

Identificação de Padrões em Ciclos Epidêmicos e Métodos Para Estimativa de Suas Durações:

Estudo de Caso COVID-19

Eduardo Atem de Carvalho¹ e Rogerio Atem de Carvalho²

1 – Bacharel, MSc e PhD em Engenharia Mecânica, Laboratório de Materiais Avançados - Universidade Estadual do Norte Fluminense

2 – Bacharel em Informática, MSc e DSc em Engenharia de Produção, Pólo de Inovação – Instituto Federal Fluminense

Resumo

Este trabalho apresenta diversos ciclos epidêmicos do COVID-19 já praticamente encerrados em países, estados e cidades e através de métodos numéricos simples e bem conhecidos, os normaliza. Fica evidente que existe um padrão praticamente universal entre eles, de forma triangular. É possível também achar os ciclos similares com escalas muito próximas e desta forma empregar os casos com ciclos já encerrados para se prever o término dos ciclos ainda em andamento. Três métodos são apresentados e discutidos e o caso da cidade do Rio de Janeiro é apresentado em mais detalhe.

1) Introdução

Com a recente Pandemia por COVID-19, uma questão que se colocou imediatamente foi a duração de seu ciclo. Como o vírus não se espalha de maneira instantânea por todo o Globo, haverá naturalmente países onde ela estará em estágio mais avançado, por ter se iniciado antes, e outros em que estará mais “atrasada”. Assim, o meio mais comum de se tentar fazer esse tipo de predição é usar o ciclo de contaminação dos lugares onde esta contaminação se iniciou antes para se tentar prever como será a evolução da doença onde se iniciou depois. Mas, para fazer comparação, sabendo que os tamanhos das populações são diferentes, algumas medidas são necessárias para que o padrão de comportamento local da pandemia seja devidamente identificado.

Assim, este trabalho se propõe a empregar a combinação do uso da normalização de ciclos epidêmicos já encerrados superpostos a valores de bancos de dados públicos para chegar a algumas conclusões e estimativas sobre durações de ciclos ainda em andamento. Todo o embasamento teórico e numérico empregado é apresentado, demonstrado e discutido antes de qualquer projeção ser feita. Neste estudo o termo “epidemia” irá se referir a um episódio local, seja ele cidade ou mesmo país. Quando o termo usado for “pandemia”, os autores se referem ao fenômeno mundial.

Neste início de século XXI diversas entidades transnacionais e órgãos governamentais se convenceram da inevitabilidade e vem se preparando para uma pandemia devastadora nos moldes da Peste Negra medieval [1]. O porquê disto em muito escapa ao escopo deste trabalho. E inevitavelmente junto com as discussões

vem as previsões, quase sempre calamitosas [2] e baseadas em “modelos matemáticos” supostamente infalíveis e oniscientes. Pois bastou ocorrer um episódio de maior magnitude, como o COVID-19, para que todas as previsões baseadas nestes modelos fracassassem [3], mas o pânico gerado e as medidas draconianas que se seguiram deixarão consequências por gerações.

A partir da observação de dados disponíveis e sem tentar elaborar qualquer modelo de natureza preditiva determinística, este trabalho analisa diversos casos de países que já enfrentaram o que no momento presente parece ser o ciclo crítico da doença, permitindo que alguns padrões mais óbvios sejam detectados.

As discussões dos fatores agravantes ou atenuantes da doença também escapam ao modesto escopo deste trabalho.

E por fim, este trabalho irá somente trabalhar com o número de mortes registrados. O número real de casos de COVID-19 não fatais, nas suas diversas intensidades, é impossível de se determinar de forma correta e absoluta, uma vez que isto implicaria em que a totalidade das populações fossem testadas continuamente, ou que pelo menos amostras contínuas e significativas fossem testadas ao longo do tempo de duração das epidemias locais, o que não foi feito [4], embora alguns países como a Alemanha tenham testado quantidades significativas de sua população no início da pandemia. Desta forma, o número de óbitos, embora também não seja perfeito, é menos sujeito a fatores como os acima descritos.

2) DA NATUREZA DOS DADOS OBSERVADOS

A tentativa de se descrever os diferentes ciclos epidêmicos que compõem a atual pandemia esbarra muitas vezes na qualidade dos dados que são tornados públicos. As mortes ocorrem em datas diferentes daquelas registradas diariamente pelos diversos governos e autoridades sanitárias. Isto é inevitável em um primeiro momento e será corrigido ao longo do tempo, de forma que as datas corretas quando os óbitos ocorrerem se tornará disponível. No presente momento este é um processo em andamento. Para se introduzir e comentar e apresentar a metodologia conhecida previamente citada como Método das Médias Móveis, apresenta-se o caso da Suécia, que serve de exemplo, uma vez que os primeiros dados disponíveis até o final de mês de junho de 2020, traziam as mortes nos dias de registro, desta forma, a distribuição de fatalidades sofre uma distorção que gera uma aparência de “serrote” nos gráficos dos valores apresentados (Figura 1), sendo que nos finais de semana nota-se uma clara ausência de óbitos, seguida de uma explosão desses valores nos inícios das semanas. Uma maneira de contornar este efeito é aplicar o chamado Método das Médias Móveis (MMM), no qual o valor diário de mortes é substituído pela média dos 6 dias anteriores e mais daquela data. Com a divulgação pela Suécia das datas corretas de óbito, pode-se comparar a eficácia do MMM com os valores reais.

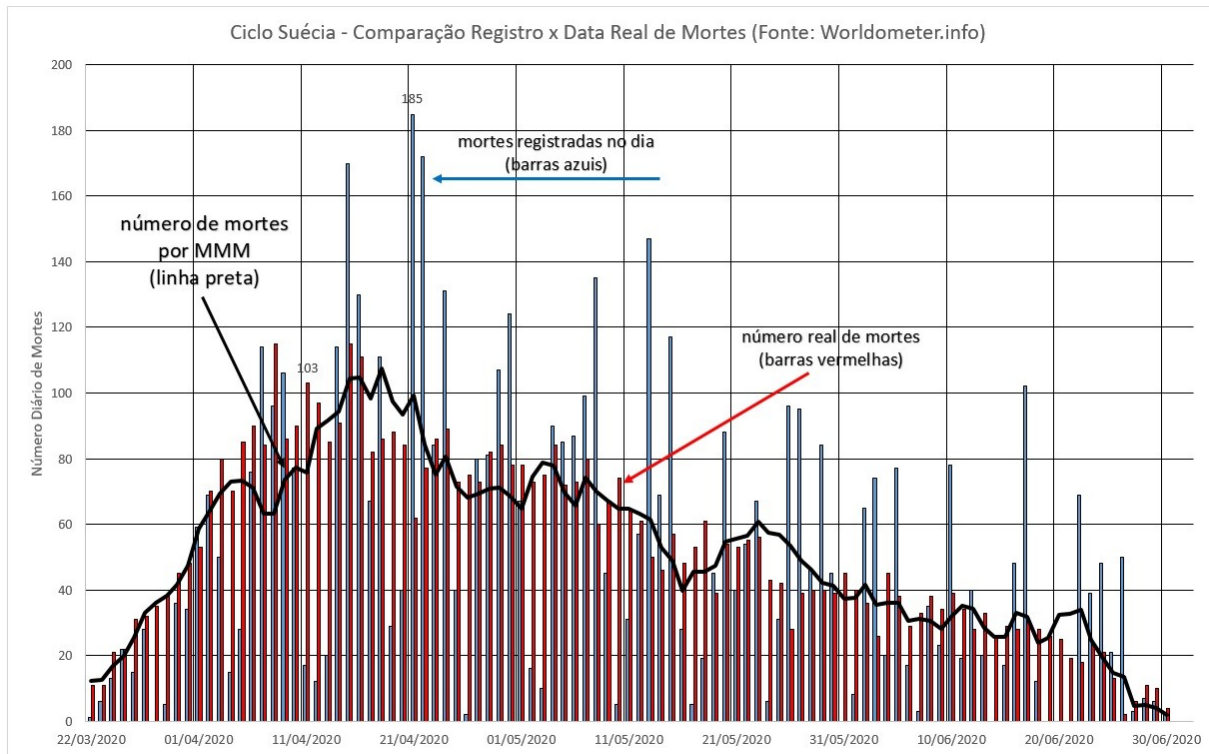


Figura 1 – Comparação entre os valores apresentados anteriormente para a Suécia (data de registro das mortes – barras azuis [6]) e os atualizados para as datas reais das mortes (barras vermelhas [7]). A curva preta foi obtida pelo MMM.

Para julgar se o MMM é um método aplicável para minimizar os efeitos da sazonalidade nas datas de registro das mortes, será usado o critério estatístico conhecido como R^2 ou R-Quadrado [5], que neste caso determina o quão bem os valores do MMM descrevem os dados verdadeiros, ou seja, caso seja usado o MMM, o quão próximo estaremos da distribuição de mortes conforme a data real ocorrida. O valor de R^2 pode variar entre 0 e 1 e de uma forma geral, quanto mais próximo de 1 o valor de R^2 está, melhor foi o desempenho do MMM. R^2 é definido da seguinte forma:

$$R^2 = \left(\frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}} \right)^2 \quad (1)$$

Onde, neste caso, x são os pontos contendo os dados referentes ao número real de mortes ocorridas em cada data do ciclo epidêmico e y são os valores gerados pelo Método das Médias Móveis e atribuídos aos dias observados. Três modelos iniciais básicos foram testados para o caso sueco:

- a) Média Móvel tendo como valor definido (aquele que aparece nos gráficos) o de posição central. Por exemplo: em uma semana de 7 dias, se iniciando no domingo e acabando no sábado, a média dos 7 valores será

computada no valor de quarta-feira. Este é o caso do Método de Média Móvel com valor central (MMMC).

- b) Média Móvel tendo como valor definido o de posição inicial. Ou seja: em uma semana de 7 dias, se iniciando no domingo e acabando no sábado, a média dos 7 valores será computada no valor de domingo. Este é o caso do Método de Média Móvel com Valor Inicial (MMMI).
- c) Média Móvel tendo como valor definido o de posição final. Ou seja: em uma semana de 7 dias, se iniciando no domingo e acabando no sábado, a média dos 7 valores será computada no valor de sábado. Este é o caso do Método de Média Móvel com Valor Final (MMMF).

A Tabela 1 apresenta os valores obtidos para os casos a, b e c acima descritos e os compara.

Tabela 1 – Variáveis estatísticas básicas para o número de mortes na Suécia no período de 22/03/2020 até 30/06/2020.

	Média	Variância Amostral	R ²	Obs.
Mortes Data Real	53	758	1,0000	Valor de referência [7]
Mortes Data de Registro	52	1974	0,2409	Valores inicialmente disponíveis [6]
MMMC	52	626	0,7241	Calculado segundo (a)
MMMI	52	648	0,8448	Calculado segundo (b)
MMMF	52	645	0,5472	Calculado segundo (c)

Obs.: As casas decimais se fazem necessárias porque muitas vezes o valor de R² se aproxima muito da unidade e só se pode perceber a sua mudança quando se opera com um número mínimo de casas decimais. Os valores de fora deste período foram excluídos dos cálculos.

Fica claro que o Método de Média Móveis Inicial (descrito no item (b) acima) é o que melhor descreve o fenômeno estudado, como seria esperado, já que as mortes só podem ser registradas após terem acontecido, criando inevitável atraso nos valores divulgados durante um ciclo epidêmico. Porém, é observável também, que o método proposto (MMMI) consegue converter os valores de registro de maneira aceitável para valores próximos aos verdadeiros. Este estudo usará o MMMI em todos os seus gráficos e simulações a seguir.

