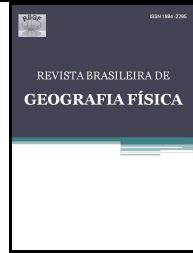




# Revista Brasileira de Geografia Física

Homepage: [www.ufpe.br/rbgfe](http://www.ufpe.br/rbgfe)



## Imputação Multivariada de Dados Diários de Precipitação e Análise de Índices de Extremos Climáticos

Rafaela Lisboa Costa<sup>1</sup>, Fabrício Daniel dos Santos Silva<sup>2</sup>, Gabriel Fonseca Sarmanho<sup>3</sup>, Paulo Sérgio Lucio<sup>4</sup>

<sup>1</sup>Centro de Previsão do Tempo e Estudos Climáticos - Instituto Nacional de Pesquisas Espaciais – INPE. Email: [rafaelalisboac@gmail.com](mailto:rafaelalisboac@gmail.com); <sup>2</sup>Instituto Nacional de Meteorologia – INMET. Email: [fabricao.silva@inmet.gov.br](mailto:fabricao.silva@inmet.gov.br); <sup>3</sup>Instituto Nacional de Meteorologia – INMETRO. Email: [gabrielfsarmanho@gmail.com](mailto:gabrielfsarmanho@gmail.com); <sup>4</sup>Professor associado, Universidade Federal do Rio Grande do Norte– UFRN. Email: [pslucio@ccet.ufrn.br](mailto:pslucio@ccet.ufrn.br).

Artigo recebido em 09/10/2012 e aceito em 19/10/2012

### RESUMO

Neste trabalho utilizou-se a técnica *MICE* (do inglês “*Multivariate Imputation by Chained Equations*”) para imputação de dados diários de precipitação em seis séries com falhas do Estado da Paraíba: Areia, Campina Grande, Monteiro, João Pessoa, Patos e São Gonçalo, entre período de 1979 e 2010, usando como fonte de informação para o preenchimento dados de pontos de grade próximos à estação meteorológica selecionada. Com a finalidade de validação do método, foram geradas falhas em dados observados para determinados anos - princípio da validação cruzada. A metodologia apresentou resultados promissores. Foram obtidos altos valores de correlações tanto entre os dados diários observados e imputados, assim como quando estes foram acumulados mensalmente. A partir dos dados imputados, utilizou-se o software R-ClimDex com o objetivo de avaliar possíveis tendências nos índices climáticos relacionados à precipitação. As análises das séries sem falhas mostraram que das cidades estudadas, Areia foi a única que apresentou tendência de redução no número de dias com chuvas acima de 1mm, enquanto as demais cidades apresentaram tendência de aumento.

Palavras - chave: dados imputados, *MICE*, R-ClimDex, tendências climáticas.

## Imputation Multivariate of Precipitation Daily Data and Analysis of Climate Extremes Index

### ABSTRACT

This work used the *MICE*(*Multivariate Imputation by Chained Equations*) technique for daily rainfall data imputation in series with gaps, for six series of Paraíba State: Areia, Campina Grande, Monteiro, Joao Pessoa, Patos and São Gonçalo, between 1979 and 2010, using as a source of information data grid points near the selected station (nearest neighbors). In order to validate the method one generates missing values in observed data for certain years– the cross-validation principle. The methodology presented promising results. One obtains high values for both correlations between daily observed and imputed data, as well as when they are monthly accumulates. Based on the imputed data, used the software R-ClimDex with the objective to evaluate possible trends in climate indices related to precipitation. The analysis of the series without flaws of the cities studied showed that Areia was the only one that showed decrease trend in the number of days with precipitation above 1 mm, while the other cities showed increase trend.

Keywords: imputed data, *MICE*, R-ClimDex, climatic trends.

### 1. Introdução

Em Meteorologia e Climatologia, diversos produtos operacionais de

monitoramento em tempo e clima dependem de uma boa rede de observações em superfície, que não apresentem ou tenham o menor número possível de falhas. Séries de dados pluviométricos completos servem tanto

\* E-mail para correspondência: [rafaelalisboac@gmail.com](mailto:rafaelalisboac@gmail.com) (Costa, R. L.).

para caracterizar a climatologia média das chuvas de uma região, como para a determinação do volume de água necessário para culturas irrigadas, como também para estudos de mudanças, tendências e extremos climáticos (Salvador, 2010). Lucio et al. (2007) observou que, frequentemente, dados de recordes climatológicos de localidades onde não há observações meteorológicas, tem que ser calculados via interpolações temporais de observações em estações vizinhas. Desta forma faz-se necessário a busca por boas técnicas de preenchimento de falhas em dados diários de precipitação. Este trabalho pretende analisar séries temporais de precipitação de seis estações meteorológicas localizadas no Estado da Paraíba, cedidas pelo Instituto Nacional de Meteorologia (INMET).

Os eventos diários de precipitação da região Nordeste do Brasil, em particular a Paraíba, estão associados a mecanismos dinâmicos de grande escala entre os quais a Zona de Convergência Intertropical (ZCIT) (Hastenrath, 1984; Moura e Shukla, 1981), Vórtices Ciclônicos de Altos Níveis (VCAN) (Kousky e Gan, 1981), a Oscilação de Madden-Julian (OMJ) (Kousky e Kayano, 1994), e, em menor grau de intensidade, a penetração de sistemas frontais oriundos de latitudes médias austrais (Kousky, 1979). Estes são os principais sistemas indutores de precipitação, principalmente nos meses de janeiro a abril, na parte oeste do Estado, e de abril a julho, na porção leste do Estado. A ocorrência destes fenômenos favorece a

ocorrência, inclusive, de eventos extremos diários de precipitação, que precisam ser resgatados para diversos estudos de natureza sinótica e/ou climatológica.

A partir de séries de dados, muitos estudos já buscaram identificar tendências de mudanças climáticas na região Nordeste do Brasil (Hastenrath e Greischar, 1993; Santos et al, 2006). No entanto, um grande fator de dificuldade está sendo a qualidade das séries usadas, sejam de temperaturas ou de precipitações, que devem estar na escala diária, e quase sempre apresentam falhas que muitas vezes inviabilizam a investigação para determinada localidade, ou para determinados índices.

Este trabalho teve então dois objetivos: o primeiro aplicar e testar a técnica *MICE* a séries pluviométricas do Estado da Paraíba, usando como variável preditora a mesma variável proveniente de análise gradeada do CPC/NOAA; e o segundo analisar possíveis mudanças nas características e no comportamento da precipitação pluviométrica nas cidades que tiveram suas séries tratadas sem falhas por meio de índices climáticos de extremos, usando o software R-ClimDex. De acordo com os resultados alcançados, torna-se objetivo futuro a execução desta pesquisa aos demais Estados da região.

## 2. Material e Métodos

### 2.1 Imputação de dados faltantes

Vários são os campos de aplicação da técnica *MICE* (Van Buuren, 2011),

principalmente a dados da área de saúde pública: arteriosclerose (McClelland et al 2008), epidemiologia (Ton et al 2009); política (Tanasoiu e Colonescu 2008); psicologia (Sundell et al 2008) e sociologia (Finke e Adamczyk. 2008), onde todos os autores utilizaram a técnica para lidar com os dados faltantes.

Essa técnica é utilizada para preenchimento de dados faltantes em series históricas, faltas essas causadas por motivos diversos, como falha no instrumento ou problemas nas estações meteorológicas. Os

dados faltantes que se deseja completar são caracterizados por “NA”, e foi usado o número de imputações múltiplas padrão do pacote *MICE* versão 2.12 (m=5 iterações) do software estatístico livre R, para versões  $\geq$  2.10. As imputações são geradas de acordo com o método padrão, que é, para dados numéricos, o método PMM-Predictive Mean Matching. A série de precipitação original com falhas da estação meteorológica é colocada lado a lado dos quatro pontos de grade mais próximos da localidade, em uma matriz exemplificada na Tabela 1.

