**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 \( \mathcal{V} = \{1, 2, \ldots, n\}, n \geq 1 \) be the set of vertices of the graph \( \mathcal{G} = \left( \mathcal{V}, \mathcal{E}, \vec{a}, \vec{A} \right) \), where \( \mathcal{E} = \{\{i, j\} : i, j \in \mathcal{V}\} \) is the set of all edges, \( \vec{a} = (a_1, a_2, \ldots, a_n) \) is the sequence of positive weights of the vertices and \( \vec{A} = (A_{ij}) \) is the matrix of non-negative weights of the edges. Assume also that \( A_{jj} = 0 \) for every \( j = 1, \ldots, n \). 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 \( \Omega \) with \( \mathcal{V} \) and \( \mathcal{E} \) fixed. So that \( \mathcal{G}(\omega) = \left( \mathcal{V}, \mathcal{E}, \vec{a}(\omega), \vec{A}(\omega) \right) \) with \( \vec{a} : \Omega \rightarrow \mathbb{R}^n \) a random vector with positive components and \( \vec{A} : \Omega \rightarrow \mathbb{R}^{n \times n} \) a random matrix with non-negative
entries, with \( A_{ij} = 0 \) and \( A_{ji} = A_{ij} \). 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 \( L^1(\Omega, \mathcal{F}) \), 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 \( \mathbb{E}(G) = \left(V, \mathcal{E}, \mathbb{E}(\vec{a}), \mathbb{E}(\vec{A})\right) \), with \( \mathbb{E}(\vec{a}) = (\mathbb{E}a_1, \ldots, \mathbb{E}a_n) \), and \( \mathbb{E}(\vec{A}) = (\mathbb{E}A_{ij} : i, j = 1, \ldots, n) \). 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 \( G(\omega) \) and the analysis provided by the graph \( \mathbb{E}(G) \). In this paper we focus on building a metric, by the diffusion method given in [1], on the graph \( \mathbb{E}(G) \). 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 $G$ defined in $\Omega$ with values on the simple undirected weighted graphs on $V = \{1, 2, \ldots, n\}$, is a random graph on $V$ with finite first moments if $G(\omega) = (V, E, \vec{a}(\omega), \bar{A}(\omega))$ with $V = \{1, 2, \ldots, n\}$, $E = \{\{i, j\}: i, j \in V\}$, $\vec{a}(\omega) = (a_i(\omega): i = 1, \ldots, n)$, $\bar{A}(\omega) = (A_{ij}(\omega): i, j = 1, \ldots, n)$ with each $a_i(\omega)$ and each $A_{ij}(\omega)$ 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 $i \in V$ and $A_{ij}(\omega) \geq 0$ for $i, j \in V$ and $\omega \in \Omega$.

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

**Proposition 2.1.** Let $G(\omega)$ and $\mathbb{E}G$ 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 \( \Omega \). In other words

\[
\Omega = \bigcup_{k \in \mathbb{Z}} \Omega_k, \quad \Omega_k \cap \Omega_{k'} = \emptyset.
\]

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

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

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

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 \( \Gamma = (\mathcal{V}, \mathcal{E}, \bar{d}, \bar{A}) \) the Laplacian on \( \Gamma \) is given by

\[
\Delta \Gamma f(i) = \frac{1}{a_i} \sum_{j=1}^{n} A_{ij} (f(i) - f(j))
\]

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

\[
\Delta \Gamma = \bar{d}^{-1} \left( \bar{A} - \bar{D} \right)
\]

with \( \bar{d}^{-1} = \text{diag} (a_1^{-1}, \ldots, a_n^{-1}) \) and \( \bar{D} = \text{diag} (\sum_{j \neq 1} A_{1,j}, \ldots, \sum_{j \neq n} A_{n,j}) \).

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

**Proof.** Since we are assuming the finiteness of $\int_{\Omega} a_i(\omega) \, d\mathcal{P}(\omega)$ 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 $f : \mathcal{V} \to \mathbb{R}$. Hence

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

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

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

$$
= \Delta_{\mathcal{G}} f(i),
$$

as desired. \hfill \Box

Once we have a Laplacian defined on $(\mathcal{V}, \mathcal{E})$ which could be $\Delta_{\mathcal{E}\mathcal{G}}$ 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 $\Delta_\Gamma$ is selfadjoint with respect to the inner product

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

b) the operator $\Delta_\Gamma$ is negative definite, i. e.

$$
\langle \Delta_\Gamma f, f \rangle_b \leq 0, \quad \text{for every } f;
$$
c) the operator \( \Delta_{\Gamma} \) 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 \( \langle \cdot, \cdot \rangle_{\bar{b}} \) such that

\[
\Delta_{\Gamma} \phi_i = \lambda_i \phi_1, \quad \text{for} \quad i = 0, 1, \ldots, n-1;
\]

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

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

is a metric on \( \mathcal{V} \).

\[\text{Proof.}\]

