

HOMOGENEIZAÇÃO DE DADOS PLUVIOMÉTRICOS DIÁRIOS: UMA CONTRIBUIÇÃO METODOLÓGICA

DAILY PLUVIOMETRIC DATA HOMOGENIZATION: A METHODOLOGICAL CONTRIBUTION

Antonio Carlos Oscar Júnior

¹ Universidade do Estado do Rio de Janeiro (UERJ), Rio de Janeiro, RJ, Brasil

Correspondência para: Antonio Carlos Oscar Júnior (antonio.junior@uerj.br)

doi: 10.12957/geouerj.2019.40957

Recebido em: 01 set. 2018 | Aceito em: 09 out. 2018



RESUMO

Diversas são as interferências não climáticas que podem produzir ruídos em séries de dados climáticos, sobretudo na unidade diária, cuja relevância para a Climatologia é notória. Em face disto, o presente artigo objetivou avaliar a eficiência do RHtest, técnica de homogeneização de dados pluviométricos diários, para o contexto tropical brasileiro. Em âmbito internacional é cada vez maior a exigência pelo controle da qualidade dos dados, entretanto, diante da possibilidade de alteração do dado observado, muitas dúvidas surgem em relação à homogeneização; por isto, se propôs a homogeneização da série de dados (1967-2015) de três postos pluviométricos no estado do Rio de Janeiro em diferentes contextos topográficos. A partir da comparação entre a série de dados brutos (observados) e homogeneizados, utilizando a correlação de Spearman, a distribuição de frequência em classes de pluviosidade e os índices de extremos climáticos (RClimDex), constatou-se o potencial de aplicação da técnica para o Rio de Janeiro. Os resultados mostraram que há uma expressiva correlação ordinal entre os dados brutos e homogeneizados (superior a 99%) e com significância de 95%. Além disso, a técnica, na medida que retirou ruídos da série, favoreceu que outros índices climáticos apresentassem tendência com significância estatística. Já a distribuição de frequência demonstrou uma possível interferência das falhas do banco de dado e do relevo nos resultados, com destaque ao posto a sotavento (Petrópolis), com 4,7% de falha no banco de dados e onde a concentração de chuva reflete em variações do limiar; entretanto, as diferenças observadas entre os dados observados e homogeneizados, em termos de extremos, não foram superiores a 21 mm para o referido posto.

Palavras-chave: RHtest; estatística; controle de qualidade; tendência.

ABSTRACT

Several non-climatic interferences can produce noise in a series of climatic data, especially in the daily unit, with relevance to Climatology. In face of this, the present paper aimed to evaluate the efficiency of the RHtest, technique of homogenization of daily rainfall data, for the Brazilian tropical context. At the international level, the demand for data quality control is increasing, however, given the possibility of altering the observed datas, many doubts arise regarding homogenization; for this reason, the data series (1967-2015) of three rain gauges in the state of Rio de Janeiro, in different topographic contexts, were submitted to the test. From the comparison of the observed and homogenized data series, using Spearman correlation, frequency distribution in rainfall classes and climatic extremes indexes (RClimDex), the potential application of the technique to the Rio de Janeiro was confirmed. The results showed that there is an expressive ordinal correlation between the observed and homogenized data (greater than 99%) and with a 95% significance. In addition, the technique, removing the noises from the series, favored that other climatic indexes presented a tendency with statistical significance. On the other hand, the frequency distribution showed a possible interference of the relief in the results in relation to the Petrópolis rain gauge (leeward), in which the concentration of rain reflects in variations of the threshold; however, the observed differences between the observed and homogenized datas, in terms of extremes, were not greater than 21 mm to Petrópolis rain gauge.

Keywords: RHtest; statistic; quality control; trend

INTRODUÇÃO

Da concepção de clima como uma sucessão habitual de tipos de tempo (SORRE, 1951; PÉDÉLABORDE, 1970) até a proposição metodológica da análise rítmica (MONTEIRO, 1971), constrói-se uma agenda de pesquisa para a climatologia brasileira voltando para o tempo atmosférico, eregindo o diário como unidade fundamental de análise. Da mesma forma, Martín-Vide (2004), para o caso da pluviosidade, discorre sobre a importância de sua análise em escala diária, haja vista seu desdobramento para vários campos da ciência ambiental e da sociedade. Entretanto, o autor conclui sobre a pouca atenção científica para esta unidade de análise.

Entender melhor as conexões entre clima e tempo é cada vez mais urgente já que muitos dos processos de tomada de decisão, bem como os impactos da variabilidade climática estão conectados com o tempo atmosférico, especialmente os eventos extremos de precipitação. Deste modo, é necessário maior atenção aos critérios de detalhamento da pluviosidade em detrimento daqueles baseados em totais anuais ou mensais, ainda mais em um momento em que cabe à climatologia fornecer melhor entendimento para aqueles fenômenos que fogem da habitualidade e à Geografia, oferecer subsídios adequados para o planejamento e organização do território de modo a diminuir os impactos negativos destes eventos sobre as organizações humanas.

Contudo, como relata diversos autores, entre eles Baptista (2009), muitas estações meteorológicas encontram-se em situação inadequada para coleta de dados, refletindo realidades microclimáticas muito singulares que inviabilizam estudos sobre a variabilidade e mudanças climáticas. Além disso, acrescentam-se os erros de coleta e registro, as constantes mudanças de localização, as mudanças de instrumentos e cálculos utilizados, a falta de manutenção ou calibração de aparelhos.

Todas as causas listadas anteriormente têm um impacto sobre a série de dados, principalmente em escala diária, podendo mascarar a variabilidade da série histórica, o que faz com que os resultados fiquem enviesados. Dessa forma, antes de qualquer análise é necessário identificar e remover tais ruídos, ou seja, homogeneizar a série.

