infections with respect to the base-case scenario, where no restrictions are in place. We consider the wearing of masks in public transport and/or in some activities (work, leisure and shopping) and the implementation of a lockdown. Our analysis shows that education, including the primary level, is one of the major activities where infections occur. We find that the wearing of masks in transportation only does not yield important impacts. The lockdown is efficient in containing the spread of the virus but, at the same time, significantly increases the length of the wave (factor of two). This is because the number of agents who are susceptible to be infected remains high. Our analysis uses the `murdasp` tool specifically designed to process the output of transport models and performs the simulation of the pandemic.

Keywords: activity-transport simulations; the dynamic of the pandemic (COVID-19); social and physical distancing

Citation: Kilani, M.; Diop, O.; Diop, N. Using Transport Activity-Based Model to Simulate the Pandemic. *Sustainability* **2023**, *15*, 2257. https://doi.org/10.3390/su15032257

Academic Editor: Xiang Li

Received: 14 December 2022
Revised: 30 December 2022
Accepted: 6 January 2023
Published: 26 January 2023

1. Introduction

The COVID-19 is a highly transmissible and pathogenic coronavirus that spread since 2019 and caused 6.65 million deaths worldwide (more than 150,000 in France) until now (https://www.worldometers.info/coronavirus/ (accessed on 2 December 2022)). Most governments were not familiar with the situation and the absence of vaccine in the early stage of the pandemic led to great difficulties in identifying the appropriate sanitary measures to impose [1,2]. Severe restrictions, including strict lockdown, were in place for several weeks and caused important disruptions in the social life and economic activities. This crisis pointed out the importance of simulation tools that can be used to analyse the dynamic of a pandemic under distinct contexts and depending on specific restrictions in place [3]. This should fit with existing transport reforms that are dealing with challenges of efficiency and sustainability (energy transition, accessibility to the low incomes). In the post-pandemic perspectives, transport policies should deal with these multiple objectives in the development of public transport services. Simulations tools, which use has dramatically increased in the last decade, need to include the possible spread of the pandemic, so that authorities can examine how public transport services should be tuned to deal with a specific epidemic problem. While this research is concerned with the COVID-19 pandemic, its extensions to any other epidemic context is straightforward.

The combination of transport simulation and epidemic models is one of the most relevant uses of mathematical and computational models. The transport model provides the population and interaction network of agents (individuals) to the epidemic model in order to produce the dynamic of infections. The epidemic model takes into account the heterogeneity of the agents' behavior, and thus, the risk of infection depending on several factors such as barrier measures and social distancing. There are several methodological approaches to model the pandemic [4]. For example, networks can be used to link distinct geographical locations. Depending on the connectivity between the nodes, infections can be more or less frequent [5]. This approach can take into account transport and can be used to study how social distancing can reduce the number of infections. To provide a more realistic and detailed description, however, agent-based simulation is more convenient. Indeed, since all the individuals' daily events are tracked, we can identify the locations and the durations of all the interactions between the agents and infer the probability that an infected agent contaminates a non-infected one. Consequently, by processing all events, all the related characteristics of the dynamic of the pandemic can be identified and described.

At the time this paper is written, Europe, including France, is facing the ninth wave of the COVID-19 pandemic. Even if the latest variants produced small and moderate fatalities by comparison to the earlier waves, public authorities are still deploying important efforts to deal with possible worst-case scenarios. There are great concerns that the medical system reaches levels of tense activity, where some health operations will need to be postponed, and even higher levels where medical services become overloaded. Our analysis contributes to understand the dynamic of a wave of the pandemic and to evaluate how effective are the sanitary measures, in connection with public transport services.

In this paper, we use an activity-based model of urban and regional transport for the North of France to produce social interactions. Then, by applying an epidemic model, we simulate the spread of the virus and characterize the dynamic with respect to time and space. Indeed, we identify the locations of infections and evaluate the frequency of their occurrence in public transport modes. This framework is used to compare some scenarios such as the wearing of masks, social distancing, and the consideration of vaccination, with respect to the base-case scenario where no sanitary measures are used.

The analysis developed below uses the `murdasp` tools, which is a set of scripts specifically written to process the output of a transport simulation model and build all interactions, on which it applies applies an epidemic model. A detailed description of the `murdasp` tool is given in [6]. The approach is transparent and flexible, allowing a fine tuning of all the model parameters. Agents' travel behavior is strongly related to their locations and to the daily schedule of their activities. A comprehensive transport model should, to some extent, take into account the nature of these activities, their durations and how they are chained during a typical day [7]. Our analysis is based on the output of `MATSim` transport simulator [8], but the used tool (i.e., `murdasp`) can be easily tuned to use the output of any simulator that produces detailed description of the daily events.

The paper is organized as follows. Section 2 reviews epidemic models and how they can be combined with transport models. The used transport model is briefly described in Section 3. The epidemic model is calibrated in Section 4 and the simulation of several scenarios is discussed in Section 5. Section 6 concludes.

2. Epidemic Models and Transportation: A Brief Review

We give a short review of the literature, focusing on recent works that have considered transport and activity-based models, and then describe an epidemic model.

2.1. An Overview on the Literature

The first epidemic model dates back to [9], who setup a compartmental model where, in case of viral infections, each agent is either susceptible (S), those who are not yet infected, infected (I) or recovered (R) after the infection. A contamination rate β reflects the intensity of contamination corresponding to transitions from (S) to (I) and recovery rate corresponding to transitions from (I) to (R). The dynamic model that builds on these transitions is referred to as `SIR` and is generally described, in continuous time, on the basis of a set of ordinary differential equations. The solution to these equations describes the evolution of the agents through the three compartments.

As illustrated in Figure 1, the SIR model has several extensions and improvements to deal with further details like incubation delays or asymptomatic agents. The plain lines (in blue) relate the three basic compartments "susceptible", "infected" and "recovered". Extensions of the basic model add further compartments to provide a finer description that takes into account more possible states like "exposed", in "Quarantine" or "isolated" which are interrelated as shown by the dotted lines in Figure 1.

