

Diffusive Metrics Induced by Random Affinities on Graphs. An Application to the Transport Systems Related to the COVID-19 Setting for Buenos Aires (AMBA)

M. F. ACOSTA¹, H. AIMAR², I. GÓMEZ^{3*} and F. MORANA⁴

Received on September 28, 2021 / Accepted on June 30, 2022

ABSTRACT. The aim of this paper is to apply the diffusive metric technique defined by the spectral analysis of graph Laplacians to the set of the 41 cities belonging to AMBA, the largest urban concentration in Argentina, based on public transport and neighborhood. It could be expected that the propagation of any epidemic disease would follow the paths determined by those metrics. Our result reflects that the isolation measures decided by the health administration helped at the attenuation of the actual spread of COVID-19 in AMBA.

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 $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \vec{a}, \vec{\bar{A}})$, 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 $\vec{\bar{A}} = (A_{ij})$ 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 undirected weighted graphs based on \mathcal{V} .

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

*Corresponding author: Ivana Gómez – E-mail: ivanagomez@santafe-conicet.gov.ar

¹Instituto de Matemática Aplicada del Litoral, CONICET, UNL, FICH, Santa Fe, Argentina – E-mail: mfacosta@santafe-conicet.gov.ar

²Instituto de Matemática Aplicada del Litoral, CONICET, UNL, Santa Fe, Argentina – E-mail: haimar@santafe-conicet.gov.ar <https://orcid.org/0000-0003-3829-7738>

³Instituto de Matemática Aplicada del Litoral, CONICET, UNL, Santa Fe, Argentina – E-mail: ivanagomez@santafe-conicet.gov.ar <https://orcid.org/0000-0003-2921-2392>

⁴Instituto de Matemática Aplicada del Litoral, CONICET, UNL, Santa Fe, Argentina – E-mail: fmorana@santafe-conicet.gov.ar

entries, with $A_{ii} = 0$ and $A_{ij} = A_{ji}$. So that $a_i : \Omega \rightarrow \mathbb{R}$ and $A_{ij} : \Omega \rightarrow \mathbb{R}$ 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 $\sum_{i=1}^n a_i(\omega) = 1$ and $\sum_{i=1}^n \sum_{j=1}^n A_{ij}(\omega) = 1$ for every $\omega \in \Omega$.

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

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

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.

