

A pandemia de SARS-CoV-2: medidas de correlações entre os Estados Brasileiros

Nascimento, J. O. do*¹, Passos C. A. C.², Moret, M. A.^{1,3}

* jeffersonascimento@gmail.com

1 Departamento de Modelagem Computacional, CIMATEC, Salvador/BA, Brasil.

2 Departamento de Física, Universidade Federal do Espírito Santo, UFES, Vitória/ES, Brasil.

3 Universidade do Estado da Bahia, UNEB, Salvador/BA, Brasil.

* e-mail: jeffersonascimento@gmail.com; carlos.passos@ufes.br

Resumo

Em 2019 o mundo foi apresentado ao SARS-CoV-2 que, de acordo com a Organização Mundial de Saúde (OMS), o primeiro caso no planeta ocorreu ao início da primeira semana dezembro de 2019 em Wuhan, na China, resultando na atual pandemia. No Brasil, o primeiro caso oficial diagnosticado é datado da última semana de fevereiro, conforme informações oficiais do Ministério da Saúde (MS). Este artigo tem como objetivo apresentar, em que nível um Estado da Região Norte (Pará) e um do Sudeste (Espírito Santo) apresentaram um comportamento de correlação na difusão do Coronavírus em território nacional, com os demais Estados Brasileiros. O percurso metodológico é por meio de métricas de correlações estatísticas. Os resultados indicaram em que medida a propagação do Coronavírus ocorreu em similaridade entre os Estados Brasileiros, em comparação com o Espírito Santo e o Pará. Por fim, é apresentado a diferença entre os índices estatísticos por meio do impacto nos resultados das análises da difusão do SARS-CoV-2 no território nacional.

Palavras-chave: SARS-CoV-2, Pandemia de Coronavírus, Covid-19, Pará, Espírito Santo.

Abstract

In 2019 the world was introduced to SARS-CoV-2 which, according to the World Health Organization (WHO), the first case on the planet occurred at the beginning of the first week of December 2019 in Wuhan, China, resulting in the current pandemic. In Brazil, the first official case diagnosed is dated to the last week of February, according to official information from the Ministry of Health (MS). This article aims to present, at what level a State in the North Region (Pará) and one in the Southeast (Espírito Santo) presented a correlation behavior in the diffusion of the Coronavirus in the national territory, with the other Brazilian States. The methodological path is through metrics of statistical correlations. The results indicated the extent to which the spread of the Coronavirus occurred in similarity between the Brazilian states, in comparison with Espírito Santo and Pará. Finally, the difference between the statistical indices is presented through the impact on the results of the analysis of the diffusion of the SARS-CoV-2 in the national territory.

Keywords: SARS-CoV-2, Coronavirus Pandemic, Covid-19, Pará, Espírito Santo.

1. Introdução

Para os Estados do Pará (PA) e do Espírito Santo (ES), como na maioria dos Estados Brasileiros e Distrito Federal (DF), as curvas de infecções e óbitos estão em crescimento acelerado. A correlação entre elas é linear, enquanto os casos de novas infecções continuarem aumentando, os óbitos também. As estimativas das curvas que representam os diversos

comportamentos na pandemia do Coronavírus são não-lineares. Cada

novo caso impacta diretamente na previsão do comportamento e no pico de cada curva e, conseqüentemente, estes cálculos preditivos devem ser realizados para cada nova atualização. No Pará, por exemplo, as curvas que representam o número de

infectados e o número de óbitos, chegou a apresentar 99% de correlação, conforme Figura 1:

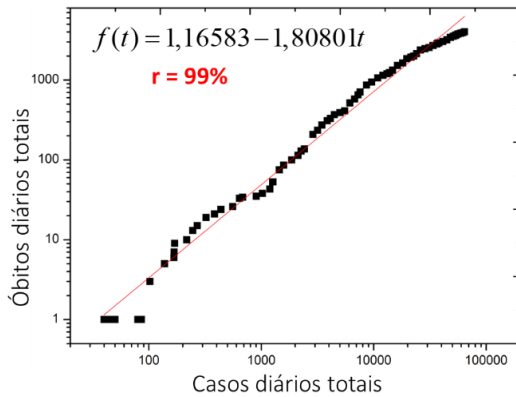


Figura 1: Medida de correlação de Pearson entre as curvas de infecções e óbitos para o Pará: Fonte: Adaptado do relatório de Pós-doutoramento em Cenários Epidemiológicos de Nascimento [1].