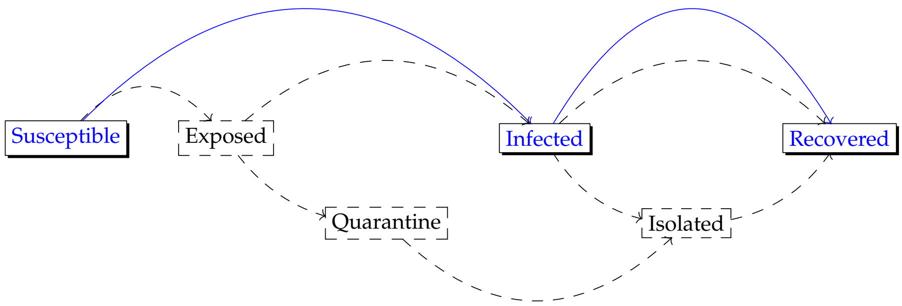


Figure 1. The SIR model and extensions. The basic three compartments are shown in blue and connected by plain lines. A Possible extension is illustrated through three further compartments connected with dashed lines.

Another main limitation to the basic model is the hypothesis of homogeneity, making all the individuals equally treated with respect to infections and recovery. Indeed, several realistic features of the pandemic are directly dependent on individual attributes. The extended models generally build on the individual centered features, called also multi-agent models. This approach is based on simulation and uses individual behaviour and attributes [10]. Interactions between agents are the source of contamination, the frequency of which depends on the duration of these interactions and the places where they occur. Sanitary measures can then be applied and tested to assess their impact on the spread of the pandemic.

For the case of COVID-19, Ref. [11] is one of the first studies that have described how the virus spreads in a population and suggested some policies to control it. The authors' results have been taken into account in Great Britain and abroad in developing some public policies to face the pandemic. They show that, in the absence of vaccine, social distancing plays an important role in reducing the transmission of the virus. Appropriate application of these measures can reduce demand for emergency health care by 2/3 and divides fatalities by half. However, the reduced contamination leaves a large part of the population exposed, and strong measures, like school closing and isolation of risky persons, are the only options to contain to pandemic on a longer perspectives.

With respect to transportation, the analyses in [12,13] combine activity-based simulation models with epidemic models to describe the dynamic of the pandemic. The transport simulator `MATSim` produces the activity schedules of all the agents [8]. This output is then processed to produce social interactions on which the epidemic model is applied to simulate infections. A dedicated tool, called `Episim`, was developed to perform the required operations [13]. The probability that an infection occurs depends on several factors, including physiological respiratory characteristics. Their study area is the city of Berlin where they notice that the behavior of the individuals has changed even before social distancing was in place. The authors show that actions, like wearing of masks, can efficiently reduce the speed of propagation of the virus. But, definitely, in a serious and risky situations, only strong actions like the lockdown can effectively reduce the number of infections.

A related study was conducted in [14] for the city of Montreal. The authors have used a similar approach based on multi-agent modeling. The main objective in this study is the contribution of the distinct activities to infections. The probabilities of infections are differentiated depending on the intensity of the contacts. The authors find that most infections occur in houses, primary schools and workplaces. By comparison, the contributions of activities like leisure or shopping are quiet limited. When we focus on workplaces, there is a significant difference between distinct economic sectors. For example, and for obvious reasons, those working in medical activities are highly exposed by comparison to the average level in the population. Related to this fact, there are differences with respect to gender, since women and man are unequally distributed among distinct sectors.

2.2. The Structure of an Epidemic Model

An epidemic model has two stages. The first one describes infections (transmission of the virus) and the second one describes progression of the disease in the infected agents. So, at a first stage we estimate the probability of a susceptible agent to be infected when exposed to a contagious agent. At the second stage, given that an agent is infected, we have estimated when he is contagious and when he recovers. Of course in real situations, like the COVID-19 pandemic, there are many other attributes that need to be accounted for.

For the first stage, the transmission of the SARS-CoV-2 virus generally occurs through aerosol infections. Assume that when a susceptible agent is in contact with an infected and contagious agent with index m for a time span τ_m the probability it is not infected is $exp(-\theta\,\tau_m)$, where the positive parameter θ reflects the transmission rate which depends on the characteristics of the virus and the context of the interaction (intensity of the contact, whether the agents are wearing masks, etc.). This probability is decreasing; when τ_m is zero the agent is not infected with probability 1, and when $\tau_m = \infty$, the agent is not infected with probability zero. This formulation is thus consistent with basic assumptions. When a susceptible agent interacts with several contagious agents $m = 1\ldots,n$, and if we assume that the probabilities of infections are independent, then the probability he is not infected is $\Pi_{m=1}^{n} \exp(-\theta\,\tau_m)$. Thus, the probability it is infected is the complement of this expression. Indeed, the probability of an infection of an agent given it was exposed to other infected agents is frequently computed accordingly and through the exponential expression (see [13]) given by:

$$\mathbf{P}_a(\text{infection} \mid n \text{ contacts}) = 1 - \exp\left(-\theta_a \sum_{m=1}^{n} \tau_m\right). \tag{1}$$

Here, we add an index to the transmission rate θ to indicate that it corresponds to activity a. It is considered constant which is a reasonable assumption if the interactions occur at the same location. For example, if we consider that interactions in leisure activities are more favorable for infections than interactions in shopping activities we can consider two parameter values θ_{leisure} and θ_{shop}, respectively, such that $\theta_{\text{leisure}} > \theta_{\text{shop}}$. By extension, the index a covers both activities and transport modes.

The exponential distribution given in Equation (1) has the property of being memoryless, i.e., $P(X > x + y|X > y) = P(X > x)$, where P denotes the probability, x and y some given positive values and X a random variable distributed according to the exponential density function. Stated in simple words, the probability that an event occurs after noon does not depend on whether it occurred before noon or not. This is useful in simulations since the events are updated at fixed intervals of time, generally equal to one second. Based on this feature, we simply compute the probabilities of infections for distinct intervals as the simulation progresses.

