Forecasting the COVID-19 Space-time Dynamics in Brazil with Convolutional Graph Neural Networks and Transport Modals

Lucas Caldeira de Oliveira[1], **Marcelo Teixeira**[1], **Dalcimar Casanova**[1]

[1]Universidade Tecnológica Federal do Paraná
Pato Branco, PR – Brazil

lucasoliveira.2017@alunos.utfpr.edu.br
{mtex,dalcimar}@utfpr.edu.br

Abstract. *This study presents a novel scalable method to forecast the numbers of cases and deaths by SARS-CoV-2 according to the influence that certain (micro) regions exert on others, predicting for specific regions while generalizing for general extents. By exploiting graph convolutional networks with recurrent networks, our approach maps the main access routes to municipalities in Brazil using the modals of transport, and processes this information via neural network algorithms. We compared the performance in forecasting the pandemic daily numbers with three baseline models, with the forecasting horizon varying from 1 to 25 days. Results show that the proposed model overcomes the baselines, being specially suitable for forecasts from 14 to 24 days ahead.*

1. Introduction

Since the appearance of the SARS-CoV-2 pandemic, in December 2019, many researches have been conducted, with epidemiological control playing a central role as it allows to project future scenarios and support predictive containment policies. However, planning when, where, and for how long to apply each action is only possible upon monitoring and understanding how the pandemic spreads and evolves [Han et al. 2020, da Silva et al. 2021], which have shown to be very hard.

The literature provides ways to predict the number of cases or deaths caused by the Covid-19, such as by exploiting the theoretical background of statistical methods [Espinosa et al. 2020, Siqueira et al. 2020], or by applying Machine Learning methods [Lima et al. 2022, Da Silva et al. 2020, Pereira et al. 2020]. These approaches focus essentially on anticipating the numbers for macro-regions [da Silva et al. 2021], such as states or countries. Although this benefits general planning, it is ineffective to capture the internal features of smaller regions, such as municipalities, which face peculiar and heterogeneous situations. Moreover, they do not fully model the dynamics of contagions since the virus spreading at the local level depends not only on its past state but also on the neighborhood context. In this sense, it is essential to consider that every region is influenced, to some extent, by its past state and the connections it may have, and that the choices made in one place influence the surroundings.

Recent studies about influenza A (H1N1) suggest that the speed in which this type of virus spreads is related to the road map that interconnects different places to the virus epicenter [Xu et al. 2019]. Similarly, distinct transport modals may play different roles

in the virus dissemination [Cai et al. 2019]. Therefore, it is reasonable to suppose that similar conditions also apply to the case of Covid-19 spreading, enforcing the relevance of considering the spatial and temporal relationships.

This study proposes a *Graph Neural Network* (GNN)-based single-model architecture to forecast the daily spatio-temporal dynamics of the pandemic, in terms of cases and deaths by SARS-CoV-2, according to the influence that certain regions exert on others. Our proposal consists in mapping the main access routes to the Brazilian cities by four transport modals: roads, railways, waterways, and air; and processing it by a convolutional GNN integrated with a recurrent network with an input time window.

2. Problem Statement

To model the problem, we consider a weighted undirected graph to represent the relationship among the 5,570 Brazilian municipalities and their particular features. Municipalities and their features are vertices, and the existing connections among them are edges. The difficulty behind setting the dataset relies on defining an appropriate neighboring criterion to connect municipalities. Human mobility should be a natural option, but this is not so easy to be established. Thus, we opted by defining mobility based on the main transport modals available between locations.

Two databases from the *Brazilian Institute of Geography and Statistics* (IBGE) were integrated: a cartographic map [IBGE 2019], and a logistics map [IBGE 2014]. From them, four connectivity meshes were constructed: roadway, railway, waterway, and air routes. Table 1 shows the coverage map.

Table 1. Municipalities covering of the meshes.