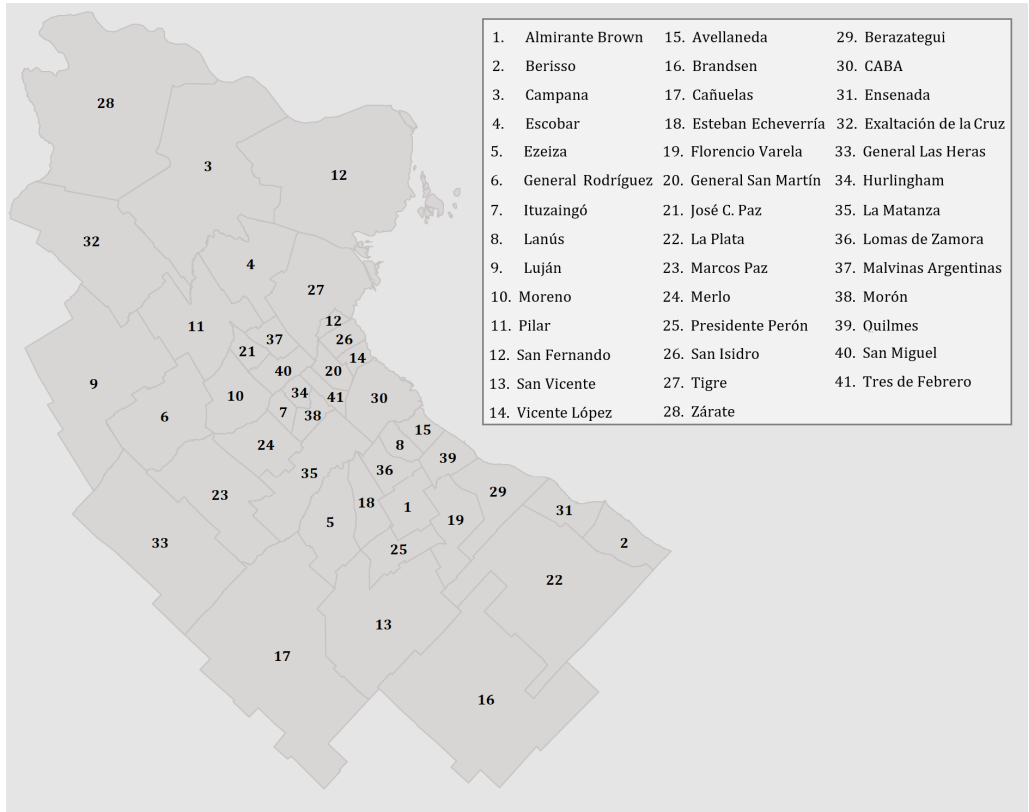


Figure 1: A map of the 41 cities of AMBA. Buenos Aires city (CABA) has the label 30.

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 $\mathcal{G}(\omega) = (\mathcal{V}, \mathcal{E}, \vec{a}(\omega), \vec{\bar{A}}(\omega))$ with $\mathcal{V} = \{1, 2, \dots, n\}$, $\mathcal{E} = \{\{i, j\} : i, j \in \mathcal{V}\}$, $\vec{a}(\omega) = (a_i(\omega) : i = 1, \dots, n)$, $\vec{\bar{A}}(\omega) = (A_{ij}(\omega) : i, j = 1, \dots, n)$ with each $a_i(\omega)$ and each $\bar{A}_{ij}(\omega)$ in $L^1(\Omega, \mathcal{F}, \mathcal{P})$. We shall also assume the probabilistic normalizations

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

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

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

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

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

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

$$\Delta_\Gamma = \vec{a}^{-1} (\vec{A} - \vec{D})$$

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

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

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

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

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

$\omega \in \Omega$, $i \in \mathcal{V}$. 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*

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

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

Proof. Since we are assuming the finiteness of $\int_{\Omega} \frac{1}{a_i(\omega)} d\mathcal{P}(\omega)$ 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 $\sum_{j=1}^n A_{ij}(\omega) (f(j) - f(i))$ for any $f : \mathcal{V} \rightarrow \mathbb{R}$. Hence

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

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*

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.*

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

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

$$\Delta_\Gamma \phi_i = \lambda_i \phi_i, \quad \text{for } i = 0, 1, \dots, n-1;$$

d) for any $t > 0$, the function $d_t : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$ 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 $B_{ij} = B_{ji}$,

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

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

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

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 have that $\lambda_0 = 0$ is the eigenvalue corresponding to the eigenfunction $\phi_0(i) = \left(\sum_{j=1}^n b_j\right)^{-1/2}$ for $i = 1, \dots, n$, which has the L^2 norm given by the inner product $\langle \cdot, \cdot \rangle_{\bar{b}}$ equal to one.

d) it is clear that d_t is nonnegative, symmetric, faithful and satisfies the triangle inequality for every $t > 0$. Let us 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 $\frac{\partial u}{\partial t} = \Delta_{\Gamma} u$. □

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.

Our basic vertex set is $\mathcal{V} = \{1, \dots, 41\}$ 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. Precisely, the unnormalized matrix A^3 is given by A^3_{ij} equals the product of the length of the shared boundaries times 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 $\vec{a}_1 = (\frac{1}{41}, \dots, \frac{1}{41})$ and a normalization of the density of the disease in each location (total number of active infections over population) by July 2020, given by

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

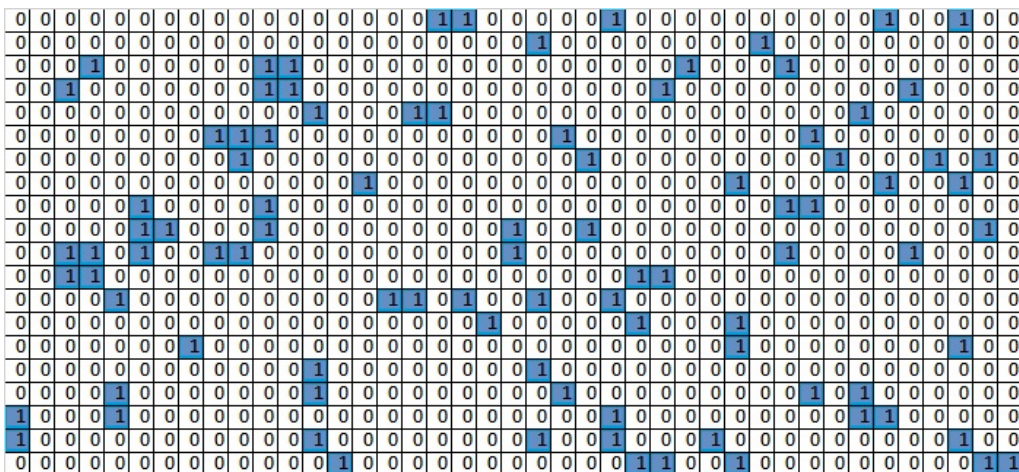


Figure 2: Unnormalized 20×41 submatrix of A^1 , the adjacency matrix provided by the neighborhood relation $A^1_{ij} = 1$ when cities i and j share points of their boundaries.