**Tabela 1.** Ilustração de dados a serem imputados, representados por NA, referentes à estação de Campina Grande (a) e dados imputados (b). Os dados originais (coluna Orig) estão ao lado das séries gradeadas mais próximas (Gr-01; Gr-02; Gr-03 e Gr-04).

| (a)  |     |     |      |       |       |       |       | (b)  |     |     |      |       |       |       |       |
|------|-----|-----|------|-------|-------|-------|-------|------|-----|-----|------|-------|-------|-------|-------|
| Ano  | Mes | Dia | Orig | Gr-01 | Gr-02 | Gr-03 | Gr-04 | Ano  | Mes | Dia | Orig | Gr-01 | Gr-02 | Gr-03 | Gr-04 |
| 1979 | 01  | 023 | NA   | 0.7   | 1.0   | 1.8   | 0.9   | 1979 | 01  | 023 | 1.7  | 0.7   | 1.0   | 1.8   | 0.9   |
| 1979 | 01  | 024 | 1.1  | 4.6   | 0.3   | 1.2   | 3.4   | 1979 | 01  | 024 | 1.1  | 4.6   | 0.3   | 1.2   | 3.4   |
| 1979 | 01  | 025 | 6.7  | 8.3   | 3.8   | 6.1   | 19.8  | 1979 | 01  | 025 | 6.7  | 8.3   | 3.8   | 6.1   | 19.8  |
| 1979 | 01  | 026 | 17.4 | 14.8  | 3.5   | 9.1   | 26.6  | 1979 | 01  | 026 | 17.4 | 14.8  | 3.5   | 9.1   | 26.6  |
| 1979 | 01  | 027 | NA   | 0.3   | 0.2   | 1.4   | 1.0   | 1979 | 01  | 027 | 0.0  | 0.3   | 0.2   | 1.4   | 0.1   |
| 1979 | 01  | 028 | 0.0  | 0.1   | 0.0   | 0.3   | 0.8   | 1979 | 01  | 028 | 0.0  | 0.1   | 0.0   | 0.3   | 0.8   |

A técnica foi utilizada para seis séries pluviométricas do Instituto Nacional de Meteorologia (INMET) referente ao Estado da Paraíba, entre 1979 e 2010. Como variáveis predictoras, usou-se séries das análises gradeada de precipitação do CPC/NOAA (Chen et al. 2008), baseadas no método de interpolação ótima (IO) de Gandin (1965).

Após a imputação dos dados faltantes, foram identificadas nas séries originais ao menos 05 anos com dados medidos sem falhas na série original, e então falhas foram

geradas artificialmente para estes anos e o método foi novamente empregado, a fim de comparar dados realmente observados com a imputação gerada, com o intuito de avaliar a habilidade do método. Estatísticas como histogramas e correlações foram calculadas para estes períodos de verificação para validação da metodologia proposta.

## 2.2 Índices climáticos

As séries de precipitação sem falhas foram utilizados para a investigação de índices de mudanças climáticas. Para tanto,

foi utilizado o software R-ClimDex 2.3.1 (Zhang e Yang, 2004). Este software é obtido gratuitamente no site <http://cccma.seos.uvic.ca/ETCCDMI/software.html>. Ao todo foram analisados 11 índices descritivos e indicadores de extremos, para precipitação. Para manuseio e execução, é necessário dispor os dados de precipitação em formato texto ASCII, com colunas correspondentes ao ano, mês, dia, precipitação, temperatura máxima e temperatura mínima (que não foram aqui analisados). Tais índices descrevem as características particulares de extremos, incluindo a frequência, amplitude e persistência. Os dados passaram por um controle de qualidade, que incluiu a consistência espacial e temporal das observações diárias utilizadas. Os índices

usados neste estudo têm sua descrição detalhada na Tabela 2, conforme a definição realizada pelo “Expert Team on Climate Change Detection Monitoring and Indices (ETCCDMI)”.

O R-ClimDex fornece para cada índice dados estatísticos, como, por exemplo, tendência linear calculada pelo método dos mínimos quadrados, nível de significância estatística da tendência (valor  $p$ ), coeficiente de determinação ( $r^2$ ), erro padrão de estimativa e todos os gráficos das séries anuais. Foram considerados como possíveis mudanças climáticas, aqueles que apresentaram tendências lineares (positivas ou negativas) superiores ao erro padrão de estimativa e com alta significância estatística ( $p < 0,05$ ).

**Tabela 2.** Definições dos 11 índices de precipitação usados no estudo.

| Índice  | Definição  | Unidades |
|---------|--|----------|
| RX1day  | Precipitação máxima em 1 dia no ano  | mm       |
| RX5day  | Precipitações máximas em 5 dias consecutivos no ano  | mm       |
| SDII    | Média de precipitação nos dias com chuva   | mm/dia   |
| R10mm   | Contagem de dias no ano com precipitação maior ou igual a 10 mm                                  | dia      |
| R20mm   | Contagem de dias no ano com precipitação maior ou igual a 20 mm                                  | dia      |
| R1mm    | Número de dias em um ano em que a Precipitação $\geq 1$ mm                                       | dia      |
| CDD     | Número máximo de dias secos consecutivos   | dia      |
| CWD     | Número máximo de dias úmidos consecutivos  | dia      |
| R95p    | Quantidade de chuva anual que excedeu o percentil 95 da precipitação durante o período analisado | mm       |
| R99p    | Quantidade de chuva anual que excedeu o percentil 99 da precipitação durante o período analisado | mm       |
| PRCPTOT | Precipitação total anual para dias úmidos  | mm       |

### 3. Resultados e Discussão

#### 3.1 Análises da Imputação Múltipla de dados

A Tabela 3 mostra as coordenadas geográficas das estações, seu respectivo código OMM, porcentagem total de dados

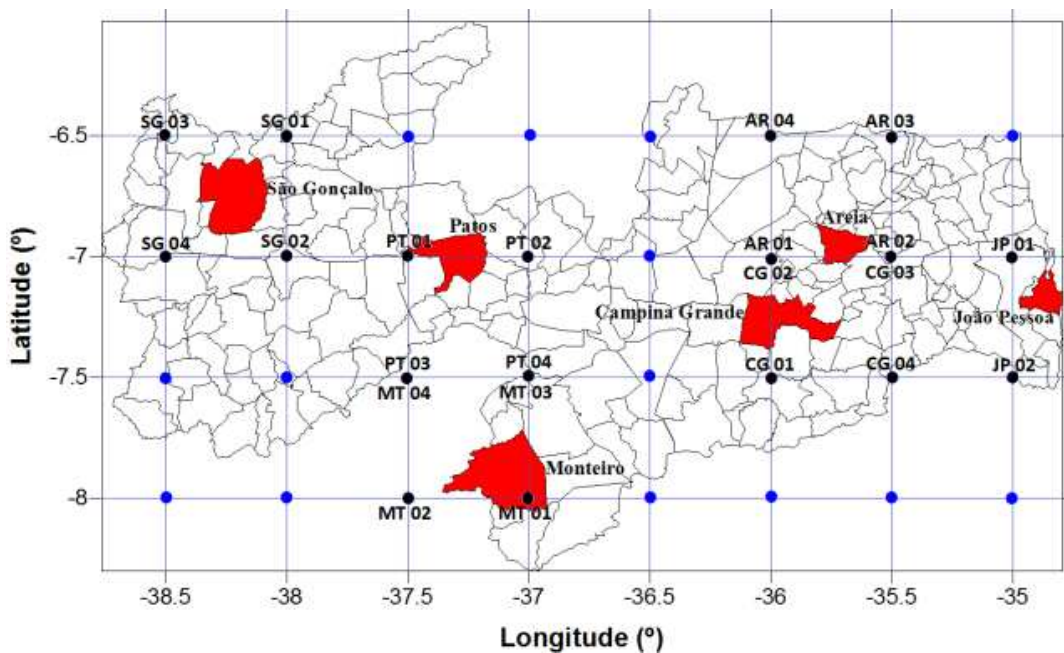
faltantes entre 1979 e 2010, e as coordenadas dos pontos de grade utilizados para o emprego da técnica *MICE* e posterior preenchimento das falhas, e a Figura 1 mostra a localização das estações de estudo

entre os pontos de grade. Dentre as estações, Areia apresentou a maior quantidade de dados faltantes e João Pessoa, a menor. Os anos escolhidos com dados observados que foram retirados para verificação e validação foram para: Areia: 1996, 1998, 2003, 2006 e 2008; Campina Grande: 1984, 1998, 2004, 2007 e 2009; Monteiro: 1983, 1997, 2001,

2003 e 2007; João Pessoa: 1985, 1996, 2001, 2004 e 2009; e Patos: 1984, 1994, 1998, 2003 e 2004, em um total de 1825 dados para serem re-amostrados. Para São Gonçalo foi possível obter apenas quatro anos com dados completos: 1983, 1984, 1985 e 2002, em um total de 1460 dados para serem re-amostrados.