3) O CICLO EPIDÊMICO CHINÊS

Logo no início da pandemia, no seu ponto de origem, a China completou o ciclo epidêmico em 8 semanas e os dados sobre infectados e fatalidades ficaram disponíveis para análise [8]. A distribuição de fatalidades por dia de epidemia parecia seguir a clássica curva gaussiana e a duração do ciclo indicar que a doença cumpria seu ciclo conforme previsto (Figura 2) e esperado por diversos modelos e fontes com credibilidade analítica [9]. Estes fatos contribuíram para diversas previsões e análises iniciais catastróficas ou otimistas demais. Ao longo do tempo, mesmo as previsões de grupos mais experientes continham margens de erro e expectativas que chegavam a 300% de variação [4]

A discussão sobre a veracidade dos números chineses escapa completamente ao escopo deste trabalho, mas denota que algumas reações foram extremadas, seja por menosprezar a epidemia – ao ver um país 1,4 bilhão de habitantes ter relativamente tão poucos óbitos, seja por prever um crescimento apocalíptico, ao aplicar a exponencial na subida da pandemia no Ocidente

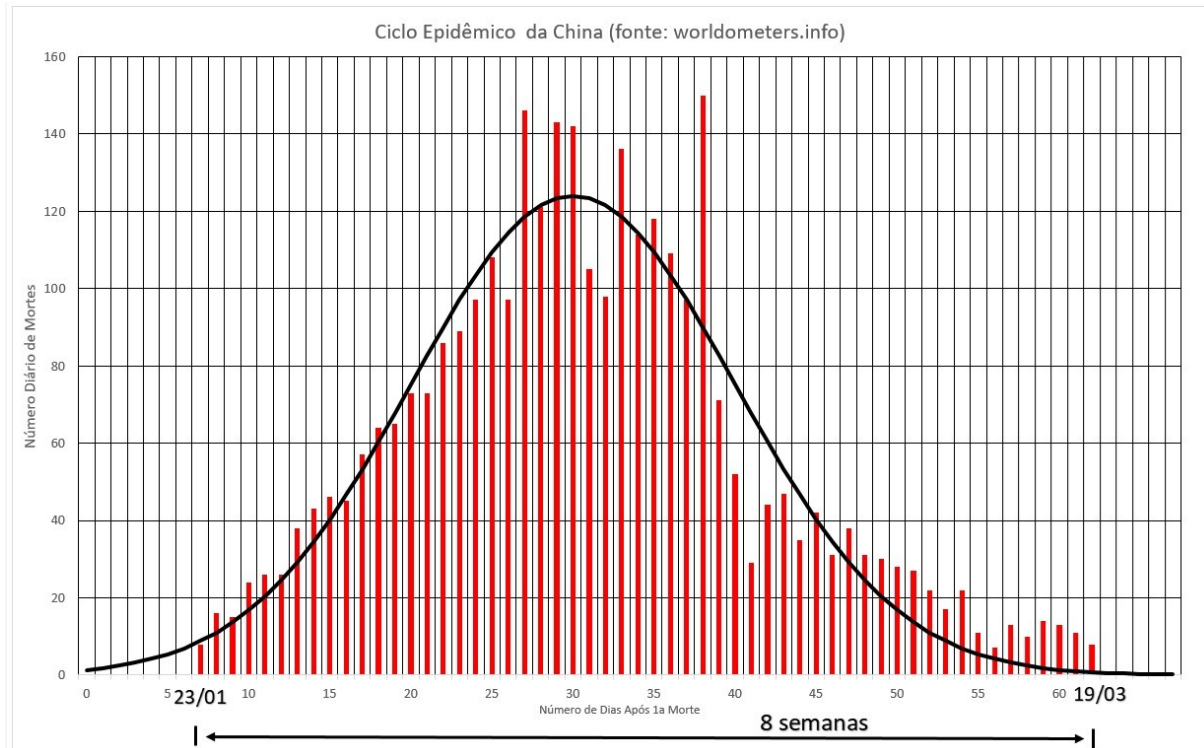


Figura 2 – Mortes causadas pelo COVID-19 na China.

4) CICLOS EPIDÊMICOS OBSERVADOS

No momento da escrita deste trabalho (julho 2020), diversas cidades, regiões, estados e países já completaram o que aqui vai ser definido como o Ciclo Mais Letal da Epidemia (CMLE), que é quando o número de mortes aumenta diariamente, na média, até atingir um pico e depois começa a decrescer continuamente até atingir um valor mínimo. Depois deste período, a ocorrência de mortes continua de forma intermitente, porém relativamente muito pequena e oscilante decaindo até certos níveis de mortes diárias, onde então se torna aparentemente crônica e apresenta valores relativamente baixos, porém não nulos. Ainda não nos é possível determinar se são fenômenos universais ou associados a condições específicas de evolução da doença, como clima, hábitos alimentares, genética, histórico de vacinação etc.

4.1) Espanha

O primeiro país europeu de grande porte a debelar seu surto epidêmico, apresentou durante o CMLE o seguinte ciclo, como descrito da Figura 3. Considera-se o final do ciclo o último dia antes do primeiro dia com zero mortes [10].

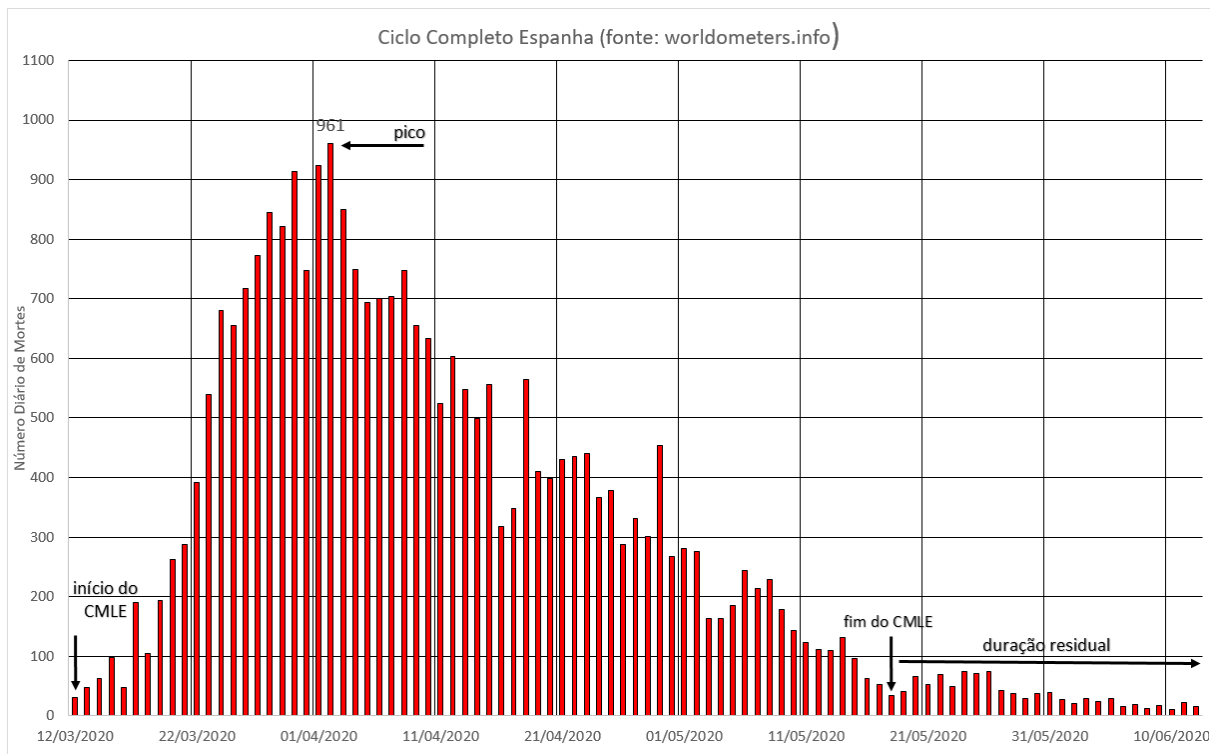


Figura 3 – CMLE da Espanha.

4.2) Holanda

A Holanda conseguiu conter a epidemia dentro de seu território e na data de atualização destes dados (03 de julho de 2020) já havia reduzido o número de fatalidades às unidades. A Figura 4 mostra os dados para a Holanda [11].

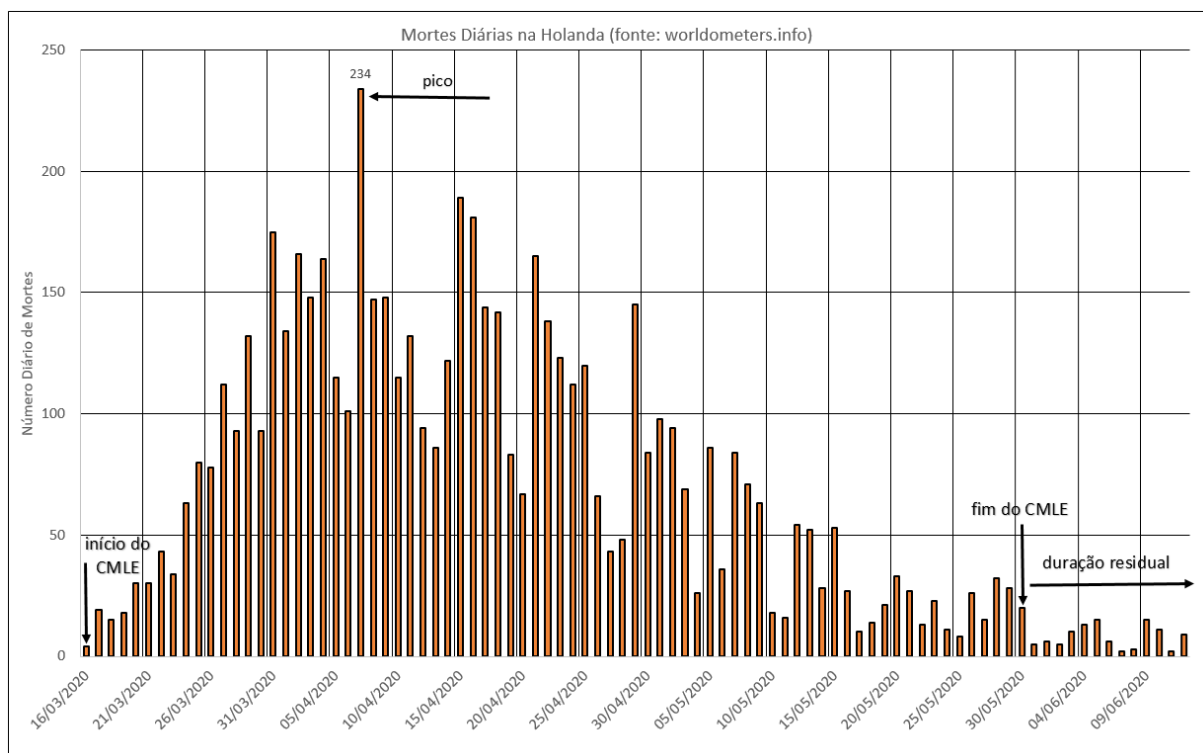


Figura 4 – Ciclo Mais Letal da Epidemia (CMLE).

4.3) Itália

País que esteve no epicentro europeu da crise, apresenta evolução no número de mortes (Figura 5) que indica a superação do CMLE [12]. Este trabalho aponta a existência dos chamados Falsos Picos. Estes são máximos locais que foram registrados durante o ciclo de subida ou descida na tendência de mortes, mas não são pontos de inflexão. Para que um ponto seja considerado como Pico, é preciso que a partir dele se note a tendência à queda no número de mortes. Esta queda não dará de forma linear, mas existe uma tendência óbvia, numérica e visual que indica tal coisa.

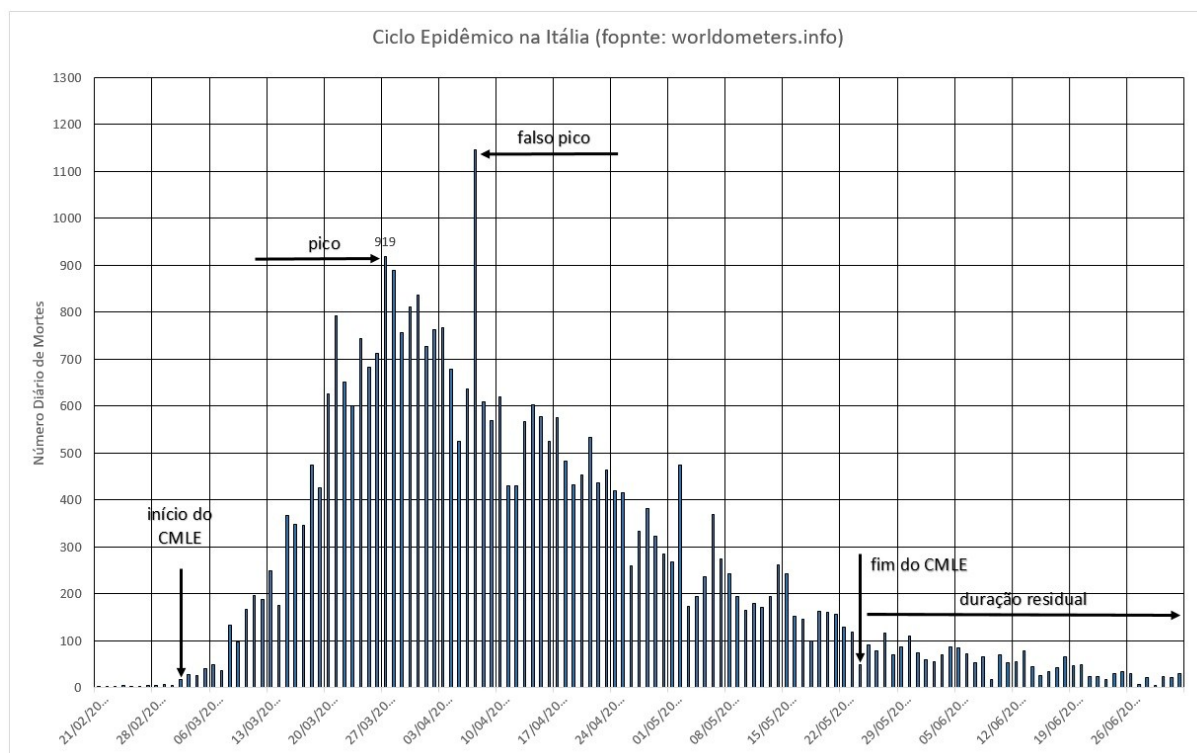


Figura 5 – Ciclo Italiano.

