Where, when and how people move in large-scale urban networks: the Grenoble saga

Ujjwal Pratap, Carlos Canudas de Wit, Federica Garin

Where, when and how people move in large-scale urban networks: the Grenoble saga

Ujjwal Pratap[a], Carlos Canudas-de-Wit[a], Federica Garin[a]

[a] *Univ. Grenoble Alpes, CNRS, INRIA, Grenoble INP, GIPSA-lab, 38000 Grenoble, France*

Abstract

This work studies the mobility of people in the metropolitan city Grenoble in France, building a supply-demand based mobility model which captures the daily movement of people between residences and places of interests called destinations, using time schedules and gating profiles which also accommodate the possibility of imposing restrictions on mobility. The goal of this paper is to build the mobility network in Grenoble answering the three main questions (WWH): Where does mobility happen? When does it happen? How many people move? We provide methods to compute the time gating functions and different parameters for this network. Further, we identify and address the issues encountered during discretization of the continuous-time mobility model.

As an application, we also study how different restrictions on the mobility effects the epidemic spread in Grenoble. Finally, we describe the online GTL-covid demonstrator which is being developed by implementing the results of this paper. It is an academic platform which can be used to simulate different scenarios of mobility and visualise its effect on epidemic spread in Grenoble area.

Keywords: `Human mobility, large-scale networks, origin-destination matrices, epidemic mobility`

1. Introduction

Human mobility, because of its practical applications in epidemic control, urban planning, traffic planning and land use patterns, has attracted the attention of researchers from transportation [1], economic geography [2], epidemiology [3], physics [4] and many other fields over past decades. Study of urban human mobility in cities, is important and facilitates better planning and policy-making. For example, in the recent events of COVID-19 pandemic, controlled and restricted mobility of people became important and a fundamental issue faced by the policymakers. In such scenarios, both fully operational mobility or a completely restricted mobility could result in negative consequences on the socio-economic status of a country or a region. Therefore, a better understanding of human mobility patterns is necessary on the level of a metropolitan area with the three mobility components, 1. 'Where' do the people go? 2.'When' do they move? and 3.'How

Email addresses: `ujjwal.pratap@gipsa-lab.fr` (Ujjwal Pratap), `Carlos.Canudas-de-Wit@gipsa-lab.fr` (Carlos Canudas-de-Wit), `federica.garin@inria.fr` (Federica Garin)

many' people move? (WWH). There are several human mobility prediction models which can be categorised mainly in two categories based on information they use: 1. mathematical models using traditional data such as population census or travel survey data 2. models based on sensed data such as mobile phone Call Detail Record (CDR) and GPS data.

At first, we give an overview of the first category which is the mathematical models using traditional datasets. The prediction of human mobility has been studied extensively by statistical physicists over the years. One of the first models proposed to predict population movement was inspired by the Newton's law of gravitation and hence called the gravity model [5]. This model proposes that the flow of people between two locations is proportional to the product of their populations and inversely proportional to the distance between them. Since then, many of its generalised versions have been proposed which have several exponential parameters and functions to be inferred from the empirical data [6]. This requires previous traffic data to fit the parameters in order to find the exact mathematical formula making it difficult to predict mobility in regions which lack systematic traffic data. Many improvements include the radiation model proposed in [4] and other models which can be found in the survey [7]. These models give an estimate of number of people moving between different places in terms of origin-destination matrices which can be useful in predicting migratory flows and large scale mobility between cities. However, they lack spatio-temporal characteristics which are relevant for urban mobility. In other words, they estimate 'how many' and to some extent 'where' but give no estimate of 'when'. Another class of literature such as [8], [9] model mobility of people between two locations and time spent at such locations. the model is given in form of some power laws. Some mover-stayer type dynamic mobility models have also been proposed in the literature such as [10] and [11] in order to see its effect on a disease spread. These models capture the evolution of mobility over time between different geographical regions. However, they capture averaged mobility between different cities with large time-scales and cannot capture daily mobility patterns and there is no distinction between origins and destinations.

Now, we give a brief overview of the second category which are data-driven models. Recently, with the vast development in the information and communication technology, various data sources have been used to facilitate travel behaviour research. Examples are smartcard records data, GPS data and roadside sensor data, and among these, mobile phone CDR are the most widely applied [12]. The mobile phone CDR can be used to construct the origin-destination matrices and mobility patterns. For example, [13] developed OD matrices using mobile phone CDR, [14] estimated OD matrices by transport mode, [15] estimated the mobility pattern along the day with OD matrices, to name a few. The main reason of extensive use of this data is ubiquitousness of mobile phones, hence the ability to track the whole population at large span of time and relatively high spatial accuracy. These techniques might cover 'how much' and 'when' to some extent provided the availability of data. However, some researchers have pointed to the biases of these data, caused by the event-triggered nature of mobile phone data [16]. These data can also be unavailable in case of bad weather or connectivity and moreover amount of mobile phone usage also varies according to the socio-economic conditions of different regions. Another widely used data source is GPS data due to its high geographical precision. It is used to study individuals' movement in order to analyse mobility patterns and give mobility

models [17]. GPS data have been used to identify stay and destinations [18], estimate mobility patterns [19], estimating mobility patterns using taxi GPS data [20], identifying transport mode [21] and updating OD matrices [22]. However, these location data come at a risk of privacy of individuals and taxi location data has a very low penetration rate. There is a serious concern over privacy of individuals as removal of personal data does not fully preserve privacy and have risk of re-identification [23, 24] of the individuals. In future, new laws can also be implemented which will restrict the mobile and location data use even more [25]. Moreover, privacy protection techniques require data aggregation or need to be recorded for short period limiting their usefulness [26].

In this paper, we extend, adapt and implement the supply-demand flow mobility model proposed in [27] to the large-scale network of Grenoble, a metropolitan city in France. In this model, every day, a certain number of people travel from their residences, which are called origins, to locations visited daily such as for work, education, shopping, hospitals etc., which are called destinations, and then return to their home again. The daily mobility patterns are captured by the time-dependent supply and demand gating profiles. The Supply Gating Functions (SGF) of each destination is controlled by its daily destination schedule, which is its opening and closing hours. The Demand Gating Profile (DGP), on the other hand, is defined on each edge of the mobility network and corresponds to the daily mobility window, which is the time interval during which people move between origins and destinations along that edge. The supply function of each destination controlled by the SGF determines the inflow allowed to that destination and depends on its operating capacity which can be tuned using a control input, for instance for epidemic attenuation. The demand function controlled by the DGP determines the outflow from one location to another. The process of urban human mobility is modeled on the network edges that connect different locations through flows. A significant advantage of this model is that it can capture mobility on a smaller time-scales with a different pattern on a given day of the week thanks to SGF and DGP. Another important advantage with respect to existing models, is that the proposed model can encompass all the three mobility components (WWH). However, [27] uses a toy example and hence lack a precise procedure for the identification of the functions, gating profiles and parameters needed to implement this model on a real large-scale network. One of the main aim of this paper is to devise methods and procedures for identifying the origin and destination nodes, their O/D relations, setting the parameters and defining the gating functions DGP and SGF for the large-scale mobility network of Grenoble.

In this paper, at first, we build the large-scale network of Grenoble using public information which makes our method free from privacy and legal issues. precisely, we build the three components of mobility: Where, When and How many (WWH). In 'where', We consider the residential areas of the communes as origins and identify different places of interests such as schools, universities, workplaces, hospitals parks etc. as the destinations. The destinations are then classified into categories and subcategories. Each origin and destination is represented by a node in the network. While the origin nodes are placed in the residential areas, the destination nodes are placed at the exact locations using open source map data. Then, we collect the population of the origins from census data and the capacities of the destinations from different sources such as their websites. The capacity of a destination is the maximum number of persons allowed to be there at any instant. If any subcategory has a large number of nodes, then they are aggregated by replacing such

nodes in a commune by a single node with its capacity being the sum of their capacities and its location being the barycenter of their locations. Having identified the origins and destinations, we make a bipartite graph corresponding to the destinations of each subcategory to decide which origins and destinations are connected. They are connected based on a form of attraction law proportional to the population and the capacities of origins and destinations and there is a threshold over distance, a person would travel to go to a destination. For this, we compute the real minimum road distance between all possible locations. In the component, 'When', we collect the information about the opening and closing time per destination category and the gating functions are computed using the mobility profiles from travel survey or inspired from the real time popularity trends. Finally, in the component 'How many', we compute the average daily number of people traveling from an origin to a destination if they are connected by using rules of proportions.

Since, the mobility-model proposed in [27] is in continuous-time, therefore, in order to extend and implement it on the real large -scale network, we redefine the model in discrete-time ensuring mass-preservation and non-negativity. These properties are also satisfied even if we take larger time-steps in simulations. In addition, since the model and the developed setup is modular in nature, it can be used for many applications such as optimal control of epidemic spread, understanding social behavior and urban planning. To tap this property, we have developed an online demonstrator called GTL-covid interface where an user can simulate different mobility control scenarios and as a case study he/she can visualise the effect on mobility on COVID-19 epidemic spread in Grenoble area as per the SIR epidemic model.

The paper has been organised as follows. In Section 2, we recall the preliminaries concerning the model [27] which we adapt and extend to implement it on a large scale network mobility network of Grenoble. In Section 3, we compute the parameters under the component 'where' that is we describe the methodology to build the network by locating the origins and destinations, identifying their population or capacities, and then defining the connections between them. In Section 4, we give procedures to compute the time schedules and the gating functions which are covered under the component 'when'. In Section 5, we compute the the average daily number of people moving between an origin and destination, which is concerned with the component 'how many'. In Section 6, we address the numerical issues encountered in order to adapt and implement the mobility model to the large scale network with the computed parameters. Finally, in Section 7, we provide an application of the developed model to analyse and study the effect of mobility on the epidemic spread. In Section 8, we give a description of the GTL-covid interface which is being developed under this work followed by conclusion in Section 9. We enlist the notations which are global and give their descriptions in Appendix A.

2. Preliminaries

In this section, we give a basic description of the urban-human mobility model proposed in [27].

2.1. Urban human mobility model

Consider the human mobility in an urban area between two types of locations: origins and destinations. The origins correspond to the residential areas and the destinations, on the other hand, correspond to locations that people visit daily for some purpose like work, shopping, education or leisure for example, companies, research centers, schools, restaurants, cinemas, etc. Every day a certain number of people visit the destinations during specified time intervals and then return later on the same day.

Consider the daily human mobility in the timescale of hours represented by a network $\mathcal{G} = (\mathcal{V}_o, \mathcal{V}_d, \mathcal{E})$ between locations of two types: origins $\mathcal{V}_o$ and destinations $\mathcal{V}_d$. For the sake of simplicity, we assume that the network $\mathcal{G}$ is bipartite, i.e., we assume that there are edges only between origins and destinations, and there is no edge between a pair of origins or a pair of destinations. Since, the mobility is happening instantly between two locations, this assumption doesn't effect the mobility pattern much as the mobility between two destinations can be understood as if a person first goes back to the origin from one destination and then moves from the origin to the other destination. Let $\mathcal{N}_i = \{j : (i, j) \in \mathcal{E}\}$ be the set of i's neighbors. The assumptions on the model are as follows:

- The total population of the city remains constant.
- The mobility occurs only between pairs of origins and destinations, and not among a pair of different origins or a pair of different destinations.
- The number of people who visit destination j from origin i during a day is equal to the number of people who return to i from j on the same day.

The destinations $\mathcal{V}_d$ are partitioned into q categories $\mathcal{C}_1, \ldots, \mathcal{C}_q$, where each category $\mathcal{C}_a$ is further partitioned into subcategories, $\mathcal{D}^a_1, \ldots, \mathcal{D}^a_{k_a}$. Denote the total population of origin $i \in \mathcal{V}_o$ by P_i which is the total number of people residing in i and the nominal instantaneous capacity of destination $j \in \mathcal{V}_d$ by C_j which is the maximum number of people who can visit j at any given time. This nominal capacity, corresponds to the capacity of the destination when it is normal operation. We introduce the possibility to describe the restrictions of capacity, such as those that were imposed by many governments during peaks of covid pandemics. Such capacity reduction is decided per category of destinations, and is described by a coefficient $u_h(t) \in [0, 1]$ for $h \in \{1, \ldots, q\}$ which determines the allowed operating capacity of the destinations in $\mathcal{D}_h$ in terms of the proportion of nominal capacity at time t. In other words, it can be considered as a policy at time t that limits the operating capacities denoted as $C^o_j(t)$ in the destinations of category h, where

$$C^o_j(t) = C_j u_h(t), \quad \text{for } j \in \mathcal{D}_h. \tag{1}$$

In (1), if for a destination $j \in \mathcal{D}_h$, $u_h(t) = 0$, then the operating capacity $C^o_j(t) = 0$, which means that no person is allowed to visit j. On the other hand, if $u_h(t) = 1$, then $C^o_j(t) = C_j$ that is the operating capacity of j is equal to its nominal capacity.

Let $N_i(t) > 0$ be the number of people in $i \in \mathcal{V}_o \cup \mathcal{V}_d$ at time $t(hr)$. Then, according to the urban human mobility model proposed in [27], the rate of change of the number of

people at any location at time t is equal to the sum of inflows to that location minus the sum of outflows from that location. In other words, if $\phi_{ij}(t, N_i(t), N_j(t))$ denotes the flow from i to j, then

$$\dot{N}_i(t) = \sum_{j \in \mathcal{N}_i} (\phi_{ji}(t) - \phi_{ij}(t)), \tag{2}$$

where

$$\phi_{ij}(t) = \min(\Delta_{ij}(t), \Psi_j(t)). \tag{3}$$

Here, $\Delta_{ij}(t, N_i(t))$ is the demand of i with respect to j which describes the flow of people who would like to travel to j and $\Psi_j(t, N_j(t))$ is the supply of j *i.e.* the inflow that can be allowed to enter j from its neighbors depending on the operating capacity C_j^oif j is a destination and P_j if j is an origin. For the sake of simplicity, here onward, we will use the notation $Z_j(t)$ for the operating capacity $C_j^o(t)$ if j is a destination or the population P_j if j is an origin unless otherwise stated. Note that, if j is an origin, $Z_j(t)$ will be constant and equal to the population P_j, however, if j is a destination then $Z_j(t)$ will be equal to the operating capacity $C_j^o(t)$ which varies according to the coefficient $u_h(t)$ if $h \in \mathcal{D}_h$ as defined in (1).

Demand The demand of i with respect to j is given by

$$\Delta_{ij}(t, N_i(t)) = \delta_{ij}(t) \cdot f_{ij}(t) \cdot 1_{N_i(t)>0}, \tag{4}$$

where $\delta_{ij}(t)$ is the Demand Gating Profile (DGP) such that

$$\begin{cases} \delta_{ij}(t) > 0 & \text{if } t \bmod 24 \in [t_{ij}, t_{ij} + \tau_{ij}) \\ \delta_{ij}(t) = 0 & \text{otherwise} \end{cases} \tag{5}$$

satisfying $\int_0^{24} \delta_{ij}(t)dt = 1$. Here, $[t_{ij}, t_{ij} + \tau_{ij}) \subseteq [0, 24)$ is called the *mobility window*, defined as the time interval in which there is mobility from i to j. We can have different Demand Gating Profiles and mobility window for different days of the week. The function δ_{ij} gives the profile of demand over the day, while the intensity of the demand is described by a factor $f_{ij}(t)$ defined as:

$$f_{ij}(t) = M_{ij} u_h(t). \tag{6}$$

Here, M_{ij} is the total number of people that would like to visit j from i, daily and $M_{ij}u_h(t)$ is the number of visitors when capacity is restricted with the coefficient $u_h(t)$. The term $M_{ij}u_h(t)$ reflects the fact that when a government policy regarding capacity reduction is announced, it not only restricts the number of people who can visit a destination but also reduces the number of people willing to go to those destinations. Finally,

$$1_{N_i(t)>0} = \begin{cases} 1 & \text{if } N_i(t) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

is an indicator function.

Supply The supply of j is given by

$$\Psi_j(t, N_j(t)) = \psi_j(t) \cdot F_j \cdot 1_{N_j(t)<Z_j(t),} \tag{8}$$

where

$$\psi_j(t) = \begin{cases} 1 & \text{if } t \bmod 24 \in [a_j, b_j) \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

is the Supply Gating Functions (SGF) with $[a_j, b_j) \subseteq [0, 24)$ called the *destination schedule* where a_j is the opening hour and b_j is the closing hour of location j. Similar to DGPs, also SGFs are defined over a day and repeated periodically, possibly with different profiles for different days of the week. For the origins $j \in \mathcal{V}_o$, $[a_j, b_j) = [0, 24)$ because they are always open. Moreover, recall that the demands, when non-zero, are constant. Hence, the maximum total inflow to j from its neighbors is equal to the sum of such constants, i.e., $F_j = \sum_{i \in \mathcal{N}_j} f_{ij}$ is the maximum, constant inflow to j from its neighbors. Finally,

$$1_{N_j(t) < Z_j(t)} = \begin{cases} 1 & \text{if } N_j(t) < Z_j(t) \\ 0 & \text{otherwise} \end{cases} . \tag{10}$$

Similar to the restrictions on capacity, we also introduce the possibility to describe the restrictions on closing time of destinations, such as those that were imposed by many governments during peaks of Covid pandemics. Such restriction is decided per category of destinations, and is described the *effective closing time* denoted by $b_h^e \in [6, 24]$ for $h \in \{1, \ldots, q\}$. Similar restriction on the opening time can also be enforced by another such coefficient to tune the effective opening time. However, if done, it must be enforced that the effective closing time is later or same as the effective opening time.

2.2. Extension to the large-scale network of Grenoble

In this paper, we adapt and extend the mobility model described above to implement it to the Grenoble area. Grenoble is metropolitan city in France located in the valley of Alpes mountain range. A *métropole* (French for"metropolis") is an administrative entity in France, in which several *communes* cooperate, and which has the right to levy local taxes. Grenoble-Alpes métropole is centered on the city of Grenoble.The considered area includes 55 *communes* including all the 49 *communes* in Grenoble-Alpes métropole and 6 *communes* in Gresivaudan, which are of economic importance to the Grenoble *métropole.* The population of this area is about 500000 and has a surface area of about 600 km^2. In order to adapt, extend and implement the mobility model (2) to a large-scale network of a metropolitan city, we need the following parameters:

1. Map of the *communes* in the area.
2. Origin nodes $\mathcal{V}_o$ and destination nodes $\mathcal{V}_d$.
   (i) Location of all the nodes in $\mathcal{V}_o$ and $\mathcal{V}_d$.
   (ii) Population P_i, $\forall\ i \in \mathcal{V}_o$.
   (iii) Capacities C_j, $\forall\ j \in \mathcal{V}_d$.
3. OD matrix $\mathcal{O}_{ij} = \begin{cases} 1 & \text{if } (i,j) \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases}$, to decide which origins and destinations are connected to each other.

4. For destination j in each subcategory $\mathcal{D}_c^\alpha$ for each day of the week.

   (i) Time schedules to define DGP and SGF
   - opening times (a_j),
   - closing times (b_j) and
   - average time spent (s_j), which will be described later and used to define DGP for some destination categories.
   - mobility window

   (ii) Demand Gating Profiles(DGPs) $\delta_{ic}(t)$ for and Supply Gating Functions(SGFs).

5. For each edge $(i, j) \in \mathcal{E}$, M_{ij}, which is defined the average number of people going from i to j daily.

We will discuss how to retrieve these parameters to build the large-scale mobility network of Grenoble from the public information in detail in the next sections. In particular, we will discuss the parameters in terms of the three components of mobility: 1. Where? 2. When? 3. How many? In 'where' we consider the map, nodes $\mathcal{V}_o$ and $\mathcal{V}_d$ and the OD matrices for each subcategory. In 'when', we compute the time schedules and the gating profiles, and finally in 'How many', we compute M_{ij} for the destinations in each subcategory.