Dada a possibilidade de tais interferências, o procedimento de homogeneização de dados é obrigatório para estudos de variabilidade climática (PETERSON et al., 1998; AGUILAR, et al., 2003; AZORIN-MOLINA, 2016), permitindo a eliminação daqueles distúrbios contidos na série em virtude de mudanças nas condições de observação. Os testes de homogeneidade utilizados em séries temporais podem ser divididos em dois grupos: testes absolutos e testes relativos. Os testes absolutos utilizam apenas a própria série, sendo eles: teste de Kruskal-Wallis, teste de Friedman, teste de Buishand, teste de Pettitt, teste de razão de Von Neumann, teste ADF (Augmented Dickey-Fuller), para citar alguns exemplos. Já os testes relativos utilizam também as informações de áreas vizinhas (áreas de referência), sendo alguns deles: técnica Bayesiana de análise de pontos de mudança, teste de Chow, teste SNH (Standard Normal Homogeneity), teste de Padrão Rítmico e RHTest (Rainfall Homogeneity Test).

Entretanto, dando atenção à escala diária que Monteiro (1971) e Martín-Vide (2004), entre outros, elegem como de fundamental importância para compreensão de processos, os métodos para homogeneização de dados diários são escassos. Kemp et al. (1983) explicam o motivo em função da maior variabilidade espacial e temporal dos valores diários. Os métodos operam melhor quando estão sendo avaliadas as médias mensais e são mais propensos a falhar quando aplicadas em escala diária, por isso, a homogeneização de dados mensais e anuais são mais frequentes.

Porém, Wang et al. (2008) desenvolveram uma técnica de homogeneização de dados mensais e anuais para aplicação em dados diários de pluviosidade, conhecido como RHtest. Os referidos autores elaboraram um pacote para o software R, disponível no sítio eletrônico: <<http://etcedi.pacificclimate.org/software.shtml>>, em que é possível aplicar a rotina de tratamento dos dados em uma interface amigável. No site, também disponibilizam um guia para usuário, dando o passo a passo para utilização da técnica.

O RHtest, amplamente divulgado pelo Serviço Canadense de Meteorologia e pela Organização Mundial de Meteorologia, tem relevante potencial aos estudos da climatologia estatística e geografia, já que oferece tratamento estatístico compatível com a unidade fundamental dos dados: diário. Em face disto, o presente artigo tem como objetivo avaliar a aplicabilidade do teste para a realidade dos trópicos

úmidos, tomando como referência três estações meteorológicas localizadas no estado do Rio de Janeiro: Manuel Ribeiro, Represa do Paraíso e Petrópolis, com série de dados de 1967 a 2015 e situadas em diferentes contextos topográficos – baixada litorânea, vertente barlavento e vertente sotavento, respectivamente..

METODOLOGIA

O RHtest é um método proposto por Wang (2008a, 2008b) utilizado para verificar a presença de pontos de mudança em séries temporais, levando em consideração a autocorrelação de primeira ordem. Esse método utiliza os testes t máximo penalizado (PMT) e F máximo penalizado (PMF). O teste t máximo penalizado (Wang et al. 2007; Wang, 2008a) assume que a série temporal que está sendo testada não apresenta tendência, enquanto que o teste F maximal penalizado (Wang, 2008a e 2008b) permite que a série tenha uma tendência linear durante todo o período. A partir dessa checagem de hipóteses, comprova-se a existência de ruídos (ou ausência) e a série de dados é homogeneizada (ou não).

A referida técnica consiste em dois algoritmos capazes de contabilizar a autocorrelação de primeira ordem na detecção de pontos de mudanças em séries temporais: o PMTred e o PMFred que utilizam os testes t máximo penalizado (PMT) e F máximo penalizado (PMF). Esses dois algoritmos são similares, uma vez que utilizam o mesmo algoritmo *stepwise*.

Nesses algoritmos, primeiramente é identificado o ponto de mudança mais provável $c_0 (t \in \{N_{min}, N_{min} + 1, \dots, N - N_{min}\})$. Então, os próximos pontos de mudança mais prováveis, c_{01} e c_{03} , são procurados a partir dos segmentos $t = \{N_{min}, \dots, (c_0 - N_{min})\}$ e $t = \{c_0 + 1 + N_{min}, \dots, (N - N_{min})\}$, respectivamente. Assim, a nova estimativa de c_0 , representada por c_{02} é identificada. Para encontrar o ponto de mudança mais significativo entre c_{01} , c_{02} e c_{03} é necessário estimar a estatística de teste PF_{max} e o valor-p para cada um deles.

Para um ponto de mudança c , o PF_{max} é calculado a partir do modelo ajustado com e sem este ponto de mudança, enquanto que para calcular o valor-p desse ponto de mudança, primeiro a série residual R_t do

modelo completo é usada para estimar a autocorrelação ϕ no lag-1 e para obter a série livre de autocorrelação através das seguintes equações:

$$\begin{aligned} R_t &= X_t - \hat{X}_t, t = 1, 2, \dots, N \\ W_1 &= R_1 \\ W_t &= R_t - \hat{\phi}R_{t-1}, t = 2, \dots, N \end{aligned}$$

Onde \hat{X}_t indica o modelo completo ajustado para a série X_t e W_t indica a série livre de autocorrelação que é usada para calcular a estatística F e o valor-p para c. Se for encontrado um ponto significativo, este é considerado como o primeiro ponto de mudança da série \hat{X}_t , sendo então necessário utilizar testes subsequentes para encontrar múltiplos pontos de mudança. Se não forem encontrados pontos significativos, a série \hat{X}_t é considerada homogênea.

