ISSN 2451-7100

IMAL preprints

www.imal.conicet.gov.ar/preprints-del-imal

Diffusive metrics induced by random affnities on graphs. An application to the transport systems related to the COVID-19 setting for Buenos Aires (AMBA)

By

María Florencia Acosta - Hugo Aimar - Ivana Gómez – Federico Morana

IMAL PREPRINT # 2021-0056

Publication date: December 7, 2021

Editorial: Instituto de Matemática Aplicada del Litoral
IMAL (CCT CONICET Santa Fe – UNL)
http://www.imal.santafe-conicet.gov.ar

Director de Publicaciones: Dr. Oscar Salinas
E-mail: salinas@santafe-conicet.gov.ar

IMAL PREPRINT # 2021-0056

ISSN 2451-710
Publication date: December 7, 2021

Diffusive metrics induced by random affinities on graphs. An application to the transport systems related to the COVID-19 setting for Buenos Aires (AMBA)

María Florencia Acosta[1], Hugo Aimar[2], Ivana Gómez[3] & Federico Morana[4]

Abstract: The aim of this paper is twofold. First we shall provide a graph metric on the set of vertices determined by the expected value of random affinities between them. This is accomplished by applying the diffusive metric defined by the spectral analysis of the Laplacian determined on the graph by the affinity. As an application we provide a metric in the set of the 41 cities belonging to the largest urban concentration in Argentina based on public transport and neighborhood. The results can be applied to predict and control the spread of COVID-19 and other pandemic diseases in such a setting.

Keywords: weighted graphs, diffusion, graph Laplacian, metrization, COVID-19

1. Introduction

Let $\mathcal{V} = \{1, 2, \dots, n\}$, $n \geq 1$ be the set of vertices of the graph $\mathcal{G} = \left(\mathcal{V}, \mathcal{E}, \vec{a}, \bar{\bar{A}}\right)$, where $\mathcal{E} = \{\{i, j\} : i, j \in \mathcal{V}\}$ is the set of all edges, $\vec{a} = (a_1, a_2, \dots, a_n)$ is the sequence of positive weights of the vertices and $\bar{\bar{A}} = (A_{ij})$ is the matrix of no negative weights of the edges. Assume also that $A_{jj} = 0$ for every $j = 1, \dots, n$. We say that $\mathcal{G}$ is a simple undirected weighted graph based on $\mathcal{V}$. Set $G(\mathcal{V})$ to denote the class of all such simple indirected weighted graphs based on $\mathcal{V}$.

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space. Let $\mathcal{G} : \Omega \to G(\mathcal{V})$ be a graph valued random variable defined in Ω with $\mathcal{V}$ and $\mathcal{E}$ fixed. So that $\mathcal{G}(\omega) = \left(\mathcal{V}, \mathcal{E}, \vec{a}(\omega), \bar{\bar{A}}(\omega)\right)$ with $\vec{a} : \Omega \to \mathbb{R}^n$ a random vector with positive components and $\bar{\bar{A}} : \Omega \to \mathbb{R}^{n\times n}$

[1]Instituto de Matemática Aplicada del Litoral, CONICET, UNL, Santa Fe, Argentina – E-mail: mfacosta@santafe-conicet.gov.ar
[2]Instituto de Matemática Aplicada del Litoral, CONICET, UNL, Santa Fe, Argentina – E-mail: haimar@santafe-conicet.gov.ar
[3]Instituto de Matemática Aplicada del Litoral, CONICET, UNL, Santa Fe, Argentina – E-mail: ivanagomez@santafe-conicet.gov.ar
[4]Instituto de Matemática Aplicada del Litoral, CONICET, UNL, Santa Fe, Argentina – E-mail: fmorana@santafe-conicet.gov.ar

ISSN 2451-710
Publication date: December 7, 2021

a random matrix with non negative entries, with $A_{ii} = 0$ and $A_{ij} = A_{ji}$. So that $a_i : \Omega \to \mathbb{R}$ and $A_{ij} : \Omega \to \mathbb{R}$ are $n + n^2 = n(n+1)$ given random variables. Assume that all of them belong to $L^1(\Omega, \mathcal{P})$, i.e. they have finite first moments $\int_\Omega |a_i|\, d\mathcal{P} = \int_\Omega a_i\, d\mathcal{P} < \infty$ and $\int_\Omega |A_{ij}|\, d\mathcal{P} = \int_\Omega A_{ij}\, d\mathcal{P} < \infty$. We shall also assume the normalizations $\sum_{i=1}^{n} a_i(w) = 1$ and $\sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}(w) = 1$ for every $\omega \in \Omega$.

The expected graph is $\mathbb{E}(\mathcal{G}) = \left(\mathcal{V}, \mathcal{E}, \mathbb{E}(\vec{a}), \mathbb{E}(\bar{\bar{A}})\right)$, with $\mathbb{E}(\vec{a}) = (\mathbb{E}a_1, \mathbb{E}a_2, \dots, \mathbb{E}a_n)$, and $\mathbb{E}(\bar{\bar{A}}) = (\mathbb{E}A_{ij} : i, j = 1, \dots, n)$. Notice that $\mathbb{E}a_i \geq 0$ and $\mathbb{E}A_{ij} \geq 0$, and that

$$\sum_{i=1}^{n} \mathbb{E}a_i = \mathbb{E}\left(\sum_{i=1}^{n} a_i\right) = \mathbb{E}(1) = 1, \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbb{E}A_{ij} = \mathbb{E}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}\right) = 1.$$

Many interesting questions arise regarding the relation between the analysis provided by each graph $\mathcal{G}(\omega)$ and the analysis provided by the graph $\mathbb{E}(\mathcal{G})$. In this paper we focus on building a metric, by the diffusion method given in [1], on the graph $\mathbb{E}(\mathcal{G})$. For a different approach see [2].

This search is motivated by the application to the analysis of the transportation of people between the 41 cities in AMBA (Buenos Aires) in the COVID-19 context, through different ways of passengers transport. The acronym AMBA is used to name the 41 cities that concentrate one third of the total population of Argentina and is spatially concentrated around Buenos Aires City. The total population of AMBA is of about 16.7 millions. The Figure 1 depicts their distribution.