3. Where?

At first, we need the map of these areas and need tools to manipulate it. For that purpose, we use an open source software called QGIS. This software can be used to make, edit/manipulate or display maps and their features. In this work, the main source of data for maps is OpenStreetMap [28]. Openstreetmap is an open source collaborative project to create a free editable map of the world and the geodata underlying the map is considered the primary output of the project. The main source of statistical data is INSEE [29]. INSEE is a government agency which conducts census to collect socio-economic information from the people in France. The other sources if any will be indicated in the details corresponding to the different origins or destinations.

The coordinates of the boundaries of all the French *communes* can be downloaded from [30] in GeoJSON format. The GeoJSON format is useful to display different information on a map. This GeoJSON file contains the following information for each *commune:* name, coordinates. We take the coordinates data for the *communes* of our area of interest. We use GeoJSON files to locate the origin and destination nodes also as it helps us store any kind of information corresponding to any location on the map and can be easily manipulated in 'Python' or simple text editors like 'Notepad++'.

3.1. Origins

The residential areas of the communes are considered to be the origins. For the mobility, our aim is to study the movement of the population of this origins. The data for the population of *communes* is retrieved in form of a table from INSEE [31]. Note that, in the

region of our study, Grenoble is the most populated commune with 160778 residents and the Mont-Saint-Martin is the least populous commune with 79 residents. Moreover, the population of Grenoble is almost five times that of the second most populous commune, therefore, in the implementation, we consider a division of Grenoble, the *commune* which is basically located at the center of the metropolis area, into 6 parts which are called *sectors*. Therefore, the number of origins increases from 55 to 60.

3.1.1. Partition of Grenoble commune

Grenoble *commune* is divided into 6 'sectors', that regroups different neighbourhoods. The sectors defines certain rules/regulations that affect daily life of the residents such as schools, voting etc. This division is done by grouping smaller partitions called IRIS. IRIS is a partition done by the government to identify an area with population greater than 10000 or with population between 5000 to 10000. In Grenoble, there are 70 IRIS. We obtain the coordinates of these sectors by regrouping IRIS whose coordinates can be found on [32]. The grouping of IRIS has been done according to the division given on the official webpage of Grenoble [33]. In QGIS software, it can be done by taking union of polygons formed by the boundaries of IRIS. Different data for these sectors can

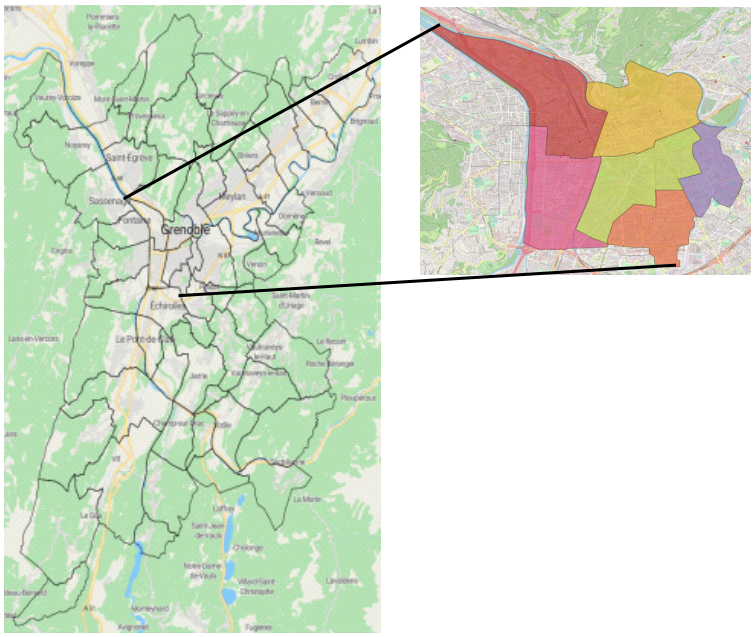


Figure 1: Zones considered in our study: 54 *communes* and the 6 sectors in which Grenoble has been partitioned. In the inset, we can see the partition of Grenoble *commune* into six sectors. This figure is a screenshot taken from the demonstrator developed as a part of this work which will be described in Section 8.

be retrieved from [33]. Now that, Grenoble *commune* has been divided into 6 sectors, from here onwards, the *communes* and these sectors of Grenoble will be referred as zones which are depicted in Figure 1.

3.1.2. Location and population of origin nodes

For the graphical representation and for implementation of the model, each zone is represented by a node which is positioned in the residential area. This positioning of the nodes

has been done manually inspired by the position of nodes on a INSEE [34] or by locating the residential areas on the online maps manually. Each origin node has a population P_i of the corresponding zone. The population of the communes has been taken from [31] and that for the partition of Grenoble from [35]. Moreover, the population at each origin according to the age groups $0-15, 16-24, 25-64$ can be obtained from [36]. We assume that these age groups are uniformly distributed. Therefore, to find the number of children at any origin between the age $5-10$, for instance, we divide the population of age group $0-15$ in that origin by 3. In this way, we obtain the population of different age groups.

3.2. Destinations

Destinations comprise of the places where people go for some time during the day or night for a short period of time and then return to their residences. We denote them by a node on the map. These include, working places, schools, hospitals etc. In the next subsection, we provide the classification of these destination nodes. Then, in the following subsection, we discuss the method for locating the destination nodes. Finally, when we move to the section of a particular destination category or subcategory, we mention the location method and discuss mainly the strategies or methods for computing the capacities of the nodes.

3.2.1. Locating the destination nodes

Unlike the origin nodes, the data for location is not always available, so we apply different strategies to locate various classes of destination nodes. We explain these methods in detail in the following:

3.2.1.1 Method 1: using Opentstreetmap data

In general, all the objects or entities on Openstreetmap are labelled correctly but many a time some objects are labelled incorrectly since anyone can edit it. It makes the retrieval of data in categories difficult. In this method, we follow the following steps:

1. Export the map data of a squared region from [28] by using export option on the website. Since the region of selection is large in our case, we need to use overpass API or planet OSM from the left panel to import the map data. On exporting, it downloads all the data for the chosen region and contains many information like, bus stops, shops, trees, parks, fountains, hospitals etc. along with many attributes such as location, opening and closing times, websites etc.

2. Since, the downloaded data is for a squared region, we need to select only the data which are under the region of our study. For that, we use 'intersection tool' of QGIS software. It is used to obtain one all the common attributes from two given layers if they share some area in common. The output is one layer with all the attributes in the intersection of these two layers. Here, we use this tool with two input layers, one being the mapdata for a square region downloaded from OSM and the other being the boundaries of our region as depicted in Figure 1.

3. Once we have the desired data for our region as an output of the previous step, we need to filter out the nodes of the desired category by using 'filter tool' in QGIS. Filter tool uses some SQL query to extract the data matching our queries.

4. Once filtered, we need to remove unwanted attributes and keep only the wanted attributes such as location or names etc. we also need to check for garbage nodes. Sometimes, there are too many garbage values because some contributors don't put the correct labels. If we find any garbage node, then we need to remove them by applying the filter again. There is also a chance of missing some nodes of the same type for the same reason that all the node are not labelled correctly.

5. Save the file in GeoJSON format using QGIS.

3.2.1.2 Method 2: Manually by creating and editing a GeoJSON file directly

If the coordinates and other data of the destinations are available from some website, we make a GeoJSON file directly by writing the coordinates manually under the geometry feature and the name and capacity etc. under properties feature. Here, the GeoJSON files have been edited using Notepad++. If needed, they can also be edited later directly in Notepad++ or by editing the attributes in QGIS.

3.2.1.3 Method 3: Locating the destinations manually in QGIS

For some destinations, neither the location coordinates nor OSM data is available. In these cases, we locate the nodes manually in QGIS using the node editing tool in a layer. While locating the node in QGIS, the location of the corresponding destination in OSM map can be taken as a reference. Once the nodes are located, we export the layer in GeoJSON format and the other attributes are edited later either in Notepad++ or in QGIS as per convenience.

3.2.2. Classification of destinations

The destination have been classified into five categories and each of them into some subcategories which can be seen in Figure 2. It can be seen that we have five categories and 15 subcategories in total. Throughout the paper, we will find the parameters for each subcategory. At first, we will discuss the location and capacities of each destination class.

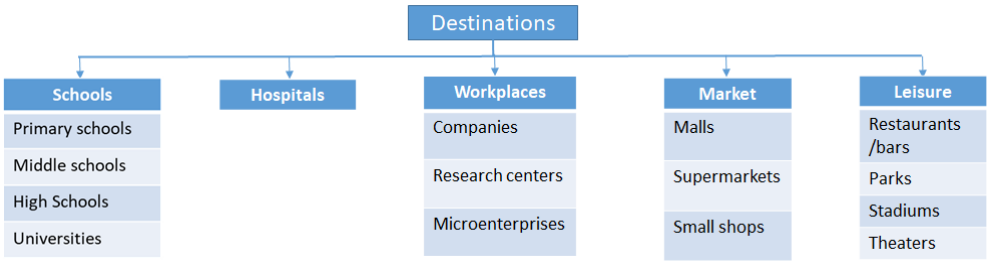


Figure 2: Classification of destinations in categories and subcategories.

3.2.3. Schools

In this category, we have all the educational institutions such as

1. Primary schools
2. Middle schools
3. High schools
4. Universities

3.2.3.1 Primary Schools

In this subcategory, we consider all the primary schools in the region. It includes *école maternelle, école élémentaire and école primaire*. There is at least one primary school in each zone. The nodes in this category have been located manually using method 2 as described in 3.2.1.2. For the capacity, we take the total number of students studying there. The location and capacity for primary schools has been taken from [37] by searching for 'école' in each zone. For each zone, this website provides the complete list of such schools with the link to their websites. The websites of each of these schools contain relevant information such as name, address, location coordinates, total number of students etc. We input these data manually directly in the GeoJSON file meant for primary schools.

3.2.3.2 Middle and high schools

In these two subcategories, we consider all the middle schools*(collège)* and high schools *(Lycée)* in our region. They have been located by method 2 as described in 3.2.1.2. For the capacity, we take the total number of students studying there. The location and capacity for the colleges and Lycee are retrieved from the website [38] by following the same procedures as that for the primary schools.

3.2.3.3 Universities

Here, we consider the sites of University Grenoble Alpes namely, Campus in Saint Martin d'Hères, Grenoble INP and IUTs in Grenoble downtown. We locate them by method 3 as described in 3.2.1.3. Here, the total capacity is defined as the total number of students and administrative or technical staffs which is 59600 as mentioned on the university website [39]. For the capacity of different sites, let C_u = total capacity = 59600. We divide the total capacity to obtain the capacities for different sites as follows:

1. Campus- 85% of C_u,
2. Grenoble INP- 10% of C_u,
3. IUTs- 5% of C_u.

This division has been done keeping in mind that most of the students visit the campus daily. Finally, Figure 3 depict the nodes in different subcategories of schools.

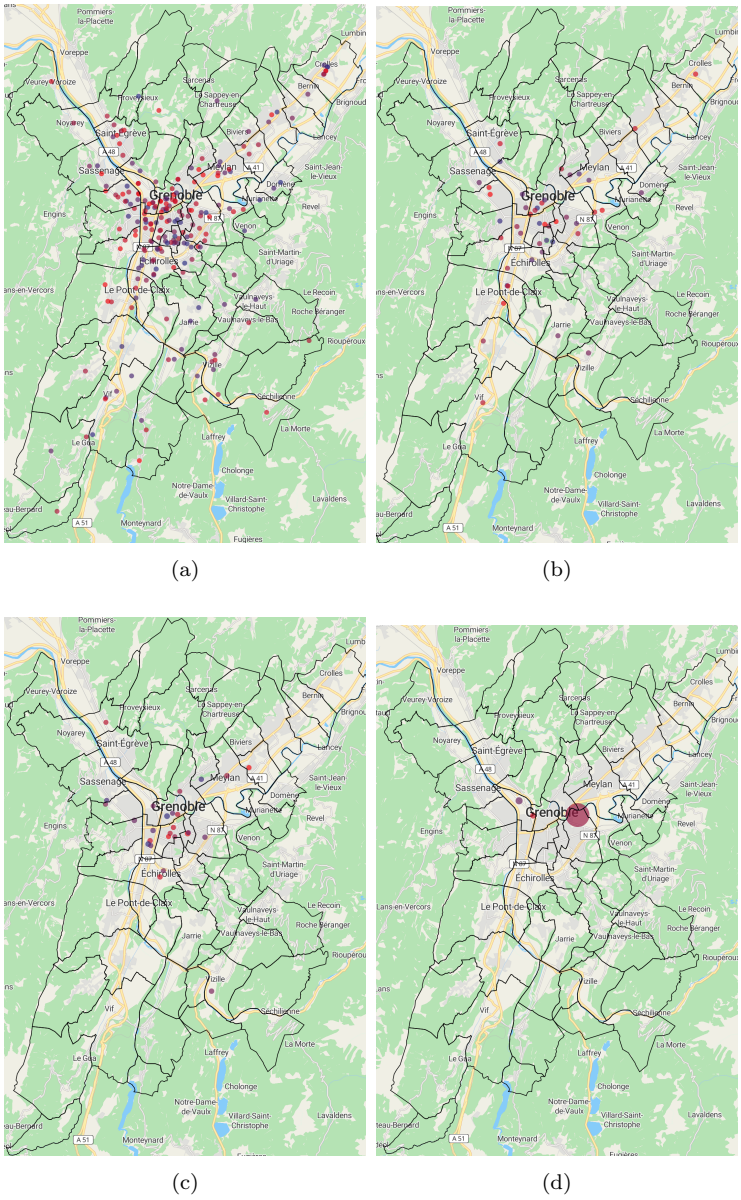


Figure 3: Nodes for destinations in school category. The figures have been taken from the online demonstrator which will be described in 8. (a) Primary school nodes. (b) Middle school nodes. (c) High school nodes. (d) University nodes.

3.2.4. Hospitals

In this category, we have considered *Centre hospitalier universitaire (CHU) nord, CHU sud, Clinique belledonne, Clinique des Cèdres* and *Clinique mutualiste.* We locate each of these hospitals by method 3 as described in 3.2.1.3. The capacity of each hospital is considered to be the sum of the following:

1. total number of available beds.
2. total number of staffs including the doctors.
3. total number of daily consultations.

4. total number of visitors accompanying the patients. Information about the first three are more or less available in some form for the hospitals, which will be discussed for each hospitals subsequently. The data regarding number of visitors need to estimated based on the first three. So, here we discuss the strategy for this estimation. For the number of visitors at a hospital, we need to consider that there are three kinds of patients who visit the hospital:

   (a) consultations only.

   (b) ambulatoire - they stay at the hospital for half a day to a full day mostly for minor surgeries.

   (c) hospitalization - the patients who need to be hospitalized and stay at the hospital for six days on an average. This information has been retrieved from [40].

   Due to the lack of data regarding the number of visitors, we can follow the following rule of thumb.

   - For consultation and ambulatoire - zero visitors for 80 percent and 1 visitor for 20 percent.
   - For hospitalized patients - one or two visitors per patient. The maximum limit of visitors is 2 per patient at CHU Voiron [41]. Due to lack of this data for others, we take this rule in general for other hospitals as well. To compute the visitors against hospitalized patients, we assume that 80% of the beds are occupied.
   - Therefore, the general rule for number of visitors in a hospital is as follows:

$$\begin{aligned} &\text{number of visitors} \\ &\qquad = 0.2 * (\text{number of consultations} + \text{patients in ambulatoire}) \\ &\qquad + 1.5 * 0.8 * (\text{total number of beds.}) \end{aligned} \tag{11}$$

Now, We discuss the data for $1, 2$ and 3 in details for each hospital in the following sections.

3.2.4.1 CHU nord and sud

There are two CHUs in the region which are CHU nord located in La Tronche and the other is CHU sud located in Echirolles. These are the main and largest public hospitals in the region. Annual data for three CHUs (CHU nord, CHU sud and CHU Voiron) are available collectively on [40]. Here, we consider the latest available annual data which is for 2018. The available data is for three hospitals (CHU Voiron, CHU nord and CHU sud) but only CHU nord and CHU sud are located in the region and the number of beds in CHU voiron is given on the website [42]. So, at first we divide the data in two parts (one for CHU Voiron and the other for both CHU nord and sud) in proportion to the number of beds. Since, CHU nord is larger than CHU sud, the remaining data is divided such as two-thirds are given to CHU nord and the remaining one-third to CHU sud. The number of beds are given as follows:

Data category	All CHU	CHU nord	CHU sud
Number of beds	2133	1222	611
Number of employees	9000	5156	2568
daily consultations and *ambulatoire*	2677	1534	767
Visitors		1773	886

Table 1: Computation of data for CHU nord and CHU sud

- Total number of beds in three CHUs = 2133.
- Total number of beds in CHU Voiron = 300.
- the proportion by which other data needs to be divided in order to obtain them for CHU nord = $\frac{2}{3} * \frac{1833}{2133}$.
- the proportion by which other data needs to be divided in order to obtain them for CHU sud = $\frac{1}{3} * \frac{1833}{2133}$.

Now, we compute all the data for CHU which is depicted in Table 1:

3.2.4.2 Clinique mutualiste

Clinique mutualiste de Grenoble is a non-profit Private Health Establishment of Collective Interest (ESPIC) participating in the public hospital service. The data for this hospital has been retrieved from the website [43]. The number of visitors has been computed using (11). Table 2 contains the capacities of this hospital.

3.2.4.3 Clinique Belledonne

Clinique Belledonne is a private hospital located in Saint Martin d'Heres. It has 290 beds and around 650 employees. This data is available on [44]. The data regarding the daily consultations are not available, so this data is computed in proportion to the number of doctors by considering the data for CHU as the ground rule. There are 2677 consultations per day at CHU which has 2000 physicians. In this proportion, Clinique Belledonne, which has 150 physicians, will have 201 consultations and ambulatoire patients daily. The number of visitors has been computed using (11). The capacity of this hospital is listed in Table 2.

3.2.4.4 Clinique des Cèdres

Clinique des Cèdres is a private hospital located in Echirolles. It has 200 beds as mentioned on the website [45]. No other relevant data other than the number of beds is available to us, so we compute them in proportion to the number of beds taking the data for Clinique Belledonne as ground rule. The number of beds in Clinique Belledonne is 290 and that in Clinique des Cèdres is 200 so the proportion by which the data of Clinique is multiplied in order to get that for Clinique des Cèdres is $\frac{200}{290}$. Please see Table 2 for the complete data. Figure 4 depict the nodes corresponding to hospitals.

Hospital name	Beds	Consultations and *ambulatoire*	Employees	Visitors	Total
CHU nord	1222	1534	5156	1773	9685
CHU sud	611	767	2568	886	4832
Clinique mutualiste	436	286	1300	581	2603
Clinique Belledonne	290	201	650	348	1489
Clinique des Cèdres	200	136	442	237	1015

Table 2: Capacities of hospitals

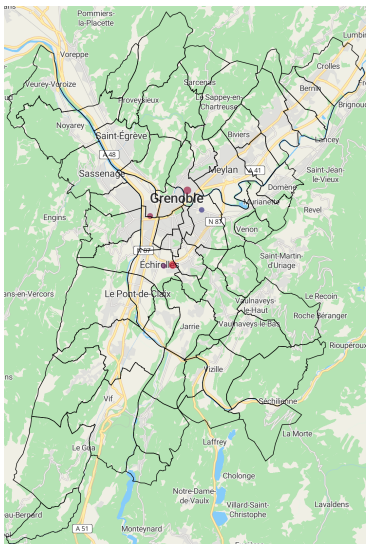


Figure 4: Hospital nodes

3.2.5. Workplaces

The category workplaces has the following three subcategories:

1. Companies
2. Research centers
3. Microenterprises and others

In France, INSEE [46] classifies the 'enterprises' on basis of number of employees as follows:

- Grandes enterprises (GE) - A GE has more than 5000 employees and 25% of the total employees work there.
- Entreprises de taille intermédiaire (ETI)- An ETI has less than 5000 but more than 250 employees and 25% of the total employees work there.
- Petites et moyennes entreprises (PME) - A PME has less than 250 but more than 9 employees and 30% of the total employees work there.

- Microenterprises (MIC) - An MIC has less than 10 employees and 20% of the total employees work there.