**Tabela 3.** Informações dos postos e pontos de grade utilizados.

| Estação        | Código | Lat(°) | Lon(°) | Dados Faltantes | Gr-01 (lat-lon) | Gr-02 (lat-lon) | Gr-03 (lat-lon)  | Gr-04 (lat-lon)  |
|----------------|--------|--------|--------|-----------------|-----------------|-----------------|------------------|------------------|
| Areia          | 82696  | -6,97  | -35,68 | 51,84%          | -7 ° ; -36 °    | -7 ° ; -35,5 °  | -6,5 ° ; -35,5 ° | -6,5 ° ; -36 °   |
| Campina Grande | 82795  | -7,22  | -35,88 | 33,63%          | -7,5 ° ; -36 °  | -7 ° ; -36 °    | -7 ° ; -35,5 °   | -7,5 ° ; -35,5 ° |
| João Pessoa    | 82798  | -7,10  | -34,87 | 8,60%           | -7 ° ; -35 °    | -7,5 ° ; -35 °  | -7 ° ; -34,5 °   | -7,5 ° ; -34,5 ° |
| Monteiro       | 82792  | -7,88  | -37,07 | 36,27%          | -8 ° ; -37 °    | -8 ° ; -37,5 °  | -7,5 ° ; -37 °   | -7,5 ° ; -37,5 ° |
| Patos          | 82791  | -7,02  | -37,27 | 36,03%          | -7 ° ; -37,5 °  | -7 ° ; -37 °    | -7,5 ° ; -37,5 ° | -7,5 ° ; -37 °   |
| São Gonçalo    | 82689  | -6,75  | -38,22 | 38,79%          | -6,5 ° ; -38 °  | -7 ° ; -38 °    | -6,5 ° ; -38,5 ° | -7 ° ; -38,5 °   |



**Figura 1.** Localização das estações, em vermelho. Pontos de grade da análise gradeada do CPC/NOAA em azul, destacados em preto os usados para preenchimento de falhas para o período 1979-2010.

Como a falta de dados pode ocorrer em qualquer parte dos dados, o algoritmo atribui uma coluna de dados incompletos onde gera valores sintéticos “plausíveis” de acordo com

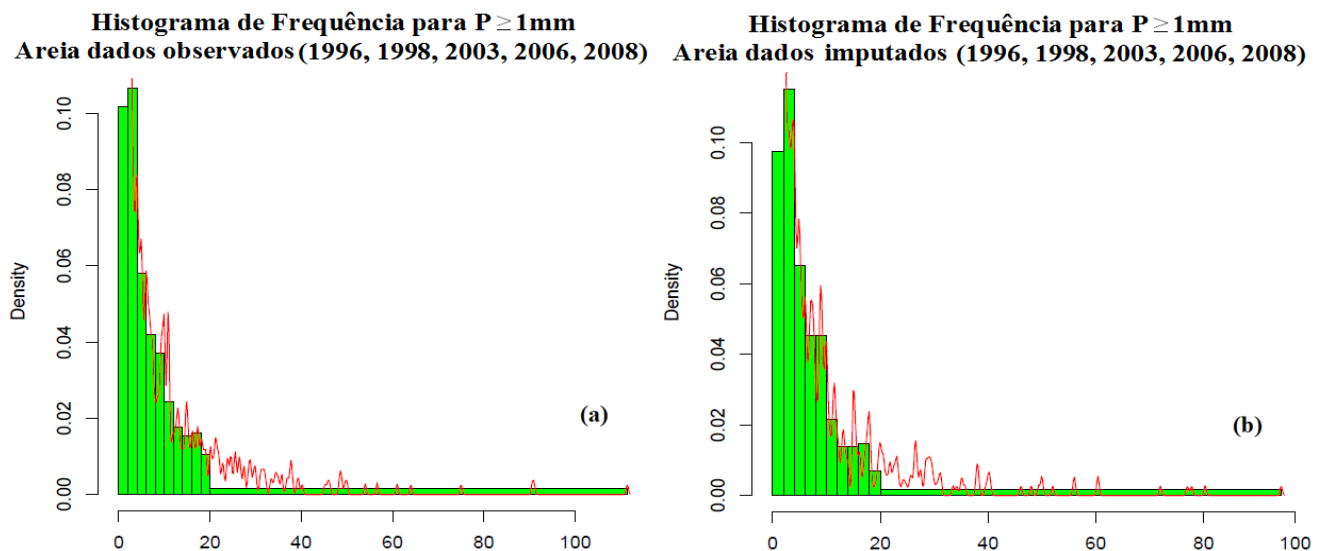
as demais colunas de dados, onde cada dado desta coluna quando coincide com a falha a ser tratada, tem seu próprio conjunto de equações baseadas em indicadores

específicos. Desta forma o método “conserva” os dados originais no arquivo de saída mesmo levando-os em consideração na geração de dados sintéticos.

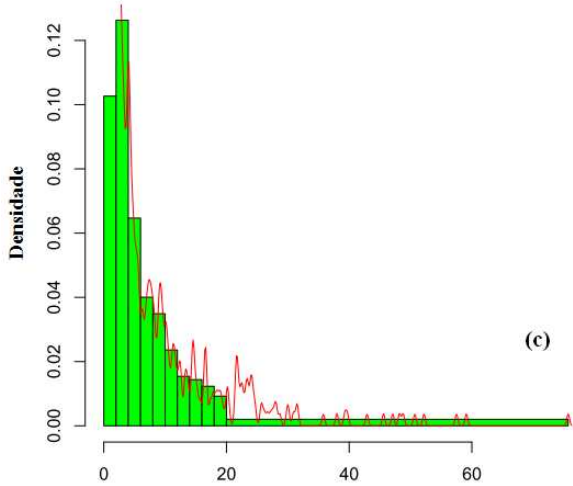
A Figura 2 mostra os histogramas de frequência para os anos de dados observados e imputados para as cidades de Areia, Campina Grande, João Pessoa, Monteiro e Patos, e para os dados imputados para São Gonçalo, já que devido à redução no número de dados observados não foi possível gerar o histograma para dados observados nesta localidade. Em todas as cidades, A distribuição de frequências mostrou-se bastante similar nos dois casos, com os dados imputados preservando a geração de valores extremos coincidentes aos dados observados, muitas vezes superiores ao maior valor indicado por um dos pontos de grade. Desde Areia, que apresentou a maior porcentagem de dados diários faltantes, 51.84%, a João Pessoa, que apresentou apenas 8.60% de dados faltantes, as escalas ilustram a fidedignidade entre valores observados e

imputados. A exceção se deu por conta de Campina Grande, cuja escala para os valores observados vai até 60 mm e para os dados imputados, vai até 100 mm, sugerindo que uma fração maior de valores mais elevados foram imputados, alongando um pouco mais as classes de dados.

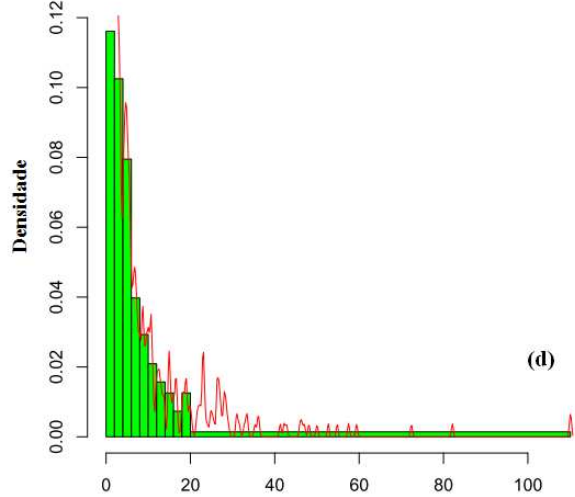
A metodologia corrige em muitos casos dois problemas comuns quando se tenta completar falhas em dados observados imputando diretamente dados do ponto de grade mais próximo: a atribuição de pequenos valores de precipitação em dias de precipitação nula (abaixo de 1 mm), e a subestimação de valores observados elevados, geralmente acima de 50 mm diários. Em alguns casos, os valores diários estimados foram superiores aos valores diários observados. Desta maneira, a metodologia foi fiel em três princípios básicos: descrição do processo que cria os dados faltantes; preserva a relação existente entre os dados e também preserva a incerteza destas relações, ao não gerar dados perfeitos.