4.4) Nova Iorque (NYC)

Cidade americana duramente atingida pela epidemia, apresentou ciclo típico encontrado por este estudo. A Figura 6 apresenta o ciclo epidêmico na cidade de Nova Iorque (NYC) [13].

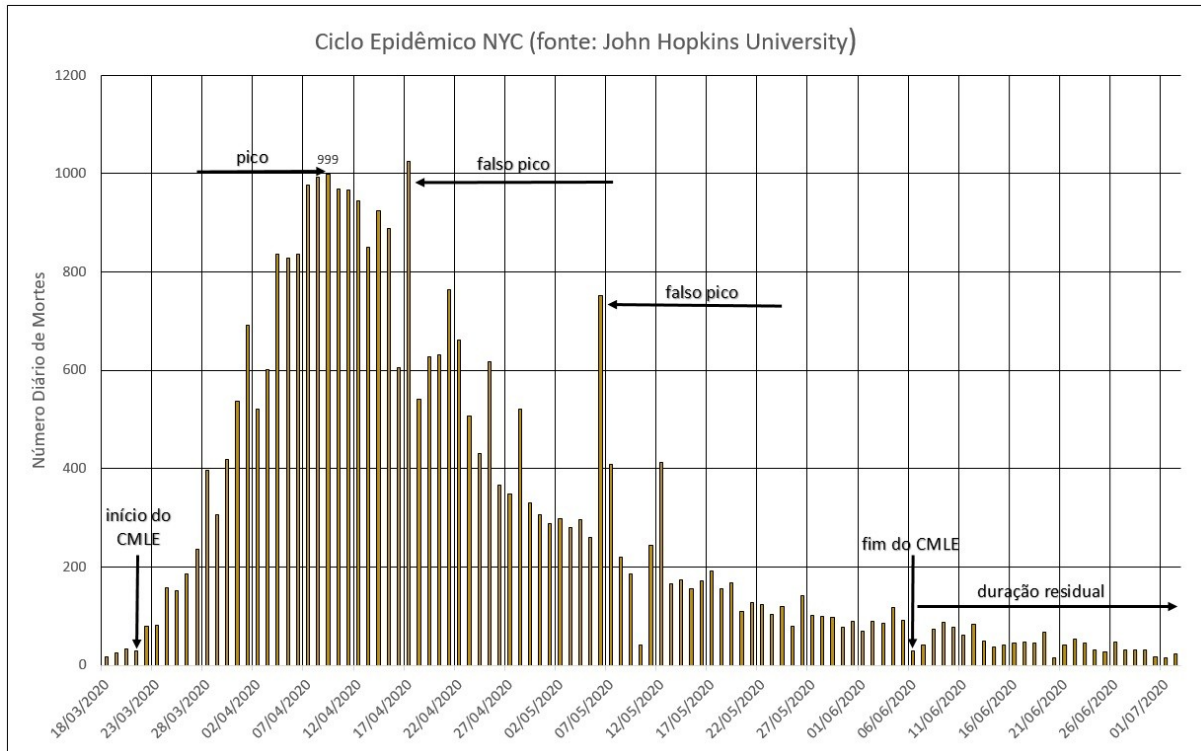


Figura 6 – CMLE da cidade de Nova Iorque (NYC).

4.5) Alemanha

A Alemanha foi descrita desde o princípio da pandemia como sendo um país que geriu a crise de forma exemplar, testando parcelas significativas de sua população e controlando e liberando o movimento em público baseando-se em quantitativos e percentuais de casos bem conhecidos. A Figura 7 apresenta a evolução das mortes na Alemanha [14].

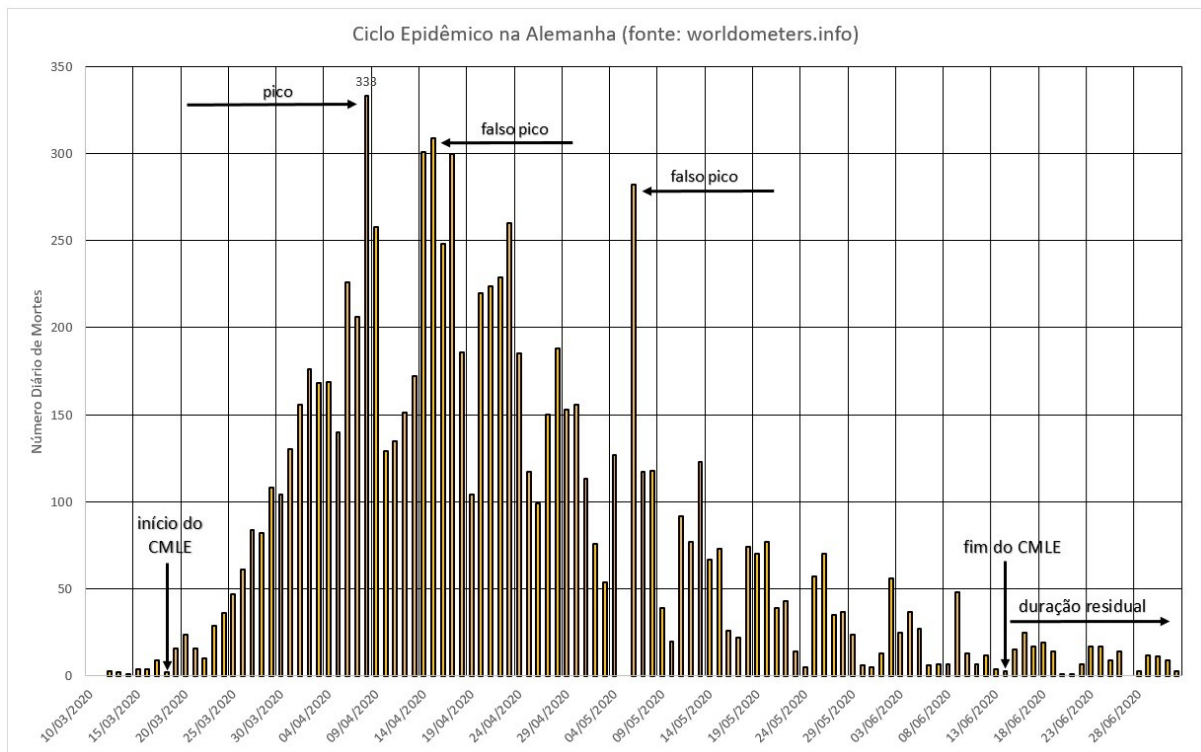


Figura 7 – Ciclo Alemão.

4.6) Suécia

País europeu que não adotou as práticas de isolamento social radical como seus vizinhos, apresenta um ciclo de aspecto não muito diferente de todos os outros países europeus [7]. A Figura 8 apresenta os valores das mortes ocorridas, já corrigidas para as datas em que efetivamente ocorreram e não a de registro.

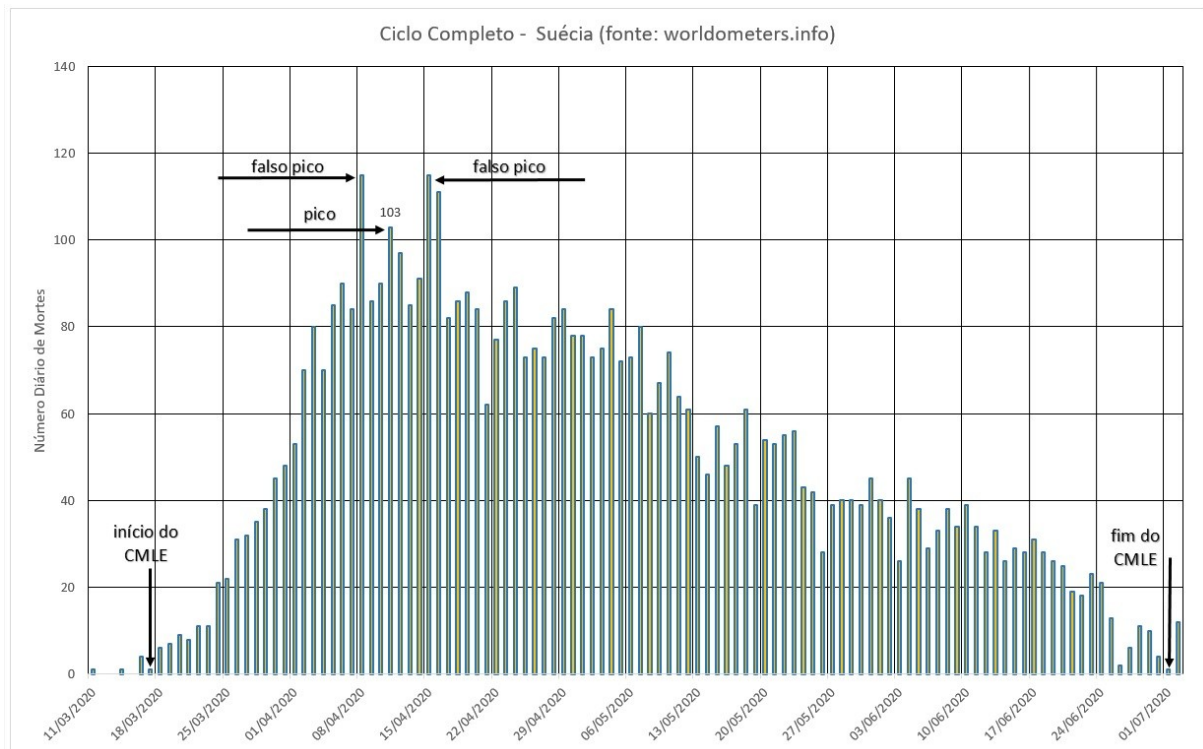


Figura 8 – Ciclo Mais Letal da Epidemia na Suécia.

5) CARACTERÍSTICAS ADIMENSIONAIS DOS CICLOS EPIDÊMICOS

De uma maneira geral os ciclos epidêmicos aqui descritos apresentam algumas características geométricas comuns, sendo a principal um aspecto triangular (Figura 9), onde se forma um lado menor, que corresponde a um aumento médio diário do número de mortes até que se atinja um pico. Este pico pode ser facilmente identificável ou requer extrapolação de uma reta, porque os valores oscilam naturalmente e alguns pontos espúrios (falsos picos) podem surgir. Após o pico se forma um período onde o número de mortes ocorridas diariamente tende a decrescer em média. Esse período, para os casos observados, é mais longo que o anterior.

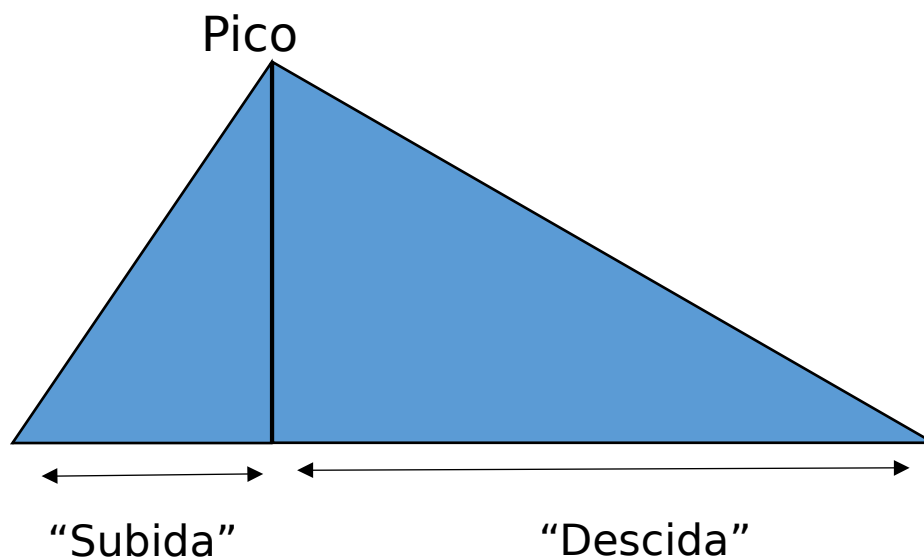


Figura 9 – Formato genérico dos ciclos letais de COVID-19 aqui descritos.