To take into account physical attributes of the places where agents meet and interact, other parameters can be added. For example, in closed areas the room size and the air exchange are important and need to be taken into account. With respect to the agents, it is usual to consider the shedding rate and the intake, which reflect more or less physical interaction, and the main means for virus transmissions between distinct individuals. At this first stage, it is important to know whether sanitary measures are in place, or not, since they significantly impact the probabilities of transmissions. For example, wearing masks can be taken into account by reducing the shedding rate. These components enter expression in (1) as multipliers of parameter θ.

At a second stage, we consider the model of the progression of the infected agents. This stage is dependent on biological attributes of the virus and the infected bodies. A compartmental approach, extending the SIR, can be used. For COVID-19, and based on the literature, some extended models distinguish between seven compartments (cf. Figure 1): susceptible, exposed (infected but not contagious), infectious, showing symptoms, seriously sick (needing health care), critical (needing intensive health care), recovered. In the activity-based model we consider here, it is not important to dive into the underlying biological details related to these distinct states, since we mainly need to know whether an agent is infected, whether he is contagious and when he recovered. Notice that the sequence of states depends mainly on the date of the infection, and this date is the main variable to keep in memory to determine the compartment where a given agent is in any period of time. Our approach follows this strategy.

3. Activity-Based Transport Models and Social Interactions

The pandemic model builds on transport model based on agents' activities. We use the transport model described in [15], and we only report in this section its structure and main features. Distinct agents have distinct schedules, which are more or less complex. Two possible schedules are illustrated in Figure 2. Simple home-to-work round trip is first shown in the left panel for an agent who uses private car only. In the second example (on the right panel), the agent has more activities and he uses the private car, but also public transport for some trips, for example those in downtown. Transport is thus a link between distinct activities. Social contact occurs in these activities and in public transport. By simulating daily transport, we construct schedules of all the agents and we can thus evaluate when each agent interacts with another one, i.e., when they are at the same location. Of course more complex schedules, where some users combine multiple modes (intermodal transport) for the same trip, are also possible and are taken into account in our framework. Indeed, we follow the schedules reported in the census data and that correspond to realistic observations.

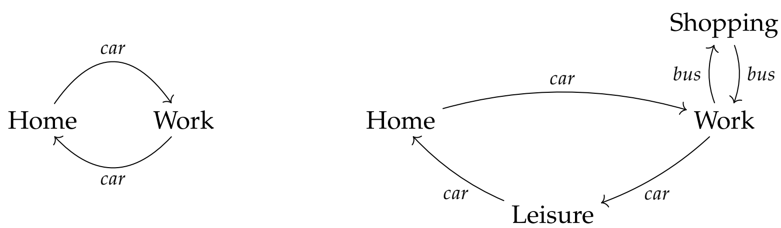


Figure 2. Examples of daily activities schedules.

The study region in given in Figure 3. It covers the departments "Le Nord" and the "Pas-de-Calais" with a total population of about four million inhabitants (2.6 million in "Le Nord" and 1.4 million in "Pas-de-Calais"). We take into account both urban, suburban and intercity trips. Lille is the largest metropolitan area, with more than 1.1 million inhabitants. Other major cities include Valenciennes, Dunkerque (Dunkirk), Calais, Boulogne-sur-Mer and Arras. The historical mining activities partly explains the trip patterns observed between the mining area (the shaded region in Figure 3) and major cities; Lille, in particular. Motorways around this city are severely congested during the peak hours. Also, the "A1" motorway, linking Lille to Paris, is a main connection for freight transport between northern Europe, Paris region and other southern cities in the south of France and Spain. The region itself is a main place for logistics activities, including warehousing, since it connects several ports (Dunkerque, Calais and Boulogne-sur-Mer) and multimodal platforms (for example, Dourge which is located in the south of Lille).

Figure 3. The region covered by the model (North of France). The two departments, "Le Nord" and "le Pas-de-Calais" are represented as well as the historical mining region.

The regional dimension of this model is particularly relevant for the spatial analysis of the pandemic. Indeed, an initial cluster of infected agents at a given agglomeration or a rural area expands, under the absence of restrictions, and spreads at the regional scale. The transmission occurs through transport (collective modes, in particular) and because several home-to-work trips are not within the same district. In particular, home-to-work trips between the metropolitan area of Lille and the mining region (shown in Figure 3) are important. Thus, a proper modeling approach should take into account urban transport (including suburban areas), inter-city trips as well as transport in rural areas. As many other regions in France, the North of France was largely impacted by the pandemic, especially by 2020 and 2021 where a lockdown was in place for several weeks to limit the spread of the virus. In the department "Le Nord", and between 1 Marsh and 20 April mortality rate increased by 20%. (See https://www.coronavirus-statistiques.com (accessed on 2 December 2022)).

We take into account four transport modes "walk", "bike", "car" and "public transport", and allow for their combinations (to some extent). Public transport includes "bus", "train", "metro" and "tramways" and each of these modes has a distinct parameter θ. Data from the Regional Household survey of 2016 is used to build the schedules of daily trips. The database covers a comprehensive description of the trips (origin, destination, departure and arrival times, purpose, modes, etc.) and the travelers (age, job qualification, household size, availability of a car, etc.). In the transport model, a "synthetic population" of agents is considered: instead of the whole population, a representative sample is considered. This practice is standard in transport simulation to speed-up computation time [8], and for case studies with millions of inhabitants the size of the synthetic population is generally between 5% and 30% of the total population. To keep the simulation consistent with respect to traffic flow dynamic [16], the capacities of the links in the network are downscaled accordingly. For our case, the transport model was calibrated for 10% and 20% of the total population. The analysis reported here is based on a synthetic population of 10%, which allows us to keep the computation time for the epidemic model within a reasonable time limit (about ten hours to run one scenario).