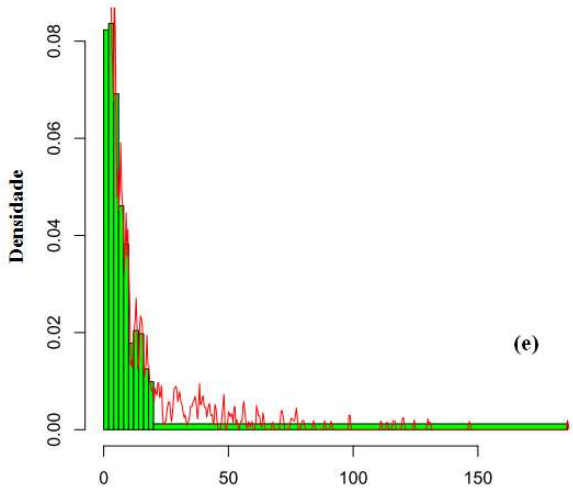
**Histograma de Freqüência para  $P \geq 1mm$**   
C.Grande dados observados (1984, 1998, 2004, 2007, 2009)



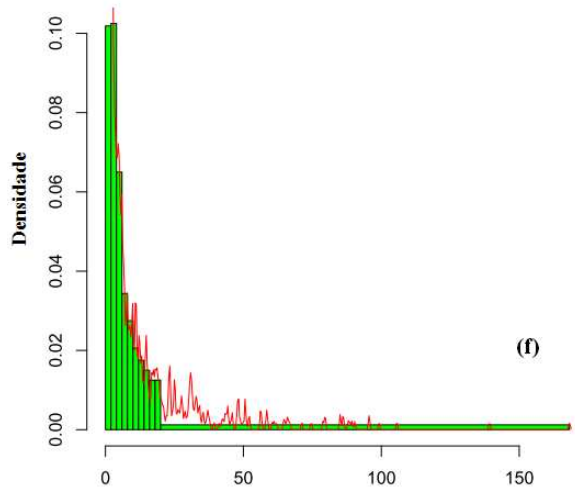
**Histograma de Freqüência para  $P \geq 1mm$**   
C.Grande dados imputados (1984, 1998, 2004, 2007, 2009)



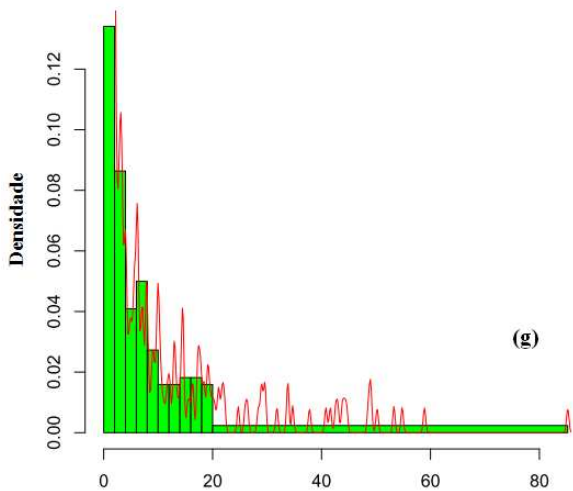
**Histograma de Freqüência para  $P \geq 1mm$**   
J.Pessoa dados observados (1985, 1996, 2001, 2004, 2009)



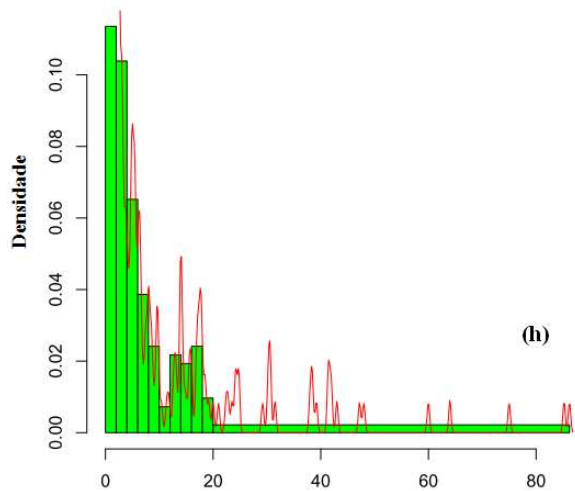
**Histograma de Freqüência para  $P \geq 1mm$**   
J.Pessoa dados imputados (1985, 1996, 2001, 2004, 2009)

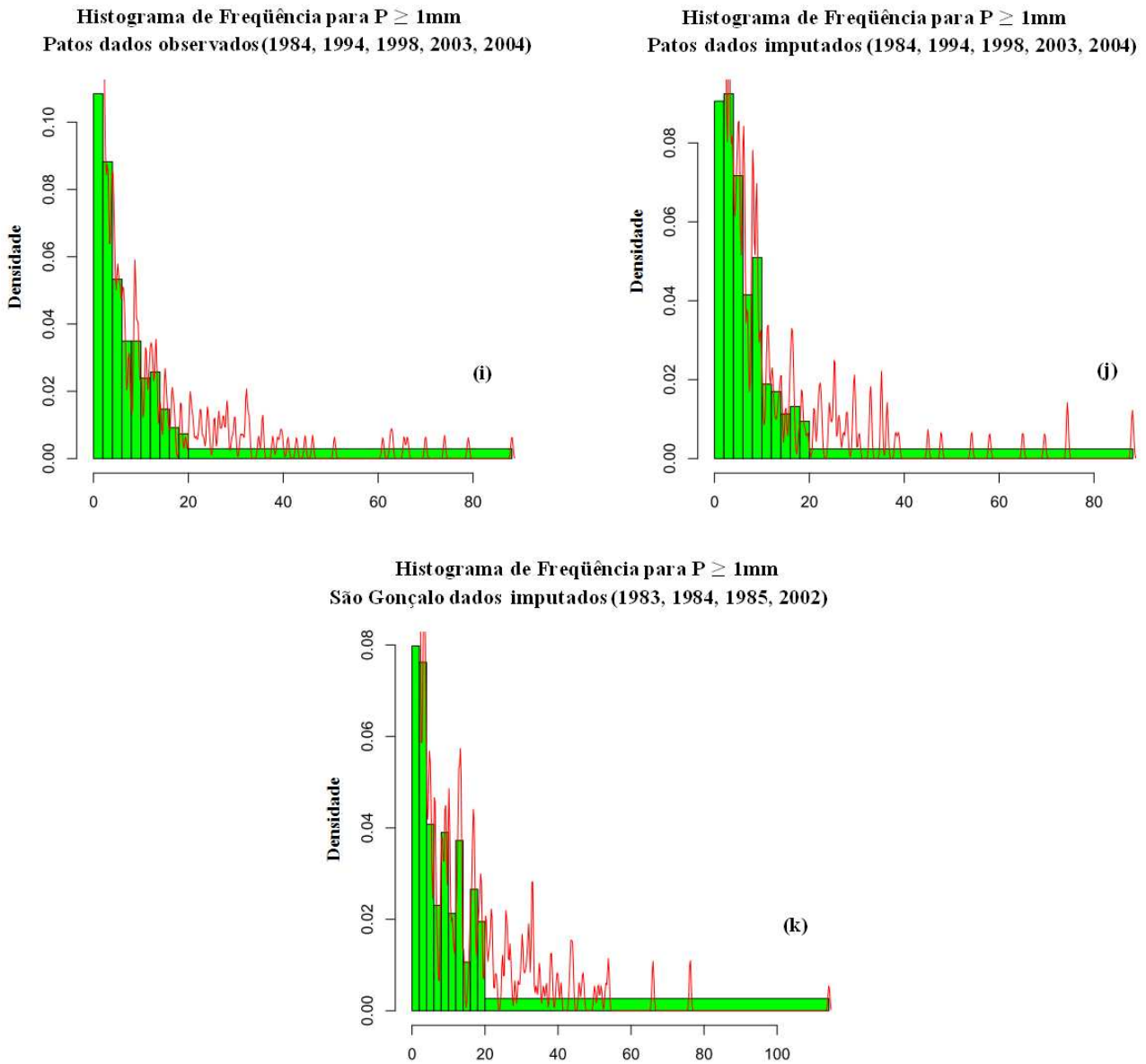


**Histograma de Freqüência para  $P \geq 1mm$**   
Monteiro dados observados (1983, 1997, 2001, 2003, 2007)



**Histograma de Freqüência para  $P \geq 1mm$**   
Monteiro dados imputados (1983, 1997, 2001, 2003, 2007)





**Figura 2.** Histogramas de frequência da precipitação diária para as cidades de areia (a) dados observados e (b) dados imputados, Campina Grande (c) dados observados e (d) dados imputados, João Pessoa (e) dados observados e (f) dados imputados, Monteiro (g) dados observados e (h) dados imputados, Patos (i) dados observados e (j) dados imputados e São Gonçalo (k) dados imputados, para os cinco anos de comparações entre dados observados e imputados. A linha vermelha é um filtro de Kernel para abordagem não-paramétrica para encaixe das densidades.

