Revista Brasileira de Climatologia
ISSN: 2237-8642 (Eletrônica)

APLICAÇÃO DE TÉCNICAS DE REPRESENTAÇÃO DE DADOS NO TRATAMENTO DE DADOS MICROCLIMÁTICOS

CRUZ, Johnnes Santos Veiga da – johnnes.santos7@gmail.com
Universidade Federal do Mato Grosso / UFMT

VENTURA, Thiago Meirelles – thiago@ic.ufmt.br
Universidade Federal do Mato Grosso / UFMT

OLIVEIRA, Allan Gonçalves de – allan@ic.ufmt.br
Universidade Federal do Mato Grosso / UFMT

RODRIGUES, Thiago Rangel – thiago.r.rodrigues@ufms.br
Universidade Federal do Mato Grosso do Sul / UFMS

FIGUEIREDO, Josiel Maimone de – josiel@ic.ufmt.br
Universidade Federal do Mato Grosso / UFMT

GOMES, Raphael de Souza Rosa – raphael@ic.ufmt.br
Universidade Federal do Mato Grosso / UFMT

NOGUEIRA, José de Souza –nogueira@ufmt.br
Universidade Federal do Mato Grosso / UFMT

RESUMO: Os sensores instalados em estações meteorológicas podem gerar uma grande quantidade de dados todos os dias. Esses dados, eventualmente, precisam ser processados antes de serem analisados. Nesse processo, a quantidade de dados pode dificultar a sua manipulação, controle e interpretação. Inclusive para técnicas computacionais, a quantidade de dados são consequências de um maior tempo de processamento. Dessa forma, espera-se que reduzir a quantidade de dados possa facilitar, do ponto de vista computacional, a manipulação desse tipo de dado. Por outro lado, a redução pode gerar perda de informação importante na série de dados. O objetivo deste trabalho foi avaliar se a utilização de técnicas de representação de dados podem reduzir a série de dados ao mesmo tempo que as características da série fossem conservadas. Testes foram realizados para avaliar a conservação das características da série de dados após a redução. Mesmo com a série reduzida em 80%, operações de tratamento de dados microclimáticos como preenchimento de falhas e detecção de outliers puderam ser realizadas com a mesma eficácia do que se fossem realizadas com a série original, indicando que as características originais da série foram mantidas. As técnicas de representação de dados com melhor desempenho foram Amostragem e Perceptually Important Points.

Palavras-chave: Amostragem, Piecewise Aggregate Approximation, Perceptually Important Points, preenchimento de falhas, detecção de outliers..

FIXING MICROCLIMATE DATA ASSISTED BY DATA REPRESENTATION TECHNIQUES.

ABSTRACT: The sensors installed in meteorological stations might generate a large amount of data every day. In some cases, it is necessary to process this data before process them. In this process, the amount of data may impose difficulties to analyze, control and interpret this data. Even for computational techniques, the amount of data is a consequence of a longer processing time. We hope that reducing the amount of data, it would facilitate the processing of microclimate data and its interpretations. However, reducing such data can leads to loss of information. In this work we evaluated the use of data representation techniques to reduce the dataset while preserving the characteristics

of the series. Tests were performed to evaluate the conservation of the dataset's characteristics after the reduction. Even with the series reduced in 80%, operations to fix microclimate data (e.g. gap filling and outlier detection) could be performed with the same effectiveness as if they were performed with the original series. The best performance data representation techniques were Sampling and Perceptually Important Points.

Keywords: Sampling, Piecewise Aggregate Approximation, Perceptually Important Points, gap filling, outlier detection

1. INTRODUÇÃO

Dados microclimáticos são fundamentais para o entendimento de fenômenos ambientais e climáticos, que por sua vez, auxiliam na compreensão dos processos físicos de ordem natural que afetam diretamente a vida em nosso planeta, como mostrado em Rodrigues et al. (2017), Alves (2017) e Querino et al. (2017). Ter a capacidade de processar dados microclimáticos significa fornecer informações úteis para pesquisadores e gestores tomarem decisões que afetam a nossa sociedade.