Considerando que não existem medicamentos específicos para esta enfermidade ou vacinas comprovadamente eficazes no tratamento dela (considerando um período até 07 de dezembro de 2020), a análise dos dados dos casos de infecção e sua interpretação matemática são essenciais para subsidiar e orientar as medidas governamentais de repressão e mitigação do impacto da COVID-19 [2]. Isso significa que são necessárias estimativas com modelos matemáticos para avaliar o potencial de desenvolvimento da transmissão humano-humana sustentada [2].

Diante do exposto, este artigo tem como objetivo investigar em que medida há correlações estatísticas entre o Estado do Espírito Santo e Pará com os demais Estados Brasileiros e o Distrito Federal, especificamente em relação às curvas que representam o número de novos casos de infecções diários e novos casos de óbitos diários. Destacamos que o espaço amostral utilizado corresponde desde o primeiro caso oficial registrado no Brasil, que ocorreu na quarta semana de fevereiro de 2020 até trinta de novembro de 2020.

Por meio de métricas estatísticas, analisamos as medidas de correlações utilizando o Coeficiente de correlação para Postos de Spearman, Pearson e o Coeficiente de correlação por Postos de Kendall, entre o Pará com os demais Estados e o DF e, de forma análoga do Espírito Santo comparando aos demais Estados e o DF, totalizando 40 semanas.

2. Metodologia

Para que pudéssemos realizar as medidas de correlações, construímos para os 26 Estados e mais o DF, uma representação matemática por meio de séries temporais. Desta forma, estabelecemos as comparações que nos interessavam, para o Pará e o Espírito Santo, em relação ao número diário de infecções e óbitos. Calculamos primeiramente a

correlação por meio do coeficiente de correlação de Pearson (r). Para tal, o nível de correlação entre as séries temporais construídas, utilizamos primeiramente o coeficiente de Pearson, proposto por Karl Pearson (1857-1936), conforme Eq. 1:

$$r = \frac{n \sum x_i y_i - (\sum x_i) \cdot (\sum y_i)}{\sqrt{[n \sum x_i^2 - (\sum x_i)^2] \cdot [n \sum y_i^2 - (\sum y_i)^2]}} \quad (1)$$

De forma simplificada, destacamos que, quando $r = +1$ a correlação é perfeitamente positiva. Se $r = -1$ a correlação é perfeitamente negativa e se $r = 0$ não há correlação entre os pontos ou há, porém, é não-linear [1-3]. Em relação ao coeficiente de Spearman (r_s), baseada em postos, é o índice de correlação mais antigo a ser utilizado. Ele exige que as variáveis, em caso de correlação X e Y, tenham suas métricas por meio de escala ordinal [4]. Caso X e Y sejam quantitativas e sua distribuição de graus seja não bivariada, Spearman é uma alternativa em substituição a Pearson. Caso correlação de Spearman, ao utilizar os dados do espaço amostral, sejam substituídos por postos ordenados, poderá ser calculada de forma semelhante ao coeficiente de Pearson. A principal vantagem do coeficiente de correlação de Spearman é a não exigência da normalidade dos dados. Reescrevendo a Eq. 1:

$$r = \frac{\sum xy}{\sqrt{\sum x^2} \sqrt{\sum y^2}}, \text{ em que teremos } x = X - \bar{X} \text{ e}$$

$y = Y - \bar{Y}$. Se X e Y correspondem a postos, teremos:

$$\sum X = \sum Y = n \frac{(n+1)}{2} \quad (2)$$

Por meio da soma dos quadrados dos postos, chegaremos na seguinte equação:

$$\sum X^2 = \sum Y^2 = n \frac{(n+1)(2n+1)}{6} \quad (3)$$

Lembre-mos de que $x = X - \bar{X}$ e, assim, teremos: $\sum x^2 = \sum (X - \bar{X})^2$, o que corresponde

$$a \sum X^2 = \frac{(\sum X)^2}{n}. \text{ Ambas as equações}$$

correspondem a: $\sum X^2 - n\bar{X}^2$.