Mode shares are shown in Figure 4 (the right panel). The private car is the dominant mode, and most short trips, less than one kilometer, are made by walk. Public transport modes are used for urban (bus and metro) and intercity (train) trips. Other modes, like tramways and intercity buses, have a very small share and are not reported here (but are used in the model). The data we use covers all the trips departing and arriving inside the study region, but also inflows and outflows with other regions in France. Unfortunately, data on cross-border traffic is not available in the used survey. Since these flows are important for the pandemic context, in particular with the neighbouring cities in Belgium, we complement the dataset by census data (from INSEE (see www.insee.fr/fr/information/2008354 (accessed on 12 December 2022))) that accounts for home-to-work trips for agents living in France and working in Belgium. The reverse flows are not available, which is a limitation for our analysis. To overcome this problem, for each trip departing from the study region and arriving to a location in Belgium a trip in the reverse direction is added. This assumption may not correspond exactly to reality, but we think it is useful to take into account spatial interaction with cross-border locations. From the survey data, the purposes of the trips are obtained from the origin and destination of each trip. As reported in the main part of Figure 4, the most important activities are "Home", "Work", "Shopping" and "Accompaniment" (family or friends).

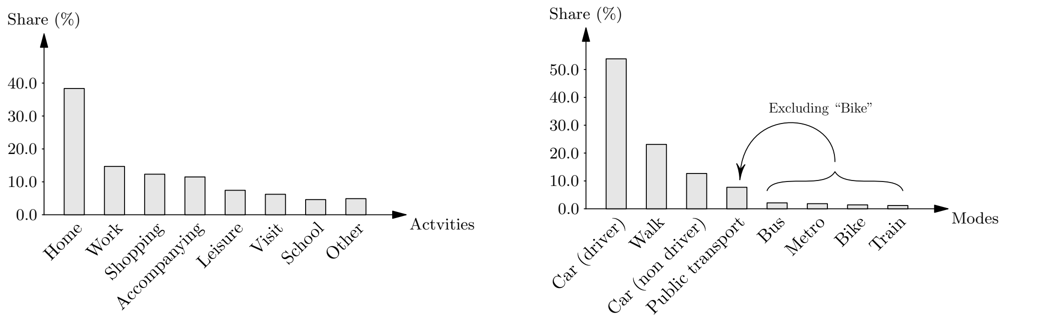


Figure 4. The proportions of activities (**left panel**) and the proportion of transport modes (**right panel**) as reported in survey data. These values are produced by the calibrated transport model.

In the simulation framework, a mode choice model is used to obtain the decisions of each agent. The mode choice includes the transport mode (or modes) used for each trip, but also the route choice (for private cars, in particular) and departure time. This mode choice model is based on a utility maximization problem. The utility reflects several facts: agents prefer to reduce time and money allocated to trips, they incur a schedule delay cost when they arrive late to work, they prefer to spend enough time at home and in leisure activities, and so on. One of the most important steps in setting the transport model is the calibration. It corresponds to a fine adjustment of the model parameters so that the simulations produce traffic flows that are comparable to those in the survey data. A detailed description of the transport model and the calibration step is given in [15].

Once the model is calibrated, it produces a report of daily events providing detailed information on all the trips: a given agent starts a trip at a given time, she arrives at the destination at a given time, which mode has been used, and so on. This output is the basis of the pandemic simulation that we develop using the model we have described in Section 2.2. By using the multi-agent approach we easily scale both the transport data and the epidemic data at the level of the agents. To simulate a wave of the pandemic, a random set of infected agents is created. The transmission of the pandemic will depend on their travel patterns and interactions with non-infected agents. Aggregation to urban or regional scale is then straightforward.

4. Implementation and Calibration of the Pandemic Model

In a first step we adjust the values of parameters θ, for the activities and transport modes so that the model produces realistic trend of the pandemic. In a second step we examine the spatial dynamic of infections over the study region.

4.1. Calibration

Transport simulation of traffic flows for a whole day produces a detailed sequence of the events that occur at each time interval (one second). The events indicate whether a given agent has entered or existed a location or a vehicle. Processing this information yields all the interactions during the day; that is, which agents have interacted together and how long each interaction was. This step was conducted with the `murdasp` tool specifically developed for this purpose. It is a set of scripts that processes the output files of a transport simulation and applies the pandemic model to simulate the spread of a virus.

The objective of the calibration is to find appropriate parameter values of the model so that the simulations produce realistic dynamic of the pandemic. In this framework, calibration comes down to finding appropriate values for parameter θ in each facility and transport mode. The main difficulty in this process is the lack of data on the location of infections. Indeed, even the number of the infected persons is not well reported, since it depends on those who made the test. But, once the test is positive, it is difficult to trace back the location where the infection has occurred. The available databases do not report information on these locations. For a given wave of the pandemic, the most reliable information is its duration, as well as the daily rate of relative change in infections. We mainly consider these variables in our calibration and use suggestive values when it comes to distinct facilities. Indeed, infection is a physical process where the virus is transmitted from an agent to another. Several studies, as in [17], predict the probability of transmission on the basis of cinematic laws. It is then possible to infer probabilities of infections from the number of susceptible persons, the number of the infected, the size of the facility where they meet and the duration of the interactions. For transport, ref [18] conducts a study on the probability of infection depending on the duration of the contact and the relative seats (in a train) of susceptible and contagious agents. The authors report that the probability highly increases when the two agents have adjacent seats, it remains high for seats on the same row, but significantly decreases as they are in distant rows. The values proposed by the authors are suggestive and were taken into account into our calibration step.

The calibration of our model requires the adjustments of ten parameters corresponding to values of θ_a in six activities (home, work, leisure, shop, education, primary) and four transport modes (bus, train, tramway and subway). Given the difficulties to get detailed data on the occurrence of infections we have adopted an aggregate approach where we focus on the length of a wave of the pandemic and set the parameters for each activity (or transport mode) to be consistent with basic intuition. For example, assuming the same number of contagious passengers per vehicles, the probability of infection in a bus is higher than that in a train since physical contact is generally stronger. This is taken into account by considering (during the calibration of the epidemic model) numerical values that satisfy $\theta_{\text{bus}} > \theta_{\text{train}}$. We start by generating a random set of values and then, at each case, perform adjustments to reach a correct duration and trend of the pandemic wave. This process was extensively repeated and we ended up with values given in Table 1. From the expression of the probability in (1) and, for example, the value of θ_{leisure} in Table 1, a non-infected agent interacting with a contagious agent for one hour in a leisure activity is infected with a probability of 1.8% (the duration τ in (1) is measured in seconds). Similarly, a one-hour bus trip yields a probability of infection equal to 3.7%. For these two cases if the non-infected agent interacts for one hour with three contagious agents each (instead of one) the probabilities of infection increase to 5.4% and 11.3%, respectively for leisure activity and a bus trip.