0	7	0	19	808	10	16	6223	5	54	24	45	1843	417	3745	55	53	4183	3551	135	22	482	12	60	5915	184	81	2	612	34339	4	0	1	30	1048	20792	32	206	8355	35	163				
7	0	0	3	0	7	0	2	1	2	11	2	11	2	11	2	0	5	31	0	0	0	7105	0	1	0	0	1	0	40	258	352	0	0	6	12	2	1	48	0	0				
0	0	0	846	0	24	1	31	31	227	17	0	1	12	30	4	5	3	0	37	211	1103	1	433	0	120	0	3	12	5	60	4	1	13	1	0	0	0	0	0	0	0	0		
19	0	846	0	6	103	23	15	29	618	5626	529	1	1635	498	1	1003	2531	222	36	12	44	4	62	3	2047	5772	122	3	4235	0	90	0	67	90	100	2631	72	16	982	102				
808	3	0	6	0	2	14	1021	1	39	8	16	178	76	498	1	1003	2531	222	36	12	44	4	46	298	54	22	2	64	7937	5	0	9	988	2685	11	49	214	12	71	2	203	146		
10	0	24	103	2	0	150	12	991	4470	1316	37	0	123	18	0	1	13	4	98	427	4	23	544	0	115	111	7	10	2198	0	26	2	130	220	31	66	267	4	203	146				
16	0	1	23	14	150	0	31	20	2938	64	28	3	142	22	1	41	11	140	104	11	81	4467	4	199	65	0	5	6070	0	1	885	1482	78	65	949	1166	40	99	4146	39	4284			
6223	7	2	15	1021	12	31	0	2	74	30	35	515	321	10006	14	60	252	1898	146	37	264	9	142	720	165	86	1	486	39521	7	3	23	971	19166	40	99	4146	39	4284					
5	0	31	29	1	991	20	0	609	266	4	0	16	8	0	0	4	3	126	0	18	18	8	1	949	0	104	1	53	38	6	15	55	38	6	15	55	38	6	15	55	38	6	15	55
54	2	31	618	39	4470	2938	74	609	0	3242	620	4	1727	115	0	9	63	28	1578	5041	40	195	7911	0	1666	2172	3	32	22871	2	35	27	885	2727	214	1309	5152	52	7601	2064				
45	1	227	5626	8	1316	64	30	266	3242	0	309	4	2188	44	0	1	25	29	430	3588	24	10	277	1	1719	2481	20	10	7229	1	493	0	227	197	77	4031	247	32	1676	430				
24	0	17	529	16	37	28	4	620	309	0	2	2223	48	0	0	36	43	271	424	20	5	74	5	7862	9176	5	0	21	5958	1	0	4	78	966	1	7	104	0	18					
1843	1	0	1	178	0	3	515	0	4	2	0	28	298	368	12	321	65	8	2	54	0	6	936	9	7	0	39	2679	2	0	0	4	78	966	1	7	104	0	18					
417	2	27	1635	76	123	142	321	16	1727	2188	2223	28	0	249	0	10	313	284	6142	1843	85	22	647	68	10087	4845	13	129	44644	5	10	3	217	1738	1071	5817	265	327	1407	1061				
3745	11	1	15	498	18	22	10406	8	115	44	48	236	249	0	7	24	988	2515	129	51	670	4	155	364	144	66	3	1994	36923	11	0	0	31	527	3538	49	106	10313	45	144				
55	2	0	1	0	0	14	0	0	0	0	0	368	0	7	0	0	5	15	0	103	0	1	30	0	0	0	6	57	0	0	0	2	35	1	0	1	0	2	0	0	0			
53	0	1	0	1003	1	2	60	0	9	1	0	12	10	24	0	0	307	10	2	0	10	35	11	18	3	2	0	5	549	0	0	1	954	161	0	13	23	0	19					
4183	5	1	22	7551	13	45	2292	4	63	25	36	321	313	988	9	307	0	460	158	25	146	8	130	299	188	84	2	218	16876	1	0	82	3783	11399	26	295	572	23	181					
3551	31	1	14	222	4	11	1898	1	28	29	43	65	284	2515	15	10	460	0	92	19	2509	1	26	72	179	60	1	6772	2192	46	0	23	403	1960	27	87	11226	26	119					
135	0	12	252	36	98	140	146	20	1578	430	271	8	6142	129	0	2	158	92	0	1295	50	24	615	30	6059	1615	12	50	26014	1	0	947	1512	553	822	1010	103	1430	12710					
22	0	30	1455	12	427	104	37	108	5041	3588	424	2	1843	51	0	0	25	19	1295	0	17	9	163	0	1314	3286	12	6	10500	0	25	0	685	284	86	4788	712	26	11265	1595				