Aside from the geographical distance between locations i and j in the map there is a valuable information given by the public transport system in AMBA. The system SUBE (unifier system of electronic ticket) keeps a great amount of information that allows to have another geometry provided by a connectivity distance built on this big data source. With the idea of considering at once a diversity of affinities between two cities i and j, such as euclidean distance, neighborhood, public transport, private transport, etcetera, we introduce a diffusive metrization of the graph that takes into account these diverse factors which all together contribute to the motion of people inside AMBA.

Section 2 is devoted to introduce theoretical background of our general setting. In Section 3 we apply the metric built in §2 to some particular cases of affinities for the graph AMBA. Here we draw the families of balls in these metrics in order to have a picture of the behavior of distance measured in terms of transport. We also give here empirical estimates of the norms of the differences between metric matrices coming from different combinations of ways of transport. In Section 4 we compare the metric maps obtained above with the actual spread of COVID-19 in AMBA during different steps of the pandemic growth in Argentina.

IMAL PREPRINT # 2021-0056
ISSN 2451-710
Publication date: December 7, 2021

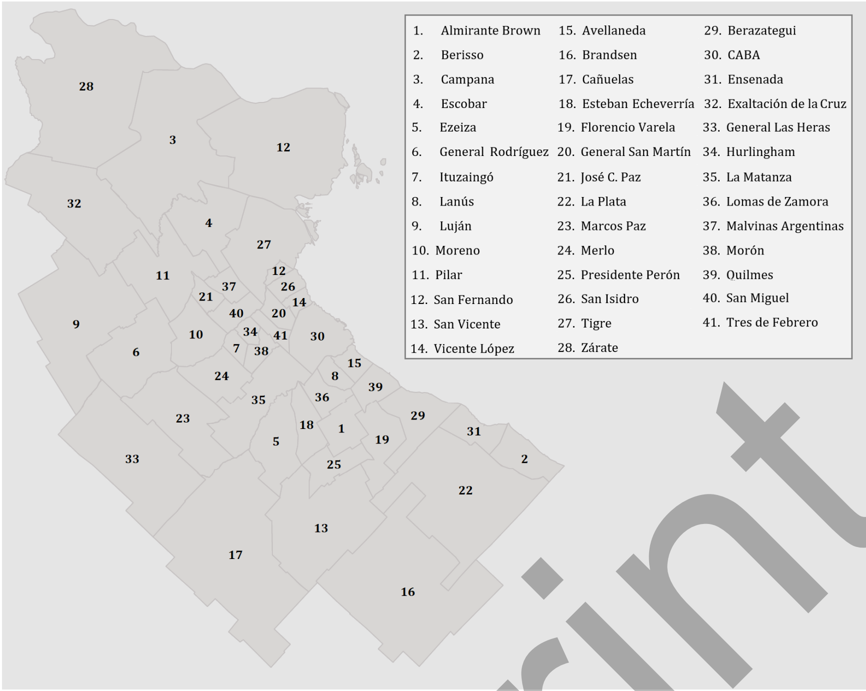


Figure 1: AMBA

2. Metrization of Random Graphs

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space. We say that a function $\mathcal{G}$ defined in Ω with values on the simple undirected weighted graphs on $\mathcal{V} = \{1, 2, \dots, n\}$, is a random graph on $\mathcal{V}$ with finite first moments if $\mathcal{G}(\omega) = (\mathcal{V}, \mathcal{E}, \vec{a}(\omega), \bar{\bar{A}}(\omega))$ with $\mathcal{V} = \{1, 2, \dots, n\}$, $\mathcal{E} = \{\{i, j\} : i, j \in \mathcal{V}\}$, $\vec{a}(\omega) = \big(a_i(\omega) : i = 1, \dots, n\big)$, $\bar{\bar{A}}(\omega) = \big(A_{ij}(\omega) : i, j = 1, \dots, n\big)$ with each $a_i(\omega)$ and each $\bar{\bar{A}}_{ij}(\omega)$ in $L^1(\Omega, \mathcal{F}, \mathcal{P})$. We shall also assume the probabilistic normalizations

$$\sum_{i=1}^{n} a_i(\omega) = 1, \qquad \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}(\omega) = 1$$

for every $\omega \in \Omega$ and that $a_i(\omega) > 0$ for each $i \in \mathcal{V}$ and $\bar{\bar{A}}_{ij}(\omega) \geq 0$ for $i, j \in \mathcal{V}$ and $\omega \in \Omega$.

With the above notation, it makes sense to consider a notion of expected graph $\mathbb{E}\mathcal{G} = \Big(\mathcal{V}, \mathcal{E}, \mathbb{E}\vec{a}, \mathbb{E}\bar{\bar{A}}\Big)$, with $\mathbb{E}\vec{a} = (\mathbb{E}a_1, \dots, \mathbb{E}a_n)$ and $\mathbb{E}\bar{\bar{A}} = (\mathbb{E}A_{ij} : i, j \in \mathcal{V})$, $\mathbb{E}a_i = \displaystyle\int_\Omega a_i(\omega)\, d\mathcal{P}(\omega)$ and $\mathbb{E}A_{ij} = \displaystyle\int_\Omega A_{ij}(\omega)\, d\mathcal{P}(\omega)$.

IMAL PREPRINT # 2021-0056

ISSN 2451-710
Publication date: December 7, 2021

Proposition 2.1. *Let $\mathcal{G}(\omega)$ and $\mathbb{E}\mathcal{G}$ as before. Then*

(i) $\mathbb{E}a_i > 0$ *for every* $i \in \mathcal{V}$*;*

(ii) $\mathbb{E}A_{ij} \geq 0$ *for every* $i, j \in \mathcal{V}$*;*

(iii) $\sum_{i=1}^n \mathbb{E}a_i = 1$*;*

(iv) $\sum_{i=1}^n \sum_{j=1}^n \mathbb{E}A_{ij} = 1$.