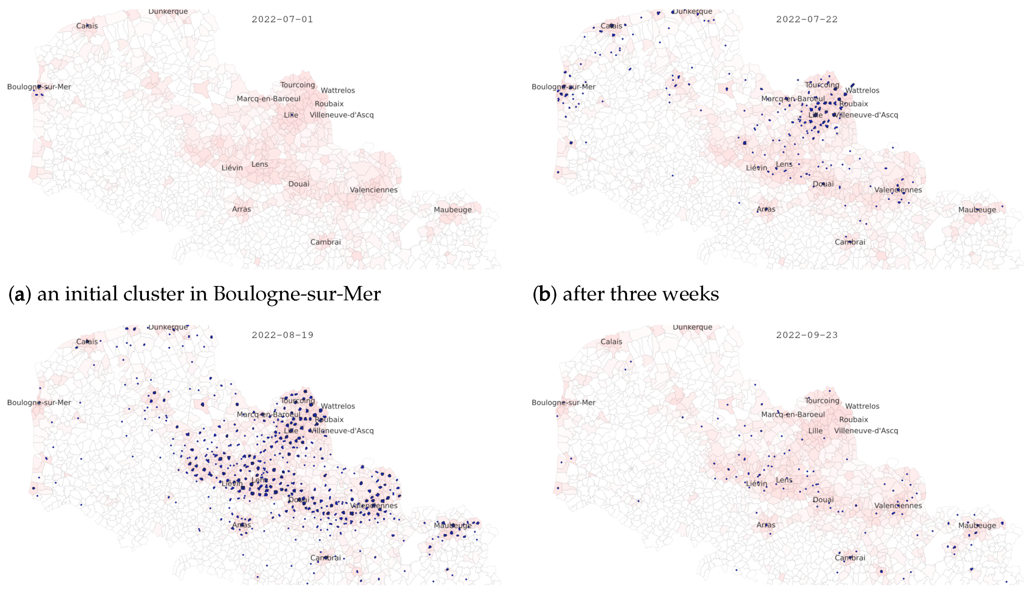
Table 1. The calibrated values of parameters θ are obtained, for each activity of transport mode by taking the value p in this table and then performing the computation $\theta = p \times 5 \times 10^{-6}$.

Home	Work	Leisure	Shop	Educ.	Primary	Bus	Train	Tram	Subway
0.1	0.3	1.0	0.08	0.009	0.15	2.1	1.2	0.9	1.75

4.2. *From a Cluster to a Large-Scale Pandemic*

Using the calibrated model obtained above, we consider, as a first application, the simulation of a wave of the pandemic on the basis of initial clusters that appear in distinct agglomerations in the study area. We compare the speeds of the spread of the virus depending on whether the initial cluster is located in a dense urban zone or whether it is located in low density area. A cluster is a set of twenty contagious individuals living in the same district.

Figure 5 illustrates the spatial progress of the pandemic. The situation starts with an initial cluster of infected agents in Boulogne-sur-Mer and reports, for some subsequent weeks, the evolution of the pandemic. On average, the growth of the pandemic is higher when a cluster is located in a large agglomeration, like Boulogne-sur-Mer. During the first week, the cluster extends to the neighbouring areas and within two weeks it reaches the metropolitan area of Lille and the nearest main agglomerations. After three weeks most infections are located in the dense area of Lille, and most locations in the study area are concerned.



(**a**) an initial cluster in Boulogne-sur-Mer (**b**) after three weeks

(**c**) after seven weeks (**d**) after twelve weeks

Figure 5. Dynamic of the pandemic spread following an initial cluster in Boulogne-sur-Mer (in the West). Background legend: a dark color corresponds to a higher population density, and vice-versa.

At the end of the first month, daily infections count in thousands and the mining area, as almost all other large urban areas (Dunkerque, Boulogne-sur-Mer, ...), are largely concerned. Home-to-work trips which are the main flows between the mining area and Lille play an important role at this stage. In the following weeks, the number of daily infections continues climbing. By the second half of the second month, the number of infections start decreasing in Lille and the other large agglomerations, but remains high in the mining area until ten weeks (from the start of the pandemic). In the following weeks, the number of infections decreases in all parts of the study region, and after three months the number of infections falls below one hundred, scattered over several rural areas where the number of susceptible agents remains relatively high.

One of the main features of this dynamic is the high correlation between the population densities and the number of infections in the corresponding districts. The low density areas (light colors in Figure 5) play a minor role in the transmission of the virus. The main risk is at the early stage of the pandemic when it may expand to neighbouring dense agglomerations and then spreads at a regional scale. In the examples we have considered we find, however, that soft restrictions in these areas can, in general, effectively isolate the infected agents and avoid reaching the pandemic level. We finally notice that clusters that appear in large agglomerations expand relatively quickly, since we reach thousands of infections within two or three weeks, while clusters that are located in small agglomerations take longer time to expand and may need four or five weeks to produce the same numbers of infections. In some rural areas, with very low densities, small clusters do not expand spatially even in the absence of restrictions. This is explained by the low level of physical contacts and traffic flows with larger agglomerations.

5. The Impact of Social and Physical Distancing

In this section, we report a set of simulations we have conducted with the transport model described above. We mainly consider imposing wearing of the mask and lockdown restrictions.