Todo esse processo é iniciado com a captação dos dados, normalmente feito de maneira automatizada nas estações micrometeorológicas. Os dados gerados nessa coleta são denominados dados de série temporal, pois são representados por meio de uma sequência ordenada de valores na qual cada um corresponde ao valor observado da variável em um determinado instante de tempo. As séries temporais são uma sequência de observações de uma variável ao longo do tempo, tomados em intervalos regulares durante um período de tempo (WOOLDRIDGE, 2000). Com o uso de estações automatizadas, o intervalo de tempo de coleta pôde se tornar menor e a coleta ser realizada o dia todo, assim, acarretando em uma grande quantidade de dados.

Em Guo (2016) é descrito que a grande quantidade de dados gerados pode levar a dificuldades de manipulação, controle e interpretação dos mesmos, tanto por limitações das técnicas utilizadas na análise ou limitações humanas em manipular a massa de dados disponíveis. Portanto, seria de grande utilidade reduzir a quantidade de dados para facilitar o processamento de dados microclimáticos, desde que as característica principais da série de dados fossem mantidas. Portanto, o desafio desta operação é manter as mesmas informações mesmo após a redução da quantidade de dados. Isso pode ser alcançado com técnicas de representação de dados.

As técnicas de representação de dados normalmente são aplicadas para facilitar a visualização dos dados. Em Fu (2011) são mostradas diferentes técnicas com este fim, no qual é comentado que a principal razão de se ter diferentes técnicas de representação de séries temporais é a redução do número de pontos das séries ou a redução de dimensionalidade.

Este trabalho propõe a utilização das técnicas de representação de dados como uma forma de pré-processamento, a fim de reduzir a quantidade de dados de entrada, agilizando modelagens e tentando manter a mesma precisão quando comparado ao processamento da totalidade dos dados. Assim, o objetivo deste trabalho é fornecer mecanismos para facilitar o processamento de dados microclimáticos, principalmente visando o tratamento dos mesmos, utilizando uma quantidade menor de dados no processamento.

ISSN: 2237-8642 (Eletrônica)

2. MATERIAL E MÉTODOS

A grande quantidade de dados microclimáticos dificulta o processamento de métodos da área. Portando, torna-se necessário o uso de métodos que diminuam a quantidade de dados sem que haja perda da sua representatividade. Esling e Agon (2012) destaca que neste tipo de procedimento deve-se buscar destacar características fundamentais da série a ser analisada. Neste trabalho serão abordadas diferentes técnicas, a fim de avaliar qual tem melhor desempenho com as características dos dados microclimáticos.

2.1 DADOS UTILIZADOS

Diferentes séries de dados foram utilizadas para avaliar as técnicas de representação de dados. Os dados utilizados foram coletados do Banco de Dados Microclimáticos para Ensino e Pesquisa (BDMEP) do INMET. No total foram utilizados 126.414 linhas de dados dos 27 estados brasileiros dos anos de 2011 até o ano de 2016, que contabilizam 5 anos de dados coletados.

Por padrão foram utilizados somente uma cidade de cada estado brasileiro, sendo preferencialmente a capital. Somente o estado de Rondônia não continha uma estação no período selecionado, sendo utilizado a base LÁBREA–AM para fazer parte da amostra por questão de proximidade ao estado de Rondônia.

Uma linha de dados nessas séries é composta por 8 colunas: dia, ano, hora, TempBulboSeco, TempBulboUmido, UmidadeRelativa, PressaoAtmEstacao e Nebulosidade. A Tabela 1 exemplifica os dados coletados no BDMEP.

Tabela 1 - Amostra de dados obtidos do BDMEP

Dia	Ano	Hora	TempBulboSeco	TempBulboUmido	UmidadeRelativa	PressaoAtmEstacao	Nebulosidade
1	2011	0	25,0	23,3	87	988,8	10
1	2011	1200	24,2	23,7	97	991,1	9
1	2011	1800	31,0	26,8	71	988,7	10
2	2011	0	25,2	23,1	83	989,9	10
2	2011	1200	23,3	23	97	992,7	10
2	2011	1800	28,3	25,2	77	989,8	10
3	2011	0	26,6	25,4	91	988	6
3	2011	1200	24,0	23,2	93	990,6	10
3	2011	1800	30,0	25,9	71	989,4	6

2.2 TÉCNICAS DE REPRESENTAÇÃO DE DADOS

As seguintes técnicas de representação de dados foram testadas: Amostragem, descrita em Astrom (1969); Piecewise Aggregate Approximation

(PAA) de Keogh et al. (2001); Segmented Sum of Variation (SSV) apresentada em Lee et al. (2003); Perceptually Important Points (PIP) que pode ser observada em Chung (2001); Random de desenvolvimento próprio para fins de comparação.