Proof. (i) Since $a_i(\omega)$ is positive for every $\omega \in \Omega$, the sets $\Omega_k = \{\omega \in \Omega : 2^{-k} < a_i(\omega) \leq 2^{-k+1}\}$ for $k \in \mathbb{Z}$ forms a disjoint partition of Ω. In other words

$$\Omega = \bigcup_{k\in\mathbb{Z}} \Omega_k, \quad \Omega_k \cap \Omega_\ell = \emptyset.$$

Hence $1 = \mathcal{P}(\Omega) = \sum_{k\in\mathbb{Z}} \mathcal{P}(\Omega_k)$. So that for some $k_0 \in \mathbb{Z}$ we have that $\mathcal{P}(\Omega_{k_0}) > 0$. Then

$$\mathbb{E}a_i = \int_\Omega a_i(\omega)\, d\mathcal{P} = \sum_{k\in\mathbb{Z}} \int_{\Omega_k} a_i(\omega)\, d\mathcal{P} \geq \int_{\Omega_{k_0}} a_i(\omega)\, d\mathcal{P} \geq 2^{-k_0}\mathcal{P}(\Omega_{k_0}) > 0.$$

The proofs of (ii), (iii) and (iv) are clear. □

Notice that under the assumptions $a_i(\omega) > 0$, $A_{i,j}(\omega) \geq 0$, $\sum_{i=1}^n a_i(\omega) = 1$ and $\sum_{i=1}^n \sum_{j=1}^n A_{ij}(\omega) = 1$ we have that each a_i and each A_{ij} belong to $L^\infty(\Omega, \mathcal{F}, \mathcal{P}) \subseteq L^1(\Omega, \mathcal{F}, \mathcal{P})$.

Given a graph $\Gamma = (\mathcal{V}, \mathcal{E}, \vec{a}, \bar{\bar{A}})$ the Laplacian on Γ is given by

$$\Delta_\Gamma f(i) = \frac{1}{a_i} \sum_{j=1}^n A_{ij}\big(f(i) - f(j)\big)$$

when $f : \mathcal{V} \to \mathbb{R}$ is any function defined on the set of vertices. In matrix notation

$$\Delta_\Gamma = \bar{\bar{a}}^{-1}\left(\bar{\bar{A}} - \bar{\bar{D}}\right)$$

with $\bar{\bar{a}}^{-1} = \operatorname{diag}\left(a_1^{-1}, \dots, a_n^{-1}\right)$ and $\bar{\bar{D}} = \operatorname{diag}\left(\sum_{j\neq 1} A_{1j}, \dots, \sum_{j\neq n} A_{nj}\right)$.

Notice now that for a given random graph on $\mathcal{V}$, $\mathcal{G}(\omega)$, as before we have at least two ways of considering an expected Laplacian. The first it to apply the above definition of the Laplace operator to $\Gamma = \mathbb{E}\mathcal{G}$. In fact

$$\Delta_{\mathbb{E}\mathcal{G}} f(i) = \frac{1}{\mathbb{E}a_i} \sum_{j=1}^n \mathbb{E}A_{ij}\big(f(i) - f(j)\big)$$

is well defined from Proposition 2.1. The second way is to ask for the existence of an expected Laplacian for the random Laplacian defined by

$$\Delta_\omega f(i) = \Delta_{\mathcal{G}(\omega)} f(i) = \frac{1}{a_i(\omega)} \sum_{j=1}^n A_{ij}(\omega)\big(f(j) - f(i)\big),$$

ISSN 2451-710
Publication date: December 7, 2021

$\omega \in \Omega$, $i \in \mathcal{V}$. It is clear that with the current hypotheses on the a_i's the expected Laplacian $\mathbb{E}\Delta_\omega$ not necessarily exists. On the other hand, it is also clear that when the a_i's are deterministic (constant) we have that $\mathbb{E}\Delta_\omega = \Delta_{\mathbb{E}\mathcal{G}}$. Actually in our application this will be the case. Nevertheless, for the sake of theoretical completeness we give some sufficient conditions on the random graph in order to guarantee the existence of the expected Laplacian and to produce a formula to compute it. This is done in the next result.

Proposition 2.2. *Let $\mathcal{G}(\Omega)$ be a random graph on $\mathcal{V} = \{1, \dots, n\}$. Assume that $a_i(\omega) > 0$ for every $i \in \mathcal{V}$ and $\omega \in \Omega$, $\sum_{i=1}^n a_i(\omega) = 1$ and $a_i^{-1} \in L^1(\Omega, \mathcal{F}, \mathcal{P})$ for every $i \in \mathcal{V}$. Assume that $A_{ij}(\omega) \geq 0$, $\sum_{i=1}^n \sum_{j=1}^n A_{ij}(\omega) = 1$ for $\omega \in \Omega$. If each $a_i(\omega)$ is independent of the random variables $A_{k\ell}(\omega)$ for every $\{k, \ell\} \in \mathcal{E}$, then with*

$$\Delta_{\mathcal{G}(\omega)} f(i) = \frac{1}{a_i(\omega)} \sum_{j=1} A_{ij}(\omega) \left(f(j) - f(i)\right), \quad \omega \in \Omega, \quad i \in \mathcal{V},$$

we have that $\mathbb{E}\Delta_{\mathcal{G}(\omega)} = \Delta_{\tilde{\mathcal{G}}}$ with $\tilde{\mathcal{G}} = \left(\mathcal{V}, \mathcal{E}, \bar{b}, \mathbb{E}\bar{\bar{A}}\right)$, $\bar{b} = (b_1, b_2, \dots, b_n)$ and $b_i = \left(\mathbb{E}\frac{1}{a_i}\right)^{-1}$.