Mesh	Initials	Edges	Covering	
Roadways	ROD	12186	5554 mun.	99.71 %
Railways	FER	1272	1209 mun.	21.71 %
Waterways	HID	1240	925 mun.	16.61 %
Air routes	AER	876	149 mun.	2.68 %

2.1. Feature selection

The time series of cases and deaths by Covid-19 were collected in the Brasil.io website [Brasil.io 2020]. They support daily bulletins informed by state health departments, per municipality, since 02/25/2020, when Brazil recorded its first contagion. In this work, records of imported cases were discarded, and we do not extend the analysis to records beyond January 25th, 2021 (336th pandemic day), because after that the vaccination of priority groups has started, possibly causing a concept drift [SE/UNA-SUS 2021].

As high concentration of people tends to accelerate the virus spreading, we add the demographic density (inhabitants per km^2) to the dataset, which is calculated based on the estimated population per municipality in 2020 [IBGE 2019]. Another relevant parameter is the municipal capacity to care for sick people, which we measured by the total number of hospital beds installed by December 2019 [IBGE 2020].

2.2. Graph composition

Consider a weighted undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{X})$ in which spatial relations are expressed by the set of edges $\mathcal{E}$, the municipalities are listed as a set of vertices $\mathcal{V}$, and

the local attributes (including time series) are organized as a set of features $\mathcal{X}$. In these structures, a graph $\mathcal{G}$ can be represented simply as adjacency $\mathbf{A}$ and feature $\mathbf{X}$ matrices.

For the feature matrix, we further split the vertices attributes into two categories. The static attributes include population, demographic density, and number of hospital beds, and are modeled by $\mathbf{X}_s \in \mathbb{R}^{|\mathcal{V}|\times 3}$. The dynamic attributes are the time series of cases and deaths, modeled by $\mathbf{X}_d(t) \in \mathbb{R}^{|\mathcal{V}|\times 2}$, with $t = 1, 2, \dots 336$ e $|\mathcal{V}| = 5570$. Then, the *general feature matrix*, for a given time t, is defined as a concatenation of terms $\mathbf{X}_t = [\ \mathbf{X}_s \mid \mathbf{X}_d(t)\ ] \in \mathbb{R}^{|\mathcal{V}|\times 5}$.

The set of edges $\mathcal{E}$ can be defined as the union of the connectivity meshes from each road mode ($\mathcal{E}_{ROD}$, $\mathcal{E}_{HID}$, $\mathcal{E}_{FER}$, and $\mathcal{E}_{AER}$). The edge weights are given by a *proximity function* $f(i, j, m)$, which considers the distance between two adjacent municipalities i, j and the referred mesh m, since the physical distance is related to the delays in the spatial propagation of the virus [Xu et al. 2019] and different means of transport can play different roles in the pandemic [Cai et al. 2019]. Let $e_{i,j}^{(m)} \in \mathcal{E}_m$ be an weighted edge that connects two distinct vertices $v_i, v_j \in \mathcal{V}$ through the modal m, and c_m be a positive real number referred to m. The weight of this edge then is given by the proximity function $f(i, j, m) = c_m \cdot dist(v_i, v_j)^{-1/2}$. And, as the distance $dist(v_i, v_j)$ does not depend on m, we can generate a *global proximity matrix* $\mathbf{P}$: $p_{i,j} = dist(v_i, v_j)^{-1/2}$ if $i \neq j$, or 0 otherwise. For the binary adjacency matrix $\mathbf{A}^{(m)} \in \mathbb{N}^{|\mathcal{V}|\times|\mathcal{V}|}$, of a modal $m \in \{ROD, HID, FER, AER\}$, we define the weighted adjacency matrix $\hat{\mathbf{A}} \in \mathbb{R}^{|\mathcal{V}|\times|\mathcal{V}|}$ as the sum of the element-wise product between $\mathbf{P}$ and $\mathbf{A}^{(m)}$, *i.e.* $\hat{\mathbf{A}} = \mathbf{P} \odot \sum c_m \mathbf{A}^{(m)}$.

3. Proposed Methodology

In this section, we show how a *Spatial-Temporal Graph Neural Network* (STGNN), using convolutional GNN, can be used to design spatial and temporal dependence.

