Analysis with SARIMA and FFT Between Two Neighboring Cities Regarding the Implementation of Restrictive Lockdown Measures

Dunfrey P. Aragão
Grad. Prog. in Elec. and Comp. Eng.
Univ. Fed. do Rio Grande do Norte
Natal, Brazil
dunfrey@gmail.com

Davi H. dos Santos
Grad. Prog. in Elec. and Comp. Eng.
Univ. Fed. do Rio Grande do Norte
Natal, Brazil
davihenriqueds@gmail.com

Adriano Mondini
School of Pharmaceutical Sciences
São Paulo State University
Araraquara, Brazil
adriano.mondini@unesp.br

Cosimo Distante
Inst. of Applied Science and Intelligent Systems
National Research Council
Lecce, Italy
cosimo.distante@cnr.it

Luiz M. G. Gonçalves
Grad. Prog. in Elec. and Comp. Eng.
Univ. Fed. do Rio Grande do Norte
Natal, Brazil
lmarcos@dca.ufrn.br

***Abstract*—In this article, we perform the SARIMA model to regress the curve of daily COVID-19 deaths and the impact of implementing restrictive measures like lockdown. For comparison, we adopt two neighboring Brazilian cities with similar characteristics and decompose the original curve of cases to extract the seasonal curve. Using Fast Fourier Transform, we noticed that restriction of human circulation had a direct impact on COVID-19 cases and deaths in Araraquara by identifying the frequencies that compose the seasonal curves during the disease's transmission period, which for future work allows analysis and identification of events and actions through the approach.**

***Index Terms*—covid-19, lockdown, sarima, fast fourier transform.**

I. Introduction

SARS-CoV-2, the virus that causes the coronavirus disease (COVID-19), was first detected at the end of 2019 and had rapidly spread around the world [1]. The disease outpaced previous outbreaks of SARS-CoV, Ebola, and Influenza H1N1 [2], [3]. Infection occurs when an individual inhales virus particles that can be present in the air and on surfaces. Crowded places, close-contact settings, and confined spaces with poor ventilation are usually sites where transmission tends to occur [4].

Governments have implemented a variety of non-pharmaceutical interventions (NPI) to fight the spread of the disease, such as reducing urban mobility and shutting down public or non-essential spaces, as well as schools. Public health authorities, on the other hand, struggled with population compliance with social distancing or crowding in communities to avoid virus spread [5]–[7].

Studies have demonstrated the importance of implementing NPIs. Such studies have expanded knowledge of how to fight and prevent COVID-19 spread [8], [9].

This work is sponsored by CAPES under grant 88887.506889/2020-00.

The first detection of SARS-CoV-2 in Brazil occurred on February 25, 2020, in the city of São Paulo, reaching neighboring counties, urban centers, and countryside cities in weeks [10].

To avoid the system health collapse and an increase in the number of cases, various cities in the country imposed urban restriction measures on different occasions of the pandemic. Araraquara lies 270 km from the state capital, the city of São Paulo, and has a population of approximately 238,000 people, with a demographic density of 235 inhabitants/km^2, according to IBGE [11]. The number of cases and the occupancy of ICU beds in 2020 pushed health authorities in Araraquara to implement red alert measures, which temporarily prohibit not only the circulation of people but forced the shut down of businesses, schools, public transportation, and malls. Only health services, grocery stores, and other essential services were maintained. The official declaration of a state of emergency was associated with a complete lockdown of the city in 2021. Due to this measure, the number of cases decreased by 43,02% in the first ten days after the beginning of red flag measures [12]. Drawing attention due to the number of cases and viral spread, São Carlos, a city neighboring Araraquara, with a population of 254,000 and a demographic density of 226 inhabitants/km^2 [11], opted for less severe restrictions measures, not implementing NPI during 2020 or 2021.

We applied regression methods and comparison analysis to evaluate and compare the effects of lockdown in two Brazilian cities, Araraquara and São Carlos, in the state of São Paulo. Furthermore, as COVID-19 seems to be seasonal, we propose tracking seasonal features in the domain of frequency with Fast Fourier Transformations (FFT).

II. Materials and Methods

A. Dataset

Data containing COVID-19 reports were obtained from daily local bulletins [13]. The dataset is organized as a time series with daily entries beginning on March 3, 2020. It also provides information on reports, city, the number of confirmed cases, the raw number of deaths, estimated population of the city, city code, percentage of confirmed cases per 100,000 inhabitants, and death rates.