We consider GE, ETI and PME in companies subactegory and MIC in Microenterprises and others along with some public offices.

3.2.5.1 Companies

Here, we consider mainly the private companies. Due to lack of information on the list of companies and number of employees working in them, we consider the list of companies provided in [47] with employees greater than or equal to 120. The destination nodes corresponding to the companies from this list have been located manually using method 3 as stated in 3.2.1.3. While locating the companies from the list, we have also added some other companies located in the same commercial zones as visible in the map on QGIS. Figure 5a depicts the companies considered. Here, the capacity is defined to be the total number of employees working in the companies. In this subcategory, we consider the GE, ETI and PMEs which employees around 80% of the total employees. Since, we consider only the companies with more than 120 employees, we assume that we have included only all GE and ETI but only half of the PMEs in our list, that is we have only considered 65% of the total employees. Therefore, we need to inflate this number which is done as follows:
Let T_e = total number of people working in private companies and A_m = sum of employees in companies from the magazine. Since, the magazine gives us 65% of T_e, therefore, T_e can be computed as

$$T_e = \frac{100}{65} * A_m. \tag{12}$$

According to INSEE, people working in GE + ETI + PME= 80%T_e. Since it is assumed that, we only have 65% of GE+ETI+PME from the magazine, we need to inflate the capacities of the companies to 80%. Since, 65% of $T_e = A_m$, then 80% of $T_e = \frac{80}{65} * A_m$. Finally, for the individual companies, we have that

$$\begin{aligned} &\text{the inflated capacity of the companies} \\ &= \frac{80}{65} * \text{ original no. of employees from the magazine.} \end{aligned} \tag{13}$$

Some Companies like ST microelectronics have more than one sites in the region. We compute the capacity of each site by dividing the total number of employees in that company by the number of sites. To the companies considered apart from the ones from [47], an average capacity of 120 has been given.

3.2.5.2 Research centers

In this category, we consider the research labs listed on [48] and some others which are associated with the University Grenoble Alpes. They have been located manually by method 3 as explained in 3.2.1.3. The capacity of each research center is the sum of the number of doctoral students, researchers and technical staffs working in each center. This information has been taken from the respective websites of each of these research centers.

3.2.5.3 MIC and others

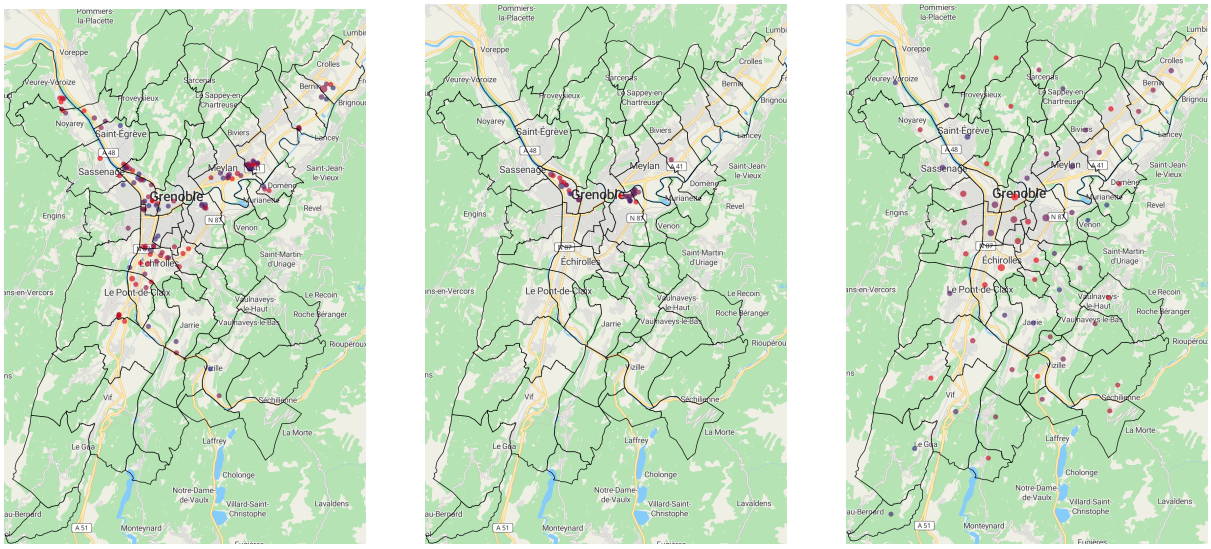
In this subcategory, we include the public offices and microenterpises (MIC). This is represented by a node in each *commune*. The capacity of each node is sum of the employees of MIC and employees in public offices in the corresponding *communes*.

Let $M_e = 20\%$ of $T_e =$ the total number of employees in microenterprises in the region and $E_p =$ the total employees of public offices in the region. Therefore, capacity of each node $= (M_e + E_p) * \dfrac{\text{population of the zone}}{\text{total population}}$.

We have the knowledge of M_e but we need to find E_p. We observed that there is a huge gap in the actual number of employees in a zone and our estimate of the number of employees taken from the magazine. Therefore, we consider to include the left ones such as public employees as others. We know from INSEE that 33% of the total employees work in public offices. Some of these employees have already been taken into account in hospitals, universities, schools, etc. so, we assume that we have already counted one-third of the total employees in them implicitly. We have also counted the employees as the capacity of companies in the previous section that is T_e in (12) which we assume to be one-third of the total. So, the remaining one-third is considered among others and hence $E_p = T_e$.

Therefore, we have that

$$\begin{aligned} &\text{the capacity of each node in MIC and others} \\ &= (0.2T_e + T_e)\frac{\text{population of the zone}}{\text{total population}} \\ &= (1.2 * T_e)\frac{\text{population of the zone}}{\text{total population}}. \end{aligned} \tag{14}$$



(a) (b) (c)

Figure 5: Nodes in the workplaces category. The figures have been taken from the online demonstrator which will be described in 8. (a) Nodes depicting companies. (b) Nodes depicting research centers. (c) Nodes depicting MIC and others.

3.2.6. Shopping centers

This category is subdivided into three subcategories:

1. Shopping malls
2. Supermarkets and
3. Small shops

The capacity in this category means the maximum number of shoppers allowed in the shop at a given time.

3.2.6.1 Shopping malls

In this category, we consider the big shopping malls such as Grand Place, IKEA and Castorama which have more than one floor. The nodes corresponding to these destinations have been located by method 3 as stated in 3.2.1.3.

For the capacity, we follow the national fire regulation [49] for maximum number of allowed shoppers in a shop depending on their surface area. From there, we know that on the ground floor, the permissible number of shoppers is 2 person per m^2 in one-third of the surface area, on first floor 1 person per m^2, on second floor 1 person every $2m^2$, in one-third of the surface area. Therefore, the capacity of the shopping mall with three floors available for shopping is given as follows:

$$\text{capacity of shopping malls} = \xi_s * \frac{7}{6} * \text{surface area} \tag{15}$$

where ξ_s is a scaling parameter to tune the capacities to a reasonable amount. The parameter ξ_s accounts for the fact that the actual available surface area of a shop also depends on a number of factors such as shopping shelves, walking area, billing area etc. It can also account for the unavailable spaces because of people moving with their shopping trolleys. Here, $\xi_s = \frac{1}{3}$. The surface area for these shopping malls have been taken from [50], [51] and [52] for Grand place, IKEA and Castorama respectively.

3.2.6.2 Supermarkets

For the supermarkets, we consider the shops with a surface area of more than $300m^2$. Therefore, we consider mainly Carrefour hypermarkets, Carrefour markets, Carrefour Contact, Geant casino, Casino supermarché, Monoprix, Biocoop, Lidl, Intermarché super, Super U and a few others. On the commercial websites of the supermarket groups such as carrefour, they specify how different outlets within their group are classified based on the surface area or location. The classification of supermarkets and their average surface areas have been taken from different webpages corresponding to different supermarkets which are Carrefour [53], Casino [54], Monoprix [55], Super U [56], Intermarché [57], Lidl [58], Biocoop [59].

These supermarkets have been located by method 3 as described in 3.2.1.3. For the capacity, we consider the nation wide rule regarding the permissible number of shoppers according to fire-safety regulations [49]. For the sake of simplicity, we assume that the

supermarkets and small shops have only one floor for shopping. Since, on the ground floor, the permissible number of shoppers is 2 persons per square meter on one- third of the surface area, the capacities of the supermarkets are given by

$$\text{Capacity of supermarkets or small shops} = \frac{2}{3} * \text{surface area of the shop.} \tag{16}$$

similar to the shopping areas, supermarkets and small shops are also never full to this capacity, so we reduce the capacity by a scaling factor ξ_s. Therefore,

$$\text{Capacity of supermarkets} = \xi_s * \frac{2}{3} * \text{surface area of the shop.} \tag{17}$$

Table 3 contains the average surface areas and the capacities of different supermarkets. Apart from the shops listed in Table 3, we also include Satoriz, Naturalia and Spar

Supermarket name	Average Surface area (m^2)	Capacity ($\xi_s = 1$, fire regulation)	Capacity $\xi_s = 1/3$)
Carrefour Hypermarket	10000	6667	2222
Carrefour Market	2000	1333	444
Carrefour contact	600	400	133
Geant casino	7400	4933	1644
Casino supermarché	1700	1133	378
Monoprix	1800	1200	400
Hyper U	4985	3323	1108
Super U	2016	1344	448
Intermarché Super	2000	1333	444
Lidl	1000	667	222
Biocoop	300	200	67
Supermarkets(national average)	6347	4320	1410

Table 3: Surface area and capacities of supermarkets

markets for which we consider the surface area and capacities of Carrefour contact. We also consider some other supermarkets and for that we take the surface area same as the national average from the website [60].

3.2.6.3 Small shops

We retrieve their location by filtering the data from openstreetmap using QGIS software as explained in 3.2.1.1. In this category, we consider the shops with average surface area less than $300m^2$. INSEE specifies the shop type and national average of their surface area on [60]. We consider that the shops located are of the types which have been enlisted in Table 4. It can be seen that these shops encompasses almost all types of small shops.

We compute the average of surface area of these shops to obtain a common surface area for all the small shops in our region and their capacity is computed by (17) with $\xi_s = 1/3$. The average surface area is $147m^2$ and hence the average capacity is 33. All the nodes in this subcategory have been given the the same capacity.

Shop type	Average Surface area (m^2)
Supérettes	216
Commerces d'alimentation générale	57
Commerces de détail de produits surgelés	263
Alimentation spécialisée et artisanat commercial	62
Habillement et chaussures	168
Culture, loisirs, sport	209
Produits pharmaceutiques et articles médicaux et orthopédiques	111
Autres équipements de la personne	79
Équipements de l'information et de la communication	128
Autres commerces de détail	180

Table 4: Small shop type and their average surface area as mentioned on [60].

The nodes corresponding to the nodes in this category have been depicted in Figure 6.

3.2.7. Leisure

The category leisure has four different subcategories which are:

1. Restaurants and bars,
2. Parks,
3. Stadiums and
4. Theaters.

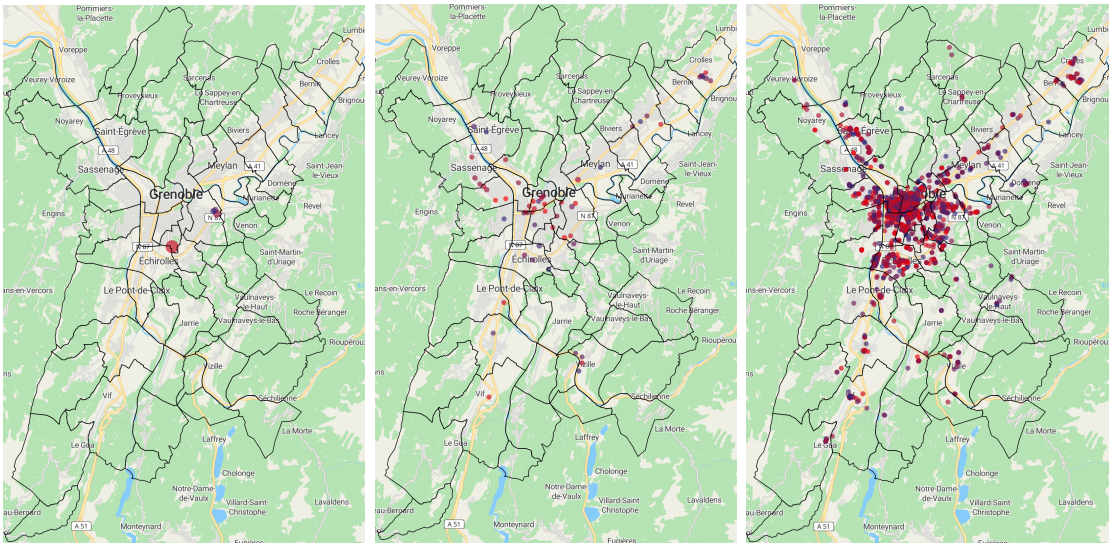
3.2.7.1 Restaurants and bars:

The restaurants and bars have been located by method 1 as described in the 3.2.1.1. We have located around 700 restaurants and bars in the considered zones. Here, the capacity is defined as the total number of seats available in a restaurant. Due to lack of data regarding available seats for all these restaurants, at first we consider the total number of seats in 55 restaurants in Grenoble downtown from the website [61]. In our sample of 55 restaurants, the mean is 46 and variance is 1320. In order to impute the capacities for the remaining restaurants, at first, we find which probability distribution fits this sample data best using "distributionFitter" app in Matlab as can be seen in Figure 7a.

We find that the desired distribution is Lognormal distribution with the following probability density function

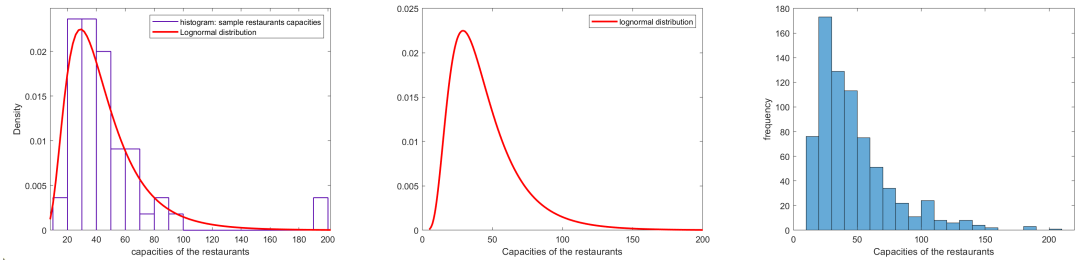
$$f_X(x) = \frac{1}{x\chi\sqrt{2\pi}} \exp(-\frac{(\log x - m)^2}{2\chi^2}). \tag{18}$$

where m and χ is computed from the mean $mean$ and variance var of the our sample as



(a) Nodes depicting shopping malls (b) Nodes depicting supermarkets (c) Nodes representing small shops

Figure 6: Destination nodes in the shopping category. The figures have been taken from the online demonstrator which will described in 8.



(a) (b) (c)

Figure 7: Curve fitting of the 55 sample capacities of restaurants to find the probability distribution which fits them the best. (a) Curve fitting for the total number of seats in the restaurants using 'distributionFitter' app in MATLAB. (b) probability density function of the Lognormal distribution which fits our sample data. (c) histogram of the capacities of all the restaurants in our region generated using Lognormal distribution with mean m and variance σ computed using (19).

follows:

$$m = \log\left(\frac{mean^2}{\sqrt{var + mean^2}}\right) \tag{19}$$

$$\chi = \sqrt{\log\left(1 + \frac{var}{mean^2}\right)}. \tag{20}$$

For our sample $mean = 46$ and $var = 1320$. The pdf of the distribution which we follow is given in Figure 7b. We generate data sets of 800 capacities using Lognormal distribution given by (18) and then we distribute these capacities to the restaurants randomly. Figure 7c shows the histogram of the capacities generated. We consider the capacities of the restaurants up to 200 only.

3.2.7.2 Parks:

We locate the parks by method 3 as described in 3.2.1.3. We retrieve the surface area of the parks on their respective websites, or the website of Grenoble-Alpes métropole or on the website of Isere tourism. For the parks for which we don't have the surface area, we allocate them an average surface area of $50000m^2$. For the capacities of the parks, we consider the following thumb rule:

$$\text{Capacity of a park} = \frac{\text{total surface area of the park}}{15}. \tag{21}$$

Here, we assume that in general, in a not very crowded situation, there can be one

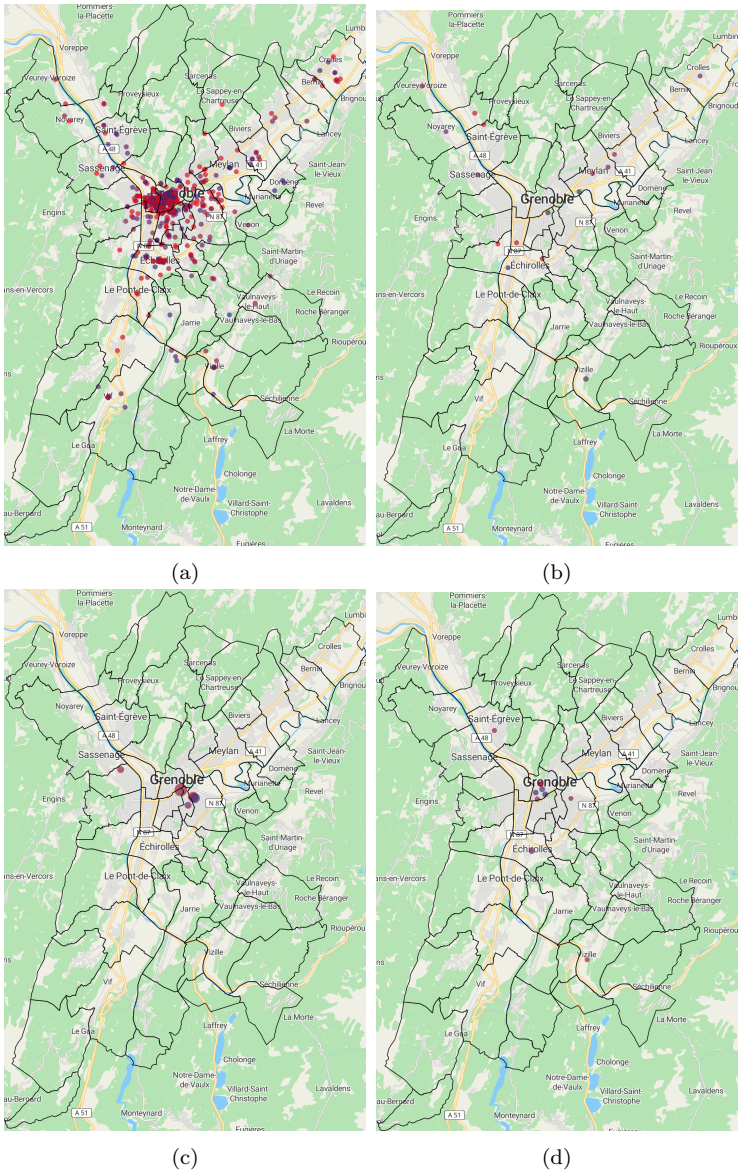


Figure 8: Destination nodes in the leisure category. The figures have been taken from the online demonstrator which will described in 8. (a) Nodes representing restaurants. (b) Nodes depicting parks. (c) Nodes representing stadiums. (d) Nodes depicting theaters.

person in an area of $15m^2$. In case of any event, this thumb rule won't be valid any more

because there would be many persons in a small area of the parks. However, in general, this capacity is not filled up and people go there either just to walk or run. Therefore, we scale the above mentioned capacity by ξ_p to have a reasonable number of daily visitors. Here, we take $\xi_p = 0.02$, therefore, we have that

$$\text{Capacity of a park} = 0.02 * \frac{\text{total surface area of the park}}{15}. \tag{22}$$

3.2.7.3 Stadiums and theaters:

They nodes in these two subcategories have been located using method 3 as described in 3.2.1.3. The capacities of the stadiums and the theaters are the total number of available seats. This information is taken from their respective websites or from the website of Grenoble métropole [62] or the booking websites for the theaters.