The model receives as input the general feature matrix $\mathbf{X}_{t-1}$ from the last time step and the weighted adjacency matrix $\hat{\mathbf{A}}$; aggregates the signals passed through the neighborhood in $\mathbf{M}_{t-1}$; and generates a *trend matrix* $\mathbf{Z}_{t-1} \in \mathbb{R}^{|\mathcal{V}|\times 2}$, which can be seen as a representation of the influence of regional situations in the municipalities. The functions $\mathcal{F}$ (Equation 1), and $\mathcal{G}$ (Equation 2), define the aggregation and update operators, respectively. As $\mathcal{G}$ considers the current local state $\mathbf{X}_{t-1}$ in order to update and set the new local state, the equations 1 and 2 define the *spatial recurrence stage*. The results of the last T days, i.e., $\mathbf{Z}_{t-1}, \cdots, \mathbf{Z}_{t-T}$, are then concatenated on the feature axis to form the expanded matrix $\mathbf{U}_{t-1}$ (Equation 3), with dimension proportional to the time window T. This matrix serves as input to the *temporal recurrence stage*, formed by a recurrent network (Equation 4) and a dense layer (Equation 5). The output is then the forecasted values for the next day, i.e., $\mathbf{Y}_t = \mathbf{X}_d(t)$.

$$\mathbf{M}_{t-1} = \mathcal{F}\left(\mathbf{X}_{t-1}, \hat{\mathbf{A}}\right) ; \tag{1}$$

$$\mathbf{Z}_{t-1} = \mathcal{G}\left(\mathbf{X}_{t-1}, \mathbf{M}_{t-1}\right) . \tag{2}$$

$$\mathbf{U}_{t-1} = [\ \mathbf{Z}_{t-1}\ \ \mathbf{Z}_{t-2}\ \ \dots\ \ \mathbf{Z}_{t-T}\ ] \in \mathbb{R}^{|\mathcal{V}|\times 2T} . \tag{3}$$

$$\mathbf{H}_t = rec\left(\mathbf{U}_{t-1}, \mathbf{H}_{t-1}\right) ; \tag{4}$$

$$\mathbf{Y}_t = dense\left(\mathbf{H}_t\right) . \tag{5}$$

Figure 1 shows the proposed STGNN model for one day ahead forecasting with a time window of T days. The inputs are the adjacency matrix and the feature matrices of T' past time steps, with $T' \geq T$ to allow gathering sufficient historical values. The output is the forecasting for the next day, and can be concatenated with the static feature matrix to retro-feed the model, when considering bigger forecasting horizons.

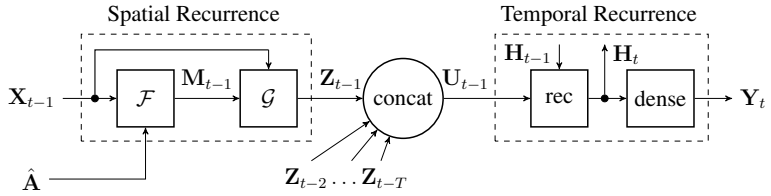

Figure 1. Schematic of the STGNN with time window.

3.1. Implementation details

The network layers are formed by a *GatedGraphConv* module [Li et al. 2017] with 1-hop in the spatial recurrence stage and one GRU module followed by an 3 layered *Multi Layer Perceptron* (MLP) in the temporal recurrence stage, using ReLU activation function in all hidden layers and Tanhshrink at the output one. This configuration, using convolutional GNN architecture and *Gated Recurrent Unit* (GRU) modules for recurrence in both spatial and temporal axes, was called *Double Gated GCN*. Also, the constant c_m was set to one, as this value showed to be the most stable during the training cycle.

The selected time window to compose the expanded matrix $\mathbf{U}_{t-1}$ was $T = 14$ days ($t-1, \ldots t-14$), which returned the best results among the tested range (1 to 21 days). This number corresponds to the average time of the events related with the disease: SARS-CoV-2 virus incubation period [Lauer et al. 2020] and time between the first symptoms and respective death [Galzo 2021, Garcia 2020, Ruan et al. 2020, Baud et al. 2020]. Also, this last one can still vary between 10 and 28 days, depending on the patient state, medical treatment conditions, and demographic features. It is important to highlight that the choice of the time window impacts on the network learning capacity.