462	7105	4	14	44	4	11	264	5	40	24	20	54	85	670	103	10	146	2059	50	17	0	2	50	58	55	47	2	3040	8465	4337	0	2	17	195	455	46	88	2410	21	43				
12	0	5	3	4	23	81	9	3	195	10	5	0	22	4	0	35	8	1	24	9	2	0	2196	1	17	10	0	3	1101	0	0	123	41	646	13	3	257	1	15	92				
60	1	3	62	46	544	4467	142	126	7911	277	74	6	647	155	1	11	130	26	615	163	50	2196	0	11	436	235	0	32	26468	0	4	139	671	8217	293	308	12490	63	311	1672				
5915	0	0	3	298	0	4	720	0	7	1	5	936	68	364	30	18	299	72	30	0	58	1	11	0	39	20	0	39	5330	1	0	0	94	1273	5	18	136	7	35					
184	0	37	2047	54	115	199	165	18	1666	1719	7862	9	10087	144	0	3	188	179	6059	1314	55	17	436	39	0	7521	6	74	17645	1	6	0	539	1017	506	3462	531	156	1060	859				
81	0	211	5772	22	11	65	86	18	2172	2481	9170	7	4845	66	0	2	84	60	1615	3286	47	10	235	20	7521	0	59	31	12837	1	18	0	310	519	394	6661	299	75	3068	441				
2	1	1103	122	2	7	0	1	8	3	20	5	0	13	0	2	0	12	12	0	0	0	0	6	59	0	2	465	0	2	465	0	0	5	6	10	2	1	4	6					
612	40	1	3	64	10	5	486	1	32	10	21	39	129	1594	6	5	218	6772	50	6	3040	3	32	39	74	31	2	0	14064	46	0	8	138	549	27	69	11521	10	65					
34539	238	433	4235	7337	2198	6070	39821	949	22871	7259	5958	2679	44844	36923	57	549	16876	2192	26014	10500	44064	1101	26468	5330	17843	12837	465	14064	0	132	98	47	7238	8724	45769	8749	21285	28156	13912	32448				
4	352	0	0	5	0	0	7	0	2	1	2	0	5	11	0	0	46	1	0	0	0	0	4337	0	0	0	0	0	0	0	0	4	9	1	0	46	0	0	0	0	0			
0	0	120	90	0	26	0	3	104	35	493	1	0	10	0	0	0	0	0	0	25	0	0	4	0	6	18	10	0	0	0	0	0	2	1	10	4	0	1	16	4				
1	0	0	0	0	2	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	0			
30	0	67	9	130	885	23	53	885	227	34	4	217	31	0	1	82	23	947	685	17	41	671	0	539	310	0	0	0	0	0	0	0	47	0	0	0	0	17	1	0	3			
1048	6	12	90	958	220	1482	971	38	2727	197	197	73	1738	527	2	954	3783	403	1512	284	199	646	8217	94	1017	519	5	198	3527	4	1	11	1090	0	3635	176	1829	476	473	8905				
20782	12	0	100	2885	6	210	18166	6	214	77	124	966	1071	3538	35	161	11399	1960	553	455	13	293	1273	506	394	6	549	45769	9	1	0	75	3635	0	123	351	5449	96	1019					
32	2	60	2631	11	66	65	40	15	1309	4031	622	1	5817	49	1	26	27	822	4783	46	3	308	5	3462	661	10	27	8749	1	10	0	237	176	123	0	291	25	7029	414					
206	1	4	72	49	267	6420	99	55	5152	247	63	7	265	106	0	13	295	87	1010	712	88	257	12490	18	531	299	2	69	21285	0	4	17	5857	8829	351	291	0	108	1696	4405				
8335	48	1	16	214	4	18	4146	6	52	32	45	104	327	10313	1	23	57	11126	103	26	2410	1	63	136	156	75	1	11521	28156	46	1	14	476	5449	25	108	0	31	129					
35	0	13	982	12	203	429	39	41	7601	1676	529	0	1407	45	0	2	26	1430	11265	21	15	311	7	1060	3068	4	10	13912	0	16	0	2567	473	96	7029	1696	31	0	2408					
163	0	1	102	71	146	484	226	10	2064	430	98	18	1061	144	2	19	181	119	12210	1595	43	92	1672	35	859	441	6	65	32648	0	4	3	1604	8905	1019									