Desta forma, a Tabela 2 reúne valores dos ciclos ditos triangulares apresentados no item 4.

Tabela 2 – Proporções entre os tempos de subida até o pico dos óbitos e de descida até o fim do ciclo mais grave da doença.

Local	Início	Pico	Fim	Dias até Pico	Dias até Fim	Proporção entre Subida e Descida
Espanha	12/03	02/04	18/05	21	45	2,1
Itália	07/03	27/03	24/05	20	57	2,9
Suécia	17/03	11/04	01/07	25	81	3,2
Alemanha	18/03	08/04	14/06	21	69	3,3
Nova Iorque	21/03	09/04	06/06	19	55	2,9
Holanda	16/03	07/04	31/05	22	53	2,4

Os valores listados na Tabela 2 indicam que o período de ascensão da doença, em países de tamanhos relativamente pequenos ou cidades, é de cerca de 21 dias, em uma faixa que vai de 19 a 25 dias até atingir ao chamado pico. A partir daí até o encerramento deste período crítico se passam cerca de 60 dias, em uma faixa que vai de 45 a 81 dias, a razão entre os dois períodos oscila entre 2,1 e 3,3, com uma média de 2,8. A Tabela 3 apresenta agora os valores do número de mortes nos períodos descritos acima.

Tabela 3 – Proporções entre o número de mortes associadas ao ciclo de subida até o pico e de descida até o fim do ciclo mais grave da doença.

Local	Início	Pico	Fim	Mortes até Pico	Mortes até Fim	Proporção entre Subida e Descida
Espanha	12/03	02/04	18/05	10293	16999	1,6
Itália	07/03	27/03	24/05	8937	24082	2,7
Suécia	17/03	11/04	01/07	1255	4141	3,3
Alemanha	18/03	08/04	14/06	2323	6521	2,8

NYC	21/03	09/04	06/06	9859	20442	2,1
Holanda	16/03	07/04	31/05	2081	3855	1,9

Os valores listados na Tabela 3 indicam que o número de mortes durante o período de ascensão da doença, em países de tamanhos relativamente pequenos ou cidades, é de cerca de 5791 mortes, em uma faixa que vai de 1255 a 10293 óbitos até atingir ao chamado pico. A partir daí até o encerramento deste período crítico ocorrem cerca de 12673 mortes, em uma faixa que vai de 4141 a 24082 óbitos. A razão entre os números de mortes oscila entre 1,6 e 3,3, com uma média de 2,4.

6) SIMILARIDADES ENTRE CICLOS

Esse estudo parte do princípio que uma vez removidos os efeitos de escala, o que sobra é um espectro de proporções do ciclo epidêmico. Então, ao se submeter os dados ao Método das Médias Móveis Inicial, ocorre uma minimização do efeito da sazonalidade nos registros das mortes, causada pelos finais de semana, feriados e outras peculiaridades locais.

Após se dividirem todos os valores já anteriormente transformados pelo pico da série (pico agora determinado pelo MMMI), os valores passam a ficar sem dimensão e a se situarem entre 0 e 1. Desta forma, todos os ciclos epidêmicos podem ser comparados entre si, uma vez que o que resta são as proporções entre a subida, o pico e descida do ciclo. O período de tempo não se altera.

Ocorre então a hipótese de que locais diferentes podem, obedecendo a regras comportamentais diferentes, apresentar mesmo comportamento. É o que se vê a seguir, indicando possivelmente que os ciclos epidêmicos são indiferentes às políticas públicas sanitárias vigentes [15,16].

7) PREVISIBILIDADE POR SIMILARIDADE

Alguns ciclos epidêmicos observados foram submetidos aos métodos numéricos acima descritos, tendo seus dados sido submetidos a uma primeira transformação que é a aplicação do MMMI, Método das Médias Móveis Inicial, onde a média dos valores observados ao longo de 7 dias, é atribuída ao primeiro deles e assim sucessivamente. Isto permite que a sazonalidade dos registros dos óbitos seja atenuada. A segunda transformação é a normalização, onde todos os dados de mortalidade de um dado país são divididos pelo seu número de pico, isto vai fazer com que a maioria dos seus valores se situe entre 0 e 1 (a menos dos falsos picos). Essas duas transformações consecutivas permitem que seja feita uma comparação de comportamentos entre ciclos, bem como a comprovação de que diversos ciclos epidêmicos, dentro da pandemia, apresentam semelhança entre si.

Com estas ferramentas triviais é possível estimar a duração e comportamento geral de um episódio local, ainda que este, em termos absolutos, não apresente o mesmo número de óbitos ou dias de duração que seu semelhante. O que se mantém aproximadamente constante são as proporções de cada ciclo. São estas que guardam semelhanças entre si. Um exemplo de aplicação desta técnica, de

forma bastante refinada, com largo sucesso é a previsão de desempenho de atletas profissionais e times feitas pelo grupo de estudo liderado pelo matemático Nate Silver e conhecido como CARM-elo [17].

A seguir serão apresentados alguns exemplos de comparações de ciclos epidêmicos semelhantes.

7.1) Holanda e Reino Unido (UK)

A Holanda e o Reino Unido (UK) apresentam uma concordância de proporções de seus ciclos epidêmicos bastante razoável desde o princípio até o pico, durante o patamar no auge da doença e mesmo nos 2/3 iniciais da queda. A Holanda apresenta uma queda mais rápida na parte final deste ciclo epidêmico analisado, conforme pode ser observado na Figura 10.

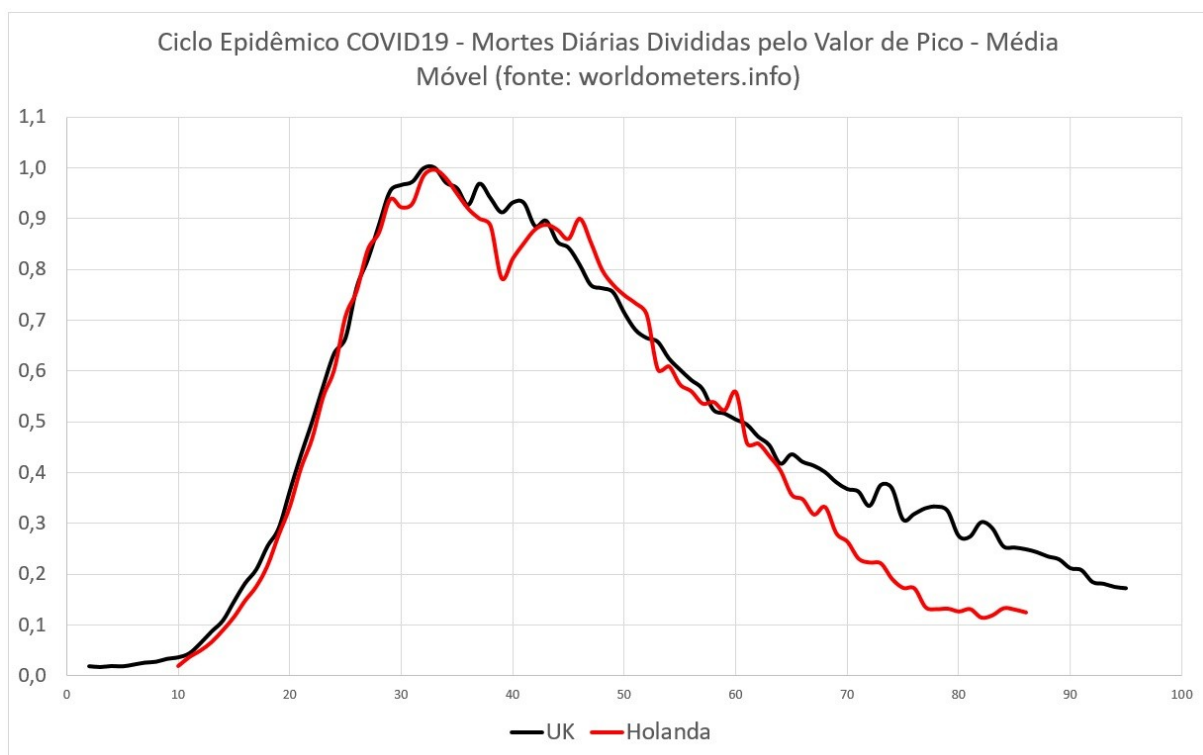


Figura 10 – Comparação entre os ciclos epidêmicos do Reino Unido (UK) e Holanda, após seus dados serem submetidos ao MMMI e normalização.

7.2) Nova Iorque (NYC) e Espanha

Atendo-se apenas às proporções e não aos números absolutos, como aqui proposto, fica clara a grande semelhança entre os ciclos epidêmicos da cidade de Nova Iorque e a Espanha. Observa-se então a existência de países que apresentam ciclos com proporções semelhantes ao de cidades, de acordo com a Figura 11.

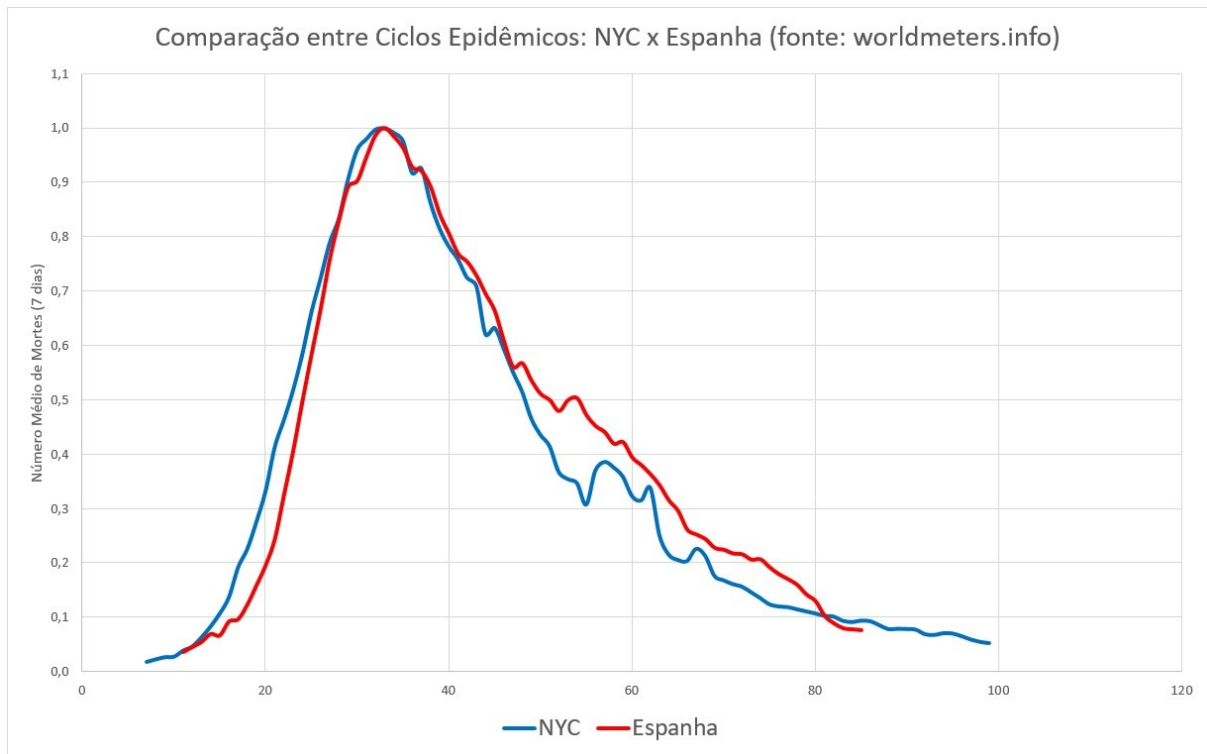


Figura 11 – Comparaç o entre os ciclos epid micos da cidade Nova Iorque (NYC) e Espanha, ap s seus dados serem submetidos ao MMMI e normalizaç o.

7.3) Nova J rsei (New Jersey) e B lgica

Donos de ciclos epid micos bastante peculiares, pelo menos nos registros apresentados, B lgica e o estado americano de Nova J rsei apresentam bastante similaridade se comparados entre si. Figura 12.