The destination nodes of the leisure category can be seen in Figure 8.

3.2.8. Aggregation of nodes

In some subcategories such as restaurants and small shops, the number of nodes is very large. So, in order to reduce the complexity in the mobility model, we aggregate the nodes of the destinations in such subcategories. This aggregation is done per zone by replacing all the nodes belonging to a subcategory in a zone by a single node. The location of such aggregated node is the barycenter of the locations of all the nodes of a subcategory in a zone and the capacity of such a node is the sum of the capacities of all the nodes of the same destination category in that zone. The destination subcategories for which this aggregation has been done are primary schools, supermarkets, small shops, companies ,research centers and restaurants. Figure 9a depicts how the the aggregation of nodes belonging to the same subcategory is done per zone and Figure 9b shows all the nodes in the region where as Figure 9c depicts the aggregated nodes.

3.3. O-D connections

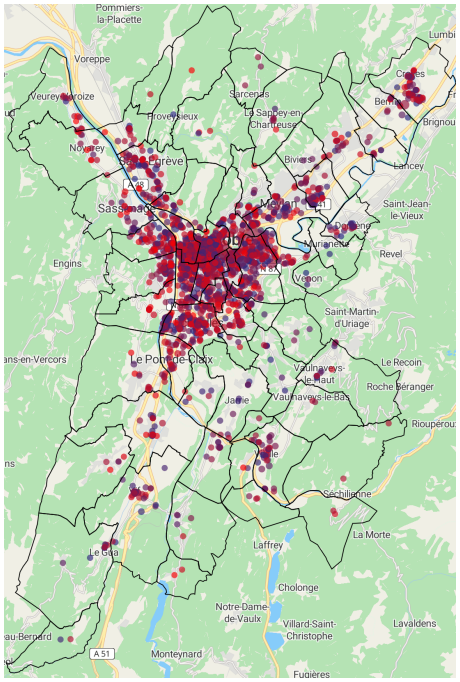
Each origin/destination node i is connected to at least one destination/origin node j. For each subcategory, we have different rules to decide if the pair $(i, j) \in \mathcal{E}$, where $i \in \mathcal{V}_o$ and $j \in \mathcal{V}_d$. These rules have been listed in Table 5.

3.3.1. Attraction based rules

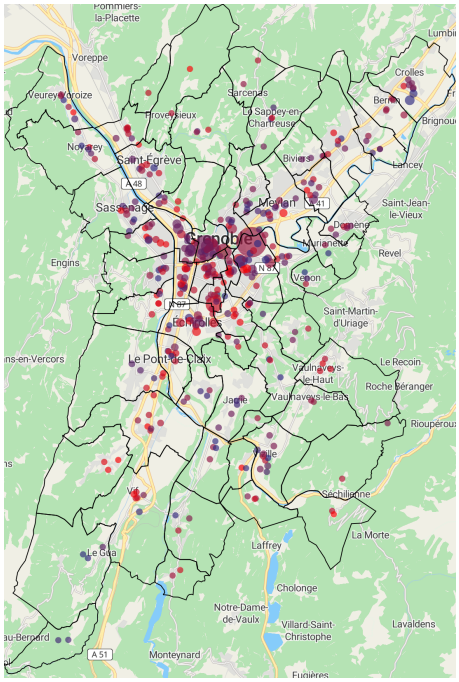
Let d_{ij} represent the road distance between origin i and destination j and σ_j, the maximum distance that ν% of people of i travel 'daily' to visit j. In our case, we have assumed $\nu = 95\%$ which means that the rest 5% of the people in i can choose to travel a distance more than σ_j to visit the destination j. Given the location of origins and destinations, we compute the minimum real minimum road distance d_{ij} between all possible $i \in \mathcal{V}_o$ and $j \in \mathcal{V}_d$ using OSMnx python library [66]. The number of individuals per unit area who travel a distance from home is inversely proportional to square of the distance travelled [67]. Inspired by the gravity models [7], we compute Q_{ij}, the attraction between i and j as

$$Q_{ij} = P_i C_j e^{-|\ln(1-\nu)|\left(\frac{d_{ij}}{\sigma_j}\right)^2}.$$

(a)



(b)



(c)

Figure 9: Aggregation of nodes in the subcategories: primary schools, supermarkets, small shops, companies, research centers and restaurants. The capacities and location of the aggregation nodes are the sum of the capacities of the non-aggregated nodes the barycenter of their locations respectively. (a) Aggregation of nodes belonging to the same subcategory in a zone. (b) Non- aggregated destination nodes in our region. (c) The aggregated destination nodes.

For each destination j, normalize the attraction Q_{ij}

$$A_{ij} = \frac{Q_{ij}}{\sum_h Q_{hj}} = \frac{P_i e^{-|\ln(1-\nu)|\left(\frac{d_{ij}}{\sigma_j}\right)^2}}{\sum_h P_h e^{-|\ln(1-\nu)|\left(\frac{d_{hj}}{\sigma_j}\right)^2}}.$$

where the dependence of C_j cancels out from numerator and denominator. Then, the OD matrix $\mathcal{O}_{ij}$ is computed as

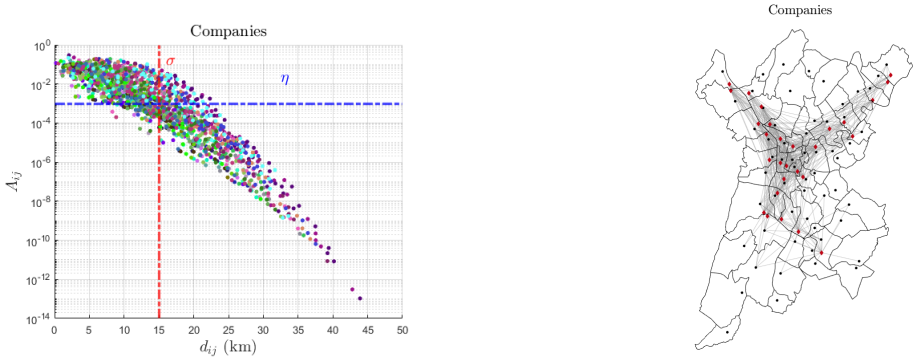
$$\mathcal{O}_{ij} = \begin{cases} 1 & \text{if } d_{ij} \leq \sigma_j \ \& \ A_{ij} \geq \eta_j \\ 0 & \text{otherwise} \end{cases}$$

Note that, here σ_j is a threshold on the maximum distance traveled by the most people 'daily' to go from an origin to destination j and η_j is a threshold on the attraction between i and j. These thresholds are different for destinations in different subcategories.

Subcategory	Rules for connecting origin and destination nodes
Primary schools	connected to the origin node of the same region.
Middle schools	connected to the corresponding govt. defined college sector.[63]
High schools	govt. defined rules based on one's address [64][65]
Universities/ Hospitals	connected to every origin node
Workplaces/ Shopping/ Leisure	Attraction based rules to be defined in Section 3.3.1.

Table 5: OD connection rules for different subcategories

This threshold reflects that the fact that a person prefers to go to a nearer destination if available. For example, in general, we visit the shops which are close to our residences. Figure 10 depicts how the thresholds σ and η_j are applied to compute the OD connections. In Figure 10a, we consider the upper-left quadrant after applying the thresholds. Based



(a) (b)

Figure 10: Computation of OD connection matrix using attraction law. (a) plot of Attraction vs distance for companies. (b) OD connections for companies with $\sigma = 15km$.

on the rules for connecting two locations corresponding to each subcategory, according to Table 5 and choosing thresholds for each subcategory, we have the connections as depicted in Figures 11-15.

Finally, we have the mobility network of Grenoble area as depicted in Figure 16. It can be seen that the city center has the highest connections which is a fact as most of the destinations are situated in the city center and it attracts a large number of people daily.

4. When?

In this section, we find the time schedules for different subcategories. In particular, we find the destination schedule, average time spent, mobility windows and Demand Gating Profile (DGP) for each subcategory.

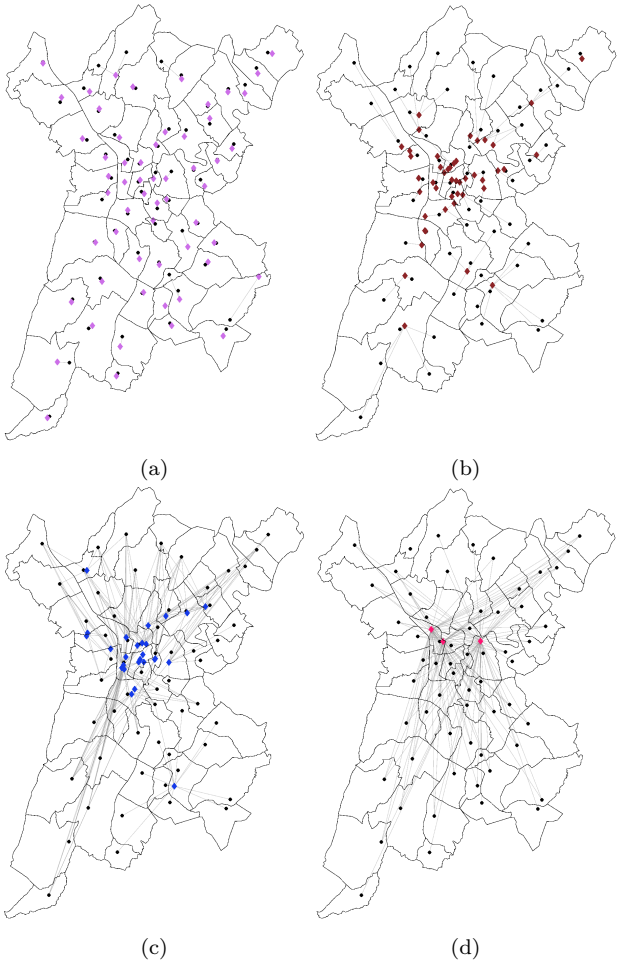
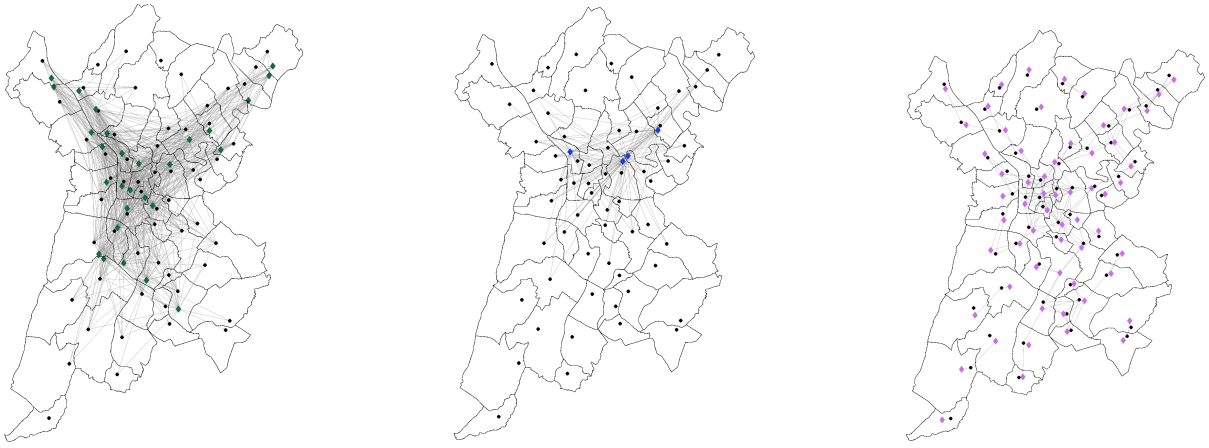


Figure 11: Connections for schools based on the government rules. The black square nodes represent the origins and the colored diamond shaped nodes represent the destinations in the corresponding subcategory. (a) OD connections for primary schools. (b) OD connections for middle schools. (c) OD connections for high schools. (d) OD connections for universities.

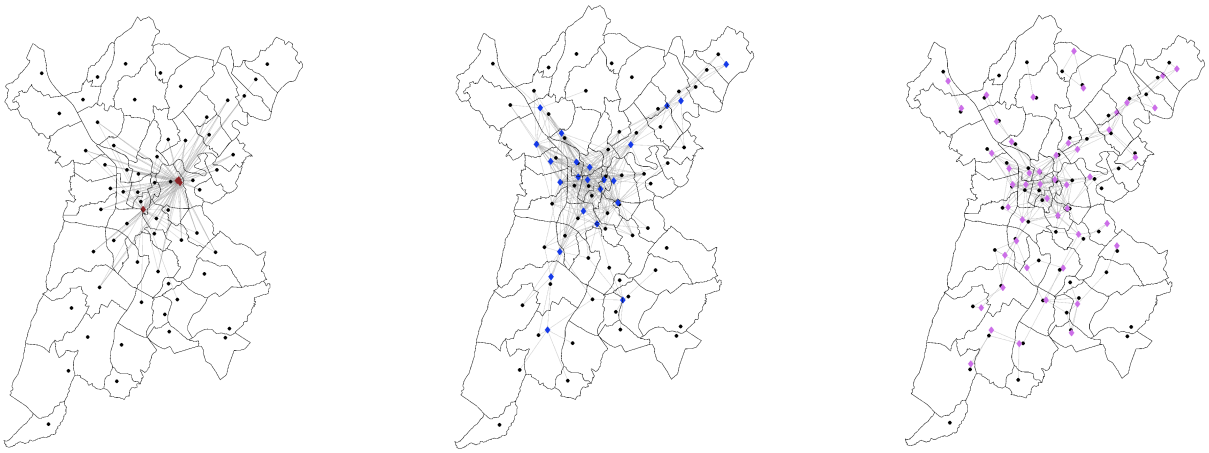
4.1. Destination schedule, mobility window and average time spent

Recall that the destination schedule of a destination $j \in \mathcal{V}_d$ is the daily time interval $[a_j, , b_j)$, where, $0 \leq a_j < b_j \leq 24$ during which j is open. Here, a_j and b_j are the opening time and closing time of the destination $j \in \mathcal{V}_d$ respectively. Mobility window is defined as the time interval $[t_{ij}, t_{ij} + \tau_{ij}]$, during which there is mobility from i to j. Here, t_{ij} is the time when the mobility starts from i to j and τ_{ij} is the duration for which mobility keeps happening between i and j. For the sake of simplicity, here we consider that $\tau_{ij} = \tau_{ji}$. It means that the mobility window for the mobility from an origin i to a destination j is equal to the return mobility window from j to i. Average time spent s_j is the duration for which a person stays in the destination j on an average. We set these parameters for each subcategory. Therefore, we will follow the notations a_g, b_g, s_g for opening time, closing time and the average time spent of the destination subcategory g, where $g \in \{1, \ldots, 15\}$ and the mobility window from the origins $i \in \mathcal{V}_o$ to the destination subcategory g will be denoted by $[t_{ig}, t_{ig} + \tau_{ig}]$. The mobility window for the return from the destinations to the origins, will be denoted as $[t_{gi}, t_{gi} + \tau_{ig}]$. Table 6 contains different



(a) (b) (c)

Figure 12: OD connections for the destinations in workplaces category using the attraction based laws. (a) OD connections for the companies. Here $\eta_j = 0.001$ and $\sigma = 15km$. (b) OD connections for the research centers. Here $\eta_j = 0.001$ and $\sigma = 15km$. (c) OD connections for the MICs. Here $\eta_j = 0.001$ and $\sigma = 5km$.



(a) (b) (c)

Figure 13: OD Connections for destinations in shopping category using attraction based laws. The black square nodes represent the origins and the colored diamond shaped nodes represent the destinations in the corresponding subcategory. (a) OD connections for the shopping malls. Here $\eta_j = 0.001$ and $\sigma = 15km$. (b) OD connections for the supermarkets. Here $\eta_j = 0.001$ and $\sigma = 8km$. (c) OD connections for the small shops. Here $\eta_j = 0.001$ and $\sigma = 5km$.

time-schedules parameters for different subcategories. The opening time, closing time and the average time spent for primary schools, middle schools and high schools has been taken from [68], [69] and [70] and that for working places from [71]. For shopping and Leisure, we referred to the timings of some of the destinations on web or OSM maps and set that for all the other destinations in that subcategory. On the weekends, schools and workplaces are closed. Note that, we have different time schedules for each day for the implementation to take into account different time schedules for each day. This is important since destinations in some subcategory such as stadiums are assumed to open only on the evenings of each Friday and closed otherwise. Also for example, primary schools have different closing times on specific days of the week and restaurants and

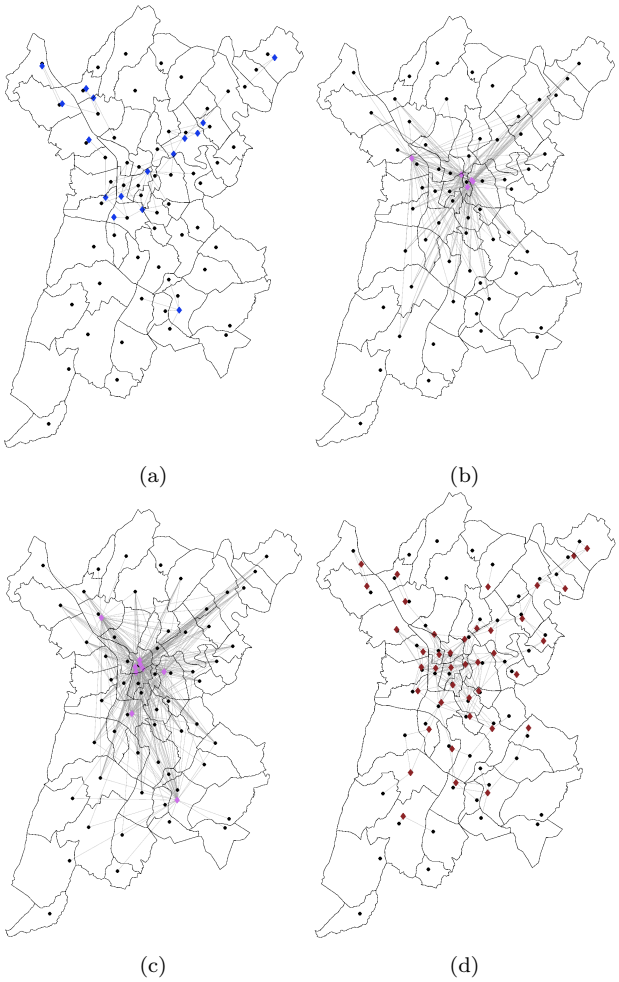


Figure 14: Connections for the destinations in the leisure category. The black square nodes represent the origins and the colored diamond shaped nodes represent the destinations in the corresponding sub-category. (a) OD connections for the parks. Here $\eta_j = 0.001$ and $\sigma = 5km$. (b) OD connections for the stadiums. Here $\eta_j = 0.001$ and $\sigma = 20km$. (c) OD connections for the theaters. Here $\eta_j = 0.001$ and $\sigma = 10km$. (d) OD connections for the restaurants. Here $\eta_j = 0.001$ and $\sigma = 5km$.

bars are open till very late on weekends. Moreover, it has been mentioned earlier that the hospitals have different types of capacities, hence there are also different purposes to visit hospitals. These persons can be: employees denoted by E_h, new hospitalized patients H_h, patients going for consultations H_c, and visitors V_h. We also assume that a constant population is present in the hospitals during night time. This includes the occupied number of beds and a proportion of employees who do night shifts. We compute the occupied number of beds B_h by the following formula:

$$B_h = H_h \times \text{average time of hospitalization}$$

where H_h is the average daily number of newly hospitalized patients. Let us assume that the proportion of employees who do night shifts is 10%, then the constant population present during the night time is given by $B_h + 0.1E_h$. The average time spent at hospitals by a person is computed as

$$s_h = \frac{E_h s_E + H_c s_C + V_h s_V}{E_h + H_c + V_h} \tag{23}$$

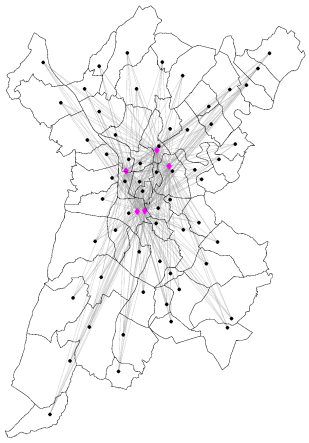


Figure 15: Connections for hospitals. The black square nodes represent the origins and the pink diamond shaped nodes represent the hospitals. Hospitals are connected with all the origins.

where s_E, s_C, s_V denote the employees work duration, average duration of consultation, and average duration of visitors, respectively. For destinations in other categories also, there are people with different purposes but the majority of the persons have a single purpose. For example, in shops there are employees and shoppers but the number of shoppers is very high in comparison to the the number of employees so the average time spent is mainly affected by the shopper so we didn't take the average time spent by the employees into account. On the other hand, all four types of people coming to the hospitals are significant in number. Now that we have the destination schedule we can easily compute the Supply Gating Functions by (9). In the next subsection, we will find the Demand Gating Profiles for different destination subcategories.