2.2.1 AMOSTRAGEM

A técnica de amostragem consiste em reduzir a dimensionalidade dos dados de tamanho *n* em espaços de tamanho *h*, isto é, determinar quais são os dados que farão parte da amostra final considerando que os dados são identificados de 0 até *n-1*, somente os múltiplos de *h*.

A taxa de redução é determinada por um fator de *m/n* onde *m* é a quantidade de valores múltiplos de *h* até *n-1*. O valor de *m* é determinado pela função $m = f(x) = 1 + (n-1)/h$.

A Figura 1 demonstra três gráficos de forma a comparar o resultado da aplicação da técnica da amostragem em uma série de dados de tamanho 30, com taxa de redução de 50% e 20% dos dados. Neste exemplo, pode ser observado semelhanças na geração com 50% dos dados, mas o mesmo não pode ser dito na geração com apenas 20%.

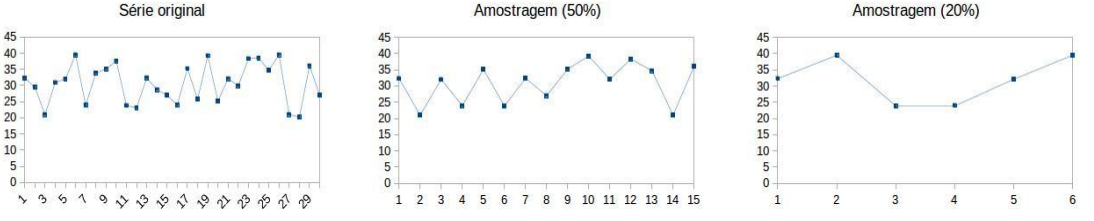

Figura 1 - Comparação da série original com o resultado da aplicação da técnica de amostragem com 50% e 20%.

2.2.2 PIECEWISE AGGREGATE APPROXIMATION (PAA)

A técnica PAA consiste em dividir a série temporal em segmentos e a partir desses segmentos calcular a média para determinar os pontos que farão parte da amostra. O conjunto original de tamanho n é dividido em N segmentos. Assim, a taxa de redução é dada por N/n, no qual o tamanho de cada segmento N é determinado por n/N.

A Figura 2 mostra uma característica importante da técnica PAA, na qual ela se comporta de forma com que o ponto gerado pela redução se mantém equivalente ao segmento. Desta forma, os gráficos tendem a terem menos picos em um curto intervalo de dados.

ISSN: 2237-8642 (Eletrônica)

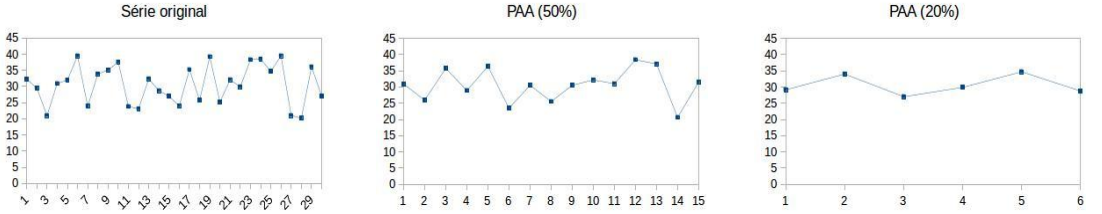

Figura 2 - Comparação da série original com o resultado da aplicação da técnica PAA com 50% e 20%.

2.2.3 SEGMENTED SUM OF VARIATION (SSV)

A técnica SSV utiliza o mesmo conceito do PPA que divide a série original em segmentos, porém, utiliza o conceito de conectividade de segmentos, onde cada segmento atual é conectado ao anterior através do último ponto do segmento anterior. Ao formar esses segmentos calcula-se a variação de cada segmento para formar o conjunto de dados reduzido.

A técnica SSV consegue manter o formato da série, porém, é pouco efetivo quando é necessário manter as faixas de valores medidos. A Figura 3 mostra de forma visível esse efeito da técnica. É possível observar que a faixa de valores da série original está entre [20,40] e na redução de 50% foi de [4,25) e de 20% foi de (14,44).