The inputs $\mathbf{X}_d(t)$ and $\mathbf{X}_s$ were normalized by z-score, with an additional 2-day moving average smoothing at the time axis, in order to eliminate noisy signal components without altering the intrinsic periodical characteristics of the curves.

3.2. Model selection and evaluation

To perform the model selection and evaluation we employ 3-way time series split cross validation. This method is a variation of k-fold which returns first k-folds as train set ($\tau_{\text{train}}^{[k]}$) and the $(k+1)$ fold as test set. The size of each k-th training subset varies in size, starting with 66 days when $k = 1$, and reaching 291 days when $k = 10$. The validation ($\tau_{\text{valid}}^{[k]}$) and testing ($\tau_{\text{test}}^{[k]}$) subsets have fixed size of 20 and 25 days per fold, respectively.

As the performance of the models can be analyzed from different perspectives (by municipality, forecasting horizon, fold, state), it will be defined two strategies to compute the errors. First, we define the *error by region*: the average error through time computed for a given place or a set of places that share common characteristics (like being from

a same state), which evaluates the quality of the models to forecast the local curves, regardless of the pandemic period. Secondly, the *error by period*: the average error among the country computed for a given time interval (like the h-th day from a forecasting horizon), which evaluates the quality of the models to forecast a specific day, regardless of the place. The models are evaluated, by region and period, with the following metrics: *Mean Absolute Error* (MAE), *Symmetric Mean Absolute Percentage Error* (sMAPE), and *Normalized Root Mean Square Error by the Standard Deviation* ($\mathrm{NRMSE_{sd}}$). The selection of these metrics took into account the comparability with similar works in the literature and a more embracing analysis to exploit different points of views.

3.3. Baselines

Implementing dedicated forecast models (i.e., specific to each locality) requires exhaustive and impracticable computational processing due to the inherent time complexity. For example, the two time series from all municipalities require at least 11,140 distinct models. Furthermore, this counting grows much more when considering that the simple task of updating the model with a new known day retrains the model from the beginning, which implies that the training/validation cycle with ten folds trains the same amount of distinct models. Finally, tuning the hyper-parameters for all models with grid search methods or other techniques becomes impracticable due to the large of models that need to be built. For illustration purposes, we implemented the models *Naive Forecaster*, *Exponential Smoothing*, and *Autoregressive Integrated Moving Average* (ARIMA). Grid Search selected the hyper-parameters in some selected state capitals – those for which the pandemic was more intense and yet had more regular records without significant distortions.

3.4. Experiment setup

The neural network was trained with τ_{train} and evaluated with τ_{valid} to adjust the learning rate and determine the early-stopping point. Then, it was trained in subsets τ_{train} and τ_{valid} using the estimated hyper-parameters, and evaluated in the test subset τ_{test}. The baselines were trained directly with $\tau_{\text{train}} \cup \tau_{\text{valid}}$ and evaluated with τ_{test}, as they do not benefit from the validation step. Also, the baselines were fitted to the non-normalized data, and we had to sum up a constant before the data input to avoid inconsistencies with some models of the [Sktime 2020] library, which does not handle zeros. The training and validation steps use the ADAM optimizer with *Mean Square Error* (MSE) for the back-propagation step. For all models, the errors computation considers the non-normalized time series.

4. Results

This section presents the performance measures of proposed and compared models, by municipality size and forecasting horizon. Also, graphical analysis of the similarities of pandemic numbers in neighboring municipalities are shown.

4.1. Performance evaluation by municipality size

The analysis of the model performance by municipality size is presented in Table 2. The first column represents the size category, while the second shows the number of municipalities in that group. We classify a municipality as small size when having fewer than 15 thousand people, medium-sized when having a population between 15 and 150 thousand people, and large size if the population if equal to or greater than 150 thousand. The

Table 2. Performance of the models to forecast Covid-19 numbers for one-day-ahead, grouped by municipality sizes.