Partindo das Eq. 2 e Eq. 3 e do lembrete anterior, teremos que:

$$\sum x^2 = \sum X^2 - \frac{(\sum X)^2}{n}, \text{ que corresponde em}$$

$$n \frac{(n+1)(2n+1)}{6} - n^2 \frac{(n+1)^2}{4n}, \text{ simplificando a expressão, teremos } n \frac{(n+1)(2n+1)}{6} - n \frac{(n+1)^2}{4}.$$

$$\text{Resolvendo a equação anterior, chegamos em } \frac{n^3 + n}{12}$$

que é igual a $\sum y^2$. Adotando $d = x - y$, d^2 , será:

$$d^2 = (x - y)^2, \text{ ou seja, } d^2 = (x^2 - 2xy + y^2).$$

Percebamos que de forma equivalente as expressões imediatamente anteriores, temos

$$\sum d^2 = \sum (x - y)^2, \text{ ou seja:}$$

$$\sum d^2 = \sum x^2 - 2\sum xy + \sum y^2 \quad (4)$$

Diante de tais resultados e relembrando a forma reescrita do coeficiente de Pearson e da afirmativa que fizemos em calcular o coeficiente de Spearman (r_s) pelo de Pearson, teremos:

$$r = \frac{\sum xy}{\sqrt{\sum x^2} \sqrt{\sum y^2}} = r_s, \text{ logo, chegaremos na Eq. 5:}$$

$$\sum xy = r = r_s \sqrt{\sum x^2} \sqrt{\sum y^2} \quad (5)$$

Substituindo na Eq. 4, a equação anterior (Eq. 5) chegaremos em:

$$\sum d^2 = \sum x^2 - 2(r_s \sqrt{\sum x^2} \sqrt{\sum y^2}) + \sum y^2, \text{ logo:}$$

$$r_s = \frac{\sum x^2 + \sum y^2 - \sum d^2}{2(\sqrt{\sum x^2} \sqrt{\sum y^2})} \quad (6)$$

Substituindo os respectivos valores de $\sum y^2$ na Eq. 6, teremos finalmente como resultado:

$$r_s = 1 - \frac{6 \cdot \sum_{i=1}^n d_i^2}{n^3 - n} \quad (7)$$

Em que $d_i = r_{Xi} - r_{Yi}$, variam de 1 a n . [5]

A Eq. 7 corresponde ao coeficiente de correlação para postos de Spearman (r_s). Destacamos algumas características: Não há problema em sua utilização mesmo com a presença de *outliers*. -1 indica perfeita correlação negativa ou inversa e 1 indica perfeita correlação positiva ou direta.

Em relação ao coeficiente de correlação por postos de Kendall (τ), corresponde a uma medida de associação, também para variáveis ordinais. Porém, este coeficiente de correlação indicará em que medida há um grau de associação entre dois conjuntos de postos [6]. O Nível da medida do relacionamento entre dois conjuntos de postos é obtido pela razão entre o total efetivo de escores + 1 e -1, em relação ao total máximo possível.

Considerando o número de pares concordantes e discordantes, em relação ao número total de pares do espaço amostral (desconsiderando os valores repetidos), encontraremos coeficiente de correlação por postos de Kendall (τ) [7]. De outra forma: Considerando o total possibilidades por S e o quantitativo máximo de escores possíveis por

$$\binom{n}{2} = n \frac{(n-1)}{2}, \text{ o coeficiente de correlação de}$$

$$\text{Kendall da seguinte maneira: } \tau = \frac{S}{n \frac{(n-1)}{2}}$$

$$\tau = \frac{2S}{n(n-1)} \quad (8)$$

Em relação aos *outliers*, o coeficiente de correlação por postos de Kendall (τ) é uma métrica estatística robusta em caso de termos no espaço amostral, com a presença de dados extremos [6].

3. Discussão dos resultados

Utilizando os materiais e métodos descritos na sessão anterior, realizamos as medidas de correlações especificamente para os estados do Espírito Santo e Pará. Desta forma, por meio das séries temporais representativas de cada Estado Brasileiro e o DF, a verificação de que, em qual nível há correlação entre PA e os 25 Estados e o DF, assim como o ES os 25 Estados e o DF, foi possível. Ratificamos que nesta pesquisa, objetivamos analisar os números de novas infecções e óbitos diários, conforme os dados oficiais disponibilizados pelo Ministério da Saúde.

Para cada par de Estados analisados, um correspondente coeficiente de correlação individualizado foi encontrado. A título de exemplo, Espírito Santo e Paraíba (ES x PB), o coeficiente de correlação de Spearman, correspondeu a $r_s = 0,75732$.