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

\[
\langle \Delta_{\Gamma} f, g \rangle_{\bar{b}} = \sum_{i=1}^{n} (\Delta_{\Gamma} f)(i)g(i)b_i
\]

\[
= \sum_{i=1}^{n} \left( \sum_{j=1}^{n} B_{ij}(f(j) - f(i)) \right) g(i)b_i
\]

\[
= \sum_{j=1}^{n} \sum_{i=1}^{n} B_{ij}(f(j) - f(i))g(i)
\]

\[
= \sum_{j=1}^{n} \sum_{i=1}^{n} B_{ij}f(j)g(i) - \sum_{i=1}^{n} \sum_{j=1}^{n} B_{ij}f(i)g(i)
\]

\[
= \sum_{i=1}^{n} \sum_{j=1}^{n} B_{ij}f(j)g(i) - \sum_{i=1}^{n} \sum_{j=1}^{n} B_{ij}f(i)g(i)
\]

\[
= \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{1}{b_i} \sum_{j=1}^{n} B_{ij}(g(j) - g(i)) f(i)b_i
\]

\[
= \langle f, \Delta_{\Gamma} g \rangle_{\bar{b}}.
\]
b) Since $B_{ij} = B_{ji}$ we have
\[
\langle -\Delta \Gamma f, f \rangle_{\bar{b}} = \sum_{i=1}^{n} (-\Delta \Gamma f)(i) f(i) b_i
\]
\[
= \sum_{i=1}^{n} \sum_{j=1}^{n} B_{ij} (f(i) - f(j)) f(i)
\]
\[
= \sum_{i=1}^{n} \sum_{j=1}^{n} B_{ij} f^2(i) - \sum_{i=1}^{n} \sum_{j=1}^{n} B_{ij} f(i) f(j)
\]
\[
= \sum_{i=1}^{n} \sum_{j=1}^{n} B_{ij} (f^2(i) - f(i) f(j))
\]
\[
= \frac{1}{2} \left[ \sum_{i=1}^{n} \sum_{j=1}^{n} B_{ij} (f^2(i) - f(i) f(j)) + \sum_{i=1}^{n} \sum_{j=1}^{n} B_{ij} (f^2(i) - f(i) f(j)) \right]
\]
\[
= \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} B_{ij} (f^2(i) + f^2(j) - 2 f(i) f(j))
\]
\[
= \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} B_{ij} (f(i) - f(j))^2
\]
\[
\geq 0.
\]

c) follows from a) and b) since we are dealing with a self-adjoint and negative definite matrix $\Delta \Gamma$. Since the constant functions are $\Delta \Gamma$-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 $i = 1, \ldots, n$, 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 us notice here the $d_t(i, j)$ is the $L^2(V, \bar{b})$ 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 $\mathcal{V} = \{1, \ldots, 41\}$ 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 into 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 $41 \times 41$ 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 $A^3$ 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 $\vec{a}_1 = \left( \frac{1}{41}, \ldots, \frac{1}{41} \right)$ and a normalization of the density of the disease in each location (total number of active infections over population) by July 2020, given by

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

![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.](image-url)
M. F. ACOSTA, H. AIMAR, I. GÓMEZ and F. MORANA

Figure 3: The full unnormalized matrix $A$ built with the data of SUBE provided by the public transport system in AMBA.
Figure 4: Unnormalized $20 \times 20$ submatrix of $A^2$, the adjacency weighted matrix provided by the length of the shared portions of the boundaries of the two cities.

Figure 5: Unnormalized $20 \times 20$ 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 $\mathcal{V} = \{1, 2, \ldots, 41\}$ 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 $d^{i, \theta, j}$ to denote the metric provided...
Figure 6: Unnormalized $20 \times 20$ 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 $b = \bar{a}_j$. We shall use the standard notation for balls keeping the above notation, precisely

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

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 $\bar{a}_1$, the uniform distribution ($a_i = 1/41$) 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

$$\varepsilon^{i,\theta}_t = \frac{d^{0,0;1}_t - d^{i,\theta;1}_t}{d^{0,0;1}_t},$$
Figure 7: Diffusion distances to CABA (30) for $t = 0.25$ with affinity matrix $A^0$ and node uniform weights given by $\bar{a}_1$.

where $d^{0,0.1}_t$ is the metric matrix associate to the public transport only and $d^{i,\theta,1}_t$ are the metrics defined above. The norm considered here is the Euclidean one, i.e.

$$\left\| d^{i,\theta,1}_t - d^{i,0.1}_t \right\|^2 = \sum_{k,\ell=1}^{n} \left( d^{i,0.1}_t(k, \ell) - d^{i,\theta,1}_t(k, \ell) \right)^2$$

and

$$\left\| d^{0,0.1}_t \right\|^2 = \sum_{k,\ell=1}^{n} \left( d^{0,0.1}_t(k, \ell) \right)^2.$$  

Table 1: Relative differences.

<table>
<thead>
<tr>
<th>$\varepsilon^{1.0}_t$</th>
<th>$\varepsilon^{1.0.5}_t$</th>
<th>$\varepsilon^{2.0}_t$</th>
<th>$\varepsilon^{2.0.5}_t$</th>
<th>$\varepsilon^{3.0}_t$</th>
<th>$\varepsilon^{3.0.5}_t$</th>
<th>$\varepsilon^{4.0}_t$</th>
<th>$\varepsilon^{3.0.5}_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.12035607</td>
<td>0.0609088</td>
<td>0.17173178</td>
<td>0.091446</td>
<td>0.0644136</td>
<td>0.0306021</td>
<td>0.09062579</td>
<td>0.04661433</td>
</tr>
</tbody>
</table>

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 \(x = j \cdot \frac{1}{10}, j = 1, 2, \ldots, 20\).

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).](image-url)
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