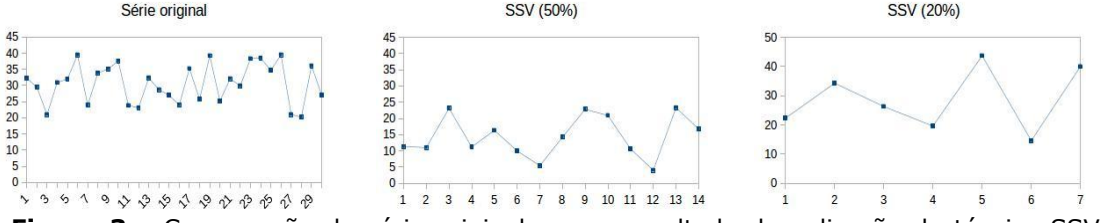

Figura 3 - Comparação da série original com o resultado da aplicação da técnica SSV com 50% e 20%.

2.2.4 PERCEPTUALLY IMPORTANT POINTS (PIP)

A técnica PIP consiste em selecionar os pontos mais importantes da série original utilizando a estratégia de cálculo da distância de um ponto a uma reta, que é formada por dois pontos vizinhos à ele que já foram selecionados para fazer parte do conjunto de dados final.

A técnica inicia com dois pontos: o primeiro e o último ponto da série original. A seguir, é feito um processo iterativo de cálculo da distância de cada ponto e, por fim, é selecionado o ponto que contém a maior distancia. Existem diferentes formas de calcular as distâncias, mas em Kaiser (2010) é descrito que a distância vertical apresenta os melhores resultados.

O conjunto de dados original de tamanho n é reduzido a um conjunto de tamanho N pré-determinado, tornando a taxa de redução da técnica dada por N/n. A Figura 4 mostra o comportamento da técnica e é possível verificar que a vantagem da técnica está em identificar os picos da série.

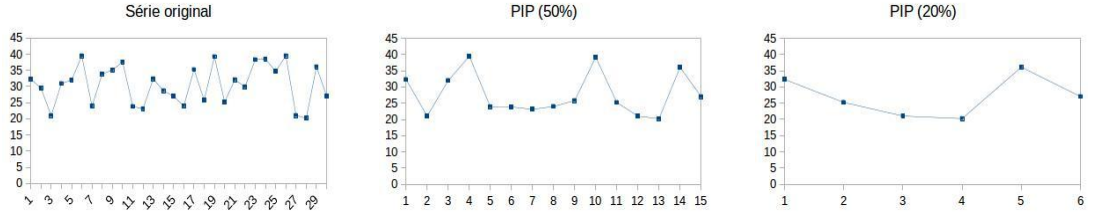

Figura 4 - Comparação da série original com o resultado da aplicação da técnica PIP com 50% e 20%.

2.2.5 RANDOM

A técnica Random consiste em enumerar o conjunto de dados original e selecionar os dados que farão parte do conjunto de dados final de forma aleatória, mantendo a sua ordem.

O conjunto de dados original de tamanho n é reduzido em um valor N que é pré-determinado. São selecionados aleatoriamente N dados para formar o conjunto de dados final. Portanto, a sua taxa de redução se dá por N/n.

Pela característica da técnica os resultados gerados podem ser completamente diferentes a cada execução. A Figura 5 exemplifica seu comportamento. Esta técnica foi desenvolvida para efeito de comparação com as demais.

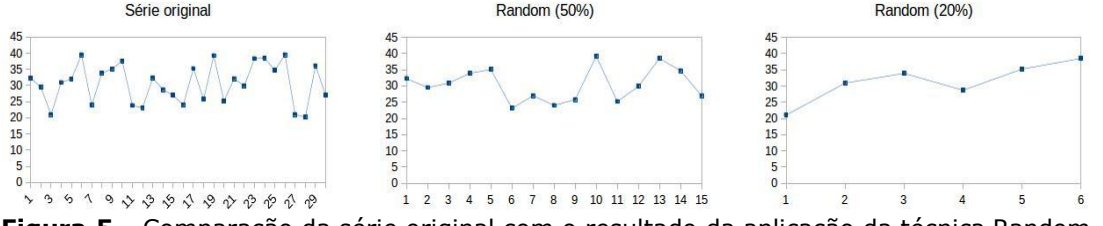

Figura 5 - Comparação da série original com o resultado da aplicação da técnica Random com 50% e 20%.