Apenas por uma questão de representação nas figuras, das métricas de correlações, adotaremos as seguintes nomenclaturas: coeficiente de correlação de Pearson (r), Spearman (s) e coeficiente de Kendall (k). Nossa opção neste artigo, é de apresentar os gráficos e tabelas (de forma não tão didáticas), para fazermos ao fim, deste tópico, nossas considerações sobre os resultados encontrados. Em relação ao

número de novas infecções diárias, para o Estado do Espírito Santo temos os seguintes resultados, conforme Figura 2 e Figura 3:

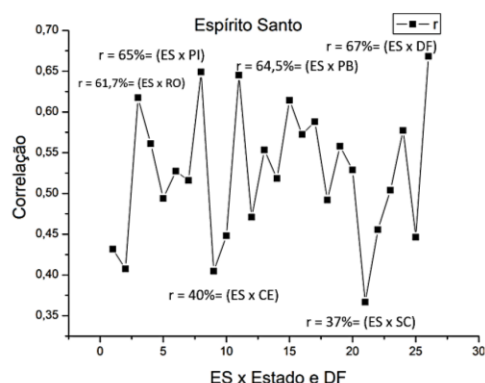


Figura 2: Correlação de Pearson entre ES e os demais Estados (e o DF), para o número de novos casos diários de Covid-19 (valores aproximados). Fonte: Dos autores.

Conforme Mukaka [8], para os coeficientes de Pearson e Spearman as correlações fortes são aquelas em que os valores estão compreendidos entre 0,71 e 0,9 e/ou -0,71 e -0,9. Já as correlações muito fortes, correspondem para os valores maiores que 0,9 ou menores que -0,9 [8]. Desta forma, apresentaremos nas Tabela 1 e Tabela 2, índices de correlações presentes na pesquisa, para o Espírito Santo. Pelo exporto em Mukaka [8], adotamos o coeficiente de correlação para postos de Spearman (r_s), para valores a partir de 0,7, como ponto de corte na apresentação das informações nas tabelas a seguir. Apresentamos na Tabela 1

Tabela 1: Coeficiente de correlação de Pearson (r), Spearman (r_s ou s), Kendall (τ ou k), calculados entre ES e os demais Estados (e o DF), para o número de novos casos diários de Covid-19. Fonte: Dos autores.

Correlações entre Estados		r	r_s		τ	
ES	DF	0,66839	RO	0,7734	RO	0,60342
ES	PI	0,64897	BA	0,76653	BA	0,60121
ES	PB	0,64524	PB	0,75732	PB	0,59521
ES	RO	0,61733	SC	0,7567	SC	0,5876
ES	BA	0,61455	DF	0,75103	DF	0,58105
ES	PA	0,58824	MT	0,74775	PI	0,57626
ES	MT	0,57742	PI	0,74355	MT	0,57012
ES	MG	0,57234	SE	0,7236	SE	0,55293
ES	AC	0,56106	RR	0,7157	RR	0,54794
ES	SP	0,55804	PR	0,70327	PR	0,53727

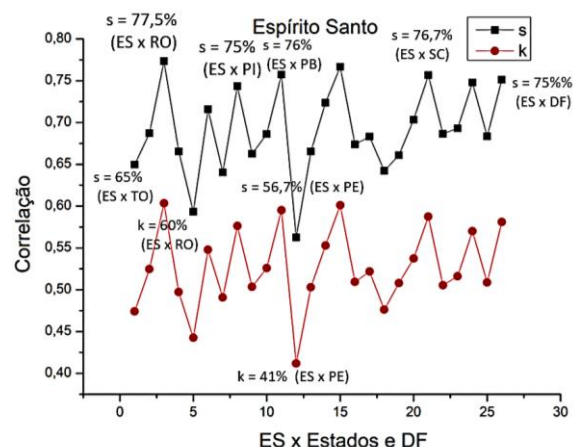


Figura 3: Correlação de Spearman e coeficiente de Kendall entre ES e os demais Estados (e o DF), para o número de novos casos diários de Covid-19 (valores aproximados). Fonte: Dos autores.