Como dito anteriormente, a maioria dos métodos utilizados para verificar pontos de mudança pressupõem que a série avaliada é normalmente distribuída. Porém, essa suposição nem sempre é válida quando se trabalha com séries climáticas, principalmente quando se trabalha com séries de precipitação, uma variável contínua. Por isso, Wang *et al.* (2010) propuseram um método capaz de detectar pontos de mudança em séries que não apresentam distribuição normal. Esse método é chamado transPMFred ou ajuste IBC e consiste em integrar uma transformação Box-Cox (Box e Cox, 1964) no algoritmo PMFred de forma a diminuir os efeitos da falta de normalidade nos resultados do teste.

Como alternativa ao algoritmo transPMFred, Wang *et al.* (2010) também propuseram o algoritmo quantile-matching (QM). O objetivo do algoritmo QM é ajustar a série de modo que as distribuições empíricas de todos os segmentos da série sem tendência correspondam entre si. Sendo assim, a forma da distribuição é frequentemente ajustada (incluindo, mas não se limitando ao ajuste à média), ao passo que no algoritmo transPMFred faz basicamente um ajuste médio, provocando pouca mudança no formato da distribuição dos dados.

Para fins de avaliação da aplicabilidade do teste supracitado, tomou-se a série histórica de três postos pluviométricos localizados no estado do Rio de Janeiro como referência: Manuel Ribeiro, Represa do Paraíso e Petrópolis (Tabela 1 e Figura 1). A escolha desses postos ocorreu em virtude da

disponibilidade de uma série de dados longa (1967-2015), com falhas inferiores a 5% do total de dados e, principalmente, situadas em contextos topográficos distintos – baixada litorânea, vertente barlavento e vertente sotavento – de extrema importância para abarcar a variabilidade temporoespacial da pluviosidade na área de estudo, mormente influenciada pelo relevo (BERNARDES, 1953; NIMER, 1971; BRANDÃO, 1987)..

Nome do posto	Latitude	Longitude	Elevação (m)
Manuel Ribeiro	-22,910	-42,730	0,0
Petrópolis	-22,510	-43,170	890,0
Represa do Paraíso	-22,500	-42,910	60,0

Tabela 1. Informações dos postos utilizados no estudo

É fundamental que se considere que a técnica foi elaborada e é amplamente aplicada em um contexto climático diferente do nosso – Canadá (clima temperado), em que menor peso é dado a variabilidade pluviométrica nos diferentes tipos de tempo. Dessa maneira, testes complementares foram aplicados em ambas as séries de dados: dados brutos e dados homogeneizados, a fim de avaliar as mudanças provocadas pelo ajuste do RHtest.

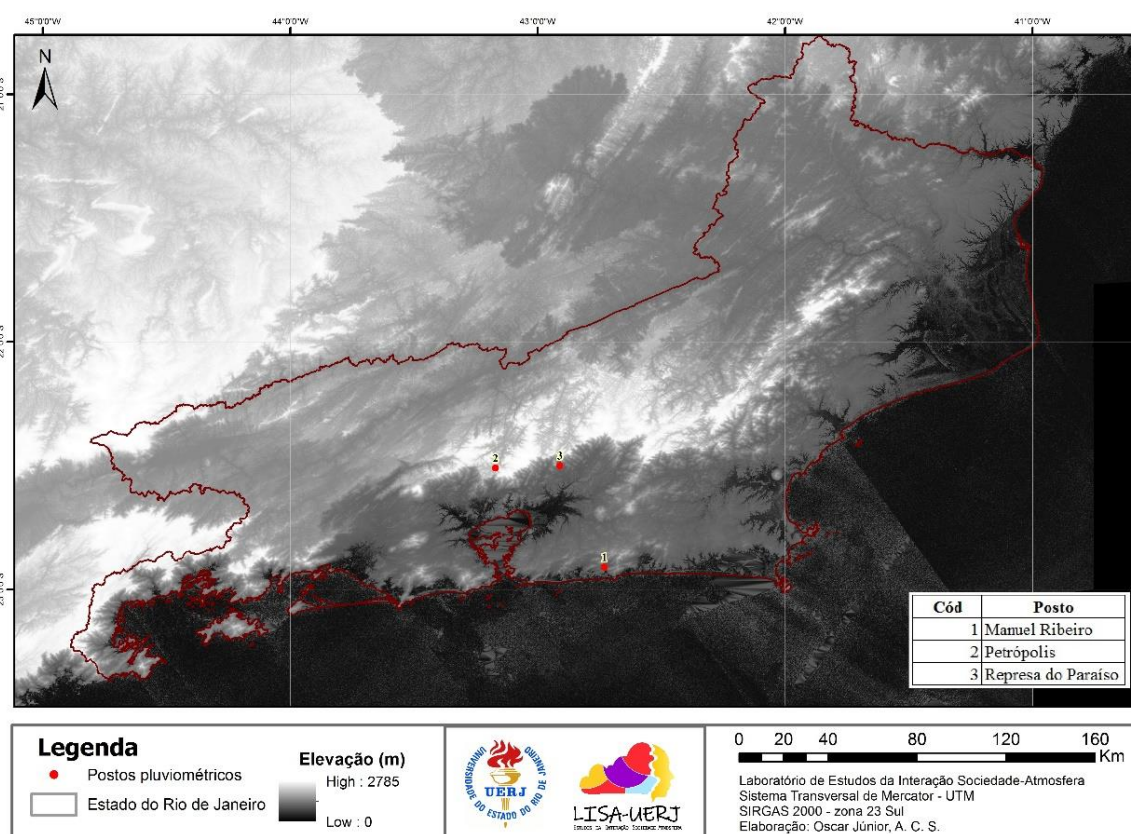


Figura 1 Postos pluviométricos utilizados no estudo.