Size	$\|V\|$	Model	MAE		sMAPE		$\mathrm{NRMSE}_{\mathrm{sd}}$	
			cases	deaths	cases	deaths	cases	deaths
Large	201	ARIMA	47.146	1.527	0.325	**0.162**	0.685	**0.557**
		Double Gated GCN	**35.002**	**1.166**	0.318	0.203	**0.509**	0.561
		Expon. Smoothing	55.332	1.851	0.330	0.202	0.929	0.691
		Naive Forecaster	48.844	1.644	**0.308**	0.166	0.684	0.614
Medium	2119	ARIMA	4.515	0.135	0.274	**0.037**	0.919	0.961
		Double Gated GCN	**2.987**	**0.096**	0.295	0.067	**0.568**	**0.628**
		Expon. Smoothing	5.525	0.189	0.299	0.047	1.479	2.737
		Naive Forecaster	4.842	0.149	**0.259**	0.039	0.993	1.146
Small	3250	ARIMA	0.879	**0.024**	**0.134**	**0.007**	0.976	1.041
		Double Gated GCN	**0.776**	0.030	0.242	0.026	**0.679**	**0.769**
		Expon. Smoothing	1.227	0.039	0.159	0.011	2.510	2.271
		Naive Forecaster	1.029	0.027	0.141	0.008	1.126	1.313
All	5570	ARIMA	3.932	0.121	0.194	**0.024**	0.647	0.542
		Double Gated GCN	**2.852**	**0.096**	0.265	0.048	**0.474**	**0.536**
		Expon. Smoothing	4.814	0.161	0.218	0.031	0.897	0.749
		Naive Forecaster	4.205	0.132	**0.192**	0.025	0.651	0.601

other columns show the error by region – in this case, the region is defined by a set of municipalities with the same size category.

As the population grows, the MAE also increases. The same behavior can be seen in the sMAPE measure, which gives a normalized result. On the other hand, as more populous locations have distinct curves with significant cases and deaths, bigger is the standard deviation, explaining why the $\mathrm{NRMSE}_{\mathrm{sd}}$ metric tends to decrease as the population grows. In general, the proposed model performs well to capture the dynamics of the pandemic, evidencing better values for medium to large municipalities, which reinforces the choice for graph-based modeling, as it can be checked in the first and third metrics. When comparing the error concerning the local average values, using the second metric, the dedicated baseline models seem to perform better, which is expected since them were fit specifically to each municipality, curve (cases or deaths), and period (fold).

4.2. Neighborhood virus propagation

We now analyze the real and predicted curves for some municipalities in the Metropolitan Region of São Paulo (from Portuguese, RMSP), and how adjacent locations have similar dynamics. For illustration we use the curves of Covid-19 cases, but the same analysis applies to the death curves.

In the left column pictures of Figure 2, see that the curves follow a M shape before the day 250, and diverges in behavior after that. In São Paulo (state capital), the numbers started to increase after the 25th day, achieving stability between the pandemic days 100 and 130, followed by a valley and a peak between 150th and 175th. After that, it decreases continuously until reaching the lowest levels around 240th, followed by a quick uptrend and an apparent new wave after 280th day. The surrounding municipalities share similar behavior, with an average delay of one week in the main events. Despite the different scales between the neighboring municipalities, it is possible to observe that São Paulo