Em relação ao número de óbitos diários, para o Estado do Espírito Santo temos os seguintes resultados, conforme Figura 4 e Figura 5:

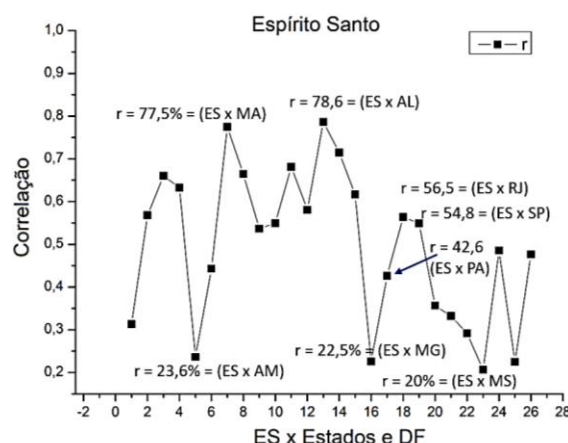


Figura 4: Correlação de Pearson entre ES e os demais Estados (e o DF), para o número de óbitos diários de Covid-19 (valores aproximados). Fonte: Dos autores.

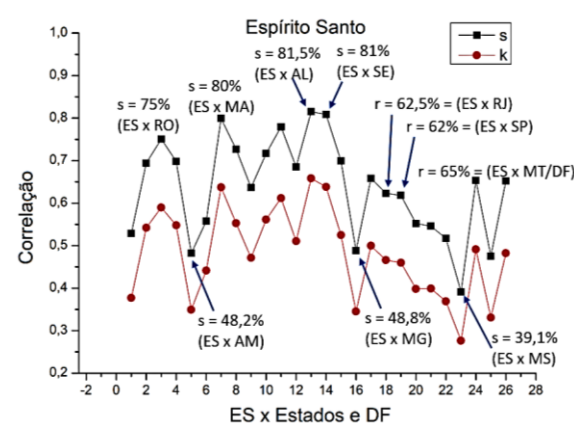


Figura 5: Correlação de Spearman e coeficiente de Kendall entre ES e os demais Estados (e o DF), para o número de óbitos diários de Covid-19 (valores aproximados). Fonte: Dos autores.

Tabela 2: Coeficiente de correlação de Pearson (r), Spearman (r_s ou s), Kendall (τ ou k), calculados entre ES e os demais Estados (e o DF), para o número de óbitos diários de Covid-19. Fonte: Dos autores.

		r		r_s		τ
ES	AL	0,78618	AL	0,81511	AL	0,65831
ES	MA	0,77499	SE	0,80885	SE	0,63835
ES	SE	0,71441	MA	0,79941	MA	0,63737
ES	PB	0,68131	PB	0,77972	PB	0,61171
ES	PI	0,66458	RO	0,75062	RO	0,58987
ES	RO	0,66019	PI	0,7264	RN	0,56116
ES	AC	0,63276	RN	0,71659	PI	0,5526

Em relação ao número de novas infecções diárias, para o Estado do Pará temos os seguintes resultados, conforme Figura 6 e Figura 7:

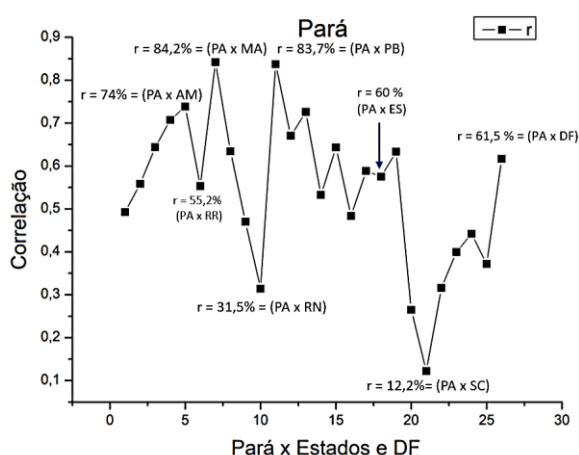


Figura 6: Correlação de Pearson entre PA e os demais Estados (e o DF), para o número de novos casos diários de Covid-19. Fonte: Dos autores.

Tabela 3: Coeficiente de correlação de Pearson (r), Spearman (r_s ou s), Kendall (τ ou k), calculados entre PA e os demais Estados (e o DF), para o número de novos casos diários de Covid-19. Fonte: Dos autores.