Proof. Since we are assuming the finiteness of $\displaystyle\int_\Omega \frac{1}{a_i(\omega)} d\mathcal{P}(\omega)$ and independence of each $a_i(\omega)$ with all the $A_{k\ell}(\omega)$, we have that $\frac{1}{a_i(\omega)}$ is a random variable which is independent of the random variable $\displaystyle\sum_{j=1}^n A_{ij}(\omega) \left(f(j) - f(i)\right)$ for any $f : \mathcal{V} \to \mathbb{R}$. Hence

$$\begin{aligned}
\mathbb{E}\left(\Delta_{\mathcal{G}(\omega)} f(i)\right) &= \mathbb{E}\left(\frac{1}{a_i}\right) \mathbb{E}\left(\sum_{j=1}^n A_{ij}\left(f(j) - f(i)\right)\right) \\
&= \frac{1}{\left(\mathbb{E}\left(\frac{1}{a_i}\right)\right)^{-1}} \sum_{j=1}^n \mathbb{E}\left(A_{ij}\right)\left(f(j) - f(i)\right) \\
&= \frac{1}{b_i} \sum_{j=1}^n \mathbb{E}\left(A_{ij}\right)\left(f(j) - f(i)\right) \\
&= \Delta_{\tilde{\mathcal{G}}} f(i),
\end{aligned}$$

as desired. □

Once we have a Laplacian defined on $(\mathcal{V}, \mathcal{E})$ which could be $\Delta_{\mathbb{E}\mathcal{G}}$ or $\mathbb{E}\Delta_\omega$ we can build the diffusive metric on $\mathcal{V}$ (see [1]). For completeness, let us state and prove the basic facts regarding the constructive of these metrics.

Teorema 2.1. *Let $\Gamma = (\mathcal{V}, \mathcal{E}, b_i, B_{ij})$ be a simple undirected weighted graph. Then*

ISSN 2451-710
Publication date: December 7, 2021

a) the operator Δ_Γ *is selfadjoint with respect to the inner product*

$$\langle f, g\rangle_{\bar{b}} = \sum_{i=1}^{n} f(i)g(i)b_i\,;$$

b) the operator Δ_Γ *is negative definite, i. e.*

$$\langle \Delta_\Gamma f, f\rangle_{\bar{b}} \leq 0, \quad \textit{for every} \quad f;$$

c) the operator Δ_Γ *is diagonalizable, i. e. there exist a sequence* $\lambda_{n-1} \leq \lambda_{n-2} \leq \cdots \leq \lambda_1 \leq \lambda_0 = 0$ *and an orthonormal sequence* $\phi_0, \phi_1, \ldots, \phi_{n-1}$ *with respect to the inner product* $\langle\ ,\ \rangle_{\bar{b}}$*, such that*

$$\Delta_\Gamma \phi_i = \lambda_i \phi_1, \quad \textit{for} \quad i = 0, 1, \ldots, n-1;$$

d) for any $t > 0$*, the function* $d_t : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ *given by*

$$d_t(i,j) = \sqrt{\sum_{\ell=0}^{n-1} e^{2t\lambda_\ell}|\phi_\ell(i) - \phi_\ell(j)|^2}$$

is a metric on $\mathcal{V}$*.*

Proof. $a)$ Let f and g be two functions from $\mathcal{V}$ to $\mathbb{R}$, then since $B_{ij} = B_{ji}$,

$$\begin{aligned}
\langle \Delta_\Gamma f, g\rangle_{\bar{b}} &= \sum_{i=1}^{n} (\Delta_\Gamma f)(i) g(i) b_i \\
&= \sum_{i=1}^{n} \left(\frac{1}{b_i} \sum_{j=1}^{n} B_{ij}\big(f(j) - f(i)\big) \right) g(i) b_i \\
&= \sum_{j=1}^{n} \sum_{i=1}^{n} B_{ij}\big(f(j) - f(i)\big) g(i) \\
&= \sum_{j=1}^{n} \left(\sum_{i=1}^{n} B_{ij} f(j) g(i) - \sum_{i=1}^{n} B_{ij} f(i) g(i) \right) \\
&= \sum_{j=1}^{n} \sum_{i=1}^{n} B_{ij} f(j) g(i) - \sum_{j=1}^{n} \sum_{i=1}^{n} B_{ij} f(i) g(i) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} B_{ij} f(j) g(i) - \sum_{i=1}^{n} \sum_{j=1}^{n} B_{ij} f(i) g(i) \\
&= \sum_{i=1}^{n} \left(\frac{1}{b_i} \sum_{j=1}^{n} B_{ij}\big(g(j) - g(i)\big) \right) f(i) b_i \\
&= \langle f, \Delta_\Gamma g\rangle_{\bar{b}}\,.
\end{aligned}$$

ISSN 2451-710
Publication date: December 7, 2021

b) Since $B_{ij} = B_{ji}$ we have

$$\begin{aligned}
\langle -\Delta_\Gamma f, f\rangle_{\bar{b}} &= \sum_{i=1}^{n} (-\Delta_\Gamma f)(i) f(i) b_i \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} B_{ij}\big(f(i) - f(j)\big) f(i) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} B_{ij} f^2(i) - \sum_{i=1}^{n}\sum_{j=1}^{n} B_{ij} f(i) f(j) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} B_{ij}\big(f^2(i) - f(i)f(j)\big) \\
&= \frac{1}{2}\left[\sum_{i=1}^{n}\sum_{j=1}^{n} B_{ij}\big(f^2(i) - f(i)f(j)\big) + \sum_{i=1}^{n}\sum_{j=1}^{n} B_{ij}\big(f^2(i) - f(i)f(j)\big)\right] \\
&= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} B_{ij}\big(f^2(i) + f^2(j) - 2f(i)f(j)\big) \\
&= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} B_{ij}\big(f(i) - f(j)\big)^2 \\
&\geq 0.
\end{aligned}$$

c) follows from *a*) and *b*) since we are dealing with a self-adjoint and negative definite matrix Δ_Γ. Since the constant functions are Δ_Γ-harmonic we hare that $\lambda_0 = 0$ is the eigenvalue corresponding to the eigenfunction $\phi_0(i) = \dfrac{1}{\sqrt{\sum_{j=1}^{n} b_j}}$ for $i = 1, \dots, n$, which has the L^2 norm given by the inner product $\langle\ ,\ \rangle_{\bar{b}}$ equal to one.

d) it is clear that d_t is nonnegative, symmetric, faithful and satisfies the triangle inequality for every $t > 0$. Let un notice here the $d_t(i,j)$ is the $L^2(\mathcal{V}, \bar{b})$ norm of the difference of the heat kernels at i and j provided by the diffusion $\dfrac{\partial u}{\partial t} = \Delta_\Gamma u$. □

As a general reference for the above see for example [3].