5.1. Wearing of Mask

Wearing masks in public places is a basic step in limiting the spread of the virus. We examine this sanitary measure by considering several scenarios, starting with public transport and then considering other alternatives. In real cases, when masks are imposed, all public transport modes are concerned. It is indeed unlikely to adopt such a sanitary measure for some modes and not for others (An exception could be air transport which is sometimes regulated by specific rules, but the scope of our analysis is limited to urban and regional transport (less than 200 km)). But, since we are interested in identifying the modes that play the most important roles, we consider a set of scenarios where wearing a mask is imposed on some modes and not imposed for users of other modes. It follows that with four public transport modes (bus, rail, tram and subway) there are sixteen possible combinations to consider.

When agents are wearing masks, the probability of an infection decreases and the exact magnitude depends on the type of the mask, among other facts. As reported in [19] several types of masks (cloth, surgical and N95) yield significantly distinct impacts on the probability of infections. For our analysis, we do not consider several types of masks and instead consider a decreases in the probability of an infection by a factor of ten, which can be considered as an average value. In our model, this measure is implemented by dividing the value of the corresponding θ_a parameter by ten for the corresponding modes. Notice that reducing the values of parameters θ_a is equivalent to reducing the shedding rate.

Figure 6 reports cumulative values of daily infections for all considered cases. In the base-case, more than 88,000 infections occur as reported by the uppermost curve. By considering wearing of masks we have two impacts: the spread of the virus is delayed and the total number of infections decreases. When only one or two public transport modes are considered, the impact is small, in particular for tram which is used by a tiny proportion of users. The buses and the subway in Lille, which are much more used, yield higher decrease. Not surprisingly, the highest impact is obtained when wearing of masks is imposed in all modes. Notice that when tramways (in Lille and in Valenciennes) are excluded from this measure, the number of infections is still the lowest, confirming the small marginal impact of infections in tramways.

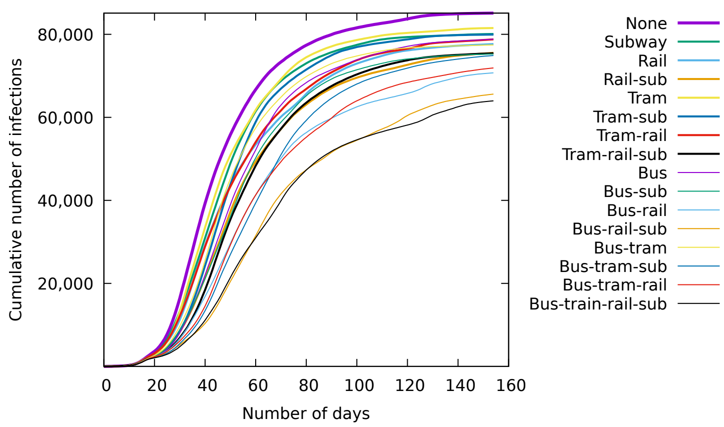


Figure 6. Number of infections depending on public transport modes where the wear of a mask is imposed. An initial cluster in Lille is considered.

When masks are not mandatory, in the base-case scenario, the total number of infections is 85,111. With the obligation to wear the mask in all public transport modes, the total number of infections decreases to 63,982, that is a decrease by 24.82%. We thus consider extending masks to other activities. As in the previous case, we consider progressive implementation in distinct activities as reported in Table 2. In the reference case, there are no restrictions and this corresponds to the first line. We then consider the case where masks are mandatory in all public transport modes, which is reported in the second line. As the wear of masks extends, respectively, to work, leisure shop and education, including primary schools, the number of infections decreases. Imposing masks in public transport and at work, the number of infections decreases by 42.80% by comparison to the base-case. Leisure activities seem also to be important since they contribute to reach a decrease in the number of infections by 57.42%, but extending masks to shopping activities produces only a small marginal impact, which is comparable to findings in [14]. Education, including primary schools, in contrast, is very important and produces a significant marginal impact. When the wear of a mask is generalized to all activities, the pandemic is almost contained and the number of infections drops to 339 only (a decrease by 99.60% by comparison to the base-case). We may notice that when sanitary measures are in place, the length of the pandemic wave increases. This is particularly the case for the before last scenario, when masks are imposed everywhere. It confirms that effective sanitary measures can limit the spread of the virus, but they should not be lifted quickly since the number of susceptible agents is still large.

Table 2. Scenarios where the wear of a mask is imposed in distinct activities.

Id	Pub. Trans.	Work	Leisure	Shop	Education	Primary	Infections	Impacts
1	no	no	no	no	no	no	85,111	-
2	yes	no	no	no	no	no	63,982	−24.82%
3	yes	yes	no	no	no	no	48,680	−42.80%
4	yes	yes	yes	no	no	no	36,238	−57.42%
5	yes	yes	yes	yes	no	no	34,138	−59.89%
6	yes	yes	yes	yes	yes	no	3652	−95.70%
7	yes	yes	yes	yes	yes	yes	339	−99.60%
8	no	yes	yes	yes	yes	yes	15,404	−81.90%

The last line in Table 2 refers to the case where the wearing of masks is in place in all activities except in public transport. This scenario is unlikely in a real situation, but it helps understand the relative role played by public transport in the spread of the virus. For our case, and from the results reported in Table 2, public transport alone does not seem to play a major role in growth of the pandemic since the number of infections still decreases by 81.90%, which is relatively large. Taking into account the second and the last lines, we can infer from these simulations that public transport account for about 18% to 25% of the total number of infections. Indeed, and *a priori*, it is not clear whether public transport plays a major or a moderate role in the spread of the virus. On one hand, physical contacts between users of public transport is intensive and makes the susceptible agents using this mode highly exposed. This is particularly important during the peak period where most vehicles are crowded. On the other hand, the share of public transport modes is relatively small in the study region (about 7%), and the duration of the interactions are, on average, short by comparison to other activities like work, or home. This explains why their contribution in the growth of the pandemic is rather limited.

5.2. The Lockdown and Vaccination

We consider a lockdown, where all activities are running at their minimum level and agents are only allowed to make basic shopping. In this case, and as shown in Table 2, the number of interactions and their durations decrease significantly. Except in houses, wearing a mask is mandatory. During, the COVID-19 pandemic, the confinement of the population was in place for several weeks and in many countries. It is perceived as the most efficient sanitary measure, especially when no vaccine is available. We use our model to evaluate the impact of lockdown by removing the interactions in all activities, except houses and shopping.