For our analysis, we highlighted the date of restrictive measures. The government of Araraquara declared red alert days from December 24, 2020, to January 6, 2021, corresponding to epidemiological week 51 of 2020 to epidemiological week 1 of 2021); from February 8 to 12 2021, which corresponds to epidemiological week 6 of 2021; from March 11 to 20, epidemiological weeks 12 to 14, 2021. Also declared lockdown (not only red alerts) on February 7th and 8th, 2021, and after from the period 21st of February to the 6th of March.

B. Seasonal Auto-regressive Integrated Moving Average - SARIMA

The p-order Auto-regressive model, $AR(p)$, is based on linear combination of past observations $X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + ... + \alpha_p X_{t-p} + \varepsilon_t$, where α is constant and ε_t a disturbance, from the white noise process, stochastic with respect to $X_{t-1}, X_{t-2}, ..., X_{t-p}$. Adjusting the parameters of an auto-regressive model means estimating the parameters α_k where $k : t - p < k < t - 1$. Therefore, we can consider:

$$Y_t = \mu + X_t = \mu + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + ... + \alpha_p X_{t-p} + \varepsilon_t \quad (1)$$

Moving Average Models (MA) are an extension of the white noise process, being a linear combination of the disturbance ε_t added to recent disturbances $\varepsilon_{t-1}, \varepsilon_{t-2}, ..., \varepsilon_{t-q}$. A q-order MA model, $MA(q)$, can be defined with the following equation:

$$X_t = \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + ... + \beta_q \varepsilon_{t-q} \quad (2)$$

where q is the order of the moving average. The combination of AR and MA models define $ARMA(p, q)$ and can be adapted to handle different forms of non-stationary time-series and, through the use of differentiation, provide the class of models $ARIMA(p, d, q)$, where d is the order of differentiation.

Seasonal Autoregressive Integrated Moving Average, SARIMA or Seasonal ARIMA, is an extension of ARIMA, supporting univariate time series, and has a second set of parameters $(P, Q, D)_s$ referring to the seasonality pattern, being s the number of periods per season. For an exogenous variable n, defined in a time step t, denoted by x_t^i, where $i \leq n$, with a coefficient β_i, we define the model $SARIMA(p, d, q)(P, Q, D)_s$ as

$$\phi_p(L)\tilde{\phi}_P(L^s)\Delta^d\Delta_s^D\varepsilon_t \& = \theta_q(L)\tilde{\theta}_Q(L^s)\zeta_t \quad (3)$$

where, ζ_t is the non-stationary time series; ε_t is the Gaussian white noise process; s is the period of the time series; the ordinary autoregressive and moving average components are represented by the polynomials $\tilde{\phi}_P(L^s)$ and $\tilde{\theta}_Q(L^s)$ of orders p and q; the autoregressive and moving average seasonal components are $\phi_p(L)$ and $\theta_q(L)$, of orders P and Q; Δ^d and Δ_s^D are the ordinary and seasonal differentiable components; L is the temporal operator.

To make predictions, 323 days of daily data were used for training. We carried out the forecast two weeks later. From the data, we used the last week as input for predictions, corresponding to the week from February 22 to 28. As a result, the model predicts the following two weeks, corresponding to the days from February 28 to March 2, 2020.

C. Fast Fourier Transform - FFT

A general function is represented by the Fourier transform as a continuous or integral superposition of complex exponential functions. The series can thus be represented in the frequency domain in this manner. The key to using this algorithm in this work is decomposition, which turns the signal into a sum of other signals, the sum of a series of sinusoids. Every nonlinear function can be represented as a sum of n-sine waves.

For discrete signals, we can use the Discrete Fourier Transformation (DFT). The Fast Fourier Transformation (FFT) is a faster algorithm to compute the Discrete Fourier Transform [14].

The Forward DFT is defined by

$$X_k = \sum_{n=0}^{N-1} x_n \cdot e^{-i\ 2\pi\ k\ n\ /\ N} \quad (4)$$

and the Inverse DFT is defined by

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} X_k e^{i\ 2\pi\ k\ n\ /\ N} \quad (5)$$

The $x_n \rightarrow X_k$ transformation is a translation from the space domain to the frequency domain that may be highly beneficial in both studying the power spectrum of a signal and changing specific issues for more efficient processing.