2.3 TESTES DE DESEMPENHO

Para avaliar o desempenho das técnicas de representação de dados foram preparados cenários de tratamento de dados microclimáticos, especificamente preenchimento de falhas e detecção de outliers, operações comumente utilizadas com dados microclimáticos (VENTURA et al., 2016; OSCAR JÚNIOR, 2015). Para preenchimento de falhas foi selecionado o método Multiple Linear Regression (MLR) e para detecção de outliers foi utilizado o método Z-score.

A ideia principal dos testes consiste em treinar os métodos de preenchimento de falhas e detecção de outliers apenas com a representação dos dados, e não com a sua totalidade, facilitando o processamento. A precisão dos métodos treinados com as representações de dados pode ser comparada com a precisão dos métodos treinados com a totalidade dos dados, como é feito tradicionalmente.

ISSN: 2237-8642 (Eletrônica)

Os testes foram aplicados para cada série de dados obtida. Os resultados foram analisados com métodos estatísticos. Para os resultados de preenchimento de falhas serão utilizadas as medidas do Erro Médio Absoluto (EMA) e Correlação (r^2), indicado por Willmott et al. (2009). O EMA consiste em calcular a média absoluta da diferença entre duas séries de dados e o r^2 em calcular um valor no intervalo de 0 a 1 onde quanto mais próximo de 1 o valor calculado mais as séries de dados estão relacionadas.

Para a detecção de outliers são utilizadas as medidas de precisão e acurácia. A precisão consiste em uma métrica determinada pelo fator de H/N, onde H é o número de detecções acertadas pelo método de detecção de outliers e N é o número de dados com outliers contidos na série temporal. Já a acurácia é determinada por (H – I) / N, onde I é o numero de total de dados detectados como outliers pelo método.

3. RESULTADOS E DISCUSSÃO

Foram realizadas simulações de falhas e outliers em todas as séries de dados descritas na seção 2.1. Após esta etapa, as técnicas de representação de dados da seção 2.2 foram utilizadas para gerar representações da série original, em um formato reduzido, formando conjuntos de dados até 87% menores que as séries originais.

Em seguida, os métodos de preenchimento de falhas e detecção de outliers da seção 2.3 foram treinados com tais representações de maneira mais rápida, uma vez que haviam menos dados para serem processados. Por fim, depois do preenchimento das falhas e da detecção dos outliers, foi possível medir o desempenho de cada ação de tratamento utilizando apenas os dados reduzidos.

A Tabela 2 mostra os resultados do preenchimento de falhas baseando-se nas representações criadas com cada uma das técnicas mencionadas. Também está sendo mostrado o resultado do preenchimento caso o treinamento do método de preenchimento de falhas fosse realizado com os dados originais (100% dos dados).

No preenchimento de falhas para a variável de temperatura de bulbo seco, o EMA com os dados originais foi de 0,24, enquanto o r^2 foi de 0,95. O mesmo resultado foi alcançado utilizando a técnica de Amostragem e o PAA. Ou seja, apenas a representação da série de dados foi suficiente para fornecer as mesmas informações e conseguir estimar valores para preencher as falhas. Resultados semelhantes podem ser vistos para as outras variáveis.

ISSN: 2237-8642 (Eletrônica)

Tabela 2 - Resultados da aplicação das técnicas de representação de dados em 5 variáveis climáticas para dados com métodos de preenchimento de falhas.

Técnica	TempBulboSeco		TempBulboUmido		UmidadeRelativa		PressaoAtmEstacao		Nebulosidade	
	EMA	r^2	EMA	r^2	EMA	r^2	EMA	r^2	EMA	r^2
Série original	0,24	0,95	0,23	0,91	1,54	0,93	1,63	0,40	1,97	0,31
Amostragem	0,24	0,95	0,23	0,91	1,59	0,93	1,66	0,40	2,02	0,30
PAA	0,24	0,95	0,24	0,91	1,59	0,93	1,74	0,40	2,06	0,31
SSV	173435,04	0,02	188129,47	0,06	1047819,28	0,06	558319,10	0,02	536634,26	0,01
PIP	2,48	0,39	1,64	0,25	12,44	0,19	2,63	0,12	2,81	0,08
Random	1,87	0,55	1,05	0,45	9,63	0,29	1,88	0,28	2,46	0,10

Uma outra observação importante é com relação ao método Random, que seleciona aleatoriamente dados da série para compor a respectiva representação. Os resultados mostram que esta técnica gerou um conjunto de treinamento ruim para o método de preenchimento de falhas. Isso mostra a importância de uma boa representação, não sendo suficiente a escolha de qualquer dado da série para representar a variável microclimática.