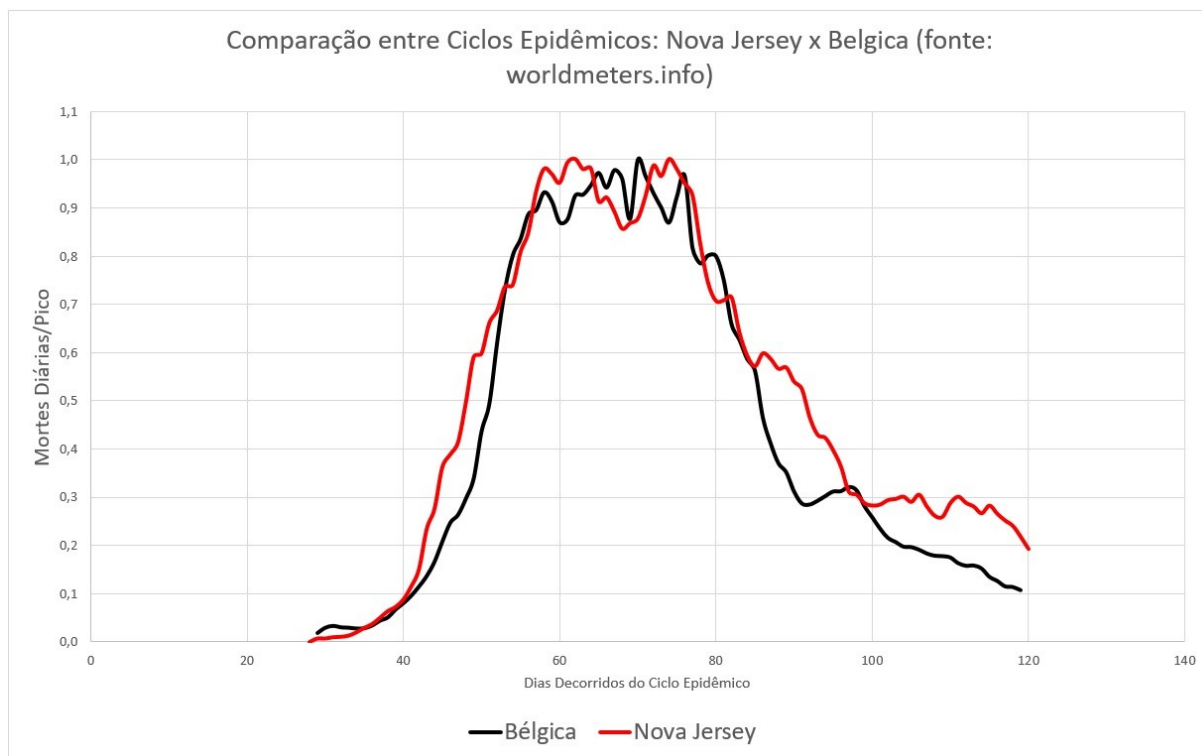


Figura 12 – Comparação entre os ciclos epidêmicos do Nova Jérsei e Bélgica, após seus dados serem submetidos ao MMMI e normalização.

7.4) EUA (USA) e Suécia

Países completamente diferentes nas suas extensões territoriais, complexidade humana, posição geográfica etc, no entanto as proporções na evolução de seus ciclos epidêmicos críticos se apresentam bem semelhantes.

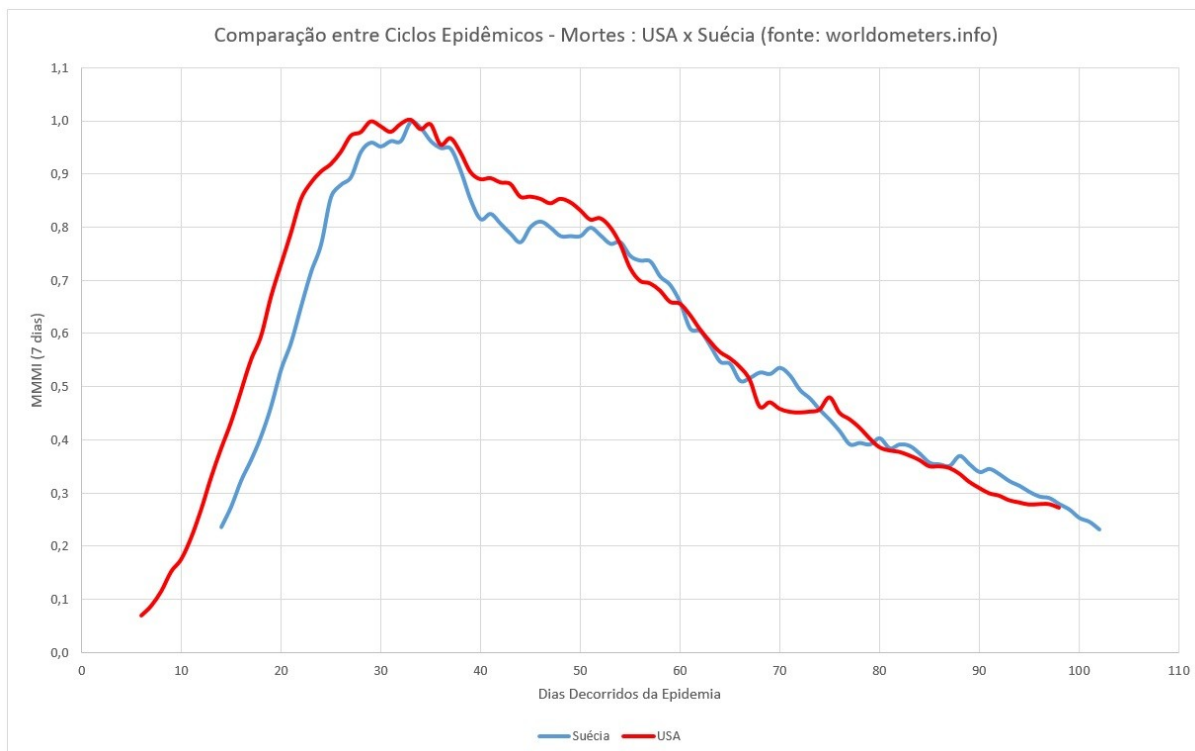


Figura 13 – Comparaç o entre os ciclos epid micos dos EUA (USA) e da Su cia, ap s seus dados serem submetidos ao MMMI e normaliza o.

8) O BRASIL, ESTADO DO RIO DE JANEIRO E A CIDADE DO RIO DE JANEIRO

O Brasil, se observado como um todo apresenta o mesmo ciclo lento de subida que a  ndia e M xico; e um ciclo longo como os EUA. S o todos pa ses com grandes extens es territoriais e concentra es populacionais. A R ssia, Austr lia e o Canad  s o pa ses grandes, mas com baixa densidade populacional, portanto com outro tipo de ciclo epid mico. Os autores entendem que pa ses como o Brasil devem ser vistos e lidados de maneira local e no m ximo regional, no tocante ao estudo dos ciclos epid micos, uma vez que regi es que iniciaram seus ciclos no in cio da pandemia, j  est o no fim ou pr ximo dele e aqueles que apenas tiveram casos residuais at  o presente momento, podem iniciar o ciclo agudo a qualquer instante. Vistos em conjunto parecem um ciclo longo e diferente de todos os apresentados aqui neste trabalho, mas se vistos individualmente, s o bastante semelhantes aos ciclos “triangulares” apresentados no item 4. Pa ses de grande popula o e dimens es continentais, podem ter sua curva de ciclo epid mico constru da a partir da convolu o das curvas das cidades de maior popula o.

As fontes de dados empregadas neste estudo relativas ao Brasil s o os s tios eletr nicos do Minist rio da Sa de [18] e chamado Portal da Transpar ncia, do CNJ [19]. O Minist rio da Sa de (MdS) divulga os dados relativos  s mortes nas datas de registro, da mesma maneira que diversos pa ses do mundo aqui apresentados, implicando no surgimento do inevit vel aspecto de “serrote” nos gr ficos. O segundo s tio apresenta em uma de suas op es o n mero de mortes nas datas em que realmente ocorreram, mas inevitavelmente sofrem de um atraso na atualiza o de

valores, devido ao tempo necessário para que os familiares registrem a morte e as informações passem a fazer parte do sistema. As ferramentas numéricas apresentadas minimizam os efeitos descritos e permitem que se tenha uma visão mais clara da epidemia no Brasil.

8.1) O Ministério da Saúde

O Ministério da Saúde divulga diariamente o somatório do número de óbitos recebidos e verificados, das secretarias de saúde estaduais e do Distrito Federal. São os números registrados naquele dia. A morte pode ter acontecido muitos dias antes, mas será contabilizada na data de registro. O efeito disto são quedas abruptas nos finais de semana e picos extremos nos meios das semanas. Mas seus números refletem na média (MMMI) os valores verdadeiros dos óbitos ocorridos no país. Se há adulterações de valores, isto ocorre nas origens dos valores e não nestas totalizações, portanto este trabalho irá considerar esta fonte como sendo válida. Confirmando a validade desta fonte, existe a referência seguinte, cujos valores se assemelham bastante com os do MdS e tendem a convergir ao longo do tempo, implicando não só na veracidade dos números apresentados, bem como na ausência de subnotificação significativa nos números do MdS.

8.2) O Portal da Transparência

Apresentado como uma iniciativa dos cartórios de Registro Civil do Brasil, esse site compila os dados de registro de óbitos no país todo, com especial atenção aos dados do COVID-19. Os dados são atualizados cerca de três vezes ao dia. Apresenta atraso em relação ao valor real, por necessidade óbvia das famílias registrarem as mortes em datas posteriores aos seus acontecimentos. Seus valores para cidades como o Rio de Janeiro variam de forma acentuada somente nos dias mais recentes (até 4 dias em média), porém considerando o Brasil como um todo e suas regiões mais distantes, esta variação ocorre por mais tempo. Desta forma existem duas variações a serem observadas neste banco de dados, sendo a primeira o quanto tempo um dado dia no calendário leva para atingir um dado valor que será considerado como tendo “convergido”, ou seja: a partir de uma dada data ele só irá mudar residualmente, não afetando qualquer estudo feito, o mesmo pode ser observado na base de dados internacional aqui referenciada [7]. A segunda variação decorre do fato de que o calendário segue normalmente e os dias avançam, com novos óbitos sendo lançados e os dias mais recentes sempre estão, por definição, subnotificados. Este trabalho avaliou quanto tempo é necessário para que um dado valor atinja a estabilidade, que foi definida como sendo quando parou de variar mais do que 1% no prazo de uma semana. O Apêndice A apresenta um estudo comparativo entre os dois portais.

8.2.1) Tempo Até Convergência de Valor – Cidade do Rio de Janeiro

Para se avaliar o tempo que uma certa curva de óbitos leva para atingir a estabilidade, sendo estabilidade aqui definida como sendo aquele patamar que varia menos de 10%, monitoraram-se os valores da cidade do Rio de Janeiro. Avaliando-se o gráfico da Figura 14, a primeira observação é que somente os últimos 4 dias

variam mais de 10% ao longo de uma semana. Ou seja: após 2 semanas não ocorre mais modificação sensível nos valores apresentados, à exceção dos últimos 4 dias. Do primeiro dia (azul, 23/jun) para o segundo dia observado (29/jun), se passam 6 dias e embora na média os valores sejam menores que 10% (6,0 calculado), existem casos isolados que variaram mais do que isto. Porém, entre o dia 29/06 e o dia 05/07 se passam de novo 6 dias e agora a média da variação é de 0,3%, se excluídos os últimos 4 dias. Desta forma, pode-se afirmar que para a cidade do Rio de Janeiro, atualmente, basta se eliminar os últimos 4 dias para se fazer uma projeção acurada da evolução do ciclo epidêmico na cidade. Da mesma forma pode-se estudar e avaliar todo e qualquer caso coberto pelo site Portal da Transparência.