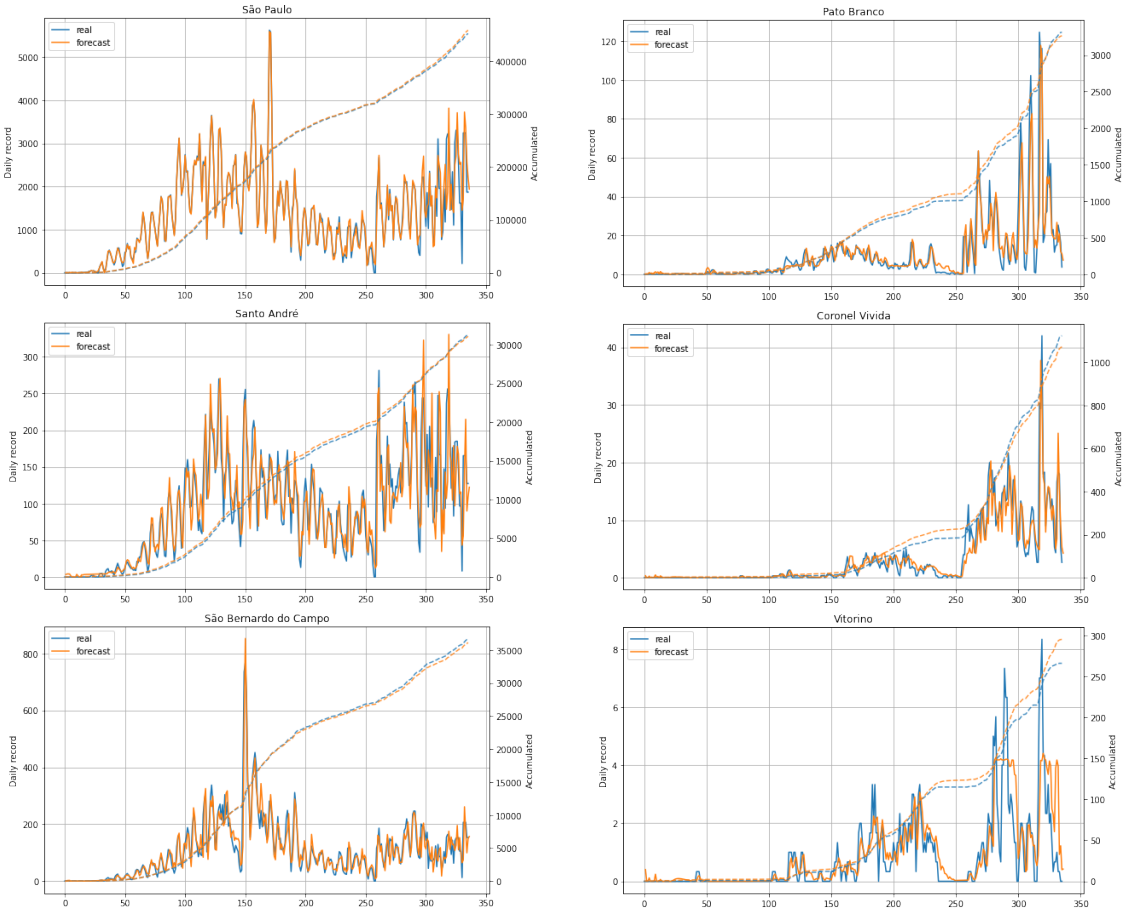

Figure 2. Daily (continuous line) and accumulated (dashed line) numbers of Covid-19 cases for RMSP (left) and southwestern Paraná (righ).

had the first growing of cases and remained at the peak of the first wave for more days, comparatively. Also, its cumulative curves show an acceleration at the 300th day, which can only be seen in the other municipalities after one week or two. For comparison, see that São Bernardo do Campo shows a slight distinct scenario, with a shrunken and abrupt first wave and a relatively smoother second wave. Santo André evidences curves similar to São Paulo, and the only apparent difference is the fact that Santo André does not have any acceleration of the cases after the 260th day, staying at a plateau – according to the almost constant slope of the cumulative curve.

Another interesting comparison can be revealed for Pato Branco, in the southwestern Paraná (PR), quite far away from RMSP. This city locates within a rural region and is surrounded by small municipalities, but it differs from its neighbors by having a large federal university, which attracts an influx of people from all over the country. There, the SARS-CoV-2 virus started to spread around the 100th day, with the first wave characterized by a flat shape and extending till the 200th day. Between days 200 and 260, one observes a stable curve, followed by a rapid increase in the number of cases. The same behavior can be seen in Coronel Vivida and Vitorino, two neighbors cities with constant interaction with Pato Branco (see right column pictures of Figure 2). However, they clearly evidence a few weeks of delay with respect to Pato Branco, reinforcing our hypoth-

esis that there exists a propagation flow that could be intercepted to support authorities in preventive measures.

In summary, it is reasonable to assume that the pandemic first hit the metropolis, and gradually propagates across the country until reaching small cities. By using the approach in this study, it is possible to forecast the propagation of the virus and anticipate some events, such as new waves and peak of local numbers, just by observing the surroundings, and this information can help to isolate and protect certain regions of interest.