The comparison between the base-case scenario, where no restriction is adopted, and the case where most activities are shut-down (the lockdown) is given in Figure 7. In the left panel the daily number of infections is reported for distinct initial clusters and without any restriction. The same clusters are considered in the second case (right panel) where strict social distancing is imposed through a lockdown. Daily infections are about ten times less frequent and we may notice two facts. First, the length of the wave significantly increases from hundred days to, for some cases, two or three hundred days. Second, some waves display multiple peaks (multimodal distributions) and reach their maximum values with a delay of two months or more.

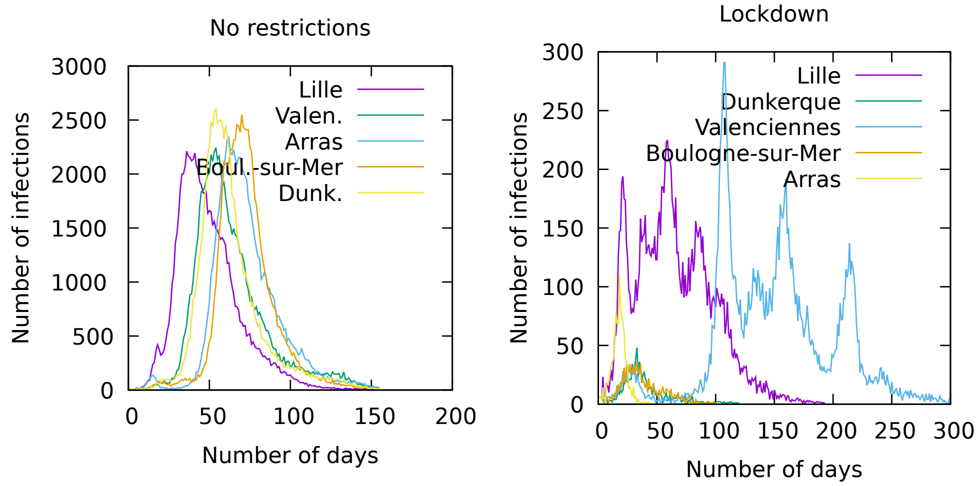


Figure 7. A comparison of infection rates between the base-case scenario (**left panel**) and lockdown situation (**right panel**). Each curve corresponds to the city (indicated in the legend) where the initial cluster appeared.

Figure 7 confirms that under a lockdown, the total number of infections remains small, and initial clusters in large agglomerations (Lille and Valenciennes, for instance) yield the highest rate of infections. Indeed, in low density areas, the implementation of a lockdown is highly efficient in keeping the number of infected agents limited. In metropolitan areas, since the population size is large, there are much more interactions in basic activities (e.g., shopping). Thus, even when some (soft) restrictions are imposed, the number of infections can remain high. The main difficulty with the lockdown alternative is that it should be maintained several weeks, since the threat of a widespread of the virus is always present. The duration of the lockdown has important economics and social costs [20,21], making its usage conceivable for few times and in especially when no vaccine is available.

6. Conclusions

We have calibrated an epidemic model for the North of France and conducted several simulations to examine the spread of the COVID-19 virus under distinct sanitary measures.

The analysis uses the `murdasp` tool and builds on the output of a multi-agent, activity-based, transport model. By collecting the daily events for the agents in the simulation, we identify all the interactions between the infected agents and the susceptible agents and, by applying a stochastic epidemic model, we identify, in each case, whether the transmission of the virus occurs or not. The probability of an infection depends on several attributes corresponding to the location where the physical contact occurs (large or small building, closed or open area, etc.) and the duration of the interaction. The simulation of the pandemic is a repetition of this process for all the agents along the whole simulation period, which extends from the date when the first infected agents appear (a first cluster) to the end of the pandemic (no more infections). More than three months are generally required for a wave to dissipate.

Our simulations confirm a strong relationship between the densities of the population and the number of infections. Rural areas with low densities play only a minor role in the spread of the virus. By imposing sanitary restrictions, particularly the wearing of masks, the number of infections decreases significantly, but the length of the wave increases, on average, by a factor of two. Also, we find that the wearing of masks is significantly more effective when it is not limited to public transport but extended to other activities. There are important differences in the protection provided by medical masks and ordinary ones. Conducted simulations, that are not reported in Section 5, confirm the intuition that the agents wearing medical (or N95) masks are less concerned by infections (either as exposed or transmitter).

This is a first application of the `murdasp` tool. The set of scenarios we have examined can be enlarged in several directions to address other issues. For example, a more detailed examination of the characteristics of the infected (and non-infected) agents and the structure of their activities can be considered. We have used only limited statistics on these facts because, for instance, there are no comprehensive databases that report the infections in relation with the nature of the activities. Indeed, when an agent is infected, it is difficult to trace back the exact location where the infection has occurred. Of course, inferences on the locations are possible, but, due to the lack of data, their accuracy cannot be evaluated by comparison to observed values. Another potential application that we plan to consider in the near future is how public transport services can be adjusted to limit the spread of the virus. Reducing the number of buses and trains may seem a natural response to the decline in demand, but can increase the occupancy rate of the vehicles and, by so, reinforces the occurrence of infections. By increasing the service frequencies at their usual level, or keeping, the same level of frequencies we may obtain a better result, but this should be balanced against the increase in operating costs.

Our analysis is based on the `murdasp` tool which is a set of scripts that process transport simulation outputs to collect physical contacts between the agents, and then applies an epidemic model to simulate the transmission of the virus. It is a flexible and transparent tool, but several improvements can be considered. In particular, the computation speed is relatively long and requires, for instance, about ten hours for each of our scenarios. Computation speed can be improved by the parallelization of several parts in source code. These issues will be considered in the future improvements.

Author Contributions: Conceptualization, M.K., O.D. and N.D.; Literature review, O.D.; Analysis and results, M.K.; Data preparation, N.D. (transport model) and O.D. (epidemic model); writing—original draft preparation, M.K., O.D. and N.D.; writing—review and editing, M.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research has benefited from financial support from University of Littoral and ANR project MURDASP (num. ANR-21-HDF1-0014).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data Availability Statement: Data available in a publicly accessible repository. The data presented in this study are openly available at https://murdasp.univ-littoral.fr (accessed on 14 December 2022).