A Figura 3 mostra a comparação dos acumulados mensais oriundos da observação e da imputação, para as seis estações escolhidas. Os valores acumulados do processo de imputação diária apresentam o mesmo comportamento da observação para os meses secos e chuvosos, sem apresentar

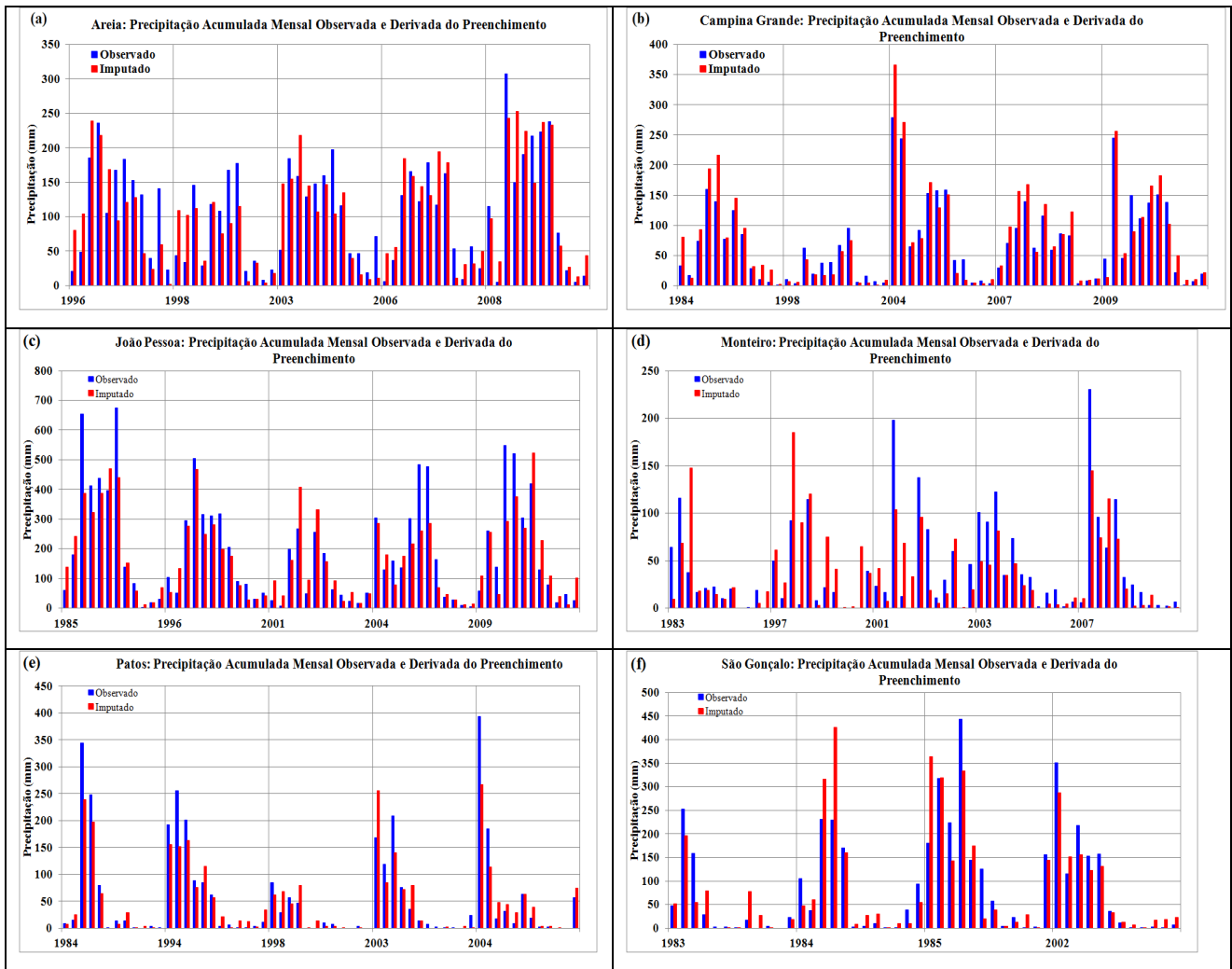
um viés preferencial de superestimar ou subestimar a observação, pois pequenas diferenças ocorrem mês a mês, a cada ano.

A Figura 4 mostra as correlações entre os dados diários e os acumulados mensais. Estes resultados evidenciam a habilidade do método em imputar valores faltantes que

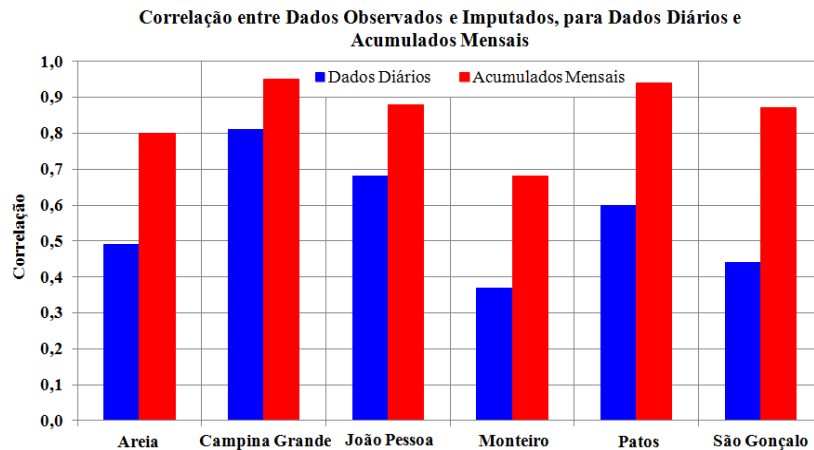


foram coincidentes com as observações. Todas as correlações obtidas foram

estatisticamente significantes, para a análise diária e mensal.



**Figura 3.** Comparação entre acumulados mensais de precipitação com dados observados (azul) e imputados (vermelho), para as cidades de Areia (a), Campina Grande (b), João Pessoa (c), Monteiro (d), Patos (e) e São Gonçalo (f).