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	0	18,3	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	0	0	11,1	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 20×20 submatrix of A^2 , the adjacency weighted matrix provided by the length of the shared portions of the boundaries of the two cities.

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	793751	4272745	0	0	0	0	0	0	0	0	0
0	0	3379775	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	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	0	3492642	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	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	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 20×20 submatrix of A^3 , the weighted matrix provided by the product of the lengths of the shared boundaries times the minimum of their population.

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 $d_i^{i,\theta;j}$ to denote the metric provided

0	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	0	105552	0	0	0	0	0	0	105552	105552	0	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	0	77161	0	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	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	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	0	109695	180914	0	0	0	378167	0	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	0	105552	174883	0	0	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	0	267655
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	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	62921	0	0	0	0	0	0	62921	0	0	0	0	0	0	0	0
370900	0	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	0	0	267655	0	0	0	0	0	0

Figure 6: Unnormalized 20×20 submatrix of A^4 , the matrix provided by the minimum of the population of any two neighboring cities.

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

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

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

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 Figure 7 and Figure 8 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.

Some global comparison of the different metrics are in order. In Table 1 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 \vec{a}_1 , the uniform distribution ($a_i = \frac{1}{41}$) of the vertices of the graph. Here we compute the relative deviations with respect to the metric induced just by public transport. Let us precise the above. Set

$$\varepsilon_i^{i,\theta} = \frac{\|d_i^{0,0;1} - d_i^{i,\theta;1}\|}{\|d_i^{0,0;1}\|},$$

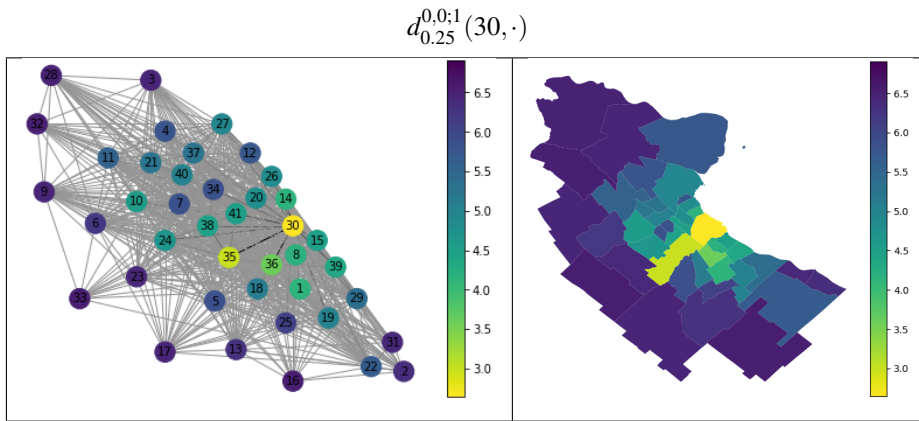


Figure 7: Diffusion distances to CABA (30) for $t = 0.25$ with affinity matrix A^0 and node uniform weights given by \vec{a}_1 .

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 1: 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 1 we observe that, as it could be expected and as it reflected by the colored maps in Figure 8, 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 .