4.2. Demand Gating Profile (DGP)

Recall that DGP denoted by $\delta_{ig}(t)$ is a function such that $\int_0^{24} \delta_{ig}(t)dt = 1$. It captures the mobility pattern from i to g. In general, we have two types of destinations.

1. One in which mobility happens during a window when people come to the destinations from the origins in the morning and stay there for a long period of time and the mobility in the opposite direction starts at a specified time in the evening. This is the case with all the destinations in school category, workplaces category and stadiums. In this case, the DGP is defined as

$$\delta_{ig}(t) = \begin{cases} \frac{1}{\tau_{ig}} & t \bmod 24 \in [t_{ig}, t_{ig} + \tau_{ig}) \\ 0 & \text{otherwise} \end{cases}$$

   and

$$\delta_{gi}(t) = \begin{cases} \frac{1}{\tau_{ig}} & t \bmod 24 \in [t_{gi}, t_{gi} + \tau_{gi}) \\ 0 & \text{otherwise} \end{cases}$$

   Notice that, if we have different destination schedules for different days of the week, then we will have different profiles for different days of the week.

2. The other kind is the one in which mobility keeps happening throughout the day. People keep coming here, stay for some time and then return and it keeps happening till the destination is closed. All the destinations of the categories hospitals,

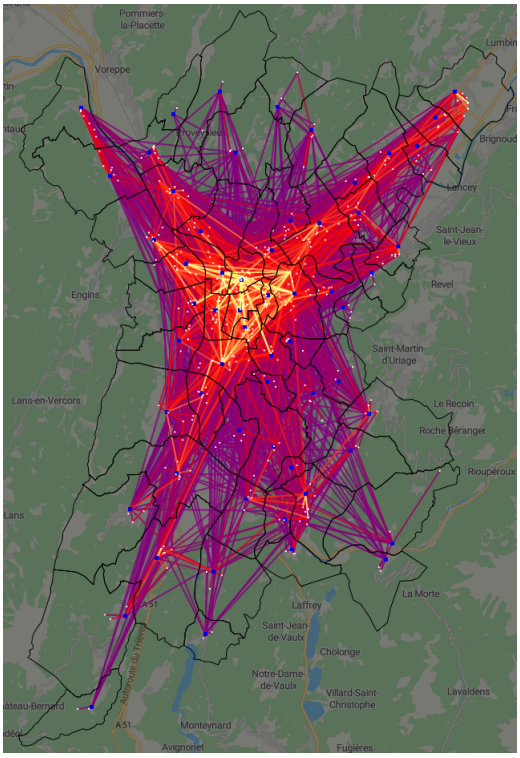


Figure 16: Mobility network in Grenoble area. The destinations are denoted by yellow nodes and origins by black. The nodes with higher connections are brighter than the ones which have low connections. This figure has been taken from the demonstrator which will be described in 8.

shopping and leisure are of this type. In this case, if we find the demand gating profile $\delta_{ig}(t)$ corresponding to the mobility from origins to the destinations then the profile for return mobility is given by $\delta_{gi}(t) = \delta_{ig}(t - s_g)$, where s_g is the average time spent per category. For the hospitals, we have

$$\delta_{ig}(t) = \begin{cases} \frac{1}{\tau_{ig}} & t \bmod 24 \in [t_{ig}, t_{ig} + \tau_{ig}) \\ 0 & \text{otherwise} \end{cases}$$

For the shopping and leisure category, we inferred this mobility pattern $\delta_{ig}(t)$ from the household travel survey of Grenoble region done by INSEE in 2010 [72]. The table 'ED20A' in the file EMD'10 contains the number of people traveling with different motives on a given hour of a day. Since, the data points are given for each hour and in absolute numbers, we obtain the required DGP for destinations in these two categories by interpolation and normalising the values by the integral of the curve. For each of the destinations, we compute these profiles for each day of the week. This information can also be retrieved by observing the popularity trends of destinations in each subcategory on google or OSM. Here, in this work, we observed the google popularity trend of some destinations and computed the mobility profiles on the weekend by manipulating the ones obtained through EMD data. In particular, the profiles on a week day has been kept different from the ones on a weekend reflecting the high popularity of destinations in leisure category. Figure 17 depicts the different types of DGPs we have for destinations in different categories on a week day.

Subcategory	$\mathbf{a_g}$	$\mathbf{b_g}$	$\mathbf{s_g}$	$\mathbf{t_{ig}}$	$\tau_{\mathbf{ig}} = \tau_{\mathbf{gi}}$	$\mathbf{t_{gi}}$
Primary Schools	7h30	16h30	$(b_g - a_g)$	7h30	1h	16h
Middle schools		18h				17h
High schools		18h30				17h
Companies	8h	18h	$(b_g - a_g)$	8h	2h	16h30
Research centers						
MIC and others						
Hospitals	7h	20h	using (23) with $s_E = 8, s_C = 1$ and $s_V = 1$	7h	$(b_g - a_g)$	$t_{ig} + s_g$
Shopping areas	10h	20h	2h	10h	$(b_g - a_g)$	$t_{ig} + s_g$
Supermarkets	7h 30	23h	$\frac{3}{4}h$	7h 30	$(b_g - a_g)$	$t_{ig} + s_g$
Small shops	9h	19h	$\frac{1}{3}h$	9h	$(b_g - a_g)$	$t_{ig} + s_g$
Restaurants and bars	8h	22h	$2h$	8h	$(b_g - a_g)$	$t_{ig} + s_g$
Parks	0h	24h			$(b_g - a_g)$	$t_{ig} + s_g$
Theaters	14h	23h	3h	14h	$(b_g - a_g)$	$t_{ig} + s_g$
Stadiums						

Table 6: Time schedules for different subcategories on a typical week-day. a_g, b_g and s_g are the opening time, closing time and average time spent for destination in subcategory g. t_{ig} is when mobility from i to destination in subcategory g starts. $\tau_{ig} = \tau_{gi}$ is the size of mobility window between i and destination in subcategory g.

5. How many?

In this section, we will find how many people travel from i to j if they are connected.

5.1. Daily capacity of the destinations

In Section 3.2, we found the instantaneous capacities of different destinations denoted as C_j. Now, the daily capacity of a destination can be defined as the total number of people who can visit there. Let us denote the daily capacity of j by C_j^d which can be computed as

$$C_j^d = C_j \frac{(b_j - a_j)}{s_j}.$$

Notice that, we have $\frac{(b_j - a_j)}{s_j} \geq 1, \ \forall j \in \mathcal{V}_d$. Therefore, the destinations such as schools where people stay for the whole day, the daily capacity C_j^d is equal to the instantaneous nominal capacity C_j. On the other hand, for destinations such as shops where the mobility is continuously happening, in that case $C_j^d >> C_j$.

5.2. Average daily number of people from origins to destinations

If for an origin-destination pair $(i, j) \in \mathcal{E}$, $\mathcal{O}_{ij} = 1$ as computed in Section 3.3 for the subcategory $\mathcal{D}_c^a$, then the average daily number of people traveling from i to $j \in \mathcal{D}_c^a$ is denoted by M_{ij} and computed per subcategory $\mathcal{D}_c^a$ as

$$M_{ij} = \min\{X_1, X_2\}, \tag{24}$$

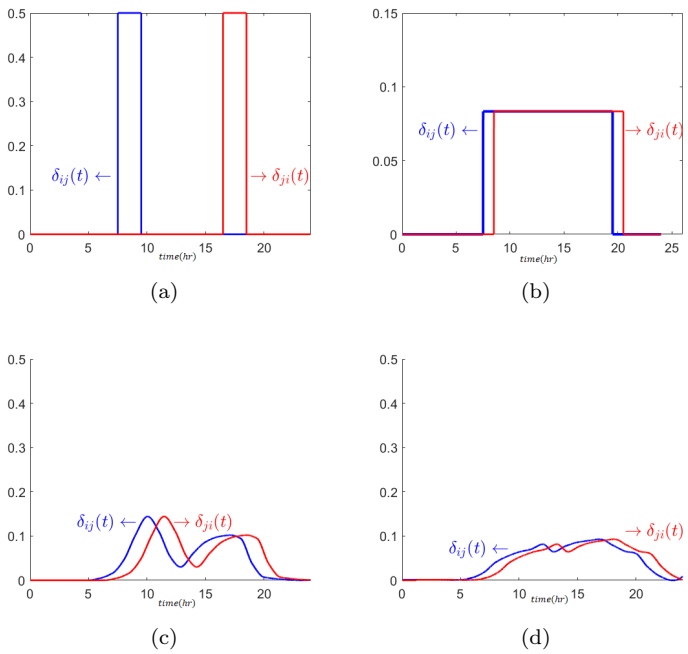


Figure 17: DGP on a weekday for different destination categories. (a) DGP for the categories schools and workplaces. (b) DGP for hospitals. (c) DGP for shopping. (d) DGP for Leisure.

where

$$X_1 = P_{ic} * f_c^p * \frac{C_j^d}{\sum\limits_{\substack{k\in\mathcal{D}_c^a,\\ k\in\mathcal{N}_i}} C_k^d} \quad \text{and} \quad X_2 = \frac{P_{ic}}{\sum_{h\in\mathcal{N}_j} P_{hc}} * C_j^d * \frac{1}{\delta_{ic}^* \tau_{ic}}.$$

Here, P_{ic} is the population of origin i which is eligible to go to destination subcategory $\mathcal{D}_c^a$, and f_c^p is the fraction of the eligible population of origin i which, in general, goes to $\mathcal{D}_c^a$ (this accounts for the fact that not all the eligible population goes every day to destination c). Also, C_j^d is the daily capacity of destination j, τ_{ic} is size of the mobility window, and δ_{ic}^* is the maximum of the gating function $\delta_{ic}(t)$. This definition of M_{ij} gives daily demands that are consistent with the population sizes and capacities. First, it takes into account the fact that in normal operating situation (i.e., with nominal capacities, without restrictions), the demand from i to j usually does not exceed population size of i and does not exceed capacity of j. Second, it reflects the fact that largest demands are from largest origins towards largest destination (as in the attraction law). The first term of (24), X_1, considers the population that is typically going from origin i to category c (this has size $f_c^p P_{ic}$) and then splits it among the destinations j that belong to category c and are neighbors of i, proportionally to their daily capacities (representing their sizes). The second term, X_2, instead, considers the daily capacity of a destination j, weighted by a factor $1/(\tau_{ic}\delta_{ic}^*)$, and splits it among origins i that are neighbors of j, proportionally to their eligible populations. The weighting factor $1/(\tau_{ic}\delta_{ic}^*)$ is equal to 1 when the gating function $\delta_{ic}(t)$ has a rectangular shape, i.e., when $\delta_{ic}(t)$ is constantly equal to $1/\tau_{ic}$ over an interval of size τ_{ic} and zero elsewhere; for other shapes, this correction terms accounts for the fact that the computation of daily capacity assumes that the destination is filled at full capacity over all its opening time, while the demand profile δ is such that the demand reaches full capacity only at its peak δ^*, and is lower at other times.

Notice that the model assumes that people travel from origin i to destination j and then returns back from j to i on the same day. This is expressed by defining $M_{ji} = M_{ij}$.

6. Model discretization

In the previous sections, we built the large-scale network of Grenoble and computed all the parameters required to implement the mobility model (2)-(9) on this network. Notice that in this mobility model, the number of people at any location $N_i(t)$ is always non-negative i.e. $N_i(t) \geq 0$ and it also remains below the population of P_i or the operating capacity $C_i^o(t)$, if it is an origin or a destination respectively i.e. $N_i(t) \leq Z_i$. Recall that $Z_i(t)$ denotes the population of i if i is an origin and the operating capacity C_i^o if it is a destination. Moreover, the total mass is also preserved. In other words, the total number of people in the system remains constant which is equal to the sum of all population i.e. for any time t, we have $\sum_{k\in\mathcal{V}_o\cup\mathcal{V}_d} N_k(t) = \sum_{i\in\mathcal{V}_o} P_i$. To simulate the mobility model, we need to discretize it. The discrete-time version of the continuous-time model (2)-(9), should preserve the same mathematical properties than its continuous counterpart. Unfortunatelly, trival discretization may lead to loss of such properties. This section analyzes these problems and provides a solution.

6.1. Discrete-time model

Let $N_i(k)$ be the number of people at location i at time step k and Δt be the time-step size, then with forward Euler discretization of (2), we have the discrete-time model as

$$N_i(k+1) = N_i(k) + \Delta t \sum_{j\in\mathcal{N}_i} (\phi_{ji}(k) - \phi_{ij}(k)), \tag{25}$$

where, the flow $\phi_{ij}(k)$ is the flow from i to j as described in (3). For any flow, $\phi_{ij}(k) \geq 0$, the discrete-time model (25) preserves the total mass. It can be seen in the following proposition.

Proposition 1. *Consider system* (25) *with any flow* $\phi_{ij}(k) \geq 0$, *the total number of people in the system remains constant, i.e., for all time* k, $\sum_{i\in\mathcal{V}_o\cup\mathcal{V}_d} N_i(k) = \sum_{i\in\mathcal{V}_o\cup\mathcal{V}_d} N_i(0)$.

Proof. We prove it recursively, i.e. we show that $\mathbf{1}^T N(k+1) = \mathbf{1}^T N(k)$, where, $\mathbf{1} \in \mathbb{R}^n$ is the vector of all ones and $N(k) = [N_i(k)] \in \mathbb{R}^n$ for $n = |\mathcal{V}_o \cup \mathcal{V}_d|$.
Let us define, $\Phi \in \mathbb{R}^{n\times n}$ as

$$[\Phi]_{ij} = \begin{cases} \phi_{ij}(k) & \text{if } j \in \mathcal{N}_i \\ 0 & \text{otherwise} \end{cases},$$

and rewriting (25) in matrix form, we have

$$N(k+1) = N(k) + \Delta t \left(\Phi^T - \Phi\right) \mathbf{1}.$$

By left-multiplication with $\mathbf{1}^T$, we get

$$\mathbf{1}^T N(k+1) = \mathbf{1}^T N(k) + \Delta t(\mathbf{1}^T(\Phi^T - \Phi)\mathbf{1}).$$

We will obtain the desired result by showing that $\mathbf{1}^T(\Phi^T - \Phi)\mathbf{1} = 0$. To do so, we notice that $\mathbf{1}^T\Phi^T\mathbf{1} = (\mathbf{1}^T\Phi\mathbf{1})^T = \mathbf{1}^T\Phi\mathbf{1}$, where the last equality is true since $\mathbf{1}^T\Phi\mathbf{1}$ is a scalar. Therefore, we have

$$\mathbf{1}^T N(k+1) = \mathbf{1}^T N(k).$$

□

Now, we discuss how to define the flows $\phi_{ij}(k)$ to be used in (25). For time-step k, one option is to have flows $\phi_{ij}(k)$ which can be computed naively by plugging $N(k)$ in the same equations as in continuous time, namely,

$$\begin{aligned} \phi_{ij}(k) &= \min(\Delta_{ij}(k), \Psi_j(k)) \\ \Delta_{ij}(k) &= \delta_{ij}(k\Delta t) f_{ij}(k\Delta t) 1_{N_i(k)>0} \\ \Psi_j(k) &= \psi_j(k\Delta t) F_j 1_{N_j(k)<Z_j(k\Delta t)} \end{aligned} \tag{26}$$

where the functions $\delta_{ij}(k\Delta t), \psi_j(k\Delta t)$ and the $f_{ij}(k\Delta t)$ and F_j are defined in section 2.1. The discrete-time system (25) along with the flows defined as in (26), could result in two undesired scenarios, which are $N_i(k) < 0$ or $N_i(k) > Z_i(k\Delta t)$, for some i and k . Indeed, in continuous-time, the indicator functions that are present in the definition of Δ_{ij} and Ψ_i are enough to ensure that the demand Δ_{ij} is zero as soon as N_i is zero and that the supply Ψ_i is zero as soon as N_i reaches the capacity (or total population) $Z_i(k\Delta t)$. However, in discrete time, they are not enough to ensure that the total demand over a sampling interval does not exceed $N_i(k)$, nor that the supply over a sampling interval does not exceed the remaining capacity $Z_i(k\Delta t) - N_i(k)$, as we can easily understand by considering the following small example.

Figure 18: A network of two nodes.

Example 1. *Let us consider a network of two nodes depicted in Figure 18. The equations for mobility between node* 1 *and node* 2 *are given by*

$$\begin{aligned} \dot{N}_1 &= \phi_{21} - \phi_{12} \\ \dot{N}_2 &= \phi_{12} - \phi_{21} \end{aligned}$$

Let us consider a simple case where $\phi_{21} = 0$ *and* $\phi_{12} = f_{12}1_{N_1>0}$*. Therefore, we have*

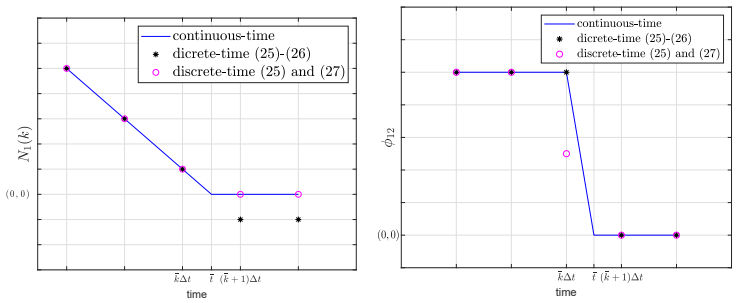
$$\begin{aligned} \dot{N}_1 &= -f_{12}1_{N_1>0} \\ \dot{N}_2 &= f_{12}1_{N_1>0} \end{aligned}$$

This system, starting from some $N_1(t) > 0$*, will have* N_1 *decreasing up to some time* $\bar{t}$ *where* $N_1(\bar{t}) = 0$*, and then remaining equal to zero, as depicted in Figure 19a. We consider the non-trivial case where* $\bar{t}$ *is not a multiple of the sampling time* Δt*, i.e.,* $\bar{k}\Delta t < \bar{t} < (\bar{k}+1)\Delta t$ *for some* $\bar{k}$*. Now, since* $N_1(\bar{k}\Delta t) > 0$ *the discrete-time flow computed naively as in* (26) *will be* $\phi_{12}(\bar{k}) = \Delta t f_{12}$*, which leads to* $N(\bar{k}+1) < 0$ *as depicted in*

Figure 19a using starred black dots. Therefore, we need to introduce a correction. We saturate the flow so as to ensure that during a sampling time, we do not extract more flow than the available number of people. Hence, we set

$$\phi_{12}(k) = \min\left(f_{12}, \frac{N_1(k)}{\Delta t}\right) \tag{27}$$

so that $\Delta t\phi_{12}(k) \leq N_1(k)$*. Therefore, in this example, we get* $\phi_{12}(\bar{k}) = N_1(\bar{k})$*, and hence* $N_1(\bar{k}+1) = 0$ *as depicted in Figure 19a using red circles. This correction is on the flow and not on the* N_1 *itself, which ensures that the total mass in the system is preserved, i.e.,* $N_1 + N_2$ *remains constant. Figure 19b depicts the different flows chosen according to flows defined in* (26) *and* (27)*. It can be seen in Figure 19b that the corrected flow given by* (27) *adjusts the value of* ϕ_{12} *according to the available number of people.*



(a) (b)

Figure 19: (a) Evolution of number of people at node 1 in Example 1 and (b) The outflows from 1 to 2 in case of continuous-time model, depicted by blue line, and in discrete time when the flows are defined naively by using (26), depicted by black stars, versus when the flow is defined as proposed in the paper depicted by red circles.