4.3. Forecasting weeks ahead

In addition to the ability of forecasting next-day indexes of Covid-19, it is equally essential to forecast for days or weeks later. Instead of training directly using weekly data, in this study we benefit from the daily behavior of the pandemic, and forecast weeks in the future by retro-feeding the proposed model itself with the last forecasted day. By repeatedly applying this method it is possible to predict more than one day ahead. Due to the form of the adopted time series split, the maximum period that can be properly tested is 25 days, which corresponds to the length of the testing subset for a single fold.

In Figure 3 we present the progression of the MAE through the forecasting horizon of 25 days. The curves follows the error computed by period, being the average of the loss obtained across the 10 folds and the 5570 municipalities.

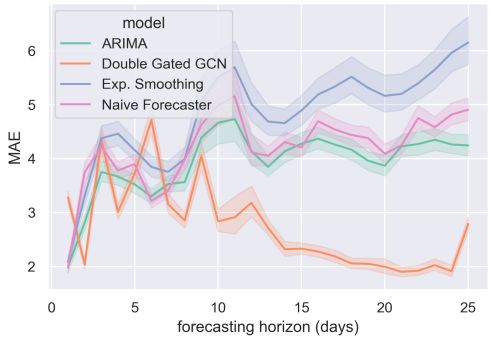

Figure 3. Average error in forecasting daily cases of Covid-19, by forecasting horizon, with 95% confidence interval.

Despite eventual oscillations due to the respective momentum of the pandemic, the computed error for the GNN shows a slight increase during the first week and decreases until reach its lower level, after two weeks, when remains quite stable. The baseline models follow other behavior, significantly increasing during the first five days and a lower growing in the subsequent days. The same occurs for the death curve, and with the $\mathrm{NRMSE}_{\mathrm{sd}}$ metric for both curves. For the sMAPE metric, the error of the proposed method begins higher than the baseline ones, and they converges to an average value between the 15th and 20th day. The performance improvement of Double Gated GCN in forecasting more than one day ahead makes the model suitable for long range estimates, being capable of keep stability when forecasting two weeks ahead for small, medium and large municipalities.

Motivated by the performance of the Double Gated GCN in long-term forecasting, we tested it over national-wide numbers for February 25th, 2021, one month after the last

day considered in the dataset, and 55 days after the last known day by the neural network (i.e., present in the training and validation subset). By informing the 336 days of the dataset, and retro-feeding the model with the following 30 days, we obtained a relative error of the accumulated Covid-19 cases and deaths of 3.72% and -2.12%, respectively.

5. Conclusions

This work proposes a STGNN model to forecast the Brazilian curves of cases and deaths by Covid-19 at a local level, based on a graph modeling of the transport modals as an indicator for flow of people between municipalities. The model, named Double Gated GCN, uses a convolutional graph neural network to capture the spatial dependencies and learn how the the pandemic propagates through the graph, followed by a recurrent neural network to handle the temporal dynamics along a time window.

The model has shown to be able to capture a wide range of pandemic scenarios, forecasting pandemic time series for small, medium, and large municipalities. To some extent, this reveals the appeal of graph-based approaches over real life problems. By learning with data from distinct domains, without relying on previously known pandemic dynamics, the model becomes useful for contexts where the epidemiological variables are unknown or uncertain. Likewise, the use of transport meshes has proven to be a suitable alternative for data sensitiveness and representativeness, besides to enabling analysis in places where urban mobility data is not available. Future researches aim to improve the edge weighting, filtering the meshes to remove possible invalid routes (*i.e.* disabled roads, unnavigable rivers, etc.) and take into account the curve of vaccination.

References

Baud, D., Qi, X., Nielsen-Saines, K., Musso, D., Pomar, L., and Favre, G. (2020). Real estimates of mortality following COVID-19 infection. *Infectious Deseases*, 20.

Brasil.io (2020). Dataset covid-19. Accessed: 2021-01-25.