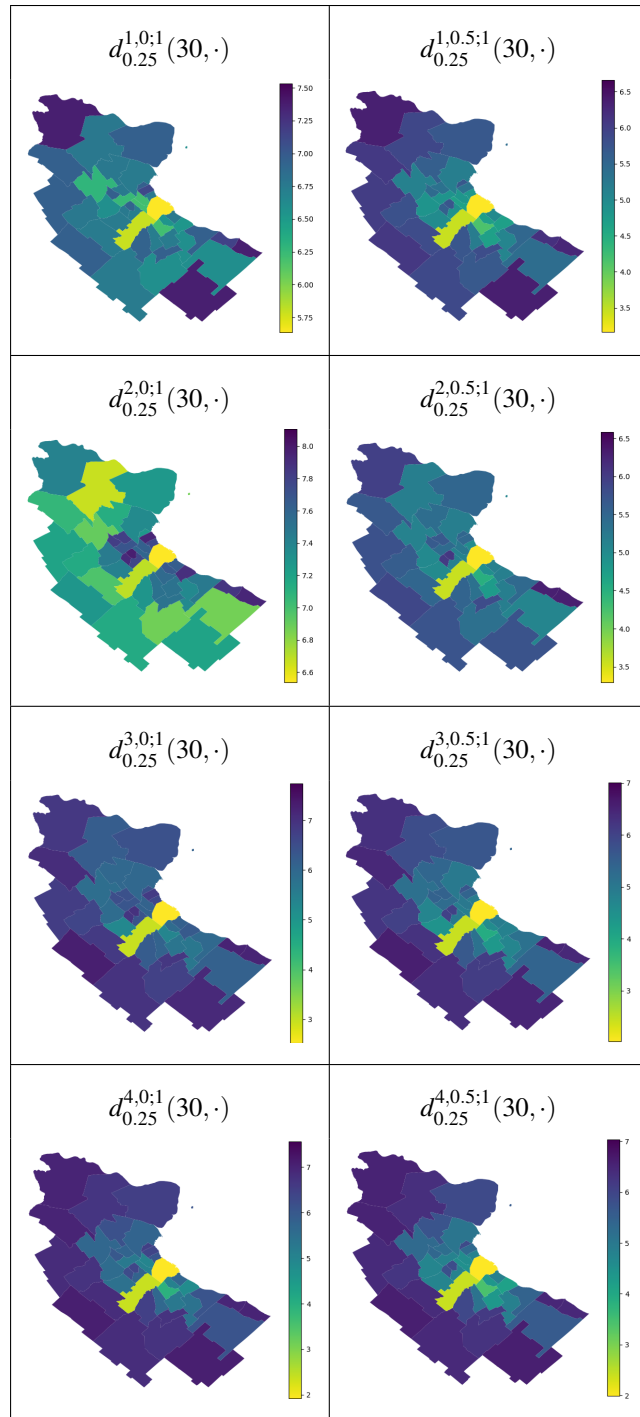


Figure 8: Diffusion distances to CABA (30) for $t = 0.25$ in eight different instances of affinities $(\theta A^0 + (1 - \theta)A^i$ with $i = 1, 2, 3, 4$ and $\theta = 0, \frac{1}{2}$) and node uniform weights given by \vec{a}_1 .

4 DISCUSSION

As we show in Section 3 all the above considered versions of the diffusive metric, provide in some way a measure of closeness of any given pair of the cities of AMBA. These metrics take into account some classical notions of proximity such as neighboring and size of the shared boundaries.

Nevertheless, each of the above considered metrics take into account the public transportation in AMBA as central contribution to their definitions. Let us notice that in any of the above considered metrics the cities of Buenos Aires and La Matanza, labeled 30 and 35 respectively can be considered as a urban unity of $3075000 + 2280000 = 5355000$ people. They share 10 kilometers of densely populated boundaries, and they have an intense people traffic through public transportation by buses and trains. The above statement can be seen in Figure 7 and the eight maps in Figure 8 that show quite close colors for the cities 30 and 35. In what follow we shall contrast these, let us say, purely geometrical considerations with the actual dynamics of the spread of COVID-19 taken from public data in [4]. We shall provide two different approaches for this comparison.

First, 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 $x = j \cdot \frac{1}{10}$, $j = 1, 2, \dots, 20$.