It can also be seen that in a complex network, no matter how small the step size be, this phenomenon can occur. Therefore, following the intuition within the toy example, we define suitable saturated flows, $\phi_{ij}(k)$ that are equal to the continuous-time flows whenever possible but ensures that the resulting model has the following properties

$$\Delta t \sum_{j \in \mathcal{N}_i} \phi_{ij}(k) \leq N_i(k)$$
$$\Delta t \sum_{j \in \mathcal{N}_i} \phi_{ji}(k) \leq Z_i(k\Delta t) - N_i(k)$$

for all k and for all i and hence ensuring $0 \leq N_i(k) \leq Z_i(k\Delta t)$ for all k.
The $\phi_{ij}(k)$ can be defined as

$$\phi_{ij}(k) = \min(\Delta_{ij}(k), \Psi_j(k)), \tag{28}$$

where, $\Delta_{ij}(k)$ is defined as

$$\Delta_{ij}(k) = \min\left(\delta_{ij}(k\Delta t) f_{ij}(k\Delta t), \frac{N_i(k)\alpha_{ij}(k)}{\Delta t}\right), \tag{29}$$

where, $\delta_{ij}(k\Delta t)$ is the Demand Gating Profile(DGP) and $f_{ij}(k\Delta t)$is defined as in (6) and the $\psi_j(k\Delta t)$is the Supply Gating Functions defined as in (9). Here, $\alpha_{ij}(k)$ in the second term is the proportion in which the number of people will move to its neighbors. It can be computed as

$$\alpha_{ij}(k) = \frac{\delta_{ij}(k\Delta t)f_{ij}(k\Delta t)\psi_j(k\delta t)}{\sum_{q\in\mathcal{N}_i}\delta_{iq}(k\Delta t)f_{iq}(k\Delta t)\psi_q(k\Delta t)}$$

such that

$$\sum_{j\in\mathcal{N}_i}\alpha_{ij}(k) = 1.$$

The supply of j with respect to i $\Psi_{ij}(k)$ is defined as

$$\Psi_{ij}(k) = \psi_j(k\Delta t)\frac{(Z_j(k\Delta t) - N_j(k))\mu_{ij}(k)}{\Delta t} \tag{30}$$

Here, $\psi_j(k)$ is defined as in (9) and $\mu_{ij}(k)$ is the proportion in which the remaining capacity of the the location j will be divided among its neighbors defined as

$$\mu_{ij}(k) = \frac{\delta_{ij}(k\Delta t)f_{ij}}{\sum_{q\in\mathcal{N}_j}\delta_{qj}(k\Delta t)f_{qj}} \tag{31}$$

such that

$$\sum_{i\in\mathcal{N}_j}\mu_{ij}(k) = 1.$$

Notice that the second term in (29) ensures that if the number of people N_i at location i is less than the sum of outflows from i, then this sum is taken equal to the number of people available at i. If i has more than one neighbor then, as a natural choice, α_{ij} divides N_i among its neighbors in proportion to what they might have received proportional to their demands. Similarly, if at k, the supply of j is less than the demand of i and number of people at i, then the flow ϕ_{ij} will be equal to the remaining capacity at j divided proportionally by μ_{ij} among its neighbors what they might have received proportional to their demands. The discrete-time model (25) with the flows as in (28)-(31) is same as the continuous time model (2) whenever possible and ensures that the number of people at any location i, $N_i(k)$ is non-negative and remains bounded by $Z_i(k\Delta t)$. Now, we propose the following result to support this claim.

Proposition 2. *Consider the discrete-time system* (25),(28)-(31)*,with the initial condition* $0 \le N_i(0) \le Z_i(0)$ *for all* i*. Then, for all time* k*, we have* $0 \le N_i(k) \le Z_i(k\Delta t)$ *for all* i*.*

Proof. We prove it recursively that is we show that if $0 \le N_i(k) \le Z_i(k\Delta t)$, then $0 \le N_i(k+1) \le Z_i((k+1)\Delta t)$.
First, notice that from (28)-(31), $0 \le N_i(k) \le Z_i(k\Delta t)$, ensures $\phi_{ij}(k) \ge 0, \forall i, j$ since it can be seen from (29) that $\delta_{ij}(k\Delta t)f_{ij}(k\Delta t) \ge 0$ and $\frac{N_i(k)\alpha_{ij}(k)}{\Delta t} \ge 0$.

Now, from the equations (28)-(30), we have that

$$\phi_{ij}(k) \leq \frac{Z_j(k\Delta t) - N_j(k)}{\Delta t}\mu_{ij},$$
$$\phi_{ij}(k) \leq \frac{N_i(k)\alpha_{ij}(k)}{\Delta t}.$$

Therefore,

$$\Delta t \sum_{i\in\mathcal{N}_j} \phi_{ij}(k) \leq \sum_{i\in\mathcal{N}_j} \mu_{ij}(Z_j(k\Delta t) - N_j(k)),$$
$$\Delta t \sum_{j\in\mathcal{N}_i} \phi_{ij}(k) \leq \sum_{j\in\mathcal{N}_i} \alpha_{ij}(N_i(k)).$$

Since, $\sum_{i\in\mathcal{N}_j} \mu_{ij}(k) = 1$, and $\sum_{j\in\mathcal{N}_i} \alpha_{ij}(k) = 1$, we have

$$\Delta t \sum_{i\in\mathcal{N}_j} \phi_{ij}(k) + N_j(k) \leq Z_j(k\Delta t), \tag{32a}$$
$$\Delta t \sum_{j\in\mathcal{N}_i} \phi_{ij}(k) \leq N_i(k). \tag{32b}$$

Now from (25), we have the discrete model with the corrected

$$N_j(k+1) = N_j(k) + \Delta t \sum_{i\in\mathcal{N}_j} \phi_{ij}(k) - \Delta t \sum_{i\in\mathcal{N}_j} \phi_{ji}(k)). \tag{33}$$

From (33),(32a) and since $\phi_{ji}(k) \geq 0$, we have

$$N_j(k+1) \leq Z_j((k+1)\Delta t), \forall j.$$

Moreover, from (33), (32b), and since $\phi_{ji}(k) \geq 0$, we have that

$$N_i(k+1) \geq 0, \forall i.$$

Therefore, we have that for all $i \in \mathcal{V}_o \cup \mathcal{V}_d$

$$0 \leq N_i(k+1) \leq Z_i((k+1)\Delta t).$$

□

As an illustration, Figure 20 shows evolution of number of people at each origin and destination in the large scale network of Grenoble for a week(Monday-Sunday). Here, we consider the discretized system (25),(28)-(31). The parameters are those of Grenoble's large-scale network described in section 3-5 and the destinations are operating at the full capacity i.e. in (1), $u_h(t) = 1$, during the whole week for all destination categories h, and the time scheduling parameters are as depicted in Table 6.

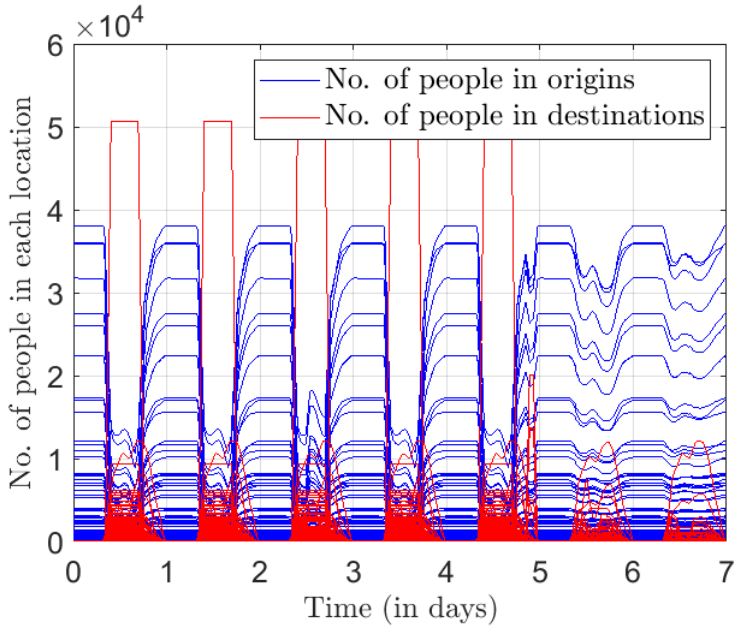


Figure 20: The evolution of number of people at different locations in large-scale network of Grenoble in a week. The blue curves correspond to the number of people in origins and the red curves correspond to the number of people in destinations.

7. Study case: using mobility model to study epidemic spread

One of the advantages of the model (2) is that it is modular. This mobility model can be used for many applications such as urban planning, studying epidemic spread etc. In the next section, we analyse a model, which incorporates mobility and an epidemic spread model, to implement it to the large scale mobility network of Grenoble.

7.1. Connection with Epidemic spread- SIR model

The mobility model (2) can be used to see the effect of mobility control on the epidemics spread. In [27], the mobility model has been incorporated with SIR epidemic model. The resulting model is called the SIR-mobility model. Here, S, I and R stands for the number of susceptible, infected and recovered people. At each location, the disease spread follows the classical SIR model, but people are also allowed to move from one location to another (irrespective of their status of infection) and such mobility happens according to the mobility model (2)-(9). This has been depicted in Figure 21.

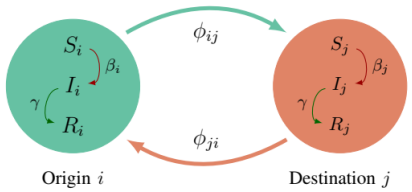


Figure 21: Mobility along the edges and the epidemic spread inside the locations [27].

Let us denote $S_i(t), I_i(t), R_i(t)$ to be the susceptible, infected and removed population at the location i at time t. Further, as defined in (2), let $N_i(t)$ be the number of people at

the location i at time time t and the flow $\phi_{ij}(t)$, the number of people moving from the location i to the location j at time t and $\mathcal{N}_i$ to be the neighbours of i, then the resulting SIR-mobility model proposed in [27] in continuous-time can be written as

$$\begin{aligned}
\dot{S}_i(t) =& -\beta_i(t)\frac{S_i(t)I_i(t)}{N_i(t)} + \sum_{j\in\mathcal{N}_i}\left(\phi_{ji}(t)\frac{S_j(t)}{N_j(t)} - \phi_{ij}(t)\frac{S_i(t)}{N_i(t)}\right)\\
\dot{I}_i(t) =& \beta_i(t)\frac{S_i(t)I_i(t)}{N_i(t)} - \gamma I_i(t) + \sum_{j\in\mathcal{N}_i(t)}\left(\phi_{ji}(t)\frac{I_j(t)}{N_j(t)} - \phi_{ij(t)}\frac{I_i(t)}{N_i(t)}\right)\\
\dot{R}_i(t) =& \gamma I_i(t) + \sum_{j\in\mathcal{N}_i}\left(\phi_{ji}(t)\frac{R_j(t)}{N_j(t)} - \phi_{ij}(t)\frac{R_i(t)}{N_i(t)}\right)
\end{aligned} \tag{34}$$

where the parameters $\beta_i(t) = \bar{\beta}_i\frac{N_i(t)}{Z_i(t)}$, where $\beta_i, \bar{\beta}_i$ and γ_i are the infection rate, nominal infection rate and the recovery rate at i respectively as defined in [27]

7.1.1. Discretization of SIR-mobility model

In order to simulate (34), we need to discretize (34). However, the implementation of this model also poses the issue that the number of susceptible, infected and removed people become negative for some time steps. Moreover, losing non-negativity might lead to losing boundedness of the variables, which is ensured in (34) and (2)-(9) where we have $N_i(t) \geq 0$ and $S_i(t) + I_i(t) + R_i(t) = N_i(t)$.First of all, we should use the discretized mobility model that ensures physically meaningful flows, as in Section 6.1. This is a key ingredient to ensure non-negativity in SIR-mobility, but is not enough, we also need a suitable discretization of the SIR dynamics. This can be done using the technique from the paper [73]. Inspired from [73], we propose a discretization of (34) which ensures that all the variables remain non-negative and the total mass is preserved. It can be called non-local as it differs from the forward Euler discretization by taking some variables at time $k+1$ on the right-hand side instead of taking them all at the local time k.

Let Δt be the time step size then the continuous-time SIR-mobility model (34) can be discretized as follows.

$$\frac{S_i(k+1) - S_i(k)}{\Delta t} = -\beta_i(k)S_i(k+1)\frac{I_i(k)}{N_i(k)} + \sum_j\left(\phi_{ji}(k)\frac{S_j(k)}{N_j(k)} - \phi_{ij}(k)\frac{S_i(k)}{N_i(k)}\right) \tag{35a}$$

$$\frac{I_i(k+1) - I_i(k)}{\Delta t} = \beta_i(k)S_i(k+1)\frac{I_i(k)}{N_i(k)} - \gamma I_i(k+1) + \sum_j\left(\phi_{ji}(k)\frac{I_j(k)}{N_j(k)} - \phi_{ij}(k)\frac{I_i(k)}{N_i(k)}\right) \tag{35b}$$

$$\frac{R_i(k+1) - R_i(k)}{\Delta t} = I_i(k+1) + \sum_j\left(\phi_{ji}(k)\frac{R_j(k)}{N_j(k)} - \phi_{ij}(k)\frac{R_i(k)}{N_i(k)}\right). \tag{35c}$$

Rearranging the terms, one can obtain the following equations that can be implemented

to actually compute $S_i(k+1)$, $I_i(k+1)$ and $R_i(k+1)$.

$$S_i(k+1) = \frac{1}{1+\Delta t\beta_i(k)\frac{I_i(k)}{N_i(k)}}\left[S_i(k) + \Delta t\sum_j\left(\phi_{ji}(k)\frac{S_j(k)}{N_j(k)} - \phi_{ij}(k)\frac{S_i(k)}{N_i}(k)\right)\right] \quad (36a)$$

$$I_i(k+1) = \frac{1}{1+\Delta t\gamma}\left[I_i(k) + \Delta t\beta_i(k)S_i(k+1)\frac{I_i(k)}{N_i(k)} + \Delta t\sum_j\left(\phi_{ji}(k)\frac{I_j(k)}{N_j(k)} - \phi_{ij}(k)\frac{I_i(k)}{N_i(k)}\right)\right] \quad (36b)$$

$$R_i(k+1) = R_i(k) + \Delta t\gamma I_i(k+1) + \Delta t\sum_j\left(\phi_{ji}(k)\frac{R_j(k)}{N_j(k)} - \phi_{ij}(k)\frac{R_i(k)}{N_i(k)}\right). \quad (36c)$$

It is well-known that in standard SIR model, the sum of susceptible, infected and recovered people remains constant, and equal to the total population. Here, in each location, the number of people is not constant because of the mobility, but we can still show that the sum $S_i(k) + I_i(k) + R_i(k)$ is equal to $N_i(k)$ for all i and k, as shown in the following proposition.

Proposition 3. *Given the discrete model* (36), (25)-(31) *with the initial condition* $S_i(0)+I_i(0)+R_i(0) = N_i(0)$, *then for all* i *and* k, *we have* $S_i(k) + I_i(k) + R_i(k) = N_i(k)$.

Proof. We prove it recursively, i.e., we prove that $S_i(k) + I_i(k) + R_i(k) = N_i(k)$, implies $S_i(k+1) + I_i(k+1) + R_i(k+1) = N_i(k+1)$.
Recall that system (36) is equivalent to system (35); yet another equivalent re-writing is the following:

$$S_i(k+1) = S_i(k) - \Delta t\beta_i(k)S_i(k+1)\frac{I_i(k)}{N_i(k)} + \Delta t\sum_j\left(\phi_{ji}\frac{S_j(k)}{N_j(k)} - \phi_{ij}(k)\frac{S_i(k)}{N_i}(k)\right)$$

$$I_i(k+1) = I_i(k) + \Delta t\beta_i(k)S_i(k+1)\frac{I_i(k)}{N_i(k)} - \Delta t\gamma I_i(k+1) + \Delta t\sum_j\left(\phi_{ji}(k)\frac{I_j(k)}{N_j(k)} - \phi_{ij}(k)\frac{I_i(k)}{N_i}(k)\right)$$

$$R_i(k+1) = R_i(k) + \Delta t\gamma I_i(k+1) + \Delta t\sum_j\left(\phi_{ji}(k)\frac{R_j(k)}{N_j(k)} - \phi_{ij}(k)\frac{R_i(k)}{N_i(k)}\right).$$

We can now sum the three equations above. Then, we use the fact that $S_i(k) + I_i(k) + R_i(k) = N_i(k)$ and $S_j(k)+I_j(k)+R_j(k) = N_j(k)$, and hence also $\frac{S_i(k)+I_i(k)+R_i(k)}{N_i(k)} = 1$ and $\frac{S_j(k)+I_j(k)+R_j(k)}{N_j(k)} = 1$. With this, we obtain

$$\begin{aligned} S_i(k+1) + I_i(k+1) + R_i(k+1) &= N_i(k) + \Delta t\sum_{j\in\mathcal{N}_i}(\phi_{ji}(k) - \phi_{ij}(k)) \\ &= N_i(k+1). \end{aligned}$$

where the last equality follows from (25).

□

Now, we show that if the initial number of susceptible, infected and recovered people is non-negative then they remain non-negative at any time-step k.

Proposition 4. *Given the discrete model* (36), (25)-(31) *with the initial condition* $S_i(0) \geq 0, I_i(0) \geq 0, R_i(0) \geq 0, N_i(0) \geq 0$, *then for all* i, k *we have* $S_i(k) \geq 0$, $I_i(k) \geq 0$ *and* $R_i(k) \geq 0$.