The FFT algorithm tries to compute the DFT into two smaller sections based on the model definition. We can expand an arbitrary function periodic on some domain as a sum of sines and cosines, as well as periods on that domain, projecting the function into orthogonal function directions.

D. Adopting Numbers of Cases or Deaths

It is possible to compare the number of cases to the number of deaths, which could demonstrate accurate accountability. This approach, however, raises the issue of ensuring that someone dies as a result of COVID-19. In these situations, the tests that were performed took time for some periods, and counting may not be reliable.

Additionally, the absence of RT-PCR tests and the general number of cases in Araraquara and São Carlos may not effectively guarantee the real scenario, which can be associated

with the disease or condition that triggers an individual's health condition.

Using a number of cases, on the other hand, is a complex problem to solve, as it is multifactorial. As a result, underreporting is to be expected. From the standpoint of epidemic establishment, at least four major descriptor profiles can be identified: susceptible, symptomatic, asymptomatic, and vaccinated. Individuals may progress from asymptomatic to symptomatic infection, from symptomatic to cured patients, or from susceptible to reinfected. As a result, such situations need immediate action by the health system to ensure the accurate identification, particularly of infected individuals, whether they are symptomatic or not.

In situations, in which there is significant efficacy in detecting such profiles via an effective surveillance system, both quantitative and qualitatively, the analysis of cases results in a better understanding of disease dynamics.

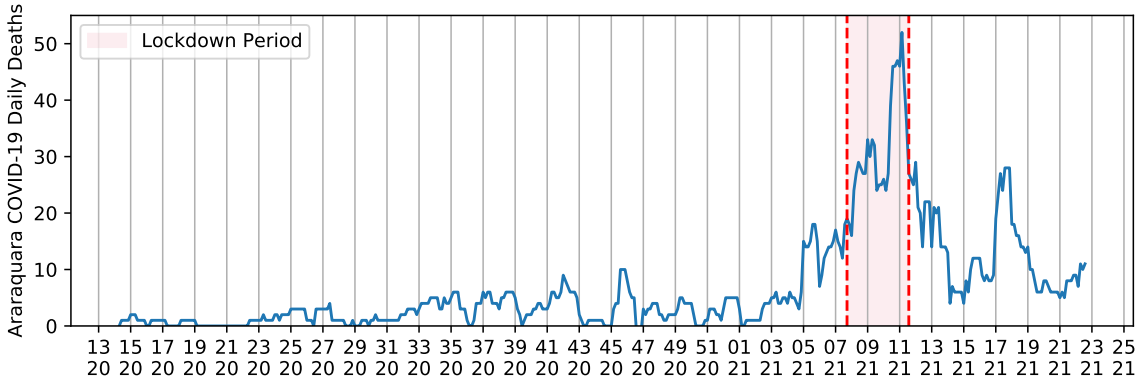
III. Experiments

A. Time-Series Comprehension about COVID-19 Daily Deaths in Araraquara and São Carlos

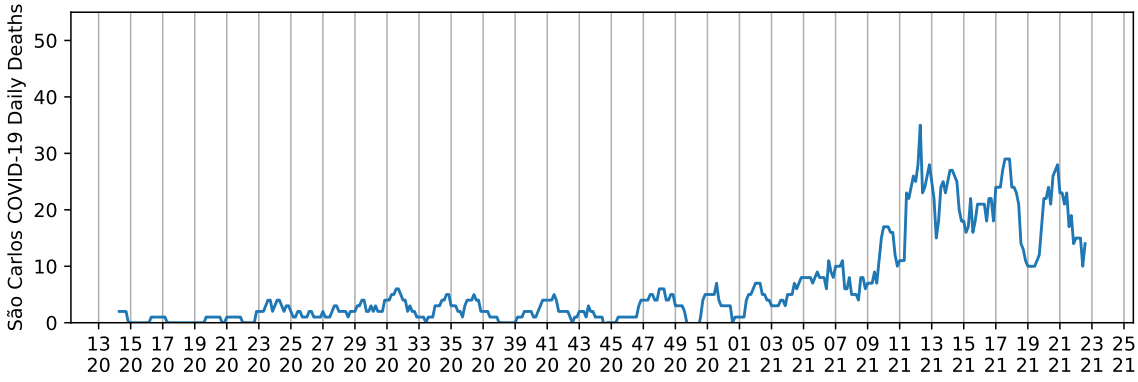
By comparing COVID-19 death curves of Araraquara (Figure 1a) and São Carlos (Figure 1b), we observed:

- The maximum peak of deaths for São Carlos and Araraquara occurs with a time difference of approximately one week;
- Araraquara has a maximum global peak of 50 daily deaths, with an average of 30 daily deaths;
- The maximum peak of daily deaths in São Carlos is in the same range as the period of highest averages, around 30 to 35 daily deaths, from week 12 to 21 of 2021;
- After the peak of deaths in Araraquara, when the lockdown was implemented, rates dropped drastically, while in São Carlos rates remained high;
- Although death rates remain high after the record of the maximum peak of daily deaths in São Carlos, reports do not exceed 35 daily deaths, unlike Araraquara, which reached 50 daily deaths;
- After a drastic drop of deaths in Araraquara, the index once again registers 30 deaths per day for a period of one week, comprising weeks 17 and 18 of 2021);
- There was a drastic drop in the average of highs in São Carlos in week 19 of 2021.

Furthermore, we must consider the possibility that, despite similar behaviors in periods preceding the maximum peaks shown in Araraquara and São Carlos, curves begin to change due to the spread of different virus variants. These dynamics



(a) Araraquara Daily Deaths.



(b) São Carlos Daily Deaths.

are inherent to each city's urban characteristics, local and individual customs, geospatial distribution, and government actions against disease spread.

As a result, to understand explicit patterns in death curves of COVID-19, which is the sum of diverse events, we use COVID-19 daily cases for data comprehension.

B. SARIMA Forecasts

Comparatively, real daily cases of Araraquara and SARIMA forecasts are shown in Figure 2. The SARIMA model presents, throughout the curve, a prediction pattern. When we changed the dates of lockdown, it is possible to observe that the curve trend was towards an increase in the number of cases.

Symptoms of an infected individual begin to appear after an estimated period of 4/5 days after infection [15]. If transmission occurs on the day t, this transmission is reflected approximately one week later, on day $t + 7$ [16], [17]. Given this information, and using the Box & Jenkins approach with ACF and PACF analysis [18], [19], we defined a SARIMA model adjusted in order (0, 1, 0), with seasonality order (7, 1, 0, 7) for the dataset (Figure 3).

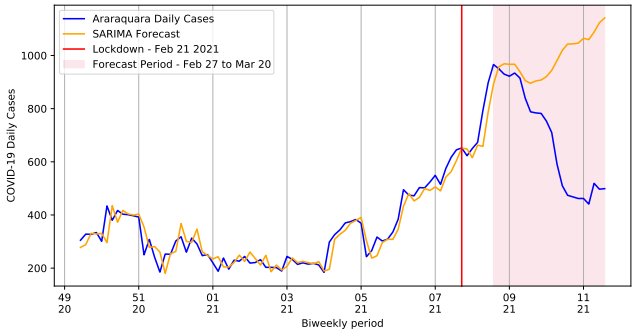


Fig. 3: Biweekly SARIMA forecasts after the first week of the determined lockdown day in Araraquara.

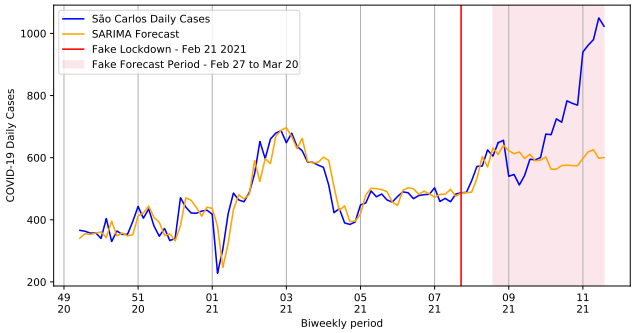


Fig. 4: Biweekly SARIMA forecasts for supposing after the first week of the determined lockdown day in São Carlos.