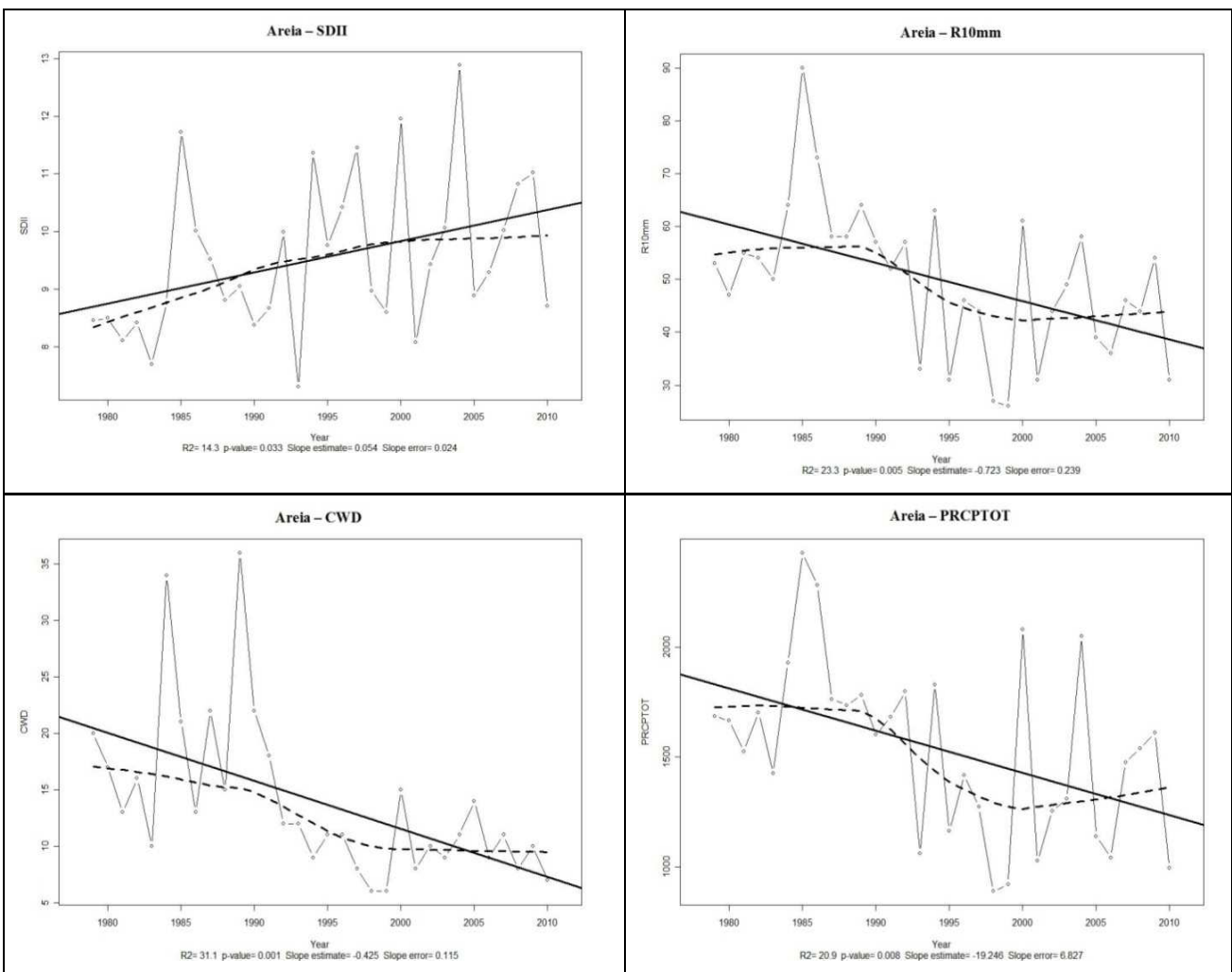
**Figura 4.** Correlações entre dados diários observados e imputados (azul) e acumulados mensais dos dados observados e imputados (vermelho).

### 3.2 Análises dos Índices climáticos

Os índices de extremos climáticos foram calculados para as seis cidades do estado da Paraíba (PB): Areia, Campina Grande, João Pessoa, Monteiro, Patos e São Gonçalo. Dentre os onze índices analisados, quatro deles não apresentaram tendências estatisticamente significativas para nenhuma das cidades, foram eles: RX1day, RX5day, R1mm e CDD.

A Figura 5 apresenta os índices que se mostraram estatisticamente significativos para Areia. Nesta localidade houve tendência

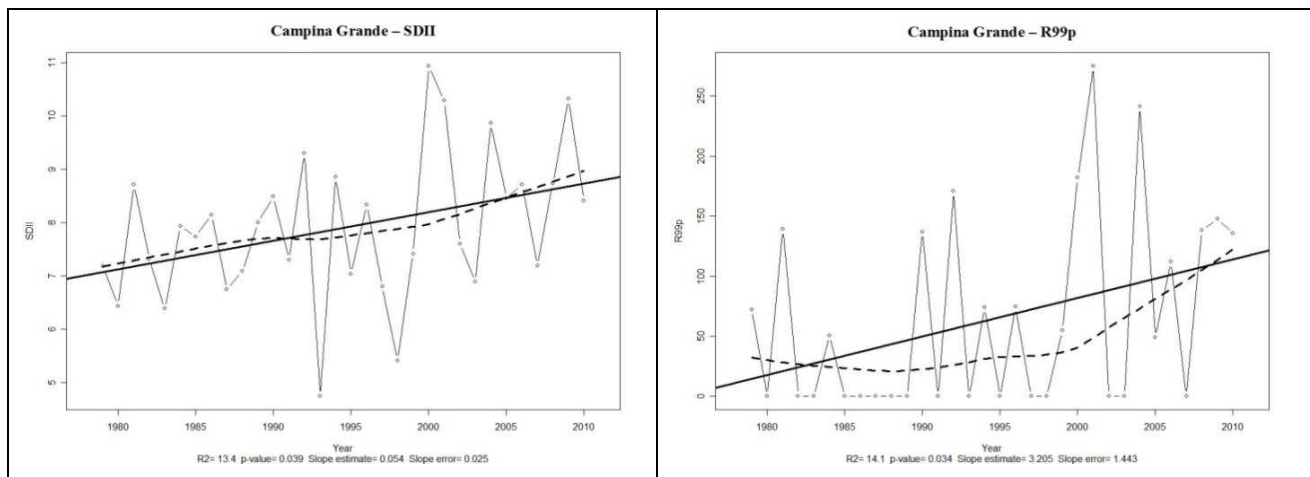
positiva para o índice SDII, que representa a média de precipitação em relação aos dias verificados com chuva no ano. Isto pode estar associado ao fato das chuvas estarem mais concentradas em menos dias do ano. Adicionalmente, esta localidade apresentou tendências negativas para os índices R10mm, CWD e PRCPTOT, ou seja, entre 1979 e 2010 houve redução gradativa do número de dias com precipitações maiores ou iguais a 20mm, redução na quantidade de dias seguidos com chuvas e consequentemente, redução na quantidade total de chuva anual.



**Figura 5.** Índices climáticos com tendência linear estatisticamente significativa ( $P\text{-valor} \leq 0,05$ ), para a cidade de Areia. (a)SDII, (b) R10mm, (c) CWD e (d) PRCPTOT.

Em Campina Grande, apenas dois índices mostraram tendências com significância estatística, o índice SDII, e o índice R99p, que mostra o total anual de precipitação que excedeu o percentil de 95%

da distribuição de precipitação diária. Tais tendências foram positivas mostrando que a quantidade de chuvas intensas no ano em Campina Grande se tornou mais frequente entre 1979 e 2010, ilustradas na Figura 6.