A ação de detecção de outliers também foi realizada com a série de dados original e as séries geradas pelas técnicas de representação de dados. Os resultados podem ser vistos na Tabela 3.

Tabela 3 - Resultados da aplicação das técnicas de representação de dados em 5 variáveis climáticas para dados com métodos de detecção de outliers.

Técnica	TempBulboSeco		TempBulboUmido		UmidadeRelativa		PressaoAtmEstacao		Nebulosidade	
	Precisão	Acurácia	Precisão	Acurácia	Precisão	Acurácia	Precisão	Acurácia	Precisão	Acurácia
Série original	0,40	0,94	0,59	0,94	0,25	0,93	0,96	0,96	0,07	0,92
Amostragem	0,40	0,93	0,59	0,94	0,25	0,93	0,96	0,96	0,07	0,92
PAA	0,40	0,93	0,59	0,94	0,25	0,93	0,96	0,96	0,07	0,92
SSV	0,28	0,93	0,36	0,93	0,13	0,92	0,49	0,94	0,02	0,92
PIP	0,42	0,94	0,60	0,94	0,27	0,93	0,96	0,96	0,09	0,92
Random	0,40	0,93	0,59	0,94	0,25	0,93	0,96	0,96	0,07	0,92

Do conjunto de dados existente, a variável mais fácil de prever outliers é a pressão atmosférica, devido a sua característica de estabilidade. Com a série original foi obtido 96% de precisão e acurácia. Os mesmos valores foram encontrados com as técnicas de Amostragem, PAA, PIP e, inclusive, Random.

O método Z-Score utiliza da média e do desvio padrão da série para realizar a detecção de outliers. Portanto, os dados utilizados para o seu treinamento não necessariamente precisam estar em ordem. Por esta razão vários métodos tiveram resultados semelhantes à série original, incluindo o Random.

A operação de detecção de outliers não é algo simples. É possível observar que em diversas variáveis a precisão foi menor que 60% usando a série original. De qualquer forma, os testes com dados gerados pelas técnicas de representação de dados tiverem desempenho semelhantes.

ISSN: 2237-8642 (Eletrônica)

Podem ser analisadas também as técnicas de maneira geral. A Tabela 4 e 5 mostram a média dos resultados em cada tipo de operação. O EMA não foi calculado por ser uma medida dependente da unidade da variável, então para o preenchimento de falhas foi calculado apenas o r^2, mantendo a precisão e acurácia para os testes de detecção de outliers.

Tabela 4 - Resultados obtidos do processamento das técnicas de representação para dados com métodos de preenchimento de falhas.

Técnica	r^2
Série original	0,70
Amostragem	0,70
PAA	0,69
SSV	0,03
PIP	0,20
Randon	0,33

Tabela 5 - Resultados obtidos do processamento das técnicas de representação para dados com métodos de detecção de outliers.

Técnica	Precisão	Acurácia
Série original	0,46	0,94
Amostragem	0,45	0,94
PAA	0,45	0,94
SSV	0,25	0,93
PIP	0,47	0,94
Randon	0,45	0,94

A técnica SSV não obteve bons resultados em nenhum teste, não sendo indicada para representar dados microclimáticos, principalmente por causa da sua característica de alterar a escala dos dados. Por outro lado, os resultados da Tabela 4 mostram que a Amostragem foi a técnica que gerou melhores representações das séries para fins de preenchimento de falhas. Para detecção de outliers a Tabela 5 mostra que a técnica PIP possibilitou o melhor desempenho na detecção de outliers.

Em alguns casos, ótimas estimativas como a apresentada na Figura 6 foram realizadas. Portanto, os resultados obtidos neste trabalho indicam que é possível aplicar técnicas de representação de dados para servir de pré-processamento para métodos de tratamento de dados microclimáticos.