Proof. First, recall that by proposition 2, $N_i(0) \geq 0$ implies $N_i(k) \geq 0$, for all k. Moreover, recall that the definition of flows in (25)-(31) together with $N_i \geq 0$ ensures (32a)-(32b). With this, we are ready to show recursively, the non-negativity of $S_i(k), I_i(k)$ and $R_i(k)$, namely we will show that $S_i(k) \geq 0$ implies $S_i(k+1) \geq 0$, and then similarly show that $I_i(k) \geq 0$ implies $I_i(k+1) \geq 0$ and $R_i(k) \geq 0$ implies $R_i(k+1) \geq 0$. From (36a), we have

$$\begin{aligned} S_i(k+1) &= \frac{1}{1+\Delta t\beta_i(k)\frac{I_i(k)}{N_i(k)}}\left[S_i(k) + \Delta t\sum_j\left(\phi_{ji}(k)\frac{S_j(k)}{N_j(k)} - \phi_{ij}(k)\frac{S_i(k)}{N_i(k)}\right)\right] \\ &= \frac{1}{1+\Delta t\beta_i(k)\frac{I_i(k)}{N_i(k)}}\left[S_i(k) - \Delta t\sum_j\phi_{ij}(k)\frac{S_i(k)}{N_i(k)} + \Delta t\sum_j\phi_{ji}(k)\frac{S_j(k)}{N_j(k)}\right] \\ &= \frac{1}{1+\Delta t\beta_i(k)\frac{I_i(k)}{N_i(k)}}\left[\frac{S_i(k)}{N_i(k)}\left(N_i(k) - \Delta t\sum_j\phi_{ij}(k)\right) + \Delta t\sum_j\phi_{ji}(k)\frac{S_j(k)}{N_j(k)}\right] \end{aligned}$$

from (32b), we have $N_i(k) - \Delta t\sum_j\phi_{ij}(k) \geq 0$, and since $S_j(k) \geq 0, N_j(k) \geq 0$ and $\phi_{ji}(k) \geq 0$ for all j, therefore,

$$S_i(k+1) \geq 0.$$

Similarly, from (32b), we have $N_i(k) - \Delta t\sum_j\phi_{ij}(k) \geq 0$, which also ensures that $I_i(k) - \Delta t\sum_j\phi_{ij}(k)\frac{I_i(k)}{N_i(k)} \geq 0$, and $R_i(k) - \Delta t\sum_j\phi_{ij}(k)\frac{R_i(k)}{N_i(k)} \geq 0$. Therefore, it can be seen from (36b) and (36c) that $I_i(k+1) \geq 0$ and $R_i(k+1) \geq 0$ respectively. □

It can be seen that the proposition 3 and the proposition 4 together ensure that S_i, I_i, and R_i remain bounded, since $0 \leq S_i(k) \leq N_i(k)$, $0 \leq I_i(k) \leq N_i(k)$ and $0 \leq R_i(k) \leq N_i(k)$. Notice that, in the right hand side of (35), the terms $S_i(k+1)$ and $I_i(k+1)$ outside the parenthesis have been taken at $k+1$ instead of k and in proposition 4, we show that with this discretization, we have non-negative $S_i(k), I_i(k)$ and $R_i(k)$. Recall that in the SIR dynamics, people are transferred from one compartment to another depending on their infection status. In the discretization (35), it is ensured that this transfer is reflected by diminishing the number in one compartment by multiplication with a fraction as can be seen in (36). On the other hand, in Euler discretization of SIR-mobility model, this transfer between compartments is done by using a subtraction term as for instance can be seen in (37) for the variable S.

$$S_i(k+1) = S_i(k) - \Delta t\beta_i\frac{S_i(k)I_i(k)}{N_i(k)} + \Delta t\sum_{j\in\mathcal{N}_i}\left(\phi_{ji}(k)\frac{S_j(k)}{N_j(k)} - \phi_{ij}(k)\frac{S_i(k)}{N_i(k)}\right) \tag{37}$$

In (37), the term $-\Delta t\beta_i\frac{S_i(k)I_i(k)}{N_i(k)}$ along with the term $-\Delta t\sum_{j\in\mathcal{N}_i}\left(\phi_{ij}(k)\frac{S_i(k)}{N_i(k)}\right)$ can sometime lead to negative S depending on the time-step size Δt. Instead, in the proposed discretization, the evolution of S is given by (36a), where we see that the transfer of people out of this compartment is reflected by diminishing the number through multiplication by a fraction. Therefore, ensuring the non-negativity of S as the non-negativity

of the terms in the the square-bracket of (36a) is ensured because of the flow defined in the section (6.1).

8. Description of the GTL-covid demonstrator

With the implementation of the mobility model (25) and (36) on the large scale network of Grenoble, development of an online web interface, GTL-covid, is under process, in our research group, in the context of ERC-ScaleFreeback project. It can be accessed at `http://gtlville.inrialpes.fr/covid-19`. This interface could serve as an academic platform where an user can choose between different scenarios of restrictions on mobility and epidemic parameters to see the mobility and epidemic evolution through different indicators such as mobility in each category, total infected, ICU occupancy etc. shown through various heatmaps and plots. To display the maps, the interface uses Mapbox which is an open source map tool. The default page of the interface has been depicted in Figure 22. In the right hand side of the interface, users can visualise the given network

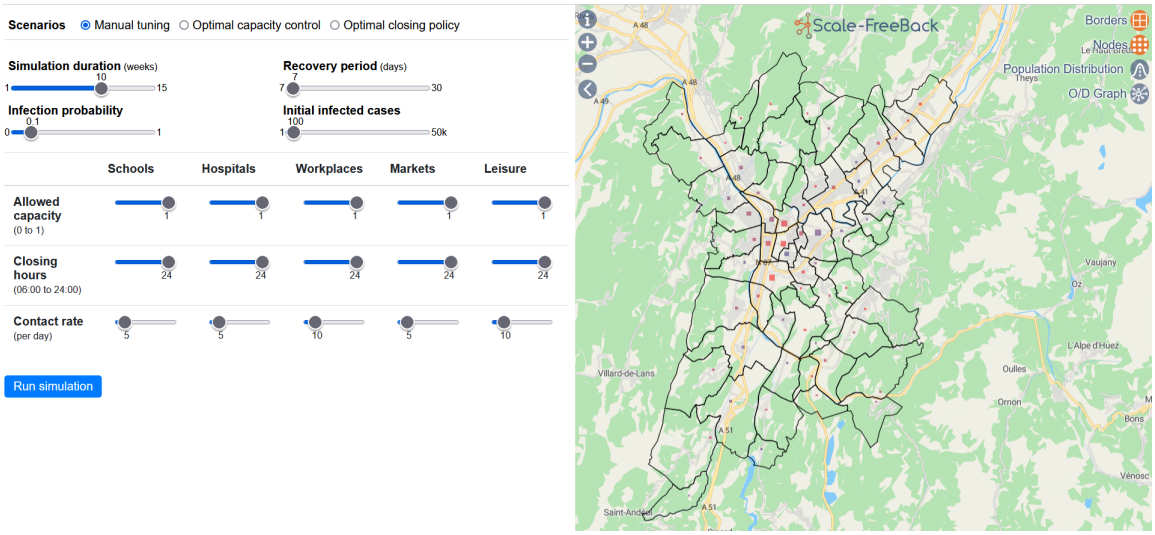


Figure 22: User interface for the GTL-covid demonstartor.

data from the menus available on the top right corner. The menu options are as follows:

- Borders : It can be used to switch the borders of the regions on and off.
- Nodes : It pops a side menu bar to view the origin and destination nodes per category or subcategory in the map. Moreover, when hovered over the nodes, it pops up information such as name of the location, population or capacity of the node. See Figure 23. There is a button called 'aggregate' in the side menu which aggregates the nodes as discussed in the section 3.2.8 and the subcategories for which the nodes are aggregated appear in italics.
- Population distribution : It also pops a side menu bar with option to choose from origin and destination subcategories. It shows the population distribution with a heatmap in when origin is selected and distribution of capacities when a particular subcategory is selected. See Figure 24a.

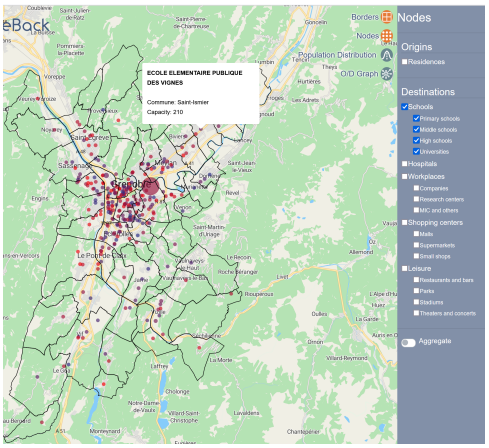
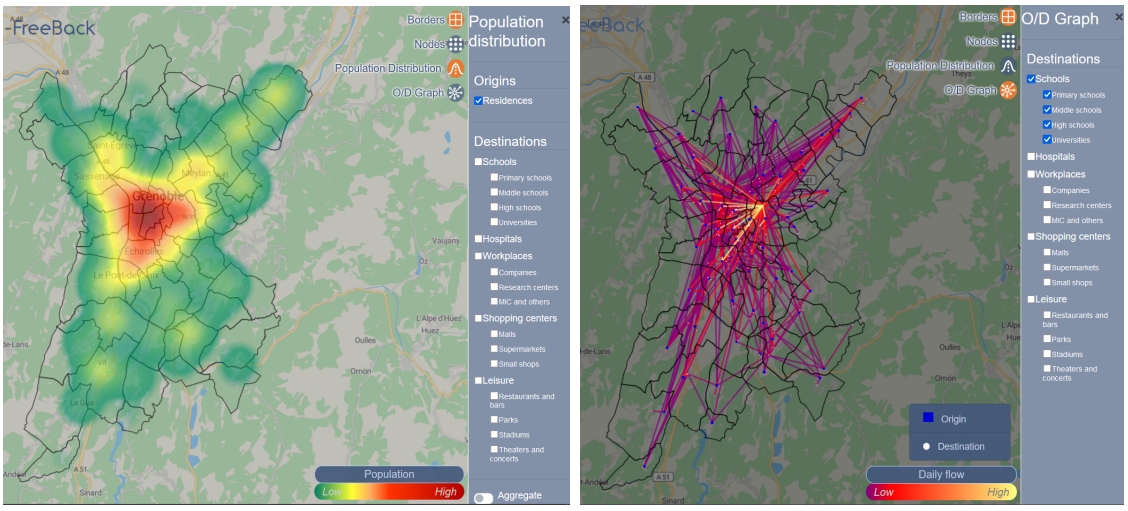


Figure 23: Nodes menu bar in the demonstrator. Here, the nodes corresponding to the school category has been selected and an instance of information about a primary school node appears when the cursor is hovered over it.

- O/D Graph : It also pops a side menu bar with option to choose from different destination categories. On selecting a subcategory, it shows the connections between different origins and the aggregated nodes of the destinations as discussed in Section 3.3. In Figure 24b, the connections between different origins and nodes of the school subcategory can be seen.



(a) (b)

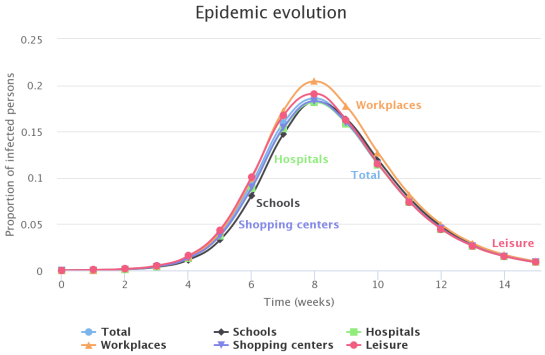
Figure 24: (a) Plot corresponding to the population distribution menu. Here, we select the origin, so it is showing the distribution of the population in the region. (b) Mobility network showing connection between schools and the origins. The origins are denoted in blue color and the destination in white color. The color of the connecting edge ranges from magenta to yellow depending on the weight of the edge which is the average daily flow M_{ij} between the origin-destination pair $(i, j) \in \mathcal{E}$.

Left hand side of the interface concerns the SIR-mobility model. The user can choose some epidemics parameters, so as to closely adapt the simulation to current data about the disease, or to compare pessimistic and optimistic scenarios. The user can also choose coefficients that correspond to mobility restrictions to study its on the epidemics spread.

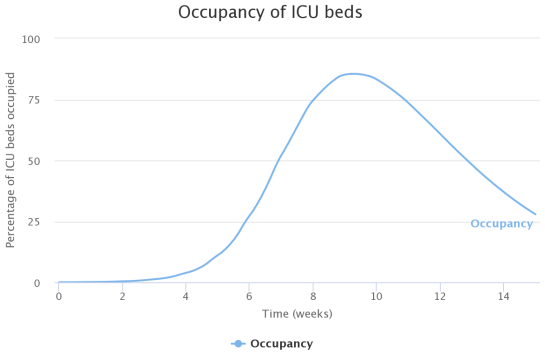
The simulation uses the model (25)-(31) to show the mobility evolution and (36) to show the evolution of epidemics. The available input options are given in form of sliders with minimum and maximum for each displayed in the left and right side of the slider respectively. The input options available on the interface are as follows:

- Simulation duration : The duration in weeks for which to run the simulation.
- Recovery period : Number of days a person takes to recover once a person gets infected from the disease. The recovery rate γ_i in (36) is computed as the inverse of the recovery period (in hours) of the infected cases.
- Infection probability : The probability by which an infected person would infect a susceptible person. It is used to compute the parameter β_i in (36).
- Initial infected cases : Initial number of infected persons i.e. $\sum_i I_i(0)$. Note that, the initial number of infections needs to be greater than zero otherwise the disease will not spread.
- Allowed capacity : With its value in the interval $[0, 1]$ for each category, it corresponds to the coefficient $u_h(t)$ as described in (1). Here, 0 means that no people can visit the destination during the simulation period and 1 means that the destinations are operating at the full capacity.
- Closing hours : With their value in the range $[6, 24]$ for each destination category, the destinations can be forced to closed before the nominal closing hours of the destinations. It can be used to simulate partial or full closure of the destinations in a category. For an example, if a user chooses the closing hours of school to be 14, then it means he is enforcing all the destinations in school category to close at $14h$ everyday for the whole simulation period.
- Contact rate : It is the average number of contact made by a person in a day in the destination of each category. It is also used to compute β_i in (36) for destinations per category.

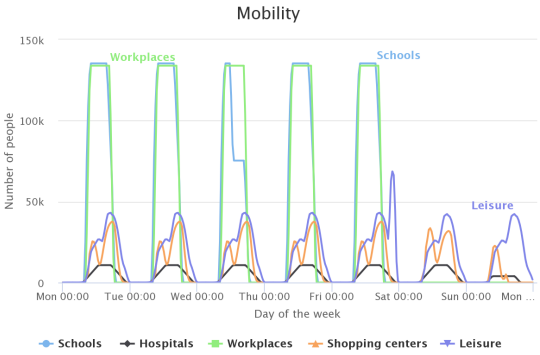
Among these options, the inputs, 'allowed capacity' and 'closing hours' are control inputs to simulate lockdowns and restricted mobility and the other options correspond to the parameters for SIR epidemic model. The button 'Run simulation' in the left hand side panel is used to run the simulation with the chosen parameters and inputs. When the simulation is run, on the left hand side of the interface, the user can see the plots for the weekly average of proportion of infected people, ICU bed occupancy and mobility of people. The plots for mobility and infected people are shown for each category and the ICU bed occupancy for the whole region. The ICU bed occupancy and the proportion of the infected people in each category is shown for the whole simulation period but the mobility patterns are shown for a week. When the simulation is run, the left hand side shows various option to show heatmaps. In the bottom right corner of the map, we have different options for heatmaps namely mobility, infected relative and infected absolute. The mobility heatmap shows the evolution of mobility, infected relative represents the fraction of infected people at each location, and the infected absolute represents the



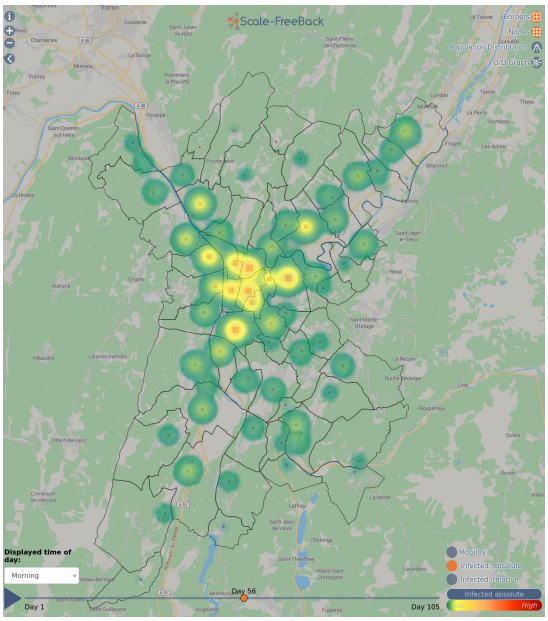
(a)



(b)



(c) Mobility of people in the region for one week.



(d)

Figure 25: Scenario 1: All destinations open. The evolution of epidemics and mobility in the region as shown on the interface for health-mobility.The scenario considers that the mobility is happening without any restrictions. (a) 7-day moving average of the proportion of the infected population in each category of destinations the case when there is no restriction on the destinations and the mobility. Every destination is running with the nominal parameters. (b) 7-day moving average of the ICU occupancy in the region in the case when there is no restriction on mobility. (c) Mobility of people per category of destinations for one week. (d) Heatmap showing the evolution of epidemic in the region. Here, the user has chosen the 'infected absolute' option.

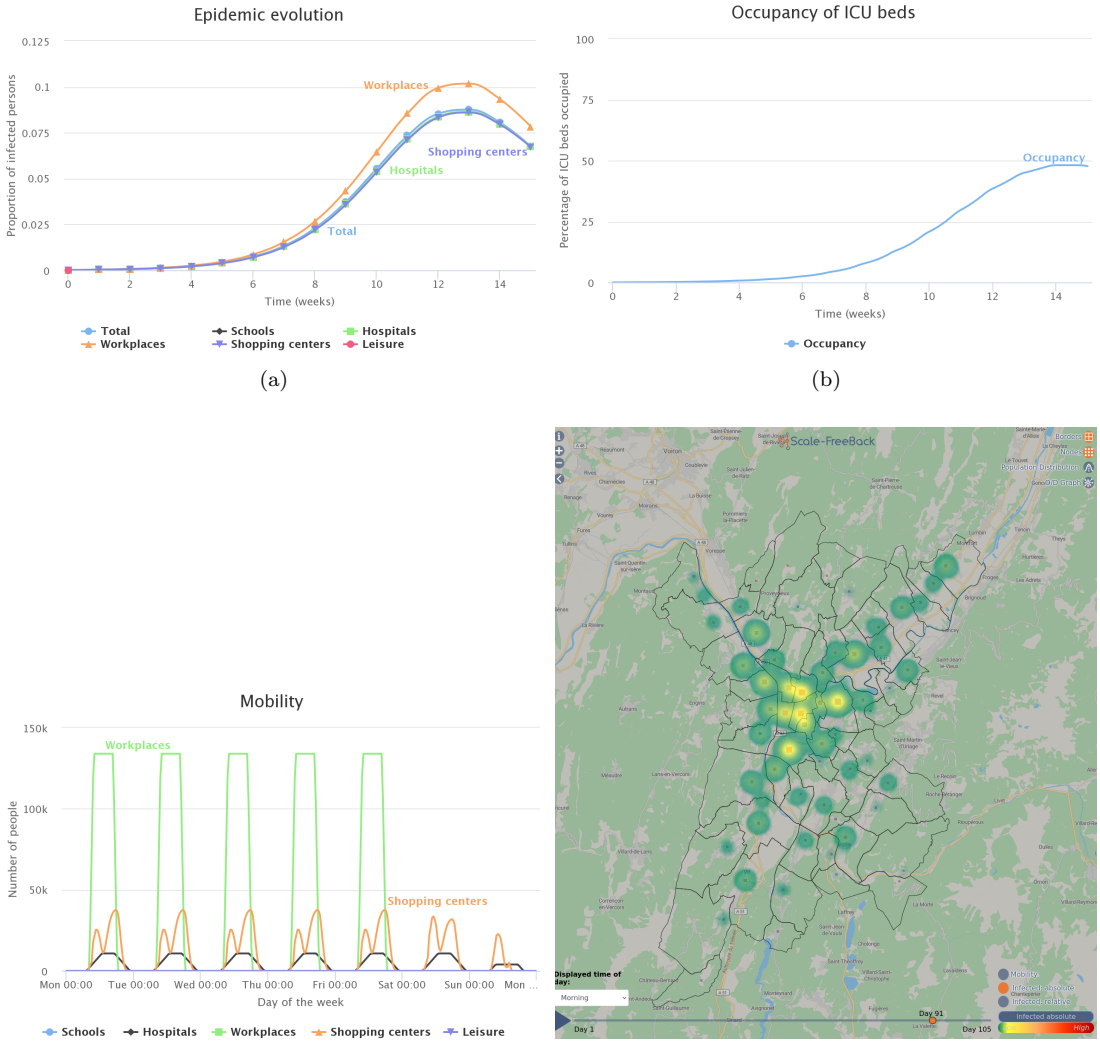


Figure 26: Scenario 2: schools and leisure closed. The evolution of epidemics and mobility in the region as shown on the interface for health-mobility for the scenario where schools and leisure destinations are closed and the other destinations are running on nominal conditions. (a) 7-day moving average of the proportion of the infected population in each category of destinations which are open. Every destination which is open is running with the nominal parameters. (b) 7-day average of the ICU occupancy in the region in the case when there is no restriction on mobility. (c) Mobility of people in the destinations which are open for one week. (d) Heatmap showing the evolution of epidemic in the region. here, the user has chosen the 'infected absolute' option.

infected population at each location. In the bottom of the heatmap, we have the time bar which shows the corresponds to the the time of the simulation corresponding to the heatmap. The user can choose to select a time of the day to see these results from an option at the bottom left corner of the heatmap. Now, we show some scenarios of mobility restrictions and infection parameters to show the results.