É importante considerar ainda, conforme nos lembra Monteiro (1971;1991), que ao contrário do tratamento estatístico convencional, em uma análise climatológica própria da Geografia, os outliers (extremos) são fundamentais; afinal podem estar relacionados a eventos de alta magnitude e baixa frequência, traduzindo-se em impactos (NUNES, 2009). Dessa maneira, para a comparação dos dados brutos e homogeneizados utilizou-se o teste de correlação de Spearman, bem como, da distribuição de frequência entre classes de pluviosidade diária.

De acordo com Silvestre (2016), o teste de Spearman é um teste não paramétrico, ou seja, adequado para os dados utilizados neste estudo já que não seguem uma distribuição normal. O coeficiente de Spearman (ρ) é usado para medir a intensidade da relação entre variáveis ordinais, assumindo, ao invés do valor observado, a ordem das observações e pode ser calculado através da equação:

$$\rho = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n^3 - n},$$

onde: n é o número de pares (x_i, y_i) e d_i (postos de x_i dentre os valores de x)-(postos de y_i dentre os valores de y).

Em seguida, procedeu-se a distribuição de frequência das classes de pluviosidade diária, seguindo Brandão (2009), a fim de avaliar as alterações na frequência das magnitudes de precipitação diária, bem como definição do 3º quartil, percentil 99 (usualmente utilizado para identificação de precipitação extrema) (CASTELLANO, 2016) e valores máximos.

Finalmente, para fins de aplicação, procedeu-se com o cálculo de índices pluviométricos através da extensão RClimDex, tanto para os dados brutos quanto homogeneizados, com vistas à comparação dos resultados. Tais índices são recomendados pelo Centro de Meteorologia do Canadá, tal como pela Organização Mundial de Meteorologia para detecção de mudança climática e vem sendo fartamente aplicado em estudos acadêmicos (HAYLOCK *et al.*, 2006; MARENGO *et al.*, 2010; SILVA *et al.*,

2015). O objetivo de sua utilização é avaliar o impacto da homogeneização na análise de tendências temporais.

Na Tabela 2 encontram-se os índices pluviométricos calculados para os postos selecionados para o estudo, tomando como referência o período histórico de 1968-1997, a opção pelo início em 1968 se justifica pela maior incidência de falha dos dados no início da série (1967), o que afetaria a média e a comparação entre estações. Como resultado, os índices remetem valores conhecidos como *slope estimate* (S), definidos na tabela 2, cujo propósito é medir a intensidade da tendência.

Índice	Nome do Índice	Definição	Unidade
PRECPTOT	Precipitação total anual nos dias úmidos	Precipitação total anual nos dias úmidos ($RR \geq 1\text{mm}$)	mm
CDD	Dias secos consecutivos	Número máximo de dias consecutivos com $RR < 1\text{mm}$	dias
CWD	Dias úmidos consecutivos	Número máximo de dias consecutivos com $RR \geq 1\text{mm}$	dias
Rx1day	Quantidade máxima de precipitação em um dia	Máximo anual de precipitação em 1 dia	mm
Rx5day	Quantidade máxima de precipitação em cinco dias consecutivos	Máximo anual de precipitação em 5 dias consecutivos	mm
R95p	Dias muito úmidos (Chuvas fortes)	Precipitação anual total em que $RR >$ percentil 95	mm
R99p	Dias extremamente úmidos (Chuvas intensas)	Precipitação anual total em que $RR >$ percentil 99	mm

Tabela 2 Índices pluviométricos

Resultados e Discussão

A partir da Tabela 3 é possível indentificar os pontos de ruptura identificados pelo RHtest para os postos pluviométricos selecionados, em que o posto Petrópolis é o que apresenta mais considerável quantidade de rupturas: 5, ou seja, o teste identificou mais interferências não climáticas neste posto que nos demais. Para uma análise mais precisa sobre estes primeiros resultados era de fundamental importância acesso aos metadados dos postos, sobretudo aquelas informações sobre o histórico: data de início das observações regulares, transferências, interrupções nas observações, mudanças na

identificação da estação e quaisquer outras alterações substanciais introduzidas no programa de observação. Entretanto, apesar de exigido pelas normas da Organização Mundial de Meteorologia, estes dados nem sempre estão disponíveis, sobretudo quando se referem à postos não administrados pelo INMET – Instituto Nacional de Meteorologia. De acordo com Vianna et al. (2017), isso confere insegurança na utilização das bases de dados.

Os produtores rurais pesquisados foram classificados em três categorias, agrupando-os segundo as rendas, pessoas envolvidas no trabalho rural, ou não, tipo de produção agropecuária, infraestrutura das propriedades (quadro 1).

Posto	Pontos de ruptura na série temporal	
	data	Pvalor
Manuel Ribeiro	08/10/1991	0,001
Represa do Paraíso	14/11/1968 04/11/1993	0,001
Petrópolis	23/04/1951 17/11/1951 04/12/1954 26/05/1956 09/12/1999	0,001

Tabela 3. Pontos de ruptura das séries de pluviosidade diária das estações da RMRJ de acordo com o RHtest.

Ainda na Tabela 3, a coluna nomeada como pvalor avalia o nível de significância estatística do resultado, ou seja, a taxa na qual o teste conclui que a data em questão é um ponto de mudança para a série temporal. Conforme a referência $\alpha=0,05$, conclui-se que as datas identificadas como pontos de mudança para os três postos pluviométricos possuem significância estatística de 95%, não devendo ser desconsiderados.

Em seguida, considerando que, a partir da identificação dos pontos de ruptura, o RHtest procede com o ajuste (modificação) dos dados pluviométricos diários de modo a suprimir as interferências não climática na série de dados, aplicou-se o teste de Spearman. Os resultados são apresentados na Tabela 4.