3. The case of AMBA (Buenos Aires)

In this section we effectively compute and sketch some families of balls, the metric provided by d_t in Theorem 2.1 for several natural instances of affinity matrices A_{ij} and some of their means and a couple of instances for the weights a_i at each node. All the underlying computations are performed in Python. In order to show our results in a compact way we shall first introduce the families of affinities A_{ij} that we shall use and the weights a_i that we consider.

ISSN 2451-710
Publication date: December 7, 2021

Our basic vertex set is $\mathcal{V} = \{1, \ldots, 41\}$ one for each city in AMBA. The first, and perhaps more relevant matrix concerning the spread of COVID-19 in this setting, is the matrix built with the data of SUBE provided by the public transport in AMBA. This matrix takes onto account buses, subte (metro), trains and even fluvial public transportation. We shall denote it by A^0. We exhibit in Figure 3 the full unnormalized form of the 41×41 matrix A^0. We shall as well consider some neighborhood matrices. With A^1 we denote the normalization of the matrix that takes the value 1 at (i, j) if the cities i and j share some points of their boundaries, and zero otherwise. In Figure 2 we show a small part of A^1 (unnormalized). With A^2 we denote a better quantified weighted approach of A^1 that takes into account the length of the shared portion of the boundary between cities i and j. See Figure 4. Since the population of different cities is in several instances quite different for two neighbor cities, we consider still another matrix that we denote A^3, which takes into account the length of the shared boundaries and also the minimum of the population of the two neighbor cities. Figure 5 depicts a part of this matrix. For last, the matrix A^4 considers only the minimum of the populations of any two neighbor cities. The matrix A^4 is partially showed in Figure 6.

Regarding the weights a_i at the nodes, we shall consider only two $\vec{a}$: the uniform $\vec{a}_1 = \left(\frac{1}{41}, \ldots, \frac{1}{41}\right)$ and a normalization of the density of the disease in each location (total number of active infections over population) by July 2020, given by

$$\begin{aligned}\vec{a}_2 = (&0.0023, 0.0009, 0.0004, 0.0014, 0.0015, 0.0009, 0.0012, 0.0030, 0.0007,\\ &0.0009, 0.0011, 0.0015, 0.0008, 0.0016, 0.0049, 0.0005, 0.0006, 0.0018,\\ &0.0015, 0.0031, 0.0013, 0.0008, 0.0012, 0.0010, 0.0019, 0.0022, 0.0014,\\ &0.0006, 0.0019, 0.0095, 0.0011, 0.0004, 0.0015, 0.0018, 0.0018, 0.0026,\\ &0.0013, 0.0018, 0.0029, 0.0018, 0.0034)\end{aligned}$$

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0
0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
0	0	1	1	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0
0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0
1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	1	1

Figure 2: Unnormalized submatrix of A^1 (20×41)

ISSN 2451-710
Publication date: December 7, 2021

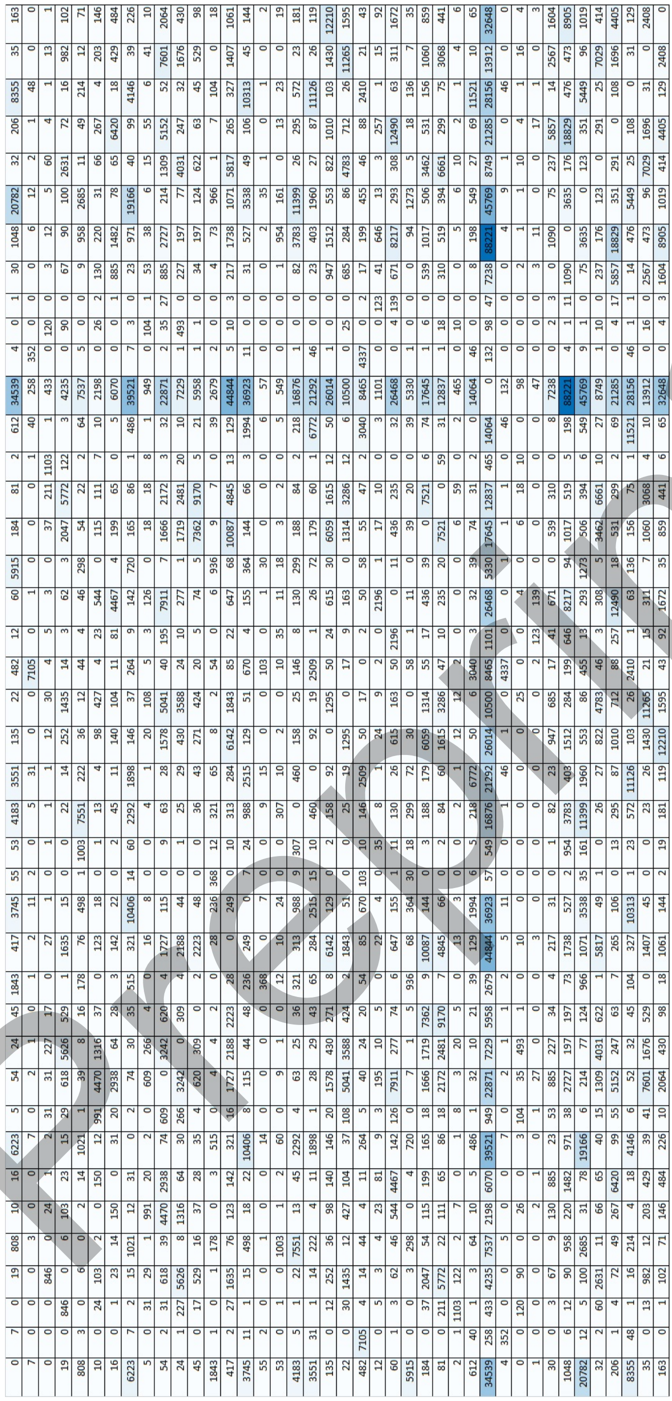


Figure 3: Matrix A^0 that reflects the data provided by SUBE before normalizing to probability