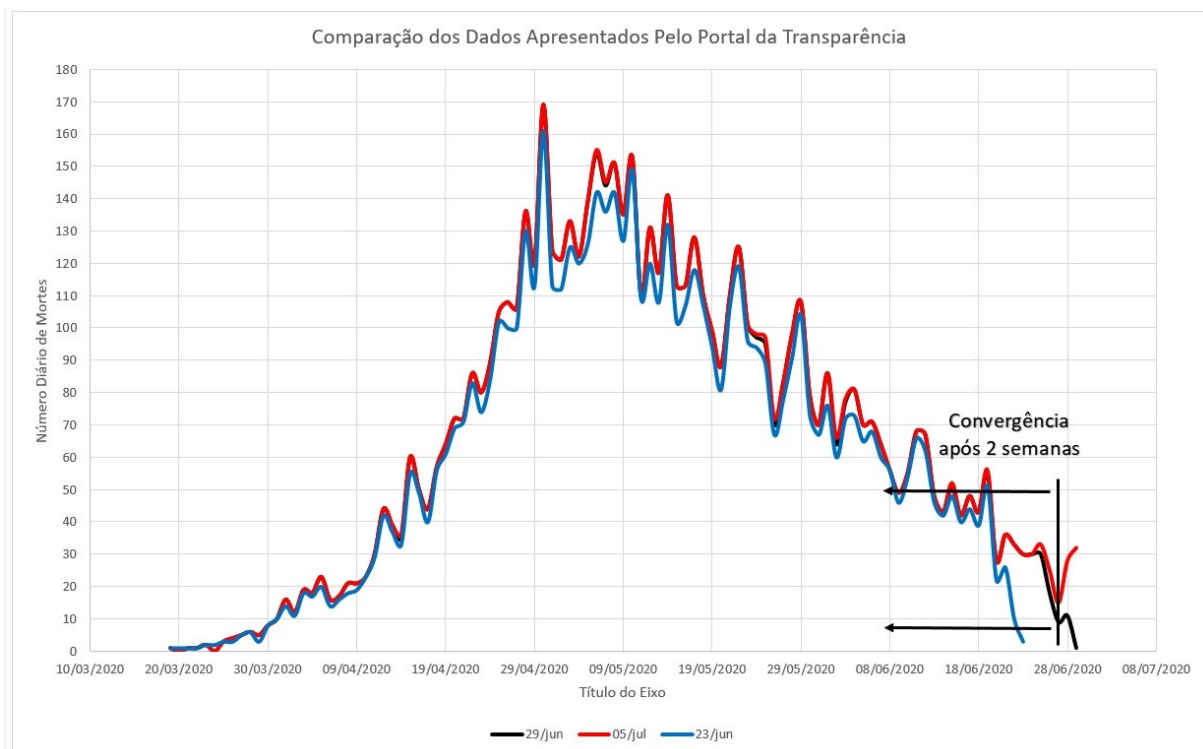


Figura 14 – Dados das mortes diárias na cidade do Rio de Janeiro em 3 datas diferentes. Não se nota mais diferença entre a 2ª e a 3ª semana observadas, a menos dos 4 últimos dias.

8.2.1) Tempo Até Convergência de Valor – Brasil

Este trabalho considerou que variações no número de mortos em um dado dia menor do que 5% são desprezíveis e que até 10% podem ser aceitos sem eliminação. Isto implica que os dados a serem eliminados serão os dos últimos 10 dias. Isto pode ser verificado na Figura 15. Comparando os dados de 05 de julho (linha preta) com 29 de junho (linha vermelha) nota-se que antes disto, do 17º dia para trás, os valores se situam abaixo de 3% e em média 1%, todos sem efeito prático em qualquer projeção que se deseje fazer (Zona A). Mesmo quando se observam os valores do dia 24 de junho (curva azul), embora a variação ocorra acima de 10% em um período acima de 10 dias (Zona B), o formato e as proporções da curva são idênticas às duas curvas em datas posteriores, isso indica que a

distribuições de notificações posteriores não altera o formato das curvas e que se elas forem submetidas aos métodos numéricos anteriormente citados, elas de fato poderão servir a predições com acurácia aceitável.

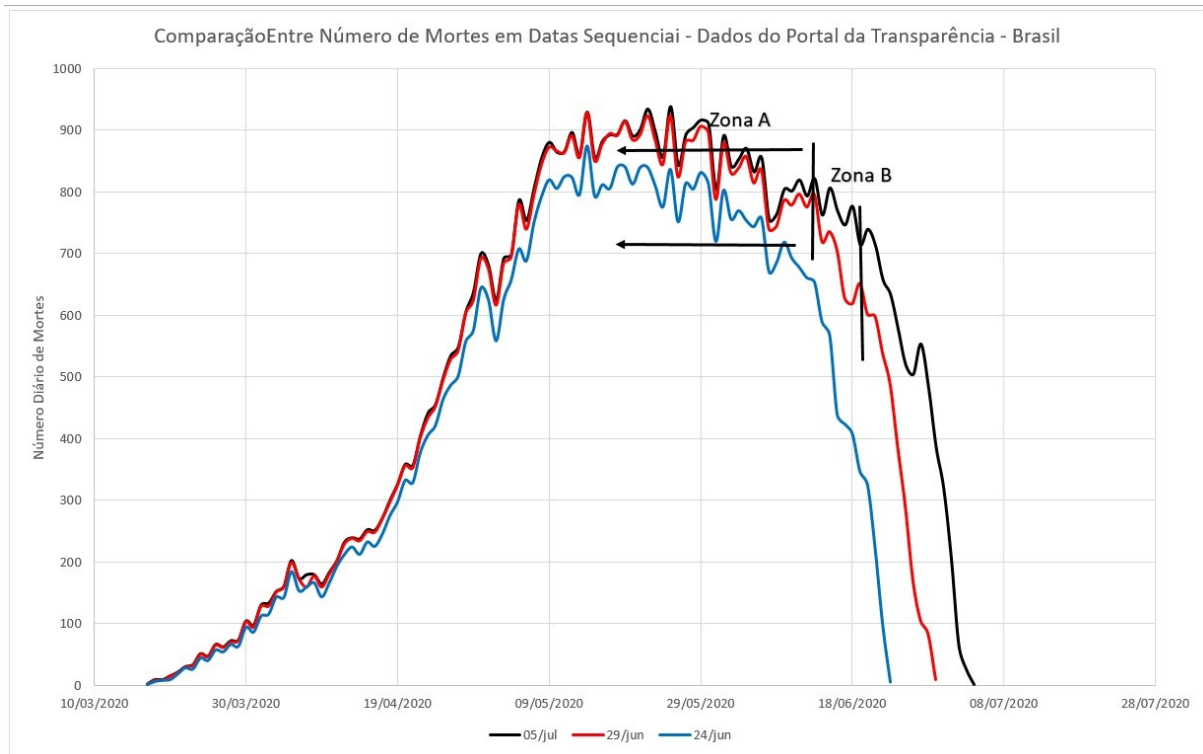


Figura 15 - Dados das mortes diárias Brasil em 3 datas diferentes. Não se nota mais diferença entre a 2ª e a 3ª semana observadas, a menos dos 10 últimos dias.

8.3) Total de Mortes Computadas

8.3.1) Brasil

O atraso inevitável entre a ocorrência do óbito físico e o registro no cartório, produz um efeito na contagem do total de mortes no Brasil e em cada localidade. Ao final da epidemia os dois valores (MdS e Portal da Transparência irão convergir, mas durante o surto, haverá sempre um atraso entre os dois). O exemplo que se segue foi realizado na tarde do dia 05 de julho de 2020. O total de mortes indicado pelo site Portal da Transparência era de 58155, enquanto o site do MdS trazia o número 64265 mortes. Uma diferença de 6110 óbitos (9,5% do total) que ocorreram, mas ainda não foram registrados.

8.3.2) Cidade do Rio de Janeiro

No caso da Cidade do Rio de Janeiro, para dados colhidos na manhã de 6 de julho de 2020, nos dois sites citados, revela que o Ministério da Saúde aponta para um total de 6898 mortes na cidade até então e o Portal da Transparência (PdT) indica 6855, uma diferença de 43 óbitos, ou seja cerca de 0,6%, indicando uma excelente coerência entre os dois valores totais, e a distribuição, após a aplicação do MMMI, indica a existência da convertibilidade dos valores do MdS para os do PdT, embora não tão completa como no caso da Suécia, anteriormente descrito.

9) PREVISÕES DE CICLOS NO BRASIL

A partir de todo o ferramental numérico e dos bancos de dados apresentados, bem como das limitações e possibilidades no seu uso, os autores passam agora a fazer projeções para o comportamento temporal dos ciclos epidêmicos e dos números de mortes associados a eles. Diversas previsões aqui apresentadas já foram divulgadas em redes sociais em datas diversas na forma de postagem [20] ou das chamadas “lives” [21], estando disponíveis nas plataformas para consultas e confirmações.

As previsões são feitas seguindo um roteiro bastante simples:

- a) Os dados de mortes diárias de uma dada localidade são anotados em uma planilha e a MMMI calculada.
- b) Normalizam-se os valores, como já explicado, aplica-se o MMMI, porque embora a base de cartórios registre as mortes nas datas corretas, para que as duas séries de dados sofram as mesmas transformações numéricas.
- c) Gera-se uma curva contínua em um gráfico tendo como eixo x o número de dias corridos do ciclo epidêmico e o eixo y, o intervalo adimensional de 0 até 1, geralmente (podendo alguns pontos, os falsos picos, passar disto).
- d) Busca-se dentre os países ou localidades, aqueles que já tenham terminado seu ciclo epidêmico crítico (que foi chamado neste trabalho de CMLE) e que seja visualmente semelhante à curva obtida em (c), embora obviamente em escala diferente.
- e) Aplica-se o MMMI
- f) Faz-se a normalização dos dados
- g) Repete-se o item (c).
- h) Como o país ou localidade de referência já terminou seu ciclo, estará posicionado anteriormente no gráfico, em relação ao local que se deseja estimar a provável data de fim de ciclo crítico. Deve-se então numericamente fazer a superposição do pico do caso que se deseja estudar, por exemplo, com o de referência.
- i) Feita a superposição, movendo sempre o caso referência, pode-se fazer uma extrapolação usando o caso referência como guia para o valor que se deseja determinar. Como a escala do caso estudado não foi alterada, basta consultar que dia seria este no futuro para se saber a data provável.
- j) Caso não exista um caso semelhante, pode-se eliminar os últimos dias, como acima discutido e fazer uma extrapolação diretamente dos valores obtidos no banco de dados do PdT.

9.1) Previsão de Ciclo da Cidade do Rio de Janeiro

A previsão do ciclo da cidade do Rio de Janeiro será feita pelos dois métodos aqui propostos, 1) da superposição de um ciclo já terminado, no caso a Itália, com o da cidade e 2) a extrapolação a partir da série de dados, com a eliminação dos últimos valores.

9.1.1) Cidade do Rio de Janeiro e Itália

Tendo sido a Itália o caso escolhido para servir de modelo de previsão, os valores previstos foram sendo atualizados a partir dos valores diários obtidos no site PdT, ou seja: valores aqui apresentados foram coletados nas datas citadas. Embora o ciclo de subida da Itália não tenha sido semelhante ao da cidade do Rio, se considerou que o de descida, nos instantes iniciais era bom o suficiente e este modelo foi adotado, o que se revelou acertado. Foi considerado no item 4 que o ciclo crítico italiano se encerrou quando o nível de mortes diários decresceu até atingir 10,6% do pico já submetido ao MMMI, a partir daí se iniciou o ciclo residual. Para a cidade do Rio de Janeiro, após a realização do MMMI, o valor de pico passa a ser 143, portanto 10,6% disto seriam 15 mortes. Então quando a cidade do Rio de Janeiro atingir 15 mortes diárias, aproximadamente, o ciclo crítico pode ser considerado encerrado. Nota-se que no dia previsto, 05 de julho, as curvas se encontraram.

Na primeira avaliação, feita no dia 01 de junho de 2020 (curva azul escuro), nota-se a concordância entre o gráfico de proporções da Itália (curva preta) e da cidade do Rio coincide até 21 de maio, depois disto existe uma grande divergência, mas é possível notar que: a) o ciclo ainda está longe do fim e b) o ciclo proposto descreve bem o episódio crítico na desaceleração, por cerca de 1/3 do período. Se o avaliador aceita este período como sendo o suficiente para caracterizar semelhança válida, basta consultar o valor final do episódio italiano (10,6%) e estimar a data onde o ciclo carioca irá atingir tal valor do seu pico, como acima descrito.

Em 10 de junho (curva roxa) bem como em todas datas usadas neste gráfico para a cidade do Rio de Janeiro, foram marcados os últimos 4 dias, que estão contidos dentro das elipses. A observação dos dados destas 4 datas finais de cada série garante que eles irão crescer e, portanto, se tornar mais similar ainda ao modelo italiano, como fica evidente ao se analisar as datas posteriores, mas no momento da coleta ainda estavam no seu mínimo. A concordância desta curva vai até 02 de junho depois diverge, mas mais da metade do ciclo de descida concorda com o italiano, de forma que a previsão se torna bastante assertiva.

Para os dados coletados no dia 17 de junho (curva verde), a concordância da curva vai até 07 de junho depois diverge suavemente até 10 de junho, mas a concordância com o italiano é bastante óbvia e a data prevista por este ciclo, 05 de julho, fica cada vez mais confirmada.

Para 23 de junho (curva amarela), os dados coincidem até 16 de junho e a curva toda concorda em mais de 4/5 de sua extensão, confirmando a estimativa inicial de 05 de julho.

Em 29/06 (curva azul), os dados se entrelaçam com o ciclo italiano até 19 de junho, sempre confirmando as estimativas iniciais.

Em 05/07 (curva vermelha) existe um forte entrelaçamento entre os dados, confirmando o final do ciclo por esta estimativa. Resta aguardar a evolução da doença na cidade e comparar a estimativa com os valores reais.