8.1. Results

In this section, we simulate two different scenarios of mobility restrictions on the interface to visualize the effect of restricted mobility on epidemic spread.

8.1.1. Scenario 1: All destinations open

Here, we simulate the scenario in which the destinations are running on full capacity with nominal closing hours. Here, the simulation is done for 15 weeks, recovery period is 8 days, infection probability is 0.2 and there are 100 initial infected cases. The contact rate is kept to be $\{5, 5, 10, 5, 10\}$ respectively for schools, hospitals, workplaces, markets and leisure respectively. This is the default scenario on the GTL-covid demonstrator. We can see the results in Figure 25 which are the screenshots taken from the demonstrator.

8.1.2. Scenario 2: schools and leisure closed

Here, we simulate the scenario in which the schools and the leisure destinations are closed and the other destinations are running on full capacity with nominal closing hours. Here, the simulation is done for 15 weeks, recovery period is 8 days, infection probability is 0.2 and there are 100 initial infected cases. The contact rate is kept to be $\{5, 5, 10, 5, 10\}$ respectively for schools, hospitals, workpleaces, markets and leisure respectively. We can see the results in Figure 26

We can see that because of the closure of destination in school and leisure, the epidemic evolves very slowly and the peak is much lower in this case.

9. Conclusion

In this work, we studied the mobility of people in the metropolitan city of Grenoble by adapting and extending a supply-demand based mobility model which captures the daily movement of people between residences and places of interests called destinations, using time schedules and gating profiles which also accommodate the possibility of imposing restrictions on mobility. We built the large-scale mobility network in the metropolitan city of Grenoble by using publicly available information to answer the questions: where do the people move? when do they move? and how many people move? We identify the origins and destinations with their population and capacities respectively. We also computed the O/D relation between these origins and destinations. We proposed methods to compute the time schedules and gating functions for the destinations which captures the mobility patterns.

The mobility model is given in continuous-time in which the flows are defined such that the model has nice properties of non-negativity, boundedness and mass conservation. For its implementation to the large-scale network, the model needs to be discretised. Therefore, We analysed the problems encountered in the discretization of this model and proposed

to redefine the flows in discrete-time which preserves the properties of non-negativity, boundedness and mass conservation.

As an application, we studied the effect of different restrictions on mobility on the epidemic spread in Grenoble by using a SIR-mobility model. Again, we analysed the problems arising due to discretization of the SIR-mobility model and proposed a discretization which along with the proposed redefined flows in the mobility ensures that the SIR-mobility model gives non-negative values and the total mass is preserved.

We gave a brief description of the GTL-covid web interface which is an online academic platform where a user can simulate and visualise different scenarios of mobility and epidemics.

Acknowledgements

This work is supported by European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC-AdG no. 694209, Scale-FreeBack, website: `http://scale-freeback.eu`). The authors would like to thank M. U. B. Niazi and Alain Kibangou for interesting discussions and useful suggestions, and Léo Sénique for his work on implementation of the GTL-covid web interface.

Appendix A

Notation	Description
$\mathcal{G} = (\mathcal{V}_o, \mathcal{V}_d, \mathcal{E})$	network
$\mathcal{V}_o$	set of origin nodes
$\mathcal{V}_d$	set of destination nodes
$\mathcal{E}$	set of Edges in the network
$\mathcal{N}_i$	neighbors of i
$\mathcal{C}_j$	categories j of the destinations
$\mathcal{D}_k^j$	subcategory k in the category j
P_i	population of the origin i
C_i	capacity of the destination i
Z_i	population of i if it is an origin or capacity of i it is a destination
$u_h(t)$	capacity control coefficient for the category $\mathcal{C}_h$
C_j^o	operating capacity of destination j
$N_i(t)$	number of people at location i at time t
$\phi_{ij}(t)$	flow from i to j at time t
$\Delta_{ij}(t)$	demand of i with respect to j at time t
$\delta_{ij}(t)$	Demand Gating Profile (DGP) for mobility from i to j
a_j	opening time of j
b_j	closing time of j
b_j^e	effective closing time of j
s_j	average time spent by a person at j
t_{ij}	time at which mobility starts from i to j
τ_{ij}	size of the mobility window from i to j
$\Psi_j(t)$	supply of j at time t
ψ_j	Supply Gating Functions
M_{ij}	average daily number of people who would like to visit j from i daily
F_j	maximum inflow to j
d_{ij}	road distance between i and j
σ_j	maximum distance traveled daily by most people to go to j from an origin
A_{ij}	normalised attraction between i and j
η_j	threshold on attraction between i and j
P_{ic}	population of origin i which is eligible to go to destination subcategory $\mathcal{D}_c^a$
f_c^p	fraction of the P_{ic} which, in general, goes to $\mathcal{D}_c^a$ every day
$\mathbf{1}$	vector of all ones
Δt	time step size
$S_i(t)$	number of susceptible people at i at time t
$I_i(t)$	number of infected people at i at time t
$R_i(t)$	number of removed people at i at time t

References

[1] Z. Huang, X. Ling, P. Wang, F. Zhang, Y. Mao, T. Lin, F. Y. Wang, Modeling real-time human mobility based on mobile phone and transportation data fusion,

Transportation Research Part C: Emerging Technologies 96 (2018) 251–269. doi:https://doi.org/10.1016/j.trc.2018.09.016.

[2] J. R. Roy, J.-C. Thill, Spatial interaction modelling, Springer Berlin Heidelberg, Berlin, Heidelberg, 2004. doi:10.1007/978-3-662-07223-3_15.

[3] L. Sattenspiel, K. Dietz, A structured epidemic model incorporating geographic mobility among regions, Mathematical Biosciences 128 (1) (1995) 71–91. doi:https://doi.org/10.1016/0025-5564(94)00068-B.

[4] F. Simini, M. González, A. Maritan, A.-L. Barabasi, A universal model for mobility and migration patterns, nature 484 (2012) 96–100. doi:https://doi-org.gaelnomade-1.grenet.fr/10.1038/nature10856.

[5] S. Erlander, N. Stewart, The Gravity Model in Transportation Analysis: Theory and Extensions, VSP, 1990.

[6] K. Gautier, C. Francesco, R. Carlo, B. V. D, Urban gravity: a model for inter-city telecommunication flows, J. Stat. Mech. (2009).

[7] H. Barbosa, M. Barthelemy, G. Ghoshal, C. R. James, M. Lenormand, T. Louail, R. Menezes, J. J. Ramasco, F. Simini, M. Tomasini, Human mobility: Models and applications, Physics Reports 734 (2018) 1–74, human mobility: Models and applications. doi:https://doi.org/10.1016/j.physrep.2018.01.001.

[8] F. Ekman, A. Keränen, J. Karvo, J. Ott, Working day movement model, in: Proceedings of the 1st ACM SIGMOBILE Workshop on Mobility Models, Association for Computing Machinery, New York, NY, USA, 2008, p. 33–40. doi:10.1145/1374688.1374695.

[9] M. Musolesi, C. Mascolo, A community based mobility model for ad hoc network research, in: Proceedings of the 2nd International Workshop on Multi-Hop Ad Hoc Networks: From Theory to Reality, REALMAN '06, Association for Computing Machinery, New York, NY, USA, 2006, p. 31–38. doi:10.1145/1132983.1132990.

[10] L. Sattenspiel, K. Dietz, A structured epidemic model incorporating geographic mobility among regions, Mathematical Biosciences 128 (1) (1995) 71–91. doi:https://doi.org/10.1016/0025-5564(94)00068-B.

[11] P. C, T. M, C. V., Human mobility and time spent at destination: impact on spatial epidemic spreading., J Theor Biol. (2013). doi:doi:10.1016/j.jtbi.2013.08.032.

[12] Z. Wang, S. Y. He, Y. Leung, Applying mobile phone data to travel behaviour research: A literature review, Travel Behaviour and Society 11 (2018) 141–155. doi:https://doi.org/10.1016/j.tbs.2017.02.005.

[13] M. S. Iqbal, C. F. Choudhury, P. Wang, M. C. González, Development of origin–destination matrices using mobile phone call data, Transportation Research Part C: Emerging Technologies 40 (2014) 63–74. doi:https://doi.org/10.1016/j.trc.2014.01.002.

[14] D. Bachir, G. Khodabandelou, V. Gauthier, M. El Yacoubi, J. Puchinger, Inferring dynamic origin-destination flows by transport mode using mobile phone data, Transportation Research Part C: Emerging Technologies 101 (2019) 254–275. doi:https://doi.org/10.1016/j.trc.2019.02.013.

[15] L. Alexander, S. Jiang, M. Murga, M. C. González, Origin–destination trips by purpose and time of day inferred from mobile phone data, Transportation Research Part C: Emerging Technologies 58 (2015) 240–250, big Data in Transportation and Traffic Engineering. doi:https://doi.org/10.1016/j.trc.2015.02.018.

[16] Z. Zhao, S.-L. Shaw, Y. Xu, F. Lu, J. Chen, L. Yin, Understanding the bias of call detail records in human mobility research, International Journal of Geographical Information Science 30 (9) (2016) 1738–1762. doi:10.1080/13658816.2015.1137298.

[17] M. Lin, W.-J. Hsu, Mining gps data for mobility patterns: A survey, Pervasive and Mobile Computing 12 (2014) 1–16. doi:https://doi.org/10.1016/j.pmcj.2013.06.005.

[18] R. Hariharan, K. Toyama, Project lachesis:parsing and modeling location histories, Geographic Information Science (2004).

[19] M. Zignani, S. Gaito, Extracting human mobility patterns from gps-based traces, in: 2010 IFIP Wireless Days, 2010, pp. 1–5. doi:10.1109/WD.2010.5657695.

[20] J. Tang, F. Liu, Y. Wang, H. Wang, Uncovering urban human mobility from large scale taxi gps data, Physica A: Statistical Mechanics and its Applications 438 (2015) 140–153. doi:https://doi.org/10.1016/j.physa.2015.06.032.

[21] S. Dabiri, K. Heaslip, Inferring transportation modes from gps trajectories using a convolutional neural network, Transportation Research Part C: Emerging Technologies 86 (2018) 360–371. doi:https://doi.org/10.1016/j.trc.2017.11.021.

[22] Q. Ge, D. Fukuda, Updating origin–destination matrices with aggregated data of gps traces, Transportation Research Part C: Emerging Technologies 69 (2016) 291–312. doi:https://doi.org/10.1016/j.trc.2016.06.002.

[23] de Montjoye Yves-Alexandre, H. C. A., V. Michel, B. V. D, Unique in the crowd: The privacy bounds of human mobility, Scientific Reports (2013). doi:https://doi.org/10.1038/srep01376.

[24] L. Yin, Q. Wang, S.-L. Shaw, Z. Fang, J. Hu, Y. Tao, W. Wang, Re-identification risk versus data utility for aggregated mobility research using mobile phone location data, PLOS ONE 10 (2015) 1–23. doi:10.1371/journal.pone.0140589.

[25] E. commission, European commision proposal for a regulation of the european parliament and of the council concerning the respect for private life and the protection of personal data in electronic communications and repealing directive 2002/58/ec (regulation on privacy and electronic commu) (2017).

[26] H. Zang, J. Bolot, Anonymization of location data does not work: A large-scale measurement study, in: Proceedings of the 17th Annual International Conference on Mobile Computing and Networking, MobiCom '11, Association for Computing Machinery, New York, NY, USA, 2011, p. 145–156. doi:10.1145/2030613.2030630.

[27] M. U. Niazi, C. C. de Wit, A. Kibangou, P.-A. Bliman, Optimal control of urban human mobility for epidemic mitigation, CDC 2021 - 60th IEEE Conference on Decision and Control, Dec 2021, Austin, United States (2021).

[28] Openstreetmap, Grenoble area on openstreetmap, https://www.openstreetmap.org/#map=11/45.1635/5.7273, accessed 6 December 2021.

[29] INSEE, Insee home page, https://www.insee.fr/en/accueil, accessed 6 December 2021.

[30] Republique-Francaise, Coordinates of the boundaries of communes, https://www.data.gouv.fr/fr/datasets/r/07b7c9a2-d1e2-4da6-9f20-01a7b72d4b12, accessed 6 December 2021.

[31] INSEE, Population of the french communes, https://statistiques-locales.insee.fr/#bbox=598577,5676525,91084,55025&c=indicator&i=pop_depuis_1876.pop&s=2017&view=map1, accessed 6 December 2021.

[32] Opendatasoft, Coordinates of the the iris in france, https://public.opendatasoft.com/explore/dataset/georef-france-iris/export/?disjunctive.reg_name&disjunctive.dep_name&disjunctive.arrdep_name&disjunctive.ze2010_name&disjunctive.bv2012_name&disjunctive.epci_name&disjunctive.ept_name&disjunctive.com_name&disjunctive.com_arm_name&disjunctive.iris_name&sort=year&q=38185&refine.arrdep_name=Grenoble&refine.com_name=Grenoble&location=12,45.19695,5.74791&basemap=jawg.streets, accessed 6 December 2021.

[33] Grenoble-metrople, Official webpage of grenoble metropole, = https://www.grenoble.fr/1459-les-secteurs-de-grenoble.htm, accessed 6 December 2021.

[34] INSEE, Statistical details of french communes, https://statistiques-locales.insee.fr/#bbox=612227,5668725,62775,37924&c=indicator&i=pop_depuis_1876.pop&s=2017&selcodgeo=38045&view=map1, accessed 6 December 2021.

[35] Grenoble-metropole, Sectors of grenoble, https://www.grenoble.fr/1459-les-secteurs-de-grenoble.htm, accessed 6 December 2021.

[36] Population of people of different age groups in different communes, https://statistiques-locales.insee.fr/#bbox=577774,5679420,119058,71925&c=indicator&i=rp.pop_3tr_ages&s=2017&view=map1, accessed 6 December 2021.

[37] Academie-Grenoble, Search engine for list of primary schools in region grenoblois per commune, https://bv.ac-grenoble.fr/carteforpub/ecole, accessed 6 December 2021.

[38] Academie-Grenoble, Search engine for list of primary schools in region grenoblois per commune, https://bv.ac-grenoble.fr/carteforpub/etab, accessed 6 December 2021.

[39] UGA, Official website of university grenoble alpes, https://www.univ-grenoble-alpes.fr/about/uga-in-numbers, accessed 6 December 2021.

[40] CHU, Collective data for chu north, south and voiron, https://www.chu-grenoble.fr/content/chiffres-cles-0, accessed 6 December 2021.

[41] CHU-Voiron, Rules for visitation in chu voiron, https://www.ch-voiron.fr/patients-visiteurs/Visites.html, accessed 15 April 2021.

[42] CHU, Public data for chus in grenoble, https://www.chu-grenoble.fr/content/sites-plans-dacces, accessed 15 April 2021.

[43] Clinique-Mutualiste, Official website of clinique mutualiste, https://www.ghm-grenoble.fr/Presentation.11.0.html, accessed 6 December 2021.

[44] Clinique-Belledonne, Clinique belledonne, https://www.clinique-belledonne.fr/presentation/, accessed 6 December 2021.

[45] Clinique-Cedres, Official website of clinique des cedres, https://www.cliniquedescedres.com/la-clinique, accessed 6 December 2021.

[46] INSEE, Classification of enterprises in france, https://www.insee.fr/fr/statistiques/4277836?sommaire=4318291, accessed 6 December 2021.

[47] Presences, Official website of the magazine presences: les magazine des enterprises du sud isere, https://www.presences-grenoble.fr/, issue: January, 2021.

[48] Grenoble-INP, Research centers associated with grenoble inp, https://www.grenoble-inp.fr/en/research/laboratories, accessed 6 December 2021.

[49] Territorial.fr, Safety regulation for the maximum number of persons allowed according to surface area, https://www.territorial.fr/PAR_TPL_IDENTIFIANT/17172/TPL_CODE/TPL_OVN_CHAPITRE_FICHE/3848-consultation-le-guide-pratique-du-dst.htm, accessed 6 December 2021.

[50] Wikipedia, Grand place(grenoble), https://fr.wikipedia.org/wiki/Grand'Place_(Grenoble), accessed 6 December 2021.

[51] Wikipedia, Liste des magasins ikea en france, https://fr.wikipedia.org/wiki/Liste_des_magasins_Ikea_en_France, accessed 6 December 2021.

[52] Wikipedia, Castorama, https://www.e-marketing.fr/Marketing-Magazine/Article/Castorama-4963-1.htm, accessed 6 December 2021.

[53] Carrefour, Different format of stores in carrefour group, https://www.carrefour.com/en/group/stores, accessed 8 December 2021.

[54] Wikipedia, Enseignes du groupe casino, https://fr.wikipedia.org/wiki/Enseignes_du_groupe_Casino, accessed 8 December 2021.

[55] Biocoop, Créer mon magasin biocoop, https://ac-franchise.com/annuaire/franchise-monoprix, accessed 15 April 2021.

[56] Wikipedia, Système u, https://en.wikipedia.org/wiki/Systeme_U, accessed 8 December 2021.

[57] Wikipedia, Intermarchè, https://fr.wikipedia.org/wiki/Intermarche, accessed 8 December 2021.

[58] T. business Journals, What's a lidl?, https://www.bizjournals.com/triad/blog/morning-edition/2015/04/whats-a-lidl-heres-the-scoop-on-huge-grocery-chain.html, accessed 8 December 2021.

[59] Biocoop, Créer mon magasin biocoop, https://www.biocoop.fr/magasins-bio/creer-mon-magasin-Biocoop, accessed 8 December 2021.

[60] INSEE, National average surface area of supermarkets, https://www.insee.fr/fr/statistiques/1281004, accessed 6 December 2021.

[61] Linternaute, Website with information on the number of seats in a restaurant in grenoble, https://www.linternaute.com/restaurant/guide/ville-grenoble-38000/, accessed 6 December 2021.

[62] Grenoble-metropole, Grenoble alpes metropole tourism, https://www.grenoble-tourisme.com/fr/faire/sorties/cinemas/?page=1, accessed 6 December 2021.

[63] Department-ISERE, Sector for different middle schools, https://carto.isere.fr/carte-interactive/index.html?_ga=2.205615123.1019776183.1552918602-1334538445.1552918602, accessed 2 february 2022.

[64] Academie-Grenoble, Eligible high schools according address in grenoble and saint-martin d'heres, https://bv.ac-grenoble.fr/secto/rec, accessed 2 february 2022.

[65] Departement-ISERE, Eligible high schools according address of a person in Isere, cache.media.education.gouv.fr/file/Scolarisation/38/6/Annexe_1_-_sectorisation_lycees-rentree_2021_1405386.pdf, accessed 2 february 2022.

[66] G. Boeing, Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks, Computers, Environment and Urban Systems 65 (2017) 126–139.

[67] M. Schläpfer, L. Dong, K. O'Keeffe, P. Santi, M. Szell, H. Salat, S. Anklesaria, M. Vazifeh, C. Ratti, G. B. West, The universal visitation law of human mobility, Nature 593. doi:10.1038/s41586-021-03480-9.

[68] Republique-Francaise, Schedules for primary schools, https://www.service-public.fr/particuliers/vosdroits/F24490, accessed 6 December 2021.

[69] Schedule for fantin latour middle school, http://www.ac-grenoble.fr/college/fantin-latour-grenoble/?page_id=71, accessed 6 December 2021.

[70] Schedule for lycee emmanuel mounnier, https://emmanuel-mounier.ent.auvergnerhonealpes.fr/le-lycee/horaires-des-cours/, accessed 6 December 2021.

[71] French official working hours, https://businessculture.org/western-europe/business-culture-in-france/work-life-balance-in-france/, accessed 6 December 2021.

[72] Enquêtes ménages déplacements (emd), https://www.data.gouv.fr/en/datasets/enquetes-menages-deplacements-emd/, accessed 6 December 2021 (2010).

[73] A. Suryanto, I. Darti, On the nonstandard numerical discretization of sir epidemic model with a saturated incidence rate and vaccination, AIMS Mathematics, 6(1): 141–155.doi:10.3934/math.2021010.