		r		r_s		τ
PA	MA	0,84203	PB	0,90119	PB	0,74368
PA	PB	0,83719	MA	0,90016	MA	0,7369
PA	AM	0,73845	AM	0,81825	AL	0,65114
PA	AL	0,7261	AL	0,81711	AM	0,63747
PA	AC	0,70731	AC	0,79959	AC	0,62456
PA	PE	0,67035	PE	0,79703	CE	0,61975
PA	RO	0,64366	SP	0,79405	PE	0,61908
PA	BA	0,64348	CE	0,78437	SP	0,61614
PA	PI	0,63416	RR	0,78126	BA	0,61259
PA	SP	0,63337	BA	0,7808	AP	0,60975
PA	DF	0,61662	AP	0,77999	RR	0,60903
PA	ES	0,58824	SE	0,77637	PI	0,60601
PA	RJ	0,57525	RO	0,77417	RO	0,60101
PA	AP	0,55841	PI	0,76711	SE	0,59602
PA	RR	0,55298	DF	0,73391	DF	0,5596

PA	SE	0,53238	TO	0,72384	TO	0,54174
PA	TO	0,49225	RJ	0,70244	RJ	0,52964

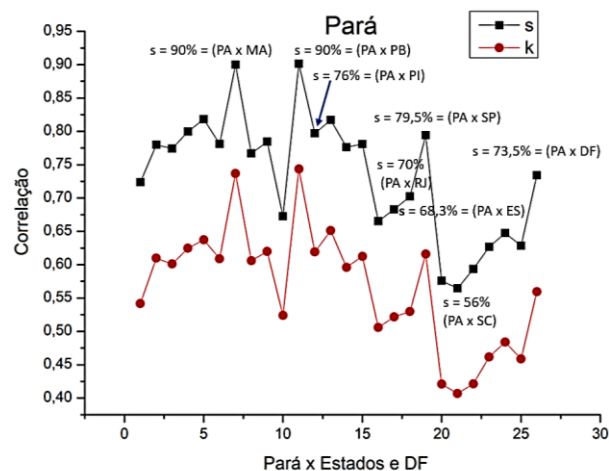


Figura 7: Correlação de Spearman e coeficiente de Kendall entre PA e os demais Estados (e o DF), para o número novos casos diários de Covid-19. Fonte: Dos autores.

Tabela 4: Coeficiente de correlação de Pearson (r), Spearman (r_s ou s), Kendall (τ ou k), calculados entre PA e os demais Estados (e o DF), para o número de óbitos diários de Covid-19. Fonte: Dos autores.

		r		r_s		τ
PA	CE	0,71093	MA	0,84872	MA	0,69022
PA	PE	0,69166	PE	0,8279	PE	0,66429
PA	MA	0,60904	AL	0,79099	AL	0,63251
PA	AP	0,59473	CE	0,75782	CE	0,5938
PA	RJ	0,58276	AP	0,72579	AP	0,57355

Em relação ao número de óbitos diários, para o Estado do Pará temos os seguintes resultados, conforme Figura 8 e Figura 9:

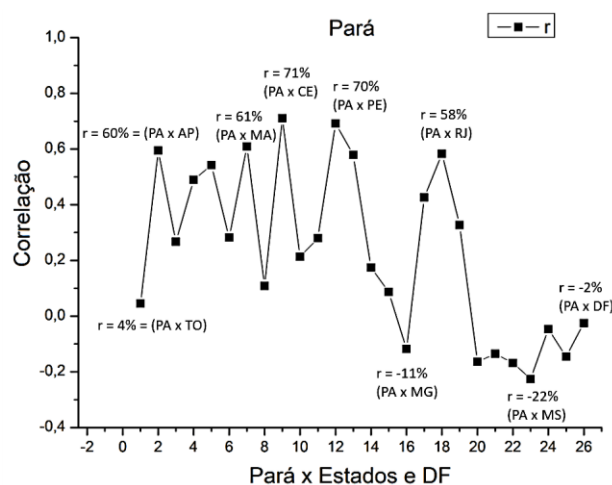


Figura 8: Correlação de Pearson entre PA e os demais Estados (e o DF), para o número de óbitos diários de Covid-19 (valores aproximados). Fonte: Dos autores.

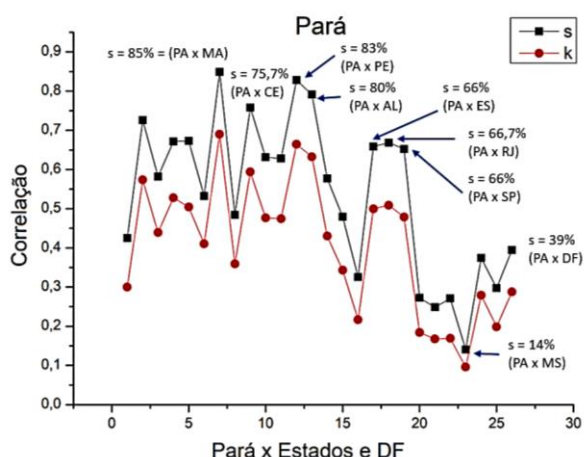


Figura 9: Correlação de Spearman e coeficiente de Kendall entre PA e os demais Estados (e o DF), para o número de óbitos diários de Covid-19 (valores aproximados). Fonte: Dos autores.