A parte crítica do ciclo de queda da cidade do Rio de Janeiro ficou então estimado como tendo sido iniciado em 04 de maio de 2020 e chegado ao fim em 05 de julho, onde teria atingido o valor onde o ciclo residual se inicia. Desta forma o ciclo de queda durou 62 dias. Juntando-se aos 42 de subida, são 104 dias no total, ou seja, um ciclo de quase 15 semanas.

junho (35 dias antes do fim), embora isto contivesse um grau de incerteza maior, já que somente 29% da curva descida concordava. Porém, a partir de 10 de junho (25 dias de antecedência), com 47% de concordância entre as curvas, já se podia estimar a data. A partir de 17 de junho, existe uma concordância bastante acentuada entre as curvas, acima de 70% e a previsão se torna mais segura ainda. Caso se descartem os últimos 4 dias das séries, como já discutido, pode-se notar que a série de 05 de julho se torna idêntica à da Itália a partir de 30 de junho.

9.1.2) Extrapolação de Dados

Para servir de exemplo, são selecionados 3 datas e dados que foram coletados naquelas datas no site PdT. As datas de coleta dos dados foram 23/06, 29/06 e 05/07/2020. Após isto, é feita uma seleção do pico como sendo aquele último ponto em que a série subiu e a partir daí passou a decrescer, mesmo que de forma irregular. Seleciona-se os pontos (incluindo este pico) excluindo os 4 últimos, pelas razões já discutidas, submete-os ao MMMI e aplica-se o processo de regressão linear, onde uma reta é forçada a passar por estes pontos. A equação desta reta é então igualada ao valor nulo e então se acha o número de dias do ciclo da epidemia, sendo o Dia Zero 19 de março de 2020, data da primeira morte ocorrida, que corresponde a uma data na planilha. Os valores obtidos são listados na Tabela 5, ao final deste item. A Figura 19 apresenta o exemplo do caso para os dados coletados no dia 05 de julho da 2020. Todos os outros casos aqui apresentados foram calculados da mesma forma, como já descrito. Nota-se pelos dados da Tabela 5 que o tempo (em dias) entre o dia em que os dados foram observados e a data prevista, se reduz conforme se aproxima o final do ciclo. Este é atingido quando a diferença entre o dia previsto e o dia dos dados empregados é zero.

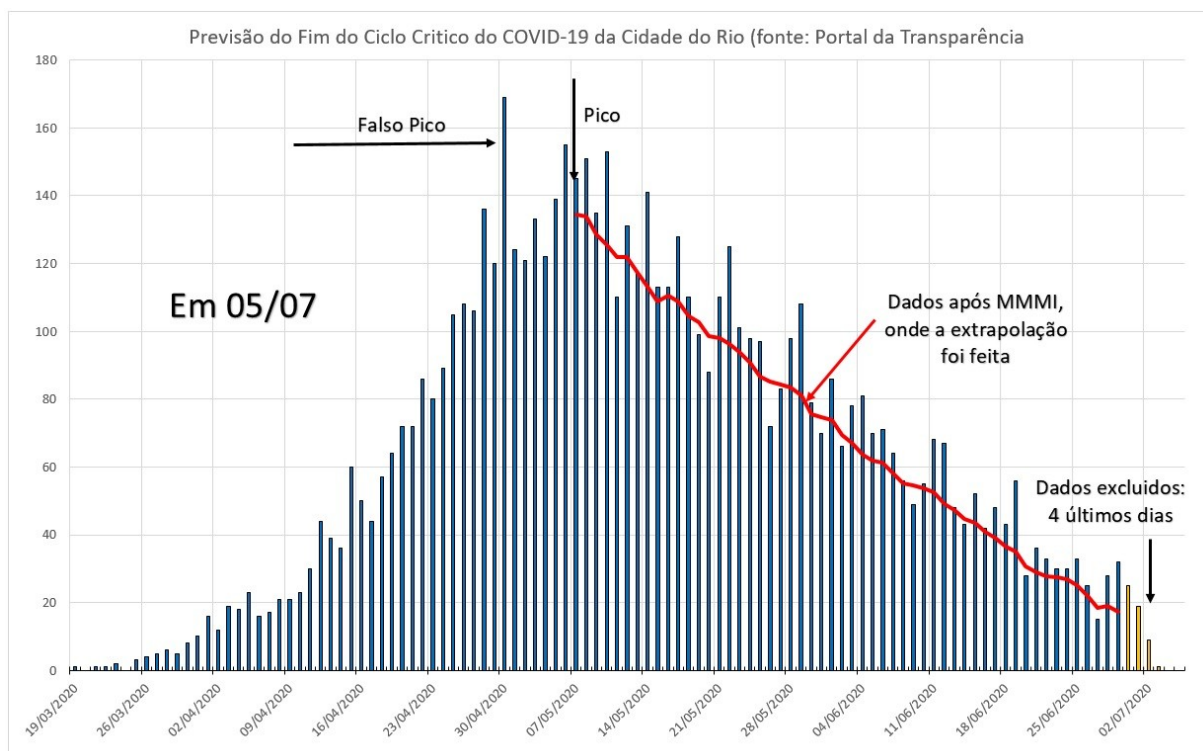


Figura 17 – Extrapolação para estimativa da data do final do ciclo crítico da doença, com os dados coletados no dia 05 de julho de 2020. A data prevista é 07 de julho.

Tabela 5 – Dados das extrapolações dos pontos de dados do método proposto.

Data	Equação	R ²	Dias de Epidemia	Data Prevista	Dias até final do Ciclo
23/06	$y = -2,2263x + 289,68$	0,9939	104	01/07/2020	8
29/06	$y = -2,3338x + 304,76$	0,9961	105	02/07/2020	3
05/07	$y = -2,2086x + 237,70$	0,9951	107	05/07/2020	1
07/07	$y = -2,1864x + 236,24$	0,9933	108	06/07/2020	0

9.1.3) Avaliação da Duração do Ciclo do Rio de Janeiro Por Comparação

No item 5 deste trabalho são compilados diversos casos cujos ciclos críticos podem, neste momento, serem considerados como encerrados. Serão usados para uma avaliação bastante simples, considerando que existe uma proporção entre o tempo decorrido no ciclo de subida, por assim dizer, da doença e o de queda dos óbitos. A Tabela 6 apresenta estas estimativas baseada nos casos apresentados. A Figura 20 apresenta o ciclo do Rio de Janeiro, que apresentou o dobro do tempo de subida que a média mundial (21 contra 42 dias). Por outro lado, a descida foi relativamente rápida, em relação ao resto do mundo, embora isto se deva à longa subida. O número de mortes reais (6918) registrada em cartório no dia 07/07/2020 se encontra dentro dos limites da Tabela 3, embora bem próximo do limite inferior. Já para o tempo previsto, seguindo as proporções listadas na Tabela 2, os valores são absurdos, uma vez que o ciclo do Rio de Janeiro, para todos os efeitos se encontra no final.

Tabela 6 – Estimativas de duração de ciclo para cidade do Rio de Janeiro, baseadas nos valores listados na Tabela 2.

Local de Referência	Fator de Multiplicação	Dias até o Pico no RJ	Dias até o Fim	Data Prevista
Espanha	2,1	42	88	02/08/2020
Alemanha	3,2		134	17/09/2020

Tabela 7 – Estimativas de número de mortes para cidade do Rio de Janeiro, baseadas nos valores listados na Tabela 3. * Não há referência a esta tabela !

Local de Referência	Fator de Multiplicação	Mortes até o Pico no RJ	Mortes até o Fim	Total de Mortes Previstas
Espanha	1,6	2588	4141	6729
Suécia	3,3		8540	11128

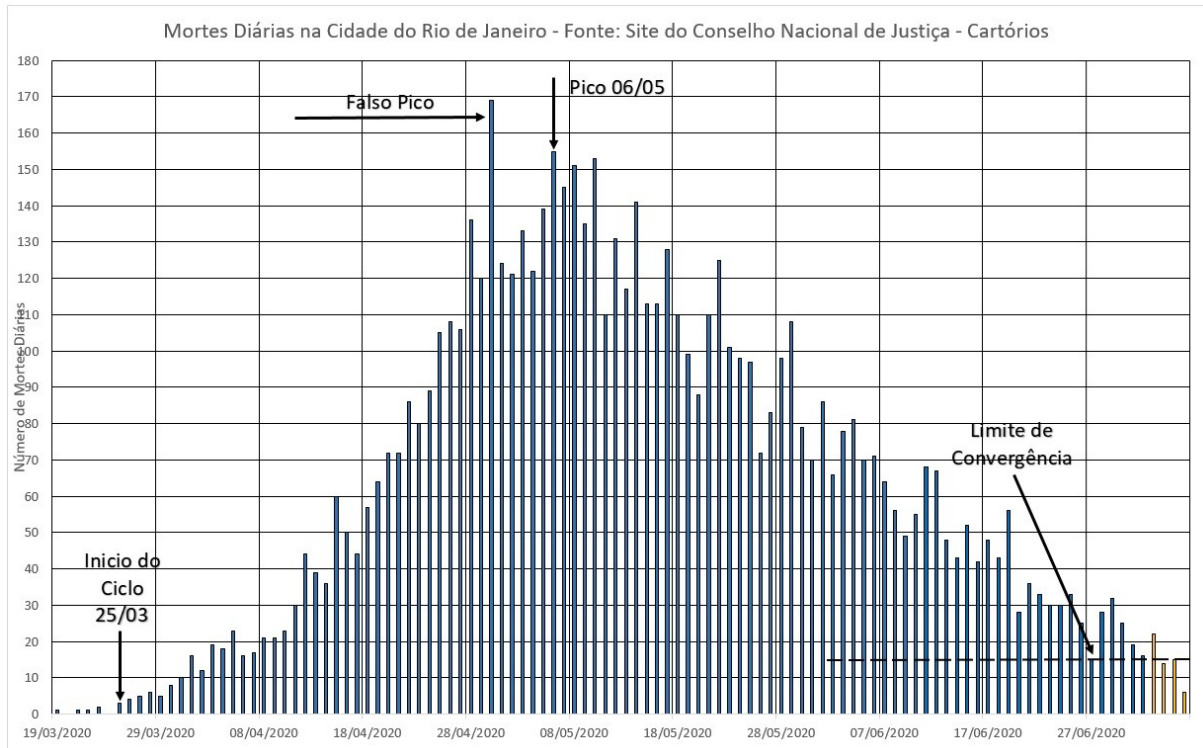


Figura 18 – O ciclo do Rio de Janeiro, já praticamente finalizado na data de preparação deste trabalho, apresentou praticamente simetria entre o período de subida e de descida.

10) CONCLUSÕES

São listadas a seguir as conclusões imediatas que foram obtidas das análises aqui demonstradas:

- Dado o grande número de subnotificações, típicas dos episódios epidêmicos, os dados considerados mais confiáveis foram os relativos ao número diário de mortes, que foram então empregados neste estudo.
- A natureza dos dados hoje disponíveis para estudos requer tratamento numérico preliminar, uma vez que grande parte deles apresentam o número de mortes diárias ocorridas nas datas em que foram registradas e não nas que efetivamente ocorreram.

c) Existe um ciclo observável em diversos países pequenos e médios, cidades e estados, aparentemente independente das medidas tomadas pelas autoridades locais. Sua aparência é triangular, com o tempo de descida que varia de 2,1 a 3,3 vezes mais longo que o de subida. E o número de mortes no período de desaceleração da doença é 1,6 a 3,3 vezes maior que no período de aceleração.

d) Existem, porém, dentre este tipo triangular, diversos ciclos que são semelhantes, uma vez que seus valores sejam tornados independentes da escala e das datas de ocorrência. A normalização permite que se use um ciclo já terminado para se estimar o comportamento de um ciclo ainda em evolução.

e) O Brasil e os outros países com territórios e populações vastas não devem ser tratados como um caso só, mas sim estudados regionalmente, de forma a que se possa entender a evolução dos ciclos da doença de forma clara.

f) O método que emprega a similaridade de ciclos conseguiu estimar o fim do ciclo com antecedência de até 34 dias o final do ciclo, porém requer que exista um ciclo semelhante.

g) O método de extrapolação de dados conseguiu estimar o fim do ciclo com 13 dias de antecedência e não depende de quaisquer outros dados além dos oferecidos pelo Portal da Transparência.

h) O método da comparação não foi um bom estimador de ciclo para a cidade do Rio de Janeiro.

i) O ciclo apresentado pela China não se reproduziu em nenhum dos outros países aqui estudados.