Cai, J., Xu, B., Chan, K. K. Y., Zhang, X., Zhang, B., Chen, Z., and Xu, B. (2019). Roles of different transport modes in the spatial spread of the 2009 influenza A (H1N1) pandemic in mainland China. *International Journal of Environmental Research and Public Health*, 16(2).

Da Silva, R. G., Ribeiro, M. H. D. M., Mariani, V. C., and dos Santos Coelho, L. (2020). Forecasting brazilian and american COVID-19 cases based on artificial intelligence coupled with climatic exogenous variables. *Chaos, Solitons & Fractals*, 139:110027.

da Silva, T. T., Francisquini, R., and Nascimento, M. C. (2021). Meteorological and human mobility data on predicting covid-19 cases by a novel hybrid decomposition method with anomaly detection analysis: A case study in the capitals of brazil. *Expert Systems with Applications*, 182:115190.

Espinosa, M. M., de Oliveira, E. C., Melo, J. S., Damaceno, R. D., and Terças-Trettel, A. C. P. (2020). Predição de casos e óbitos de COVID-19 em Mato Grosso e no Brasil. *Journal of Health & Biological Sciences*, 8(1):1–7.

Galzo, W. (2021). Tempo médio até morte por COVID-19 em UTIs de SP caiu 4 dias no último trimestre. Accessed: 2021-03-26.

Garcia, R. (2020). Estudo mapeia tempo que COVID-19 leva para matar pacientes no Brasil. Accessed: 2021-03-26.

Han, E., Tan, M. M. J., Turk, E., Sridhar, D., Leung, G. M., Shibuya, K., Asgari, N., Oh, J., García-Basteiro, A. L., Hanefeld, J., Cook, A. R., Hsu, L. Y., Teo, Y. Y., Heymann, D., Clark, H., McKee, M., and Legido-Quigley, H. (2020). Lessons learnt from easing covid-19 restrictions: an analysis of countries and regions in asia pacific and europe. *The Lancet*, 396(10261):1525–1534.

IBGE (2014). Logística dos transportes – Brasil. Accessed: 2020-11-26.

IBGE (2019). Bases cartográficas contínuas – Brasil. Accessed: 2020-12-16.

IBGE, G. (2020). Covid-19 update. Accessed: 2021-01-25.

Lauer, S. A., Grantz, K. H., Bi, Q., Jone, F. K., Zheng, Q., Meredith, H. R., Azman, A. S., and Reich, Nicholas G. an Lessler, J. (2020). The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: Estimation and application. *Annals of Internal Medicine*, 172(9):577–582.

Li, Y., Tarlow, D., Brockschmidt, M., and Zemel, R. (2017). Gated graph sequence neural networks.

Lima, C., Silva, C., Silva, E., Marques, G., Araujo, L., Junior, L., Souza, S., Santana, M., Gomes, J., Barbosa, V., Musah, A., Kostkova, P., Dos Santos, W., and Silva-Filho, A. (2022). Monitoramento dinâmico e predição espaço-temporal da COVID-19 usando aprendizagem de máquina.

Pereira, I. G., Guerin, J. M., Silva Júnior, A. G., Garcia, G. S., Piscitelli, P., Miani, A., Distante, C., and Gonçalves, L. M. G. (2020). Forecasting COVID-19 dynamics in Brazil: A data driven approach. *International Journal of Environmental Research and Public Health*, 17(14):5115.

Ruan, Q., Yang, K., Wang, W., Jiang, L., and Song, J. (2020). Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. *Intensive Care Medicine*.

SE/UNA-SUS (2021). Covid-19 update. Accessed: 2021-08-03.

Siqueira, E., Portela, C., Farias, F., and Braga, C. (2020). Temporal prediction model of the evolution of confirmed cases of the new coronavirus (SARS-CoV-2) in Brazil. *IEEE Latin America Transactions*, 100(1e).

Sktime (2020). Sktime documentation. Accessed: 2021-07-27.

Xu, B., Tian, H., Sabel, C. E., and Xu, B. (2019). Impacts of road traffic network and socioeconomic factors on the diffusion of 2009 pandemic influenza A (H1N1) in mainland China. *International Journal of Environmental Research and Public Health*, 16(7).