ISSN: 2237-8642 (Eletrônica)

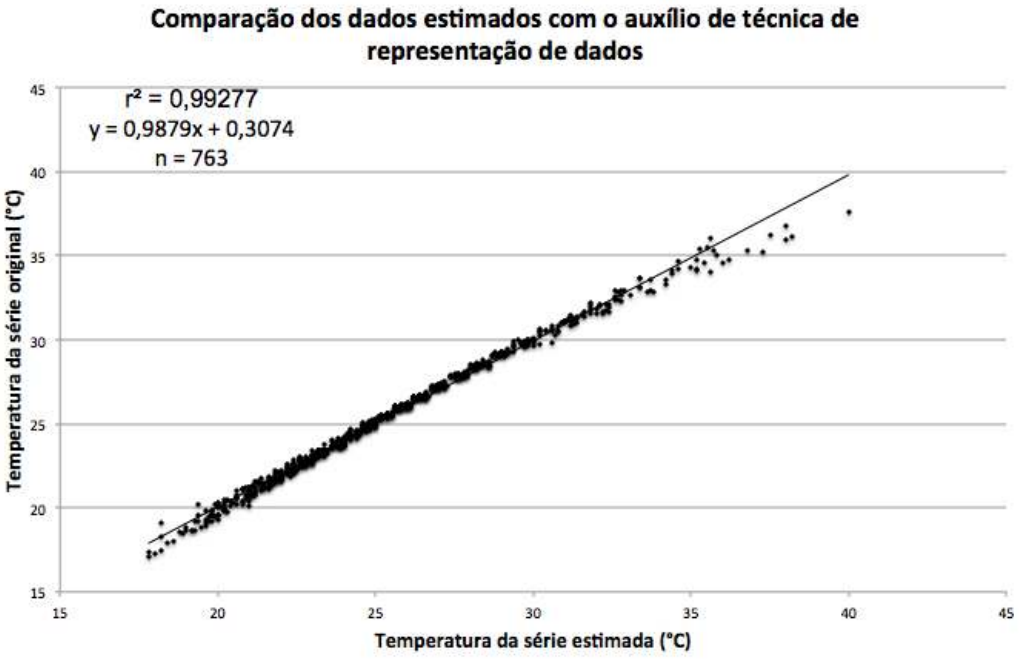

Figura 6 - Estimativa dos dados com falhas simuladas obtidos em Rio de Janeiro de temperatura de bulbo seco comparado ao dado original depois de aplicado o método de preenchimento de falhas que foi treinado apenas com a representação dos dados.

4. CONCLUSÕES

Este trabalho teve o intuito de avaliar a aplicabilidade de técnicas de representação de dados no processamento de métodos de tratamento de dados microclimáticos. Cinco técnicas de representação de dados foram avaliadas em métodos de preenchimento de falhas e detecção de outliers.

De forma geral, a técnica de Amostragem obteve maior sucesso no pré-processamento de dados para os métodos de preenchimento de falhas, enquanto o PIP foi em média a melhor técnica para os métodos de detecção de outliers. O mais importante é notar que o uso de tais técnicas não prejudicou o desempenho dos métodos de tratamento. Em contrapartida, apenas uma parcela da totalidade dos dados foi utilizada, exigindo um custo computacional inferior comparado ao processamento dos dados originais na totalidade.

Como exemplo em dados meteorológicos que exigem alto processamento, pode ser citado dados extraídas de imagens de satélite, como o LandSat, que possui necessita de 300 MB de armazenamento por amostra. Considerando sua periodicidade de 15 dias, a análise da série histórica de uma região por 10 anos resultaria em 72 GB de dados processados. Ao aplicarmos uma técnica de representação esse valor poderia ser diminuído para até 15 GB, economizando tanto em armazenamento quanto em processamento.

Como trabalhos futuros planeja-se experimentos com outras técnicas de representação de dados, assim como outros tipos de dados e modelagens climáticas. Por exemplo, dados de fluxos entre a superfície e a atmosfera são menos regulares e dependem de diversas outras variáveis, dificultando a modelagem (RODRIGUES et al., 2016). Técnicas como as de representação de

dados poderiam auxiliar no processamento dessas modelagens. De qualquer forma, este trabalho atesta a viabilidade do uso de técnicas de representação de dados em dados microclimáticos, tendo potencial para serem aplicadas em outras áreas de climatologia.

5. AGRADECIMENTOS

Os autores agradecem o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo auxílio