ISSN 2451-710
IMAL PREPRINT # 2021-0056 Publication date: December 7, 2021

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10,5	13,3	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	32	0	0	0	0	0	0	7,52	40,5	0	0	0	0	0	0	0	0
0	0	32	0	0	0	0	0	0	0	26	6,26	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	12,2	0	0	0	18,3	21,6	0	0
0	0	0	0	0	0	0	0	36,7	16,6	11,1	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	4,35	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	9,8	0	0	0	0	0
0	0	0	0	0	36,7	0	0	0	0	15,4	0	0	0	0	0	0	0	0	0
0	0	0	0	0	16,6	4,35	0	0	0	9,84	0	0	0	0	0	0	0	0	0
0	0	7,52	26	0	11,1	0	0	15,4	9,84	0	0	0	0	0	0	0	0	0	0
0	0	40,5	6,26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	12,2	0	0	0	0	0	0	0	0	0	0	29,9	33,6	0	11,1	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7,69
0	0	0	0	0	0	0	9,8	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	29,9	0	0	0	0	0	0	0
0	0	0	0	18,3	0	0	0	0	0	0	0	33,6	0	0	0	0	0	0	0
10,5	0	0	0	21,6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13,3	0	0	0	0	0	0	0	0	0	0	0	11,1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	7,69	0	0	0	0	0	0

Figure 4: Unnormalized submatrix of A^2 (20×20)

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3894450	6866849	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	3379775	0	0	0	0	0	0	793751	4272745	0	0	0	0	0	0	0	0
0	0	3379775	0	0	0	0	0	0	0	6621695	1094768	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	939049	0	0	0	1150825	4728879	0	0
0	0	0	0	0	0	0	0	4020322	1818743	1213227	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	786976	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	3492642	0	0	0	0	0
0	0	0	0	0	4020322	0	0	0	0	1846195	0	0	0	0	0	0	0	0	0
0	0	0	0	0	1818743	786976	0	0	0	3721163	0	0	0	0	0	0	0	0	0
0	0	793751	6621695	0	1213227	0	0	1846195	3721163	0	0	0	0	0	0	0	0	0	0
0	0	4272745	1094768	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	939049	0	0	0	0	0	0	0	0	0	0	927898	2112887	0	854172	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2058267
0	0	0	0	0	0	0	3492642	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	927898	0	0	0	0	0	0	0
0	0	0	0	1150825	0	0	0	0	0	0	0	2112887	0	0	0	0	0	0	0
3894450	0	0	0	4728879	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6866849	0	0	0	0	0	0	0	0	0	0	0	854172	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	2058267	0	0	0	0	0	0

Figure 5: Unnormalized submatrix of A^3 (20×20)

The result of Section 2 generate a diversity of metrics on $\mathcal{V} = \{1, 2, \dots, 41\}$ provided by any choice of $A \in \{A^0, A^1, A^2, A^3, A^4\}$ and $\vec{a} \in \{\vec{a}_1, \vec{a}_2\}$. Moreover from Proposition 2.2 in Section 2 any convex combination of matrices A provides a Laplacian and a corresponding family of metrics on $\mathcal{V}$. Sometimes we shall use a convex combination of A^0 and A^i with $i = 1, 2, 3, 4$, i.e. $A = \theta A^0 + (1 - \theta) A^i$ with $0 \leq \theta \leq 1$. In this cases we write $d_t^{i,\theta;j}$ to denote the metric provided by Theorem 2.1 with $B = \theta A^0 + (1 - \theta) A^i$ and $b = \vec{a}_j$. We shall use the standard notation for balls keeping the above notation, precisely

$$B_t^{i,\theta;j}(k, r) = \{\ell \in \mathcal{V} : d_t^{i,\theta;j}(k, \ell) < r\}$$

for $k \in \mathcal{V}$, $r > 0$, $i = 0, 1, 2, 3, 4$ and $0 \leq \theta \leq 1$.

ISSN 2451-710
Publication date: December 7, 2021

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	370900	517082	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	105552	0	0	0	0	0	0	105552	105552	0	0	0	0	0	0	0	0
0	0	105552	0	0	0	0	0	0	0	255073	174883	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	77161	0	0	0	62921	219031	0	0
0	0	0	0	0	0	0	0	109695	109695	109695	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	180914	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	356392	0	0	0	0	0
0	0	0	0	0	109695	0	0	0	0	119805	0	0	0	0	0	0	0	0	0
0	0	0	0	0	109695	180914	0	0	0	378167	0	0	0	0	0	0	0	0	0
0	0	105552	255073	0	109695	0	0	119805	378167	0	0	0	0	0	0	0	0	0	0
0	0	105552	174883	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	77161	0	0	0	0	0	0	0	0	0	0	31023	62921	0	77161	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	267655
0	0	0	0	0	0	0	356392	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	31023	0	0	0	0	0	0	0
0	0	0	0	62921	0	0	0	0	0	0	0	62921	0	0	0	0	0	0	0
370900	0	0	0	219031	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
517082	0	0	0	0	0	0	0	0	0	0	0	77161	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	267655	0	0	0	0	0	0

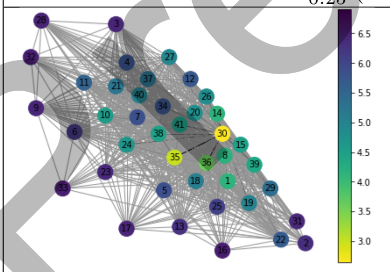
Figure 6: Unnormalized submatrix of A^4 (20×20)

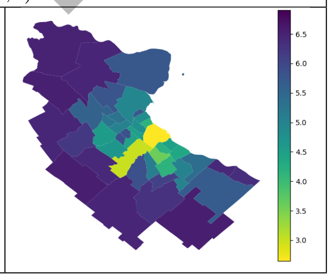
A way to schematically depict the unrestricted paths of COVID-19 propagation from the point (CABA) with higher initial concentration of diseases is to consider for each metric the balls centered at CABA (30) and growing radii.