Posto	Teste de Correlação de Spearman (dados brutos x dados homogeneizados)	
	ρ	Pvalor
Manuel Ribeiro	0,999	<0,0001
Represa do Paraíso	0,999	<0,0001
Petrópolis	0,998	<0,0001

Tabela 4. Pontos de ruptura das séries de pluviosidade diária das estações da RMRJ de acordo com o RHtest

Os resultados mostram que para todas as estações houve uma correlação positiva e significativa (com pvalor inferior à 0,05) entre os dados brutos e os homogeneizados, sendo o menor coeficiente de correlação igual a 0,989, ou seja, 98,9% de similaridade ordinal, isto para o posto Petrópolis. Tal resultado, indica, a priori, que quanto maior o número de rupturas identificadas na série de dados, mais ajustes ocorrerão e, conseqüentemente, maior será a diferença entre a série de dados brutos e a homogeneizada. Indica também que, para fins estatísticos, apesar das modificações processadas pela técnica de homogeneização, ainda é a mesma série de dados.

Dando prosseguimento à avaliação, as Tabelas 5, 6 e 7 sintetizam os resultados da distribuição de frequência de acordo com classes de pluviosidade. A partir destas, constata-se que para Manuel Ribeiro as variações mais expressivas são negativas, o que em outras palavras significa dizer houve perda significativa de dados, principalmente para o intervalo entre 200 e 250 mm.

Ainda no posto Manuel Ribeiro, o intervalo inferior a 1 mm foi o único com variação positiva, sinalizando que houve alterações nos dados pós-homogeneização que levaram a um acréscimo nesta classe; entretanto, apesar da coincidência da saída de 589 dados das demais classes e acréscimo de 589 dados na classe de inferior a 1mm, é equivocado interpretar que todos aqueles dados que saíram das classes com variação negativas se enquadram nessa classe de inferior a 1 mm, afinal as mudanças

processadas nos dados pelo RHtest não são tão abruptas, o que pode ser confirmado pelo resultado do teste de Spearman.

Enfim, em relação as medidas de extremo para o posto Manuel Ribeiro, todas foram modificadas para menos, em que a variação mais expressiva foi para o terceiro quartil que na série de dados brutos registrava 1,8 mm e na homogeneizada vai para 0,6 mm.

<i>Classes (mm)</i>	<i>Dados brutos</i>	<i>Dados homogeneizados</i>	<i>Variação %</i>
<1,0	12738	13327	4,6
1,0 - 10	2858	2494	-12,7
10 - 40	1628	1482	-9
40 - 60	208	164	-21,2
60 - 80	55	38	-30,9
80 - 100	25	12	-52
100 - 150	16	13	-18,8
150 - 200	3	3	0
200 - 250	2	0	-100
>250	0	0	0
3º quartil	1,8 mm	0,6 mm	-66,7
Percentil 99	49,6 mm	42,8 mm	-13,7
Máximo	232,0 mm	192,6 mm	-17

Tabela 5. Comparação da distribuição de frequência, 3º quartil, percentil 99 e valores máximos entre os dados brutos e homogeneizados do posto Manuel Ribeiro.

Já para o posto Represa do Paraíso, observa-se que as variações positivas são mais significativas, mormente para a classe de 40 a 60 mm. Para esse posto, a classe de inferior a 1 mm perdeu sessenta e um dados, seguido pela classe entre 1 e 10 mm que perdeu 308 dados. Pouca foi a variação nas medidas de extremo, sendo positivas, ou seja, refletindo que o RHtest aumentou os valores de pluviosidade, tanto para o terceiro quartil quanto para o percentil 99

<i>Classes (mm)</i>	<i>Dados brutos</i>	<i>Dados homogeneizados</i>	<i>Variação %</i>
<1,0	10955	10894	-0,6
1,0 - 10	2589	2281	-11,9
10 - 40	3323	3599	8,3

40 - 60	422	505	19,7
60 - 80	146	155	6,2
80 - 100	72	72	0
100 - 150	24	25	4,2
150 - 200	6	6	0
200 - 250	1	1	0
>250	0	0	0
3º quartil	9,0 mm	10,0 mm	11,1
Percentil 99	68,5 mm	68,9 mm	0,6
Máximo	204,2 mm	205,2 mm	0,5

Tabela 6. Comparação da distribuição de frequência, 3º quartil, percentil 99 e valores máximos entre os dados brutos e homogeneizados do posto Represa do Paraíso.

<i>Classes (mm)</i>	<i>Dados brutos</i>	<i>Dados homogeneizados</i>	<i>Variação %</i>
<1,0	15444	15546	0,7
1,0 - 10	3746	2587	-30,9
10 - 40	3235	3229	0,2
40 - 60	428	1018	137,9
60 - 80	177	459	159,3
80 - 100	58	176	203,4
100 - 150	32	97	203,1
150 - 200	6	12	100
200 - 250	2	4	100
>250	0	0	0
3º quartil	4,5 mm	6,6 mm	46,7
Percentil 99	64,0 mm	84,5 mm	32
Máximo	220,1 mm	220,1 mm	0

Tabela 7. Comparação da distribuição de frequência dos dados diários de chuva, 3º quartil, percentil 99 e valores máximos entre os dados brutos e homogeneizados do posto Petrópolis.

Comparando as Tabelas 5 e 6, observam-se comportamentos díspares em relação a distribuição dos dados brutos e homogeneizados, enquanto para Manuel Ribeiro o RHtest atuou na redução dos montantes de chuva diários, principalmente dos valores mais elevados (Classe 200-250, 3º quartil, Percentil 99 e máximo), para Represa do Paraíso atuou no aumento, sobretudo das classes intermediárias (10-40mm, 40-60 mm e 60-80 mm).