The maps provided by the data are of the type depicted in Figure 9

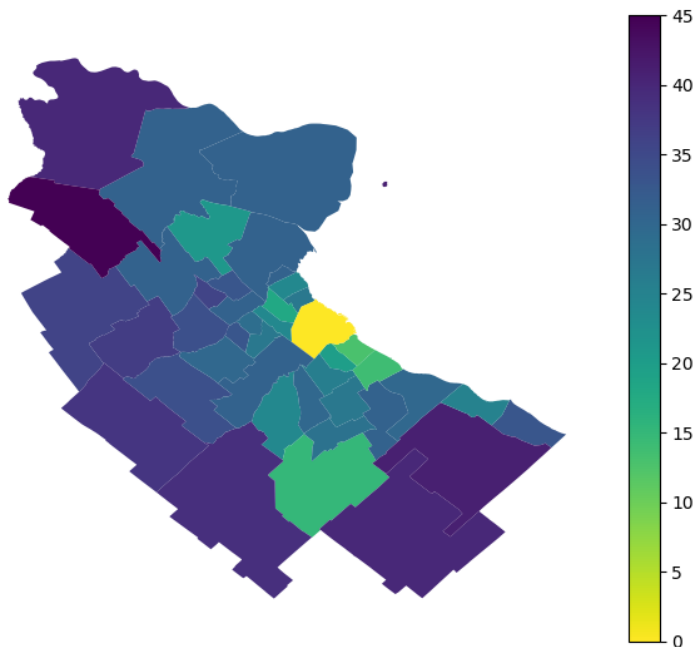


Figure 9: Days up to 0.1% of infections over the population (from 0.1% of CABA).

and Figure 10 for $j = 1$ and $j = 3$ respectively.

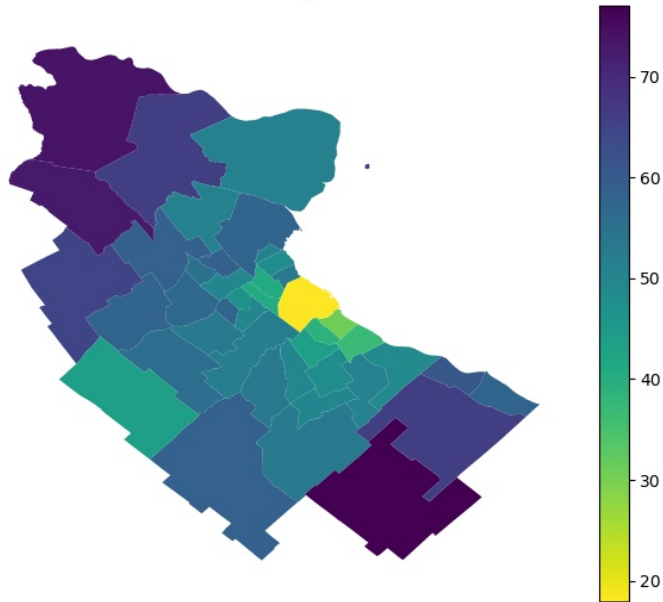


Figure 10: Days up to 0.3% of infections over the population (from 0.1% of CABA).

Second, if we measure, for fixed dates, the amount of the total infections normalized by the population of each city we obtain the patterns depicted in Figure 11.

We observe that in the six instances in Figure 11, we are using the scale of colors in such a way that, the cities with high density are depicted with the lower frequencies.

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 data regarding the spread of the pandemic in AMBA. In fact while for CABA we have the red distribution as a function of time in Figure 12, for La Matanza we have the blue one.

This lack of coincidence in the dynamics of these two large cities that share a big portion of their boundaries is certainly multicausal but it can be reasonably attributed to the government decisions regarding the social preventive isolation starting on March 20th, 2020, which in particular produced a drastic reduction of the public transportation of people in AMBA. Also some consideration has to be paid to the difference of population densities of the two largest cities of AMBA: CABA 15150 inhabitants per km^2 and La Matanza 7062 inhabitants per km^2 .