Using a prescribed scale of colors we can run our algorithm in Python in order to obtain a diversity of images for propagation due to the above described notations of neighborhood and transport and their convex combinations. With the above introduced notation we give the following illustration of the results. In Table 1 and Table 3 we use always $t = 0.25$ and $j = 1$, the other parameters are explicitly given. The center is always 30 (CABA), the growing radii are colored according to the given scale.

Table 1: Distances to CABA. Left: graph; right: map.

$d^{0,0;1}_{0.25}(30, \cdot)$





Some global comparison of the different metrics are in order. In Table 2 we shall show the comparison of the metric induced by public transport (SUBE) with the metrics induced a convex combination of the SUBE data and some of the neighborhood matrices defined above only for the case of $\bar{a_1}$, the uniform distribution $\left(a_i = \frac{1}{41}\right)$ of the vertices of the graph. Here we compute the relative deviations with

ISSN 2451-710
Publication date: December 7, 2021

respect to the metric induced just by public transport. Let us precise the above. Set

$$\epsilon_t^{i,\theta} = \frac{\left\| d_t^{0,0;1} - d_t^{i,\theta;1} \right\|}{\left\| d_t^{0,0;1} \right\|},$$

where $d_t^{0,0;1}$ is the metric matrix associate to the public transport only and $d_t^{i,\theta;1}$ are the metrics defined above. The norm considered here is the Euclidean one, i.e.

$$\left\| d_t^{0,0;1} - d_t^{i,\theta;1} \right\|^2 = \sum_{k,\ell=1}^{n} \left| d_t^{0,0;1}(k,\ell) - d_t^{i,\theta;1}(k,\ell) \right|^2$$

and

$$\left\| d_t^{0,0;1} \right\|^2 = \sum_{k,\ell=1}^{n} \left(d_t^{0,0;1}(k,\ell) \right)^2.$$

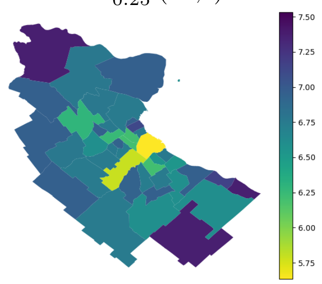
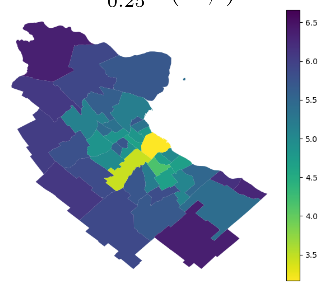
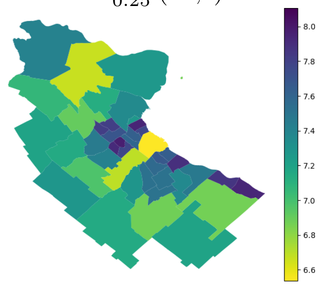
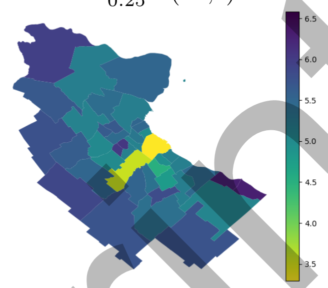
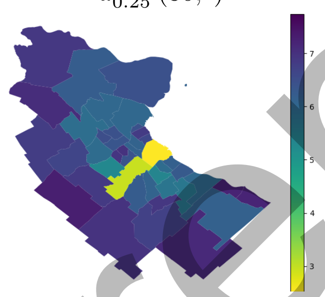
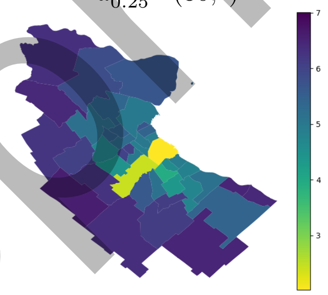
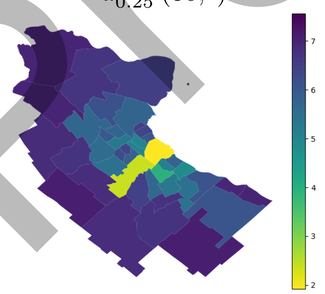
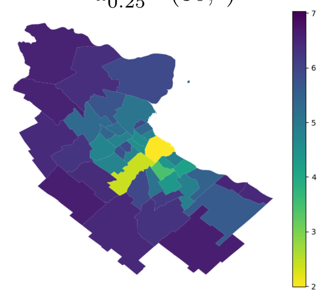
Table 2: Relative differences

$\epsilon_t^{1,0}$	0.12035607	$\epsilon_t^{1,0.5}$	0.0609088
$\epsilon_t^{2,0}$	0.17173178	$\epsilon_t^{2,0.5}$	0.091446
$\epsilon_t^{3,0}$	0.0644136	$\epsilon_t^{3,0.5}$	0.0306021
$\epsilon_t^{4,0}$	0.09062579	$\epsilon_t^{3,0.5}$	0.04661433

In Table 2 we observe that, as it could be expected and as it reflected by the colored maps in Table 3, the largest relative differences with the metric provided by the public transport are those given by matrices A^1 and A^2 which only take into account neighboring, with no reference to the sizes of populations. On the other hand, for matrices A^3 and A^4 which take into account populations, the results are closer to that of the pure public transport matrix A^0. All the interpolation cases show, at least with $\theta = 0.5$ a closer behavior to that of A^0.