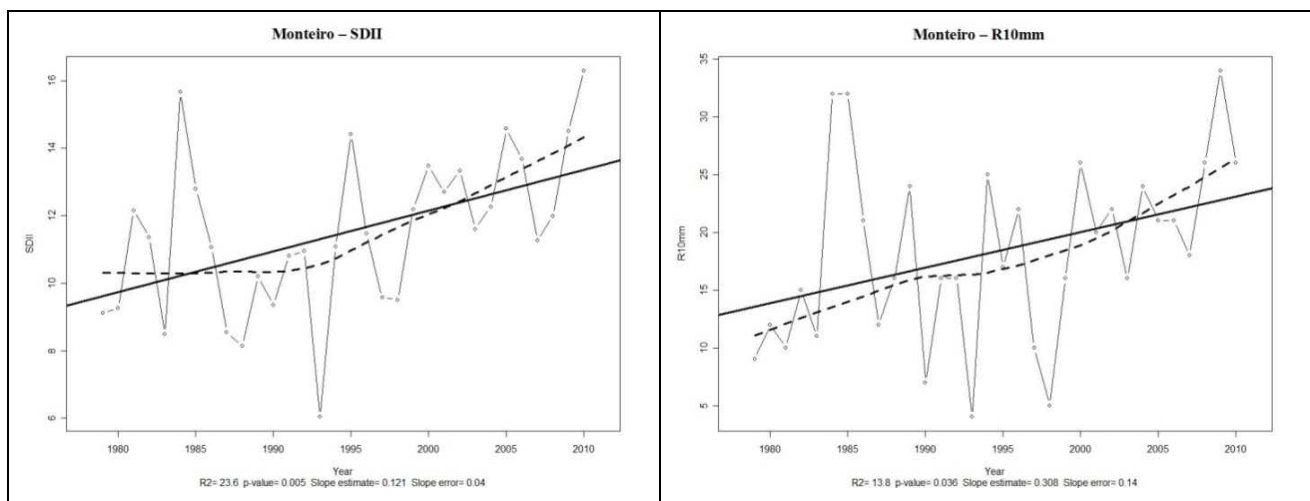


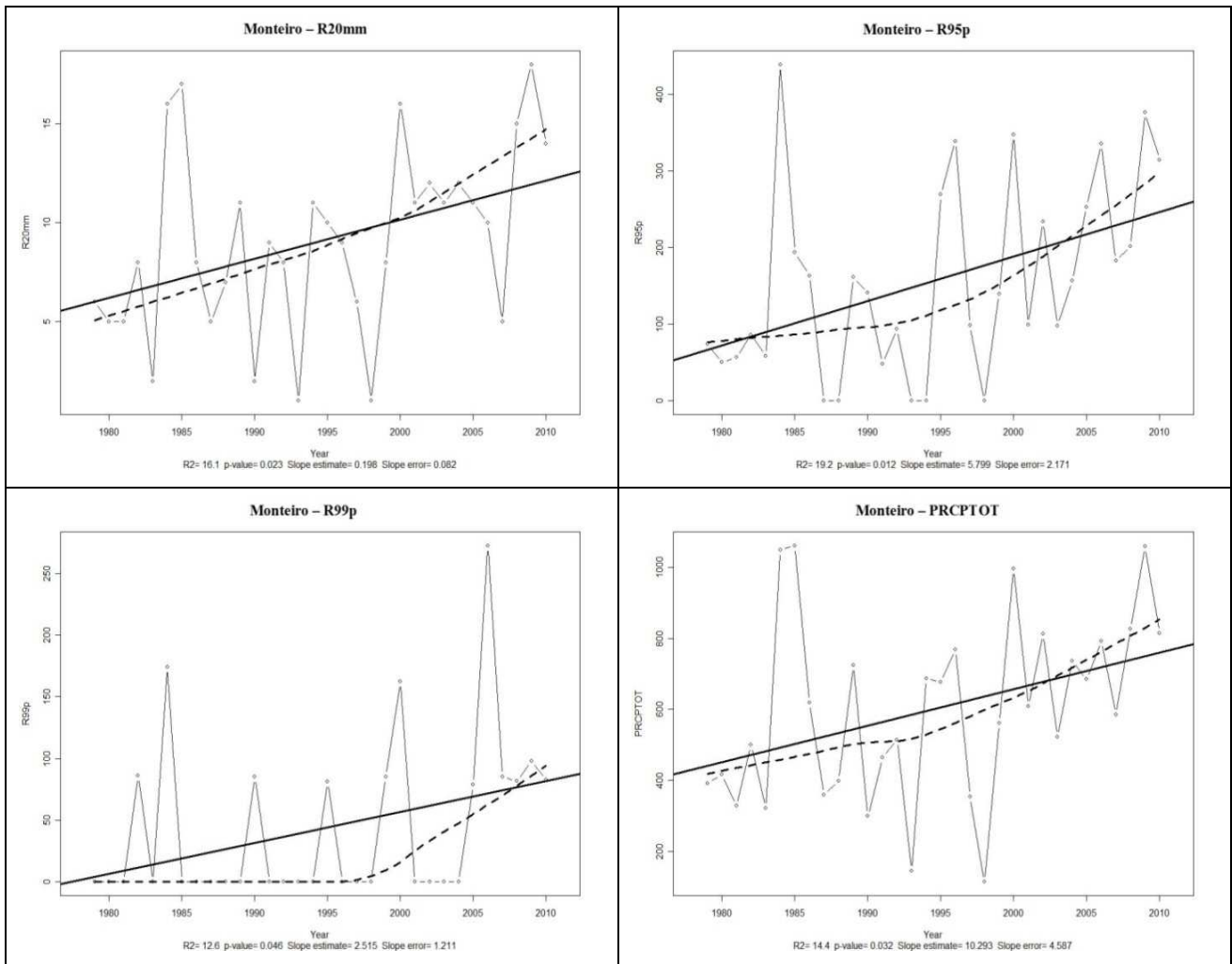
**Figura 6.** Índices climáticos com tendência linear estatisticamente significativa ( $P\text{-valor} \leq 0,05$ ), para a cidade de Campina Grande. (a) SDII, (b) R99p.

A Figura 7 evidencia que a cidade de Monteiro foi a que apresentou a maior quantidade de índices com tendências que mostraram significância estatística. Todas as tendências observadas foram positivas. Os índices observados que se mostraram estatisticamente significantes foram: SDII, R10mm, R20mm, R95p, R99p e PRCPTOT.

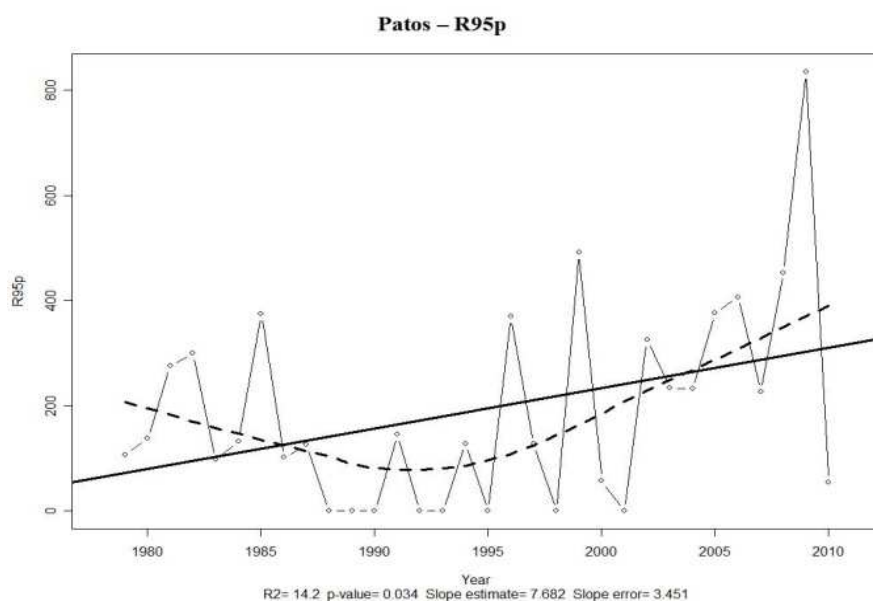
tendência positiva e estatisticamente significativa, sugerindo aumento no número de eventos extremos de precipitação. Em São Gonçalo, foram estatisticamente significantes, o R20mm e o PRCPTOT, indicando aumento no número de casos anuais com chuvas superiores a 20 mm diários. As Figuras 8 e 9 mostram estes resultados para estas cidades.

Em Patos, apenas o índice R95p teve

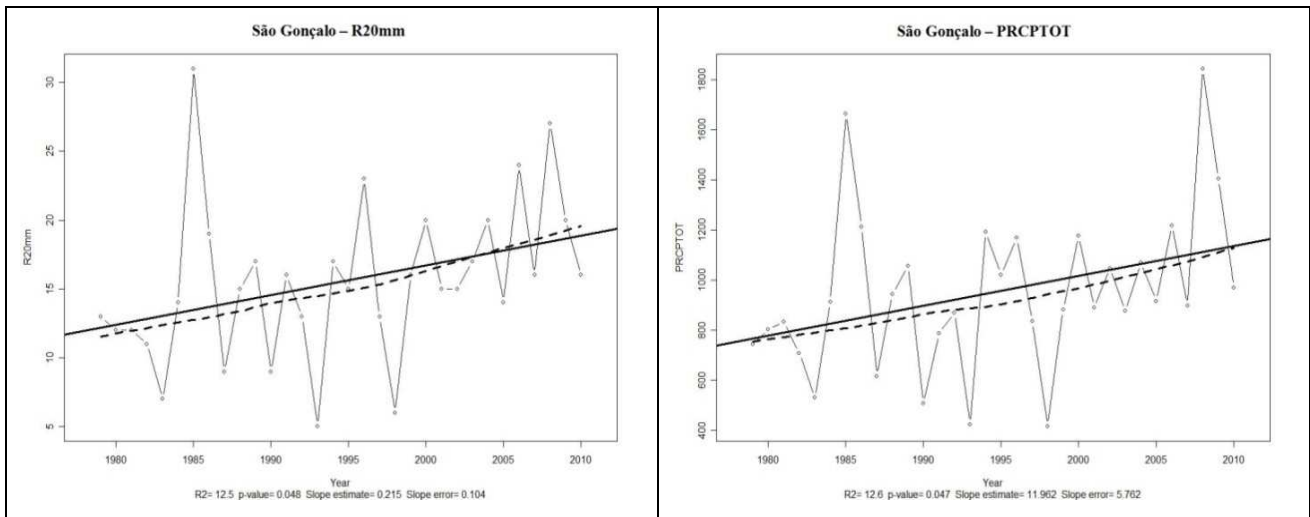




**Figura 7.** Índices climáticos com tendência linear estatisticamente significativa ( $P\text{-valor} \leq 0,05$ ), para a cidade de Monteiro. (a) SDII, (b) R10mm, (c) R20mm, (d) R95p, (e) R99p e (f) PRCPTOT.