Quanto ao posto Petrópolis (Tabela 7), para o qual o teste remeteu maior quantidade de pontos de

ruptura, observa-se uma variação positiva e expressiva (superior a 100%) em quase todas as classes acima de 40 mm, exceto para a classe maior que 250 mm. A tendência de aumento positivo também foi acompanhada pelas medidas de extremo, em que houve aumento de 2,6 mm para o 3º quartil e de 20,5 mm para o percentil 99. É interessante ressaltar que esse foi o posto com maior ocorrência de fala no banco de dados (4,7%); enquanto Manuel Ribeiro registou 1,2% e Represa do Paraíso 0,9%.

Em termos objetivos, a análise de frequência entre classes de chuva, para os três postos, mostra que a homogeneização afetou de forma mais expressiva o posto Petrópolis, a sotavento, e principalmente o centro da distribuição (com variação superior a 100%); já na cauda a variação foi inferior a 50%. O posto menos afetado foi o localizado a barlavento da Serra do Mar, Represa do Paraíso, com variações que não ultrapassam os 20%, também no centro da distribuição. Dessa maneira, conclui-se que as medidas de extremo, hodiernamente fundamentais para os estudos em climatologia, foram afetadas pelo RHtest, sendo as variações mais expressivas a do 3º quartil, tanto no posto Manuel Ribeiro (baixada litorânea), com -66,7%, quanto para Petrópolis, com 46,7%; no entanto, representam diminuição de 1,2 mm e aumento de 2,1 mm, respectivamente.

Tais resultados, além de explicados pelo banco de dados e suas falhas, haja vista que as alterações mais expressivas ocorreram para o posto com maior percentual de falha de dados (Petrópolis), já as menos expressivas para o posto com menor percentual de falhas (0,9%). Possivelmente, os contrastes topográficos também influenciem esse resultado, fundamentalmente devido a característica de concentração da chuva na área de estudo. Estudo desenvolvido por Oscar Júnior (2018), demonstra, para o Estado do Rio de Janeiro, que as chuvas diárias são mais concentradas a sotavento da Serra do Mar e melhor distribuída na baixada litorânea. Tal característica afeta o *range* da pluviosidade dada sua variabilidade mais expressiva à sotavento, podendo influenciar na identificação dos pontos de ruptura, uma vez que, refletem variações do limiar sobre um conjunto de valores pequenos.

Finalmente, para complementar a análise, alguns índices de mudanças climáticas foram obtidos para os três postos, tanto para os dados brutos quanto para os homogeneizados. Os resultados estão sistematizados nas tabelas 8, 9 e 10 e em termos gerais, demonstram que a homogeneização influenciou

a significância estatísticas das tendências encontradas pelo RCLimDex.

Considerando o nível de significância $\alpha = 0,05$, o posto Manuel Ribeiro passou de 1 índice com significância estatística (R95p) para 6 (PRECPTOT, CWD, Rx1day, Rx5day, R95p e R99p), já o posto Petrópolis a mudança é de 0 índices com significância estatística para 2 (PRECPTOT e R95p). Quanto ao posto Represa do Paraíso, na série de dados brutos 2 índices apresentavam significância estatística (PRECPTOT e CWD), já na série homogeneizada apenas 1 apresentou significância (CWD).

Manuel Ribeiro				
Índice	Brutos		Homogeneizados	
	<i>S</i>	<i>p</i>	<i>S</i>	<i>p</i>
PRECPTOT	5,724	0,076	17,68	0
CDD	0,044	0,605	-0,119	0,169
CWD	-0,009	0,765	0,06	0,045
Rx1day	0,458	0,273	1,272	0,001
Rx5day	0,985	0,101	2,195	0
R95p	4,292	0,041	9,915	0
R99p	1,252	0,42	5,365	0

Tabela 8. Comparação de índices climáticos entre os dados brutos e homogeneizados do posto Manuel Ribeiro (os destaques vermelhos sinalizam aqueles dados com significância estatística).

Represa do Paraíso				
Índice	Brutos		Homogeneizados	
	<i>S</i>	<i>p</i>	<i>S</i>	<i>p</i>
PRECPTOT	25,887	0	12,907	0,062
CDD	0,043	0,672	0,051	0,616
CWD	0,328	0,008	0,325	0,008
Rx1day	-0,551	0,159	-0,589	0,133
Rx5day	0,496	0,497	0,059	0,935
R95p	3,717	0,27	2,137	0,453
R99p	-0,727	0,676	-0,96	0,585

Tabela 9. Comparação de índices climáticos entre os dados brutos e homogeneizados do posto Represa do Paraíso (os destaques vermelhos sinalizam aqueles dados com significância estatística).

Petrópolis				
Índice	Brutos		Homogeneizados	
	<i>S</i>	<i>p</i>	<i>S</i>	<i>p</i>
PRECPTOT	-2,897	0,186	-14,104	0
CDD	0,091	0,082	0,091	0,078
CWD	-0,018	0,607	-0,025	0,434
Rx1day	0,191	0,425	0,076	0,735

Rx5day	0,44	0,17	0,002	0,996
R95p	-0,98	0,589	-6,112	0,002
R99p	-0,681	0,474	-1,463	0,197

Tabela 10. Comparação de índices climáticos entre os dados brutos e homogeneizados do posto Petrópolis (os destaques vermelhos sinalizam aqueles dados com significância estatística).

Com relação ao *slope estimate* (S), observa-se que o sinal da tendência tanto para os dados brutos quanto homogeneizados permanece o mesmo quando o índice apresentou significância estatística para ambos; contudo, para um caso em que não havia significância estatística nos dados brutos e passou a ter nos dados homogeneizados, mudança no sinal de tendência pôde ser constatado, isto ocorre para o índice CWD do posto Manuel Ribeiro, demonstrando não ser impossível uma inversão de tendência.