ISSN 2451-710
Publication date: December 7, 2021

Table 3: Distances to CABA

$d^{1,0;1}_{0.25}(30,\cdot)$	$d^{1,0.5;1}_{0.25}(30,\cdot)$
	
$d^{2,0;1}_{0.25}(30,\cdot)$	$d^{2,0.5;1}_{0.25}(30,\cdot)$
	
$d^{3,0;1}_{0.25}(30,\cdot)$	$d^{3,0.5;1}_{0.25}(30,\cdot)$
	
$d^{4,0;1}_{0.25}(30,\cdot)$	$d^{4,0.5;1}_{0.25}(30,\cdot)$
	

ISSN 2451-710
Publication date: December 7, 2021

4. Comparison of the metric closeness with the actual spread of COVID-10 in AMBA

As we show in Section 3 all the versions of the diffusive metric that we consider provide in some way the paths of propagation of COVID-19 associated only with transport of AMBA. Our model is based only in the proportional volume of people moving daily from each city to another without taking into account the restrictions imposed in each district. At this point it is important to mention that there are two different administrations in the system of the 41 cities of AMBA. One for the City of Buenos Aires, CABA, node 30 in our graph, and other administration ruling in the other 40 cities of AMBA. The restrictions imposed by both administrations during the pandemic course, were sometimes coincident and sometimes not. The current available data allows us to have a precise picture of the dynamics of the growth of infections in AMBA. For each one of the 41 cities we computed the time passed until the number of infected people surpass the threshold of $x\%$ of the population with $x = j \cdot \frac{1}{10}$, $j = 1, 2, \ldots, 20$. The maps obtained are of the type depicted in Figure 7.

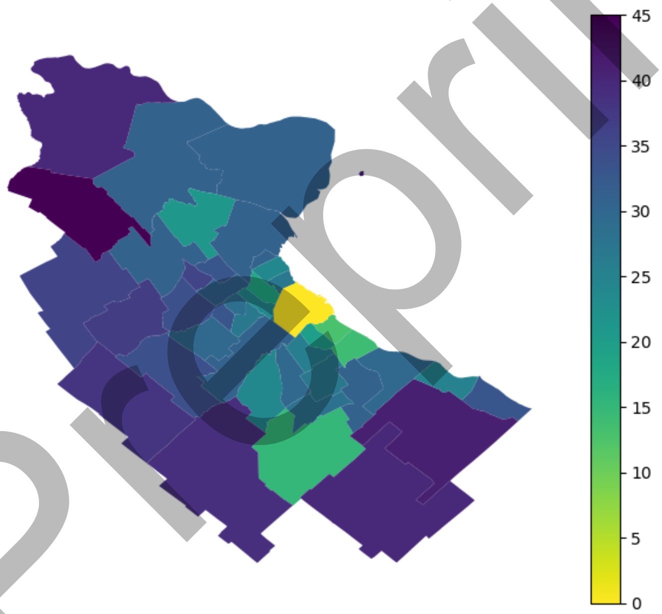


Figure 7: Days up to 0.1% of infections over the population (from 0.1% of CABA)

We shall only concentrate our analysis in the two largest cities of AMBA, CABA and La Matanza. Numbered 30 and 35 in our graph. Ciudad Autónoma de Buenos Aires (30) has a population of 3.075.000. La Matanza (35) has a population of

ISSN 2451-710
Publication date: December 7, 2021

2.280.000 people. They share a boundary of about 10 kilometers. All the metrics in the models of Section 3 place La Matanza as the closest city to CABA. This fact is by no ways reflected by the actual spread of the pandemic in AMBA based in our percentual thresholding scheme. In fact while for CABA we have the red distribution as a function of time in Figure 8, for La Matanza we have the blue one.

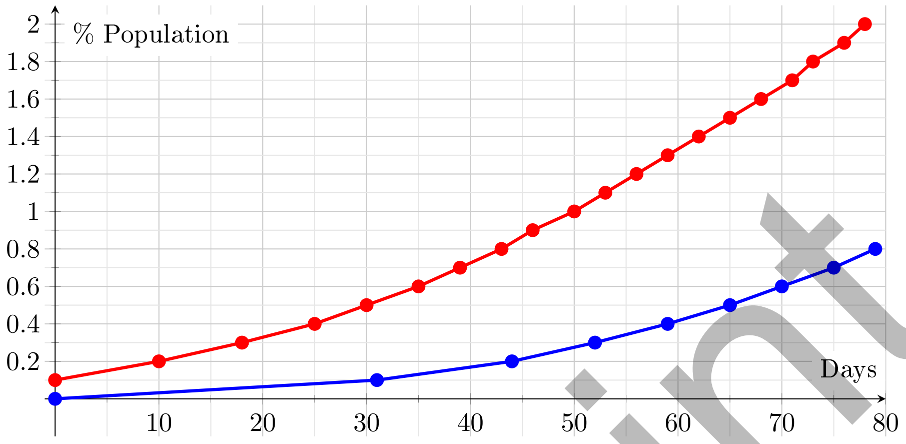


Figure 8: Evolution of cases in CABA (red) and La Matanza (blue)

At this point, it is worthy noticing that the administration of CABA has almost always been looser than the administration of La Matanza, regarding the quarantine, isolation and restriction measures associated with the pandemics COVID-19.

Declarations

Funding

This work was supported by the Ministerio de Ciencia, Tecnología e Innovación-MINCYT in Argentina: Consejo Nacional de Investigaciones Científicas y Técnicas-CONICET (grant PUE-IMAL #22920180100041CO) and Agencia Nacional de Promoción de la Investigación, el Desarrollo Tecnológico y la Innovación (grant PICT 2015-3631) and UNL (grant CAI+D 50620190100070LI).

Conflicts of interest/Competing interests

The authors have no conflicts of interest to declare that are relevant to the content of this article.

Availability of data and material

All the data used is of public access.

ISSN 2451-710
Publication date: December 7, 2021

Code availability

We use Python, free software.

References

[1] R. R. Coifman and S. Lafon, "Diffusion maps," *Applied and Computational Harmonic Analysis*, vol. 21, no. 1, pp. 5–30, 2006.

[2] M. F. Acosta, H. Aimar, and I. Gómez, "On Frink's type metrization of weighted graphs," *Asian Research Journal of Mathematics*, vol. 17, pp. 26–37, 2021.

[3] M. M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, and P. Vandergheynst, "Geometric deep learning: going beyond euclidean data," *IEEE Signal Processing Magazine*, vol. 34, no. 4, pp. 18–42, 2017.