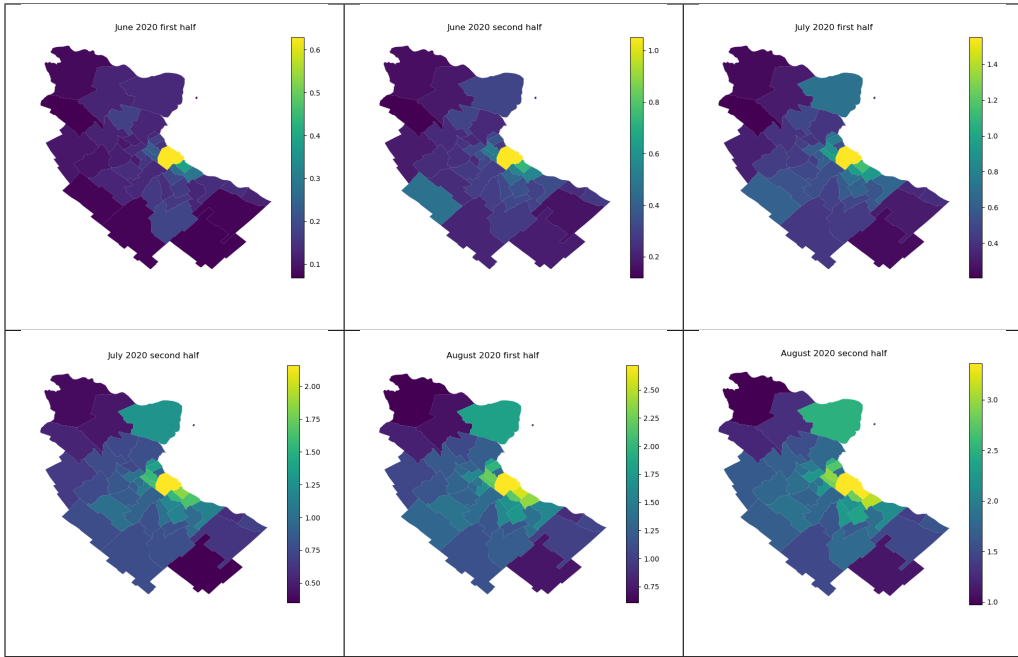


Figure 11: Percentage of total infections for the first and second halves for June, July and August 2020.

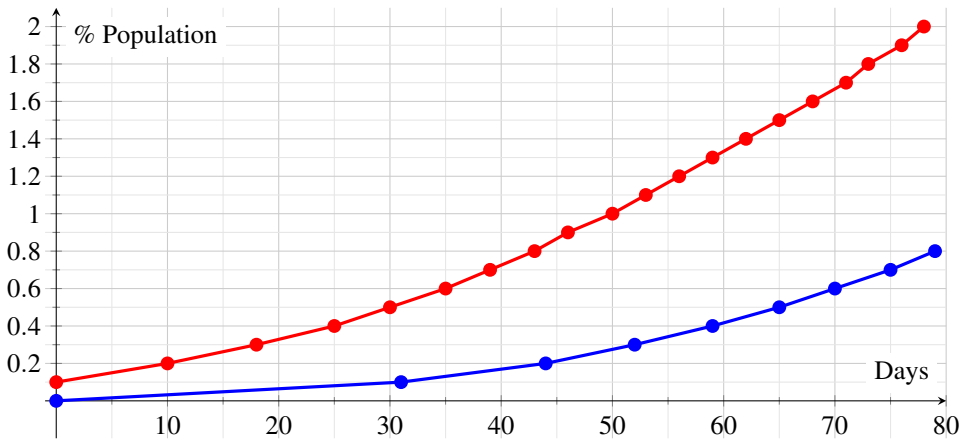


Figure 12: Evolution of cases in CABA (red) and La Matanza (blue).

Let us finally observe that among the several papers dealing with the issue of COVID and people transportation, we were unable to find the application of diffusive metrics. Neither other quantitative methods for the particular case of AMBA. Some graph based models are used for example in [5] and [6].

Acknowledgments

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 Universidad Nacional del Litoral (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.

Code availability: The Python code to create the different metric maps is openly available in the repository <https://github.com/LABRAimal/DiffusiveMetricsAMBA>

REFERENCES

- [1] R. R. Coifman & S. Lafon. Diffusion maps. *Applied and Computational Harmonic Analysis*, **21**(1) (2006), 5–30.
- [2] M. F. Acosta, H. Aimar & I. Gómez. On Frink’s type metrization of weighted graphs. *Asian Research Journal of Mathematics*, **17** (2021), 26–37.
- [3] M. M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, & P. Vandergheynst. Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine*, **34**(4) (2017), 18–42.
- [4] Argentina Public COVID-19 data. Available at <https://covidstats.com.ar>.
- [5] D. Brockmann & D. Helbing. The hidden geometry of complex, network-driven contagion phenomena. *Science*, **342** (2013), 1337.
- [6] Y. Murano, R. Ueno, S. Shi, T. Kawashima, Y. Tanoue, S. Tanaka, S. Nomura, H. Shoji, T. Shimizu, H. Nguyen, H. Miyata, S. Gilmour & D. Yoneoka. Impact of domestic travel restrictions on transmission of COVID-19 infection using public transportation network approach. *Sci Rep* **11** (2021), 3109.