Assim como na distribuição de frequência das classes de chuva, não há um mesmo comportamento para o *slope estimate* para os três postos. Ao passo que o valor de S aumenta para todos os índices climáticos, com significância estatística, da série de dados brutos para a homogeneizada no posto Manuel Ribeiro. Já para os postos Represa do Paraíso e Petrópolis, observa-se redução no valor. Dessa maneira, pode-se constatar que o processo de homogeneização dos dados diários repercutiu tanto na significância estatística da tendência quanto na sua intensidade.

Infelizmente, não se teve acesso a outro estudo com objetivos similares ao deste artigo, dificultando a comparação dos dados obtidos, o que certamente pode ser avaliado como uma limitação, já que inviabiliza uma contextualização dos resultados obtidos, mesmo que para outras realidades climáticas. O trabalho de Wang et al (2010) que também se propõem a analisar o teste, baseia-se na comparação dos resultados com os metadados das estações meteorológicas canadenses usadas no estudo, informações não disponíveis para as estações selecionadas para o caso brasileiro.

CONSIDERAÇÕES FINAIS

A verificação da qualidade e confiabilidade dos dados meteorológicos impõem novos arcahouços teóricos-metodológicos interdisciplinares, com vistas à superação das diversas limitações práticas que podem afetar de forma sistemática a aquisição dos dados e, por conseguinte, as análises. Desse modo,

considerando a premência da utilização de dados confiáveis e consistentes, surgem técnicas para homogeneização de séries de dados que se tornam cada vez mais indispensáveis para os estudos em climatologia.

Considerando a tradição da Climatologia Geográfica com a unidade diária, a partir da qual é possível estabelecer conexões entre tempo e clima, surge também uma limitação: trabalhar com dados que muitas vezes refletem realidades microclimáticas específicas e/ou são adulterados por procedimentos que acumularam erros técnicos de diversas ordens, culminando em uma interferência não climática significativa nos dados. Daí surge a importância de avaliar a eficiência do RHtest para superação dessa limitação, uma vez que a técnica se propõe a homogeneização de dados diários, ao contrário de muitas outras de focam o mensal ou anual.

Com o presente estudo, foi possível constatar a eficiência na utilização da técnica para a realidade dos trópicos úmidos, tomando como exemplo três postos pluviométricos do Rio de Janeiro em contextos topográficos diversos. A comparação ordinal entre as séries de dados brutos e homogeneizados atestam a eficiência da técnica; ademais as distribuições de frequência em classes de chuva permitiram avaliar que a técnica não afeta de forma expressiva os valores extremos, cada vez mais relevantes para a ciência climatológica.

Com o uso dos índices climáticos do RClimDex, uma técnica para avaliação de tendência de mudança em séries de dados climáticos, pôde-se constatar melhoria expressiva nos resultados homogeneizados em detrimento daqueles brutos, tomando como parâmetro a significância estatística dos índices.

Finalmente, alerta-se que este estudo não tem caráter finalístico, mas sim de apresentar uma técnica ainda “inédita” e de alto potencial de utilização pela comunidade da Climatologia Geográfica brasileira. Contudo, recomendam-se outras iniciativas de avaliação dos resultados tomando como referências estações com metadados disponíveis e testando diferentes agrupamentos de chuva diária, sobretudo em relação as estações do ano. Recomenda-se também sua aplicação em outros contextos climáticos do

Brasil; afinal, só dessa forma, será possível a construção de um consenso quanto a efetividade e eficiência da técnica para a realidade do território brasileiro.

REFERÊNCIAS

- AGUILAR, E.; AUER, I.; BRUNET, M.; PETERSON, T. C.; WIERINGA, J. Guidelines on climate metadata and homogenization. World Meteorological Organization, WMO-TD No. 1186, WCDMP No. 53, Geneva, Switzerland, p 55, 2003.
- AZORIN-MOLINA, C.; GUIJARRO, J. A.; MCVICAR, T. R.; VICENTE-SERRANO, S. M.; CHEN, D.; JEREZ, S.; ESPÍRITO-SANTO, F. Trends of Daily Peak Wind Gusts in Spain and Portugal, 1961-2014. *Journal of Geophysical Research*, c. 121, n. 3, pp.: 1059-1078, 2016.
- BAPTISTA, G. M. **Aquecimento Global: ciência ou religião**. Brasília: Hinterlândia Editorial, 2009.
- BERNARDES, L. M. C. Tipos de clima do estado do Rio de Janeiro. *Revista Brasileira de Geografia*, n.1, Rio de Janeiro: IBGE/Conselho Nacional de Geografia, 1953.
- BRANDÃO, A. M. P. M. Clima Urbano e Enchentes na Cidade do Rio de Janeiro. In: GUERRA, A. J. T.; CUNHA, S. B. (orgs), **Impactos Ambientais Urbanos no Brasil**, p.: 47-109, 5ª ed., Rio de Janeiro: Bertrand Brasil, 2009.
- BRANDÃO, A. M. P. M. **Tendências e oscilações climáticas na área metropolitana do Rio de Janeiro**. 1987. 385f. Dissertação (Mestrado em Geografia Física) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, 1987.
- CASTELLANO, M. S. **Extremos de chuva na Região Metropolitana de Campinas (SP): impactos, análise socioeconômica e políticas públicas**. 2016. 290f. Tese (Doutorado em Geografia) – Instituto de Geociências, Universidade Estadual de Campinas, 2016.
- HAYLOCK, M. R.; PETERSON, T. C.; ALVES, L. M.; AMBRIZZI, T.; ANUNCIAÇÃO, Y. M. T.; BAEZ, J.; BARROS, V. R.; BERLATO, M. A.; BIDEGAIN, M.; CORONEL, G.; CORRADI, V.; GARCIA, V. J.; GRIMM, A. M.; KAROLY, D.; MARENGO, J. A.; MARINO, M. B.; MONCUNILL, D. F.; NECHET, D.; QUINTANA, J.; REBELLO, E.; RUSTICUCCI, M.; SANTOS, J. L.; TREBEJO, I.; VINCENT, L. A. Trends in Total and Extreme South American Rainfall in 1960-2000 and Links with Sea Surface Temperature. *Journal of Climate*, v. 19, pp.: 1490-1512, 2006.
- KEMP, W.P.; BURNELL, D.G.; EVERSON, D.O.; THOMSON, A.J. Estimating missing daily maximum and minimum temperatures. *Journal of climate and applied meteorology*, v. 22, n. 9, p. 1587-1593, 1983.
- MARENGO, J. A.; RUSTICUCCI, M.; PENALBA, O.; RENOM, M. An Intercomparison of Observed and Simulated Extreme Rainfall and Temperature Events during the Last Half of the Twentieth Century: part 2: **Historical Trends**. *Climatic Change*, v. 98, pp.: 509–529, 2010.
- MARTIN-VIDE, J. Spatial distribution of a daily precipitation concentration index in peninsular Spain. *International Journal of Climatology*. N°24, p.: 959-971, 2004.
- MONTEIRO, C. A. F. **De tempos e ritmos**. Santa Catarina: Editora da UFSC, 2001.
- MONTEIRO, C. A. F. **Clima e Excepcionalismo: conjecturas sobre o desempenho da atmosfera como fenômeno geográfico**. Florianópolis: Ed. da UFSC, 1991.