**Figura 8.** Índice climático com tendência linear estatisticamente significativa ( $P\text{-valor} \leq 0,05$ ), para a cidade de Patos: R95p.



**Figura 9.** Índices climáticos com tendência linear estatisticamente significativa ( $P$ -valor  $\leq 0,05$ ), para a cidade de São Gonçalo. (a) R20mm e (b) PRCPTOT.

A única cidade que não apresentou nenhum índice com significância estatística foi João Pessoa, não apresentando qualquer indicação de mudanças no comportamento da precipitação.

#### 4. Conclusões

Os resultados da imputação de dados faltantes diários, de precipitação, foram bastante satisfatórios. Na técnica de imputação múltipla, que utiliza o pacote estatístico *MICE*, foram usados como preditores (variáveis explicativas) para a precipitação diária, dados de precipitação provenientes da análise gradeada do CPC/NOAA. Isto evitou problemas comuns quando se tenta imputar dados, do tipo: preditores também com dados incompletos; variáveis de origem ou natureza diferentes; relações não lineares entre os dados usados para geração do preenchimento de falhas, e geração de imputações sem sentido oriundas de combinações impossíveis que destroem

relações determinísticas entre os dados. Na comparação entre valores observados e imputados, nas seis cidades escolhidas, os histogramas de frequência se mostraram coincidentes, inclusive nas classes relativas aos maiores valores de precipitação. Os valores de correlações foram de moderadas a altas entre dados diários observados e imputados, e muito altos quando comparados aos acumulados mensais. Tais resultados corroboram a possibilidade de poder usar sem maiores problemas valores oriundos da imputação onde originalmente não houve dados observados.

Areia, Campina Grande e Monteiro apresentaram tendência de aumento de dias consecutivos secos, o que não ocorreu em João Pessoa, Patos e São Gonçalo, que apresentaram tendências de diminuição. Para dias consecutivos úmidos, observou-se que, apenas Patos e São Gonçalo apresentaram uma tendência de aumento. Entretanto, esse índice apresentou-se estatisticamente

significativo apenas para a cidade de Areia. O índice SDII, com significância estatística, apresentou uma tendência de aumento em todas as cidades estudadas. Em relação ao Índice de Precipitação Total Anual (PRCPTOT), esse se mostrou estatisticamente significativo apenas em Areia, Monteiro e São Gonçalo e tendência de aumento da Precipitação Total Anual em todas as cidades, com exceção de Areia, na qual esta apresentou uma tendência de diminuição. Observou-se também que todas as cidades estudadas, embora só tenham passado no teste de significância em alguns índices, apresentaram uma tendência de aumento, com exceção apenas de Areia, que apresentou tendências de diminuição nos seguintes índices: R10mm, CWD e PRCPTOT.

## 5. Agradecimentos

Agradecemos ao Instituto Nacional de Pesquisas Espaciais, ao Instituto Nacional de Meteorologia, ao Instituto Nacional de Metrologia e a Universidade Federal do Rio Grande do Norte pelo apoio ao desenvolvimento desta pesquisa.

## 6. Referências

Chen, M.; Shi, W.; Xie, P.; Silva, V. B. S.; Kousky, V. E.; Wayne Higgins, R.; and Janowiak, J. E. (2008). Assessing objective techniques for gauge-based analyses of global daily precipitation.

Finke, R.; Adamczyk, A. (2008). Cross-

National Moral Beliefs: The Influence of National Religious Context. *Sociological Quarterly*, 49(4), 617–652.

Gandin, L. S. (1965). *Objective Analysis of Meteorological Fields*. Israel Program for Scientific Translation, Jerusalem. 242 pp.

Hastenrath, S. (1984). Interannual variability and annual cycle: mechanisms of circulation and climate in the tropical Atlantic. *Monthly Weather Review*, v. 112, 1097-1107.

Hastenrath, S.; Greischar, L. (1993). Further Work on the Prediction of Northeast Brazil Rainfall Anomalies. *Journal of Climate*, v. 6, p. 743-758.

Kousky, V. E. (1979). Frontal influences on Northeast Brazil. *Monthly Weather Review*, v. 107, p. 1140-1153.

Kousky, V. E.; GAN, M. A. (1981). Upper Tropospheric Cyclonic Vortices in the Tropical South Atlantic. *Tellus*, v.33, p. 538-551.

Kousky, V. E.; Kayano, M. T. (1994). Principal modes of outgoing longwave radiation and 250 mb circulation for the South American sector. *Journal of Climate*, 7, 1131-1143.

Lucio, P. S.; Conde, F. C.; Cavalcanti, I. F. (2007). Spatiotemporal monthly rainfall reconstruction via Artificial Neural Network. Case study: south of Brazil. *Advances in*

Geosciences, v. 10, p. 67-76.

McClelland, R. L.; Kronmal, R. A.; Haessler, J.; Blumenthal, R. S.; Goff, J. (2008). Estimation of Risk Factor Associations when the Response Is Influenced by Medication Use: An Imputation Approach. *Statistics in Medicine*, 27(24), 5039–5053.

Moura, A. D.; Shukla, J. (1981). On the dynamics of droughts in northeast Brazil: Observations, theory and numerical experiments with a general circulation model. *Journal of the Atmospheric Science*, 38, 2653-2675.

Salvador, M. A.; Santos, L. S. F. C. (2010). Análise da Precipitação na Cidade de São Paulo no Período 1961-2009. *Revista Brasileira de Climatologia*, v. 7, p. 8-19.

Santos, C. A. C. (2006). Estimativas e Tendências de Índices de Detecção de Mudanças Climáticas com base na precipitação diária no Rio Grande do Norte e na Paraíba, 98 p., Dissertação (Mestrado em Meteorologia), Universidade Federal de Campina Grande, Campina Grande, 2006.

Silva, V. B. S.; Kousky, V. E.; Shi, W. and Higgins, R. W. (2007). An improved gridded historical daily precipitation analysis for

Brazil. *Journal of Hydrometeorology*, 8, 847–861.

Sundell, K.; Hansson, K.; L'ofholm, C. A.; Olsson, T.; Gustle, L. H.; Kadesjö, C. (2008). The Transportability of Multisystemic Therapy to Sweden: Short-Term Results from a Randomized Trial of Conduct-Disordered Youths. *Journal of Family Psychology*, 22 (4), 550–560.

Tanasoiu, C.; Colonescu, C. (2008). Determinants of Support for European Integration: The Case of Bulgaria. *European Union Politics*, 9 (3), 363–377.

Ton, T. G. N.; Longstreth, J.; Koepsell, T. (2009). Active and Passive Smoking and Risk of Narcolepsy in People with Hla Dqb1\*0602: A Population-Based Case-Control Study. *Neuroepidemiology*, 32 (2), 114–121.

Van Buuren, S.; Groothuis-Oudshoorn, K. (2011). MICE: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, 45(3), 1-67. <http://www.jstatsoft.org/v45/i03/>

Zhang, X.; Yang, F. (2004). R-ClimDex (1.0) User Guide. Climate Research Branch Environment Canada. Downsview (Ontario, Canada), 22p.