On the other hand, Figure 4 shows that the period of the high number of cases in Araraquara coincided with the period of significant increase of cases in São Carlos. For São Carlos, we defined a SARIMA model adjusted in order (0, 1, 0), with seasonality order (8, 1, 0, 7). Accurate reports of COVID-19 extrapolated the expectation of the SARIMA model.

As shown in Figure 3, the RMSE for February 28 to March 06, 2021, is approximately 29.59 cases. For test data, with no training data information, the model generates the upward red curve, while the blue line represents the actual data for the period. There is a rising curve with values higher than the original data. In this period, therefore, it is estimated that the expected values for March 3, 2021, would be 920.6 daily cases, a divergence of values compared to the 474 officially reported, with a recession pattern in the number of cases.

C. Frequency Events

The dynamics of infection can exhibit trends and/or seasonal effects, which are characteristic of non-stationary series. The employment of time series decomposition models is a viable method in such scenarios. The additive model considers the signal to be created by the sum of the elements Tt, St, and Rt given the series Xt.

Selecting the period from 25-05-2020 to 20-03-2021 only for Araraquara and decomposing seasonality, two different

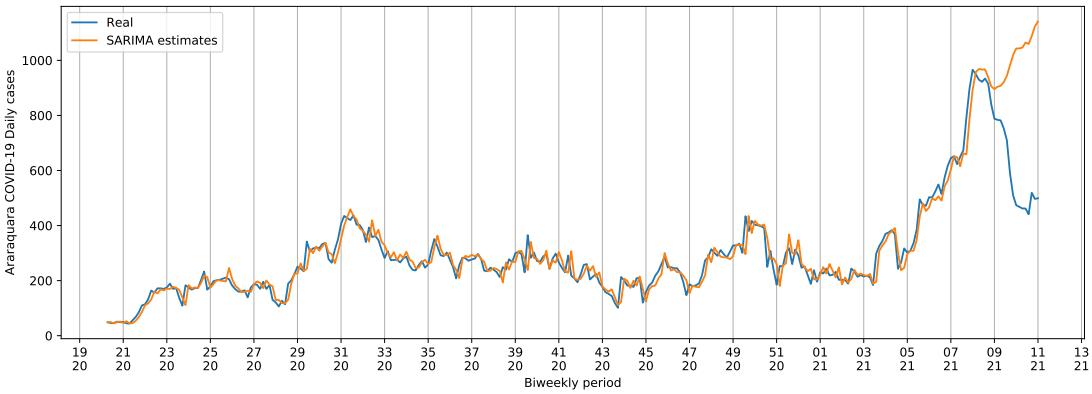


Fig. 2: Araraquara original daily cases and SARIMA forecasts.

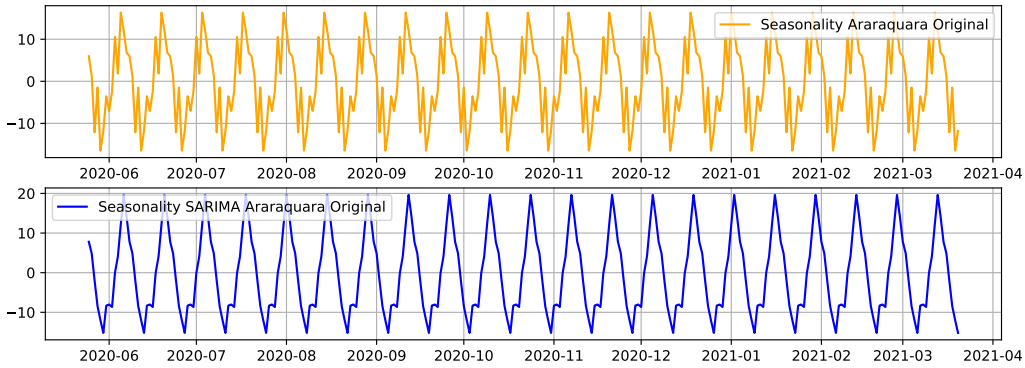


Fig. 5: Seasonal curves for Araraquara with lockdown and SARIMA forecasts.

seasonality scenarios are observed: the original curve, and SARIMA forecasts Figure 5.

When decomposing seasonality from COVID-19 deaths for Araraquara and São Carlos, a pattern is present in the two seasonal curves. Using the Dynamic Time Warping (DTW) [20] distance analysis and normalizing this distance, seasonality curves are highly correlated (Figure 6); the reference index is the Seasonal curve of Araraquara and the Query index represents São Carlos Seasonal curve.