Acknowledgments: Numerical computations presented in this paper were carried out using the CALCULCO computing platform, supported by SCoSI/ULCO (Service COmmun du Système d'Information de l'Université du Littoral Côte d'Opale).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Alfieri, C.; Egrot, M.; Desclaux, A.; Sams, K. Recognising Italy's mistakes in the public health response to COVID-19. *Lancet* **2022**, *399*, 357–358. [CrossRef] [PubMed]
2. Ferigato, S.; Fernandez, M.; Amorim, M.; Ambrogi, I.; Fernandes, L.M.; Pacheco, R. The Brazilian Government's mistakes in responding to the COVID-19 pandemic. *Lancet* **2020**, *396*, 1636. [CrossRef] [PubMed]
3. Adam, D. Simulating the pandemic: What COVID forecasters can learn from climate models. *Nature* **2020**, *587*, 533–535. [CrossRef] [PubMed]
4. Mniszewski, S.M.; Del Valle, S.Y.; Stroud, P.D.; Riese, J.M.; Sydoriak, S.J. EpiSimS simulation of a multi-component strategy for pandemic influenza. In Proceedings of the 2008 Spring Simulation Multiconference, Ottawa, ON, Canada, 14–17 April 2008; pp. 556–563.
5. House, T. Modelling epidemics on networks. *Contemp. Phys.* **2012**, *53*, 213–225. [CrossRef]
6. Kilani, M. *Murdasp: A Tool to Simulate the Pandemic Using Activity-Based Transport Model*. Available on the Git Repository. 2023. Available online: https://gogs.univ-littoral.fr/mkilani/murdasp (accessed on 13 December 2022).
7. McNally, M.G.; Rindt, C.R. The Activity-Based Approach. In *Handbook of Transport Modelling*; Emerald Group Publishing Limited: Bingley, UK, 2007; pp. 55–73. [CrossRef]
8. Axhausen, K.W.; Horni, A.; Nagel, K. *The Multi-Agent Transport Simulation MATSim*; Ubiquity Press: London, UK, 2016.
9. Kermack, W.O.; McKendrick, A.G. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. London. Ser. Contain. Pap. Math. Phys. Character* **1927**, *115*, 700–721.
10. Djidjou-Demasse, R.; Selinger, C.; Sofonea, M.T. Épidémiologie mathématique et modélisation de la pandémie de COVID-19: enjeux et diversité. *Rev. Francoph. Des Lab.* **2020**, *2020*, 63–69. [CrossRef] [PubMed]
11. Ferguson, N.; Laydon, D.; Nedjati-Gilani, G.; Imai, N.; Ainslie, K.; Baguelin, M.; Bhatia, S.; Boonyasiri, A.; Cucunubá, Z.; Cuomo-Dannenburg, G.; et al. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand. *Imp. Coll. Lond.* **2020**, *10*, 491–497.
12. Müller, S.A.; Balmer, M.; Charlton, B.; Ewert, R.; Neumann, A.; Rakow, C.; Schlenther, T.; Nagel, K. Using mobile phone data for epidemiological simulations of lockdowns: Government interventions, behavioral changes, and resulting changes of reinfections. *medRxiv* **2020**. [CrossRef]
13. Müller, S.A.; Balmer, M.; Charlton, W.; Ewert, R.; Neumann, A.; Rakow, C.; Schlenther, T.; Nagel, K. A realistic agent-based simulation model for covid-19 based on a traffic simulation and mobile phone data. *arXiv* **2020**, arXiv:2011.11453.
14. Manout, O.; Ciari, F. The contribution of everyday activities to the spread of COVID-19. In *An Agent-Based Simulation from Montreal, Canada*; Working Paper; Ecole Polytechnique de Montréal: Montreal, QC, Canada, 2021.
15. Kilani, M.; Diop, N.; Wolf, D.D. A Multimodal Transport Model to Evaluate Transport Policies in the North of France. *Sustainability* **2022**, *14*, 1535. [CrossRef]
16. Arnott, R.; Kilani, M. Social Optimum in the Basic Bathtub Model. *Transp. Sci.* **2022**, *56*, 1505–1529. [CrossRef]
17. Agrawal, A.; Bhardwaj, R. Probability of COVID-19 infection by cough of a normal person and a super-spreader. *Phys. Fluids* **2021**, *33*, 031704. [CrossRef] [PubMed]
18. Hu, M.; Lin, H.; Wang, J.; Xu, C.; Tatem, A.J.; Meng, B.; Zhang, X.; Liu, Y.; Wang, P.; Wu, G.; et al. Risk of coronavirus disease 2019 transmission in train passengers: an epidemiological and modeling study. *Clin. Infect. Dis.* **2021**, *72*, 604–610. [CrossRef] [PubMed]
19. Talic, S.; Shah, S.; Wild, H.; Gasevic, D.; Maharaj, A.; Ademi, Z.; Li, X.; Xu, W.; Mesa-Eguiagaray, I.; Rostron, J.; et al. Effectiveness of public health measures in reducing the incidence of COVID-19, SARS-CoV-2 transmission, and COVID-19 mortality: Systematic review and meta-analysis. *BMJ* **2021**, *375*, e068302. [CrossRef] [PubMed]
20. Saccomanno, S.; Bernabei, M.; Scoppa, F.; Pirino, A.; Mastrapasqua, R.; Visco, M.A. Coronavirus lockdown as a major life stressor: Does it affect TMD symptoms? *Int. J. Environ. Res. Public Health* **2020**, *17*, 8907. [CrossRef] [PubMed]
21. Chandola, T.; Kumari, M.; Booker, C.L.; Benzeval, M. The mental health impact of COVID-19 and lockdown-related stressors among adults in the UK. *Psychol. Med.* **2022**, *52*, 1–10. [CrossRef] [PubMed]