- MONTEIRO, C. A. F. Análise Rítmica em Climatologia - problemas da atualidade climática em São Paulo e achegas para um programa de trabalho. *Climatologia*, São Paulo, v.1, p.1-21, 1971.
- NIMER, E. Climatologia do Brasil. Rio de Janeiro: IBGE, 1979.
- NUNES, L. H. Mudanças climáticas, extremos atmosféricos e padrões de risco a desastres hidrometeorológicos. In: HOGAN, D. J. & MARANDOLA JR, E. (org.) **Populações e mudanças climáticas: dimensões humanas das mudanças ambientais globais**, p.: 29-52 – CAMPINAS: NEPO/Unicamp; Brasília: UNFPA, 2009.
- PÉDÉLABORDE, P. **Introduction à l'étude scientifique du climat**. Paris: SEDES, 1970.
- PETERSON, T.C.; EASTERLING, D.R.; KARL, T.R.; GROISMAN, P.; NICHOLLS, N.; PLUMMER, N.; TOROK, S.; AUER, I.; BÖHM, R.; GULLETT, D.; VINCENT, L.; HEINO, R.; TUOMENVIRTA, H.; MESTRE, O.; SZENTIMREY, T.; SALINGER, J.; FØRLAND, E.J.; HANSSEN-BAUER, I.; ALEXANDERSSON, H.; JONES, P.; PARKER, D. Homogeneity adjustments of in situ atmospheric climate data: a review. *International Journal of Climatology*, n. 18, pp.: 1493–1517, 1998.
- SILVA, W. L.; DEREZYNSKI, C.; CHANG, M.; FREITAS, M.; MACHADO, B. J.; TRISTÃO, L.; RUGGERI, J. Tendências Observadas em Indicadores de Extremos Climáticos de Temperatura e Precipitação no Estado do Paraná. *Revista Brasileira de Meteorologia*, v. 30, n. 2, pp.: 181-194, 2015.
- SILVESTRE, M. R. **Técnicas Estatísticas Utilizadas em Climatologia Geográfica: Diagnóstico e Propostas**, 2016. Tese (Doutorado em Geografia) – Faculdade de Ciências e Tecnologia, Universidade Estadual Paulista, 2016.
- SORRE, M. **Les fondements de la Geographie Humaine. (les fondements Biologiques – Le climat)** Paris, Libr. Armand Colin. 1951.
- OSCAR JÚNIOR, A. C. S. **Governança Territorial em Nível Metropolitano e Risco da Mudança Climática no Rio de Janeiro**, 2018. Tese (Doutorado em Geografia) – Instituto de Geociências, Universidade Estadual de Campinas, 2018.
- VIANNA, L. F. de N. et al. Bancos de Dados Meteorológicos: Análise dos Metadados das Estações Meteorológicas no Estado de Santa Catarina, Brasil. *Revista Brasileira de Meteorologia*, São Paulo, v. 32, n. 1, p. 53-64, Mar. 2017.
- WANG, X. L.; CHEN, H.; WU, Y.; FENG, Y.; PU, Q. New techniques for detection and adjustment of shifts in daily precipitation data series. *Journal of Applied Meteorology and Climatology*, v. 49, n. 12, pp.: 2416-2436. 2010.
- WANG, X. L. Accounting for autocorrelation in detecting mean-shifts in climate data series using the penalized maximal t or F test. *Journal of Applied Meteorology and Climatology*, n.47, pp.: 2423-2444. 2008a.
- WANG, X. L. Penalized maximal F-test for detecting undocumented mean shifts without trend-change. *Journal of Atmospheric and Oceanic Technology*, v. 25, n. 3, pp.: 368-384. 2008b.
- WANG, X. L.; Wen, Q. H.; Wu, Y. Penalized maximal t test for detecting undocumented mean change in climate data series. *Journal of Applied Meteorology and Climatology*, v. 46, n. 6, p. 916-931, 2007.
- WANG, X. L. Comments on Detection of Undocumented Change-points: A Revision of the Two-Phase Regression Model. *Journal of Climate*, n.16, pp.: 3383-3385. 2003.