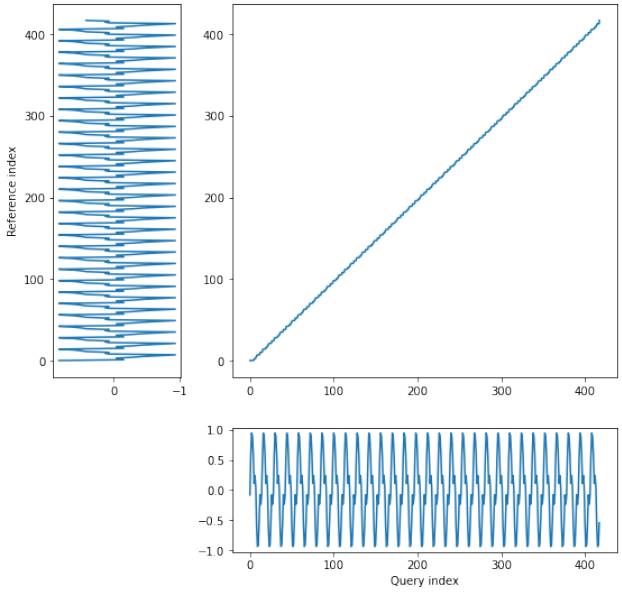


Fig. 6: DTW distance of the seasonal curves between Araraquara and São Carlos.

Now, applying the FFT, we noticed that there are significant changes in the pattern of the spectrum that is used to produce the seasonal curve of Araraquara approaches, the original curve, and SARIMA prediction model prediction (Figure 7 and Figure 8).

These findings show a significant change in the pattern of the seasonality when the lockdown, an NPI approach, is used. An important impact, a variation, occurs in the formation of

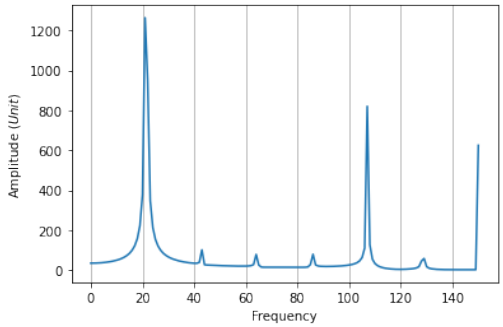


Fig. 7: FFT applied to the Araraquara Daily Cases.

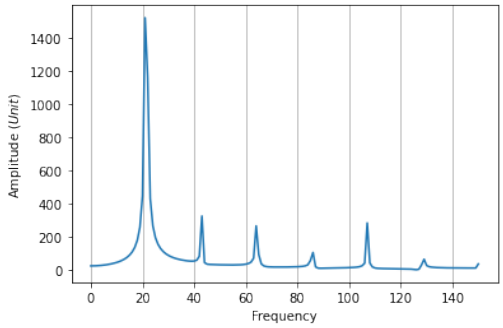


Fig. 8: FFT applied to the Araraquara Daily Cases Predicted by SARIMA model.

seasonality in the curve of COVID-19 cases, according to frequencies.

Comparing the original (without date filter) curves for Araraquara and São Carlos, despite the similarities highlighted in Section III-A, other patterns emerge and distinguish cases in each city, as shown in Figure 9 and Figure 10. Despite composition similarities at certain frequencies, there is a divergence in amplitude, such as frequencies 60, 90, 149, and 209.

IV. Conclusion

Using auto-regressive methods and comparing the period in which lockdown was implemented in Araraquara, it is possible

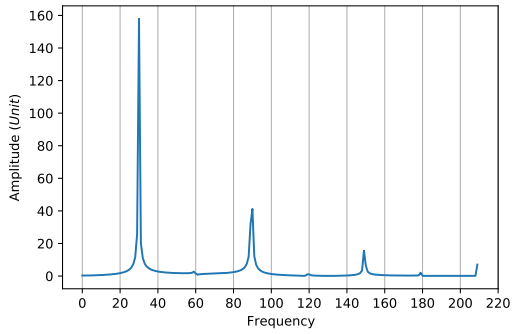


Fig. 9: FFT applied to the entire Araraquara Daily Cases.

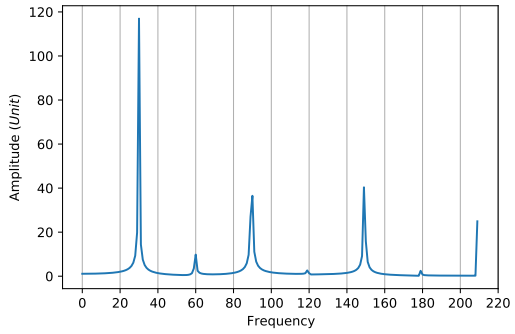


Fig. 10: FFT applied to the entire São Carlos Daily Cases.

to observe the contribution of the NPI to the decrease in cases and deaths of COVID-19, which is different from the pattern shown in São Carlos.

We observed that it is possible to use the FFT method to investigate the patterns in pandemics by measuring the cycle length of the power peak spectrum and bold in the frequency domain rather than the number of daily cases. Government and policymakers can understand and adjust their restrictive measures, such as implementing NPI measures.

Besides, we believe that the approach of using Fourier transforms to determine exactly which events influence the curve of cases of COVID-19 can be used to deepen other studies. As a result, we are conducting an analysis to identify exactly which events have a high contribution in common through comparisons with other urban centers.

REFERENCES

[1] W.-j. Guan, Z.-y. Ni, Y. Hu, W.-h. Liang, C.-q. Ou, J.-x. He, L. Liu, H. Shan, C.-l. Lei, D. S. Hui, B. Du, L.-j. Li, G. Zeng, K.-Y. Yuen, R.-c. Chen, C.-l. Tang, T. Wang, P.-y. Chen, J. Xiang, S.-y. Li, J.-l. Wang, Z.-j. Liang, Y.-x. Peng, L. Wei, Y. Liu, Y.-h. Hu, P. Peng, J.-m. Wang, J.-y. Liu, Z. Chen, G. Li, Z.-j. Zheng, S.-q. Qiu, J. Luo, C.-j. Ye, S.-y. Zhu, and N.-s. Zhong, "Clinical characteristics of coronavirus disease 2019 in china," *New England Journal of Medicine*, vol. 382, no. 18, pp. 1708–1720, 2020. [Online]. Available: https://doi.org/10.1056/NEJMoa2002032

[2] S. D. Pitlik, "Covid-19 compared to other pandemic diseases," 2020. [Online]. Available: www.rmmj.org.il

[3] A. Wilder-Smith, "COVID-19 in comparison with other emerging viral diseases: risk of geographic spread via travel," pp. 1–11, dec 2021. [Online]. Available: https://doi.org/10.1186/s40794-020-00129-9

[4] "Who - coronavirus disease (covid-19) pandemic," https://www.who.int/emergencies/diseases/novel-coronavirus-2019, accessed: 2021-06-30.

[5] T. R. Bolaño-Ortiz, Y. Camargo-Caicedo, S. E. Puliafito, M. F. Ruggeri, S. Bolaño-Diaz, R. Pascual-Flores, J. Saturno, S. Ibarra-Espinosa, O. L. Mayol-Bracero, E. Torres-Delgado, and F. Cereceda-Balic, "Spread of SARS-CoV-2 through Latin America and the Caribbean region: A look from its economic conditions, climate and air pollution indicators," *Environmental Research*, vol. 191, dec 2020.

[6] J. Verschuur, E. E. Koks, and J. W. Hall, "Global economic impacts of COVID-19 lockdown measures stand out in high-frequency shipping data," *PLOS ONE*, vol. 16, no. 4, p. e0248818, apr 2021.

[7] A. Mandel and V. Veetil, "The Economic Cost of COVID Lockdowns: An Out-of-Equilibrium Analysis," *Economics of Disasters and Climate Change*, vol. 4, no. 3, p. 1, oct 2020. [Online]. Available: /pmc/articles/PMC7304379//pmc/articles/PMC7304379/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC7304379/

[8] D. P. Aragão, D. H. dos Santos, A. Mondini, and L. M. G. Gonçalves, "National holidays and social mobility behaviors: Alternatives for forecasting covid-19 deaths in brazil," *International Journal of Environmental Research and Public Health*, vol. 18, no. 21, 2021. [Online]. Available: https://www.mdpi.com/1660-4601/18/21/11595

[9] N. Ferguson, D. Laydon, G. N. Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. C. Perez, G. Cuomo-Dannenburg, A. Dighe, I. Dorigatti, H. Fu, K. Gaythorpe, W. Green, A. Hamlet, W. Hinsley, L. Okell, S. V. Elsland, H. Thompson, R. Verity, E. Volz, H. Wang, Y. Wang, P. Walker, C. Walters, P. Winskill, C. Whittaker, C. Donnelly, S. Riley, and A. Ghani, "Impact of non-pharmaceutical interventions (npis) to reduce covid-19 mortality and healthcare demand," *Imperial College COVID-19 Response Team*, pp. 1–20, 2020.

[10] W. M. de Souza, L. F. Buss, D. d. S. Candido, J. P. Carrera, S. Li, A. E. Zarebski, R. H. M. Pereira, C. A. Prete, A. A. de Souza-Santos, K. V. Parag, M. C. T. Belotti, M. F. Vincenti-Gonzalez, J. Messina, F. C. da Silva Sales, P. d. S. Andrade, V. H. Nascimento, F. Ghilardi, L. Abade, B. Gutierrez, M. U. Kraemer, C. K. Braga, R. S. Aguiar, N. Alexander, P. Mayaud, O. J. Brady, I. Marcilio, N. Gouveia, G. Li, A. Tami, S. B. de Oliveira, V. B. G. Porto, F. Ganem, W. A. F. de Almeida, F. F. S. T. Fantinato, E. M. Macário, W. K. de Oliveira, M. L. Nogueira, O. G. Pybus, C. H. Wu, J. Croda, E. C. Sabino, and N. R. Faria, "Epidemiological and clinical characteristics of the COVID-19 epidemic in Brazil," *Nature Human Behaviour*, vol. 4, no. 8, pp. 856–865, aug 2020.

[11] D. Rousseff, M. Belchior, P. E. Pereira, N. Diretor-Executivo, S. Da, C. Côrtes, and D. W. Tai, "Sinopse do censo demográfico: 2010," *Censo 2010*, p. 265, 2011.

[12] "Prefeitura municipal de araraquara - lockdown: Araraquara tem queda no numero de casos e media movel por covid-19," http://www.araraquara.sp.gov.br/noticias/2021/marco/10/lockdown-araraquara-tem-queda-no-numero-de-casos, accessed: 2021-07-03.

[13] "Brasil.io - boletins informativos e casos do coronavírus por município por dia," https://brasil.io/dataset/covid19/caso, accessed: 2021-06-30.

[14] E. O. Brigham and R. E. Morrow, "The fast fourier transform," *IEEE Spectrum*, vol. 4, no. 12, pp. 63–70, 1967.

[15] M. A. Johansson, T. M. Quandelacy, S. Kada, P. V. Prasad, M. Steele, J. T. Brooks, R. B. Slayton, M. Biggerstaff, and J. C. Butler, "Sars-cov-2 transmission from people without covid-19 symptoms," *JAMA Network Open*, vol. 4, no. 1, pp. e2 035 057–e2 035 057, 01 2021.

[16] B. Long, B. M. Carius, S. Chavez, S. Y. Liang, W. J. Brady, A. Koyfman, and M. Gottlieb, "Clinical update on covid-19 for the emergency clinician: Presentation and evaluation," *American Journal of Emergency Medicine*, vol. 54, pp. 46–57, 4 2022.

[17] S. A. Lauer, K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, and J. Lessler, "The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: Estimation and application," *Annals of Internal Medicine*, vol. 172, pp. 577–582, 2020.

[18] Q. Song and A. Esogbue, "A new algorithm for automated box-jenkins arma time series modeling using residual autocorrelation/partial autocorrelation functions," *IEMS*, vol. 5, no. 2, 01 2006.

[19] O. D. Anderson, "The Box-Jenkins approach to time series analysis," *RAIRO - Operations Research - Recherche Opérationnelle*, vol. 11, no. 1, pp. 3–29, 1977. [Online]. Available: http://www.numdam.org/item/RO_1977__11_1_3_0/

[20] T. Giorgino, "Computing and visualizing dynamic time warping alignments in R: The dtw package," *Journal of Statistical Software*, vol. 31, no. 7, pp. 1–24, 2009.