Os autores esperam que esse breve estudo contribua para o suporte ao processo decisório na atual e nas próximas pandemias que possam surgir, trazendo racionalidade ao mesmo.

Nota: Embora os dados aqui mostrados já tenham sido testados contra alguns ciclos já se encerrando, como é o caso do Rio de Janeiro, os autores o fizeram durante o ciclo, durante “lives” e postagens em redes sociais, que podem ser consultadas, obtendo sucesso em suas previsões. A modelagem aqui apresentada visa trazer mais racionalidade à discussão em torno da Pandemia COVID-19, de maneira a colaborar para o seu enfrentamento da mesma, bem como de possíveis futuras pandemias, evitando que o pânico se instale e medidas exageradas ou mesmo completamente desnecessárias sejam tomadas.

Agradecimentos: Os autores agradecem ao Dr. Renato Nunes Teixeira, do INMETRO, pela revisão deste trabalho.

Referências Bibliográficas

[1] – Philip, Hunter, “Inevitable or Avoidable? Despite the Lessons of History, the World is Not Yet Ready To Face the Next Great Plague”, EMBO Rep. 2007 Jun; 8(6): 531–534, doi: [10.1038/sj.embor.7400987](https://doi.org/10.1038/sj.embor.7400987)

- [2] – <https://www.imperial.ac.uk/news/196496/coronavirus-pandemic-could-have-caused-40/>
- [3] – Best, Ryan and Boyce, Jay, “Where the Latest COVID-19 Models Think We're Headed — And Why They Disagree”, <https://projects.fivethirtyeight.com/covid-forecasts/> acessado em 09/07/2020.
- [4] – Ritchie, H., Ortiz-Ospina, E. et al., “Coronavirus (COVID-19) Testing”, <https://ourworldindata.org/coronavirus-testing> atualizado diariamente, acessado em 09/07/2020.
- [5] – Introduction to Statistical Quality Control, 5th ed., D.C. Montgomery, John Wiley & Sons, 2005.
- [6] – <https://www.worldometers.info/coronavirus/country/sweden/> acessado em 24/06/2020.
- [7] – <https://www.worldometers.info/coronavirus/country/sweden/> acessado em 04/07/2020.
- [8] – <https://www.worldometers.info/coronavirus/country/china/> acessado em 09/07/2020.
- [9] – Lee, T-W, Park, J.E., Jung, D., “Gaussian Statistics and Data-Assimilated Model of Mortality due to COVID-19: China, USA, Italy, Spain, UK, Iran, and the World Total”, DOI: 10.1101/2020.04.06.20055640 acessado em 09/07/2020
- [10] – <https://www.worldometers.info/coronavirus/country/spain/> acessado em 09/07/2020.
- [11] – <https://www.worldometers.info/coronavirus/country/netherlands/> acessado em 09/07/2020.
- [12] – <https://www.worldometers.info/coronavirus/country/italy/> acessado em 09/07/2020.
- [13] – <https://www.worldometers.info/coronavirus/usa/new-york/> acessado em 09/07/2020.
- [14] – <https://www.worldometers.info/coronavirus/country/germany/> acessado em 09/07/2020.
- [15] – Meunier, Thomas, A.J., “Full Lockdown Policies in Western Europe Countries Have no Evident Impact on the COVID-19 Epidemic”. doi: <https://doi.org/10.1101/2020.04.24.20078717>
- [16] – <https://www.dailymail.co.uk/news/article-8347635/Lockdowns-failed-alter-course-pandemic-JP-Morgan-study-claims.html> acessado em 09/07/2020.
- [17] – Nate Silver: <https://projects.fivethirtyeight.com/2020-nba-player-projections/> acessado em 09/07/2020.
- [18] – Site do Ministério da Saúde - <https://covid.saude.gov.br/>

[19] – Portal da Transparência - <https://transparencia.registrocivil.org.br/especial-covid>

[20] – Postagens Feitas na Rede Social Facebook nos meses de abril a julho de 2020, no nome do 1º. Autor.

[21] – Participação em “lives” postadas na página do facebook de RobertoMottaOficial.

APÊNDICE A

Comparação Entre Dados do Portal de Transparência e do MdS

Os dois bancos de dados oferecem vantagens e desvantagens, porém ambos são confiáveis e ao longo do tempo irão convergir para valores semelhantes. Para comprovar a veracidade desta afirmação, foi feito o seguinte procedimento:

- a) Os dados diários sobre as mortes, divulgados pelo MdS foram convertidos em novos valores pelo Método das Médias Móveis Inicial (MMMI), ou seja, a média dos valores de 6 dias adiante, mais a do próprio dia, assume o primeiro valor na planilha e assim sucessivamente.
- b) Foi então traçada uma curva ligando todos esses pontos (curva preta).
- c) Os dados diários sobre as mortes, coletados no site Portal da Transparência foram listados e uma curva ligando todos estes pontos foi traçada (curva vermelha).
- d) As duas curvas foram sobrepostas, para efeito de comparação, uma vez que originalmente elas ocorrem em datas diferentes (conforme esperado).
- e) Nota-se claramente que existe um período de concordância entre as duas curvas, que é o início da epidemia, seja no Brasil ou na localidade de interesse. Ao se deslocar para à direita a diferença aumenta, compatível com o fato de que os registros recentes ainda não terem sido feitos nos cartórios, mas já terem ocorrido nas secretarias de saúde de municípios e estados, então essas mortes “ocorreram” no MdS mas não nos cartórios.

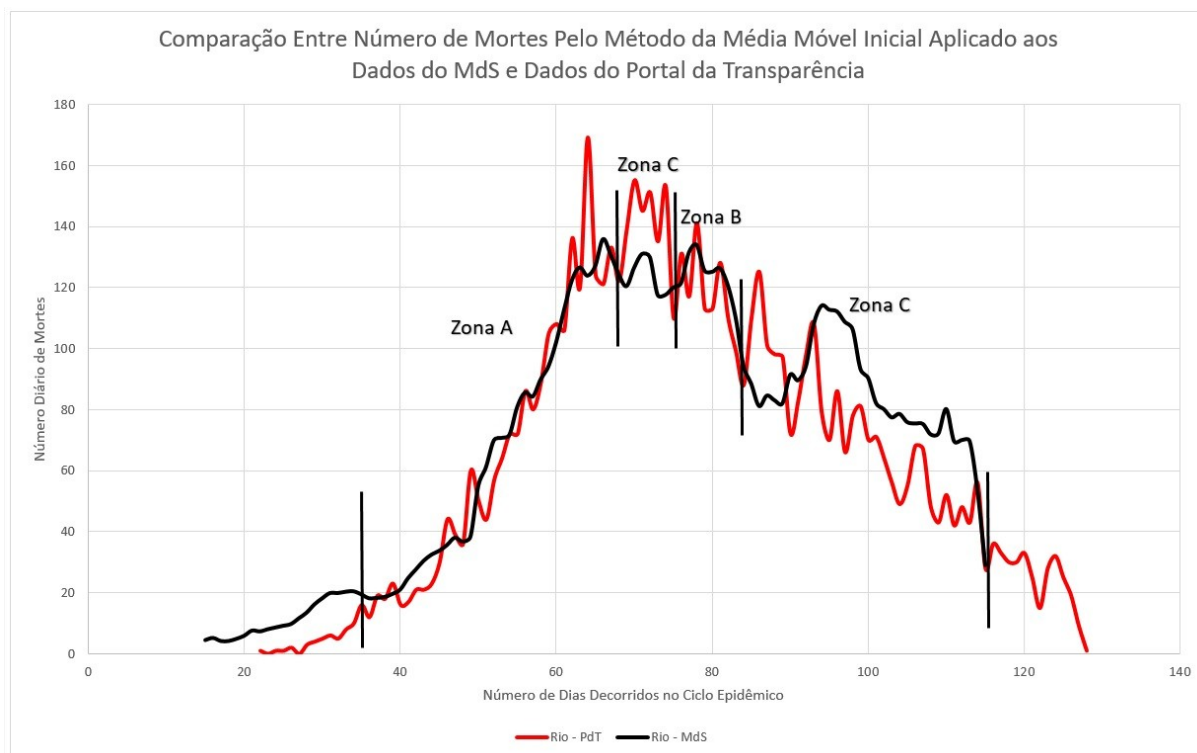


Figura A1 - Comparação entre os dados do MdS (preto) e cartórios (vermelho).

Na Figura A1 pode-se observar a Zona A, onde os valores das duas fontes diferem na média 7%, embora possam apresentar algum valor fora desta faixa, são diferenças desprezíveis. Na Zona B, este valor cresce para cerca 9%. Na Zona C o valor sobe para em torno de 15% que embora não seja desprezível, ainda é tolerável em qualquer modelagem numérica sobre eventos humanos. Não existe equivalência entre os últimos dias, o que é esperado, uma vez que as duas bases reportam valores passados, porém com atrasos diferentes. Para se realizar qualquer extrapolação usando os valores do MdS, deve se tomar o cuidado de evitar o uso de dados desta faixa final de não concordância de tempo.

A.1) O Caso do Brasil

Na Figura A2 pode-se observar a Zona A, onde os valores das duas fontes diferem na média 1%, embora possam apresentar algum valor fora da faixa unitária, são diferenças desprezíveis. Na Zona B, este valor cresce para cerca 10%. Na Zona C o valor sobe para em torno de 15% que embora não seja desprezível ainda é tolerável em qualquer modelagem numérica sobre eventos humanos. Na Zona D, a diferença atinge cerca de 20% a partir daí, que equivalente aos últimos 10 dias, para o Brasil, existe de fato uma grande diferença, então para se realizar qualquer extrapolação, deve se tomar o cuidado de evitar o uso destes valores, caso se deseje empregar a base de dados dos cartórios. Mas este fenômeno só é observável em um caso tão vasto quanto o do Brasil todo. Ver exemplo da cidade Rio de Janeiro.

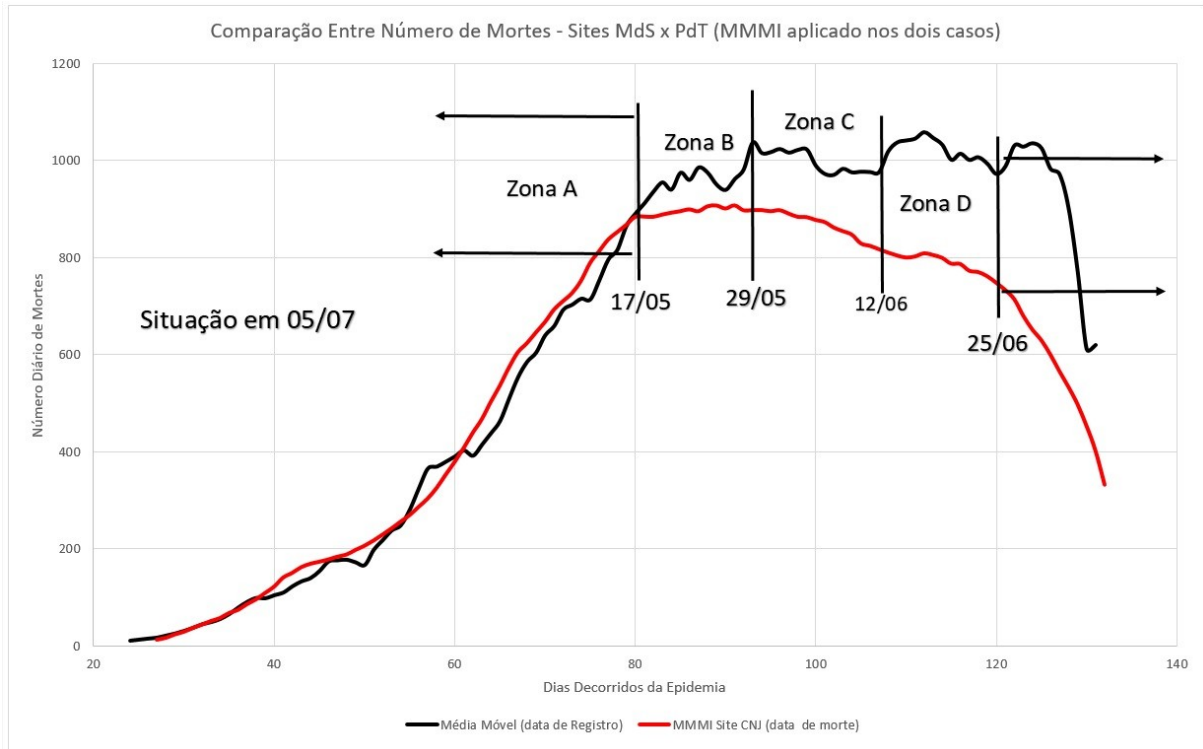


Figura A2 – Comparação entre os dados do MdS (preto) e cartórios (vermelho).