Diante das análises realizadas utilizando os diferentes coeficientes de correlação para analisar especificamente os Estados do Pará e do Espírito com os demais Estados Brasileiros e do DF. O coeficiente de correlação de Spearman (r_s), apresentou-se em destaque nas medidas realizadas. Como era de se esperar, o padrão de comportamento do o coeficiente de correlação por postos de Kendall (τ), apresentou-se “idêntico” ao coeficiente de correlação por postos de Spearman (r_s).

De uma forma geral, podemos afirmar para o Espírito Santo, em termos do número diário de infecções que, pelas figuras e tabelas, o ES está fortemente correlacionado com os Estados do Norte e Nordeste e sutilmente com o Sudeste e Centro-Oeste. Em relação às correlações específicas (correlação forte) para o número de óbitos, há uma predominância do ES com a região Nordeste e com a região Norte, respectivamente.

Semelhantemente, o Estado do Pará apresentou uma correlação fortíssima com a região Nordeste, em relação ao número diário de infectados, em que destacamos os Estados da Paraíba e Maranhão. O padrão de uma forte correlação, conforme Mukaka [8], também ocorreu com a própria região Norte e, também, Sudeste e Centro-Oeste. Para as correlações do Estado do Pará em relação ao número de óbitos diários, houve uma predominância com a região Nordeste e com a região Norte, respectivamente.

De acordo as Figura 10 e Figura 11, apresentamos as séries temporais especificamente dos coeficientes aqui utilizados, Pearson (r), Spearman (r_s), e Kendall (τ). Destacamos que os índices “i” e “o” contidos nos gráficos, representam respectivamente “número de infectados diários” e “número de óbitos diários”.

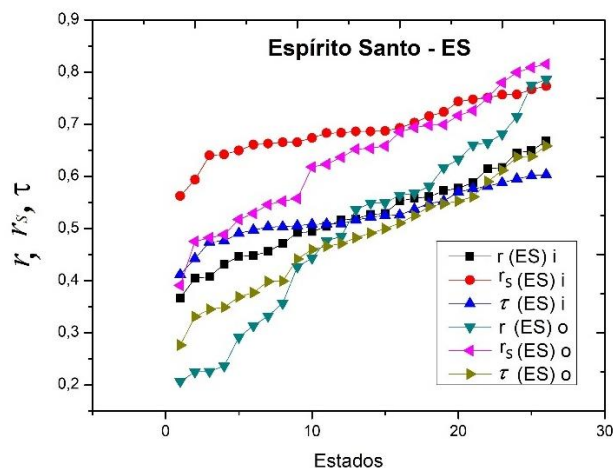


Figura 10: Padrão evolutivo dos coeficientes de correlações, para o Estado do Espírito Santo.

De forma análoga à figura anterior, para o Estado do Pará, apresentamos a Figura 11:

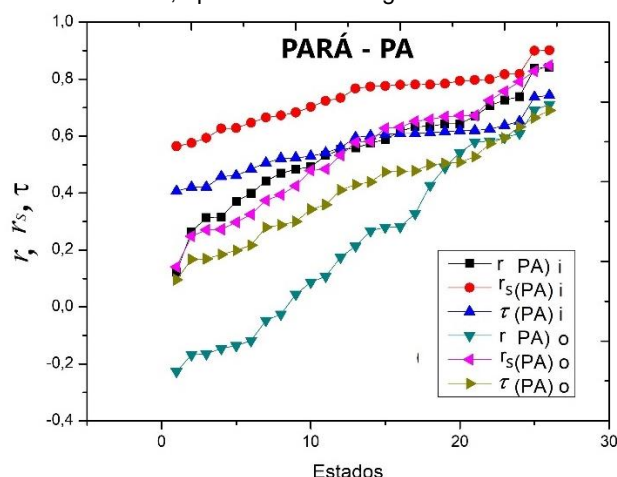


Figura 11: Padrão evolutivo dos coeficientes de correlações, para o Estado do Pará.

Tanto a Figura 10 quanto a Figura 11, ratificam toda a pesquisa realizada neste artigo. Podemos observar que, as séries temporais construídas observando a correlação entre os X e Y Estados Brasileiros, cada ponto, corresponde a um par, que seja para o Pará x (Estado Brasileiro e o DF), quer seja o Espírito Santo x (Estado Brasileiro e o DF). Com as análises destas figuras, podemos observar do menor valor de correlação entre os pares de Estados até o maior. Desta forma, diante deste cenário epidemiológico que estamos enfrentando, temos neste artigo, como avaliar em que medida ocorreu a difusão do Coronavírus nos diversos Estados, pela comparação nos padrões singulares a cada local.

4. Considerações finais

Nosso objetivo neste artigo foi o de apresentar mais uma ferramenta para ajudar a combater o Coronavírus, conforme os modelos computacionais apresentados aqui. Estes modelos representaram o

cenário epidemiológico que o Brasil está enfrentando. Analisamos desde o primeiro caso nacionalmente reconhecido até o dia 30 de novembro de 2020. Apresentamos uma concepção que podemos denominar de Físico-Matemática-Estatística, por que não? Ao tratarmos os Estados Brasileiros e o DF como séries temporais, foi possível utilizarmos os coeficientes de correlações trabalhados neste artigo, bom uma base estatística e matemática, realizando análises físicas pela observância da natureza, um fenômeno não-linear: uma pandemia.

Não foi nosso foco detalhar as nuances estatísticas aqui apresentadas, porém, demonstramos de forma didática como iríamos trabalhar neste artigo, realizando assim, nossa proposição de análise à difusão do Coronavírus, observando as métricas de correlação. Assim, pudemos perceber que tanto o Estado do Espírito Santo quanto o Pará, estão fortemente correlacionados com as regiões Norte e Nordeste e até, de uma fortíssima correlação entre, especificamente o Pará e as supracitadas regiões. Por fim, apresentamos como cada índice estatístico se comportou em seu processo de “evolução”, partindo dos menos valores de correlação entre um par de Estados até o maior.

5. Agradecimentos

Jefferson Nascimento agradece a oportunidade de ter realizado pós-doutoramento no departamento de modelagem computacional no CIMATEC/BA. Marcelo A. Moret agradece ao CNPq pelo suporte financeiro parcial oriundo de sua bolsa de Produtividade em Pesquisa (No. 305291/2018-1).

6. Referências

- [1] NASCIMENTO, J. O. do. Relatório técnico de atividades de pós-doutorado. Salvador/BA: Sistemas Complexos – Cenários Epidemiológicos, 2020. Relatório (Pós-doutorado) – Departamento/Programa de Modelagem Computacional, CIMATEC, 2020.
- [2] PASSOS, C. A. C.; VIEIRA, E. A.; LOURENCO, J. A.; NASCIMENTO, J. O. do. Characteristics and evolution of COVID-19 cases in Brazil: mathematical modeling and simulation. medRxiv - the preprint server for Health Sciences, 2020 (Preprint). Disponível em < <https://www.medrxiv.org/content/10.1101/2020.10.14.20212829v1>>. Acesso em: 09 Nov. 2020. doi: <https://doi.org/10.1101/2020.10.14.20212829>.
- [3] NASCIMENTO, J. O. do. Redes sociais e complexas: redes semânticas do Ensino de Física Brasileiro. 123f. Tese (Doutorado em modelagem computacional e tecnologia industrial) – CIMATEC, Salvador, 28 jun. 2019.
- [4] BAUER, Lidiane. Estimção do coeficiente de correlação de spearman ponderado. 2007. 95 f.

Dissertação (Mestrado em Epidemiologia). Curso de PósGraduação em Epidemiologia, Faculdade de Medicina, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2007.

[5] CONOVER, W. J. Practical nonparametric statistics. 3rd . ed. New York: John Wiley & Sons; 1998

[6] Miot, H. Análise de correlação em estudos clínicos e experimentais. *Jornal Vascular Brasileiro*, 17, 275-279, 2018.

[7] SANTOS, S.S. Estudo comparativo de medidas de dependência e aplicações em dados de expressão gênica. Universidade de São Paulo, USP, São Paulo, 2012.

[8] MUKAKA, M. M. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med J.* 2012;24(3):69-71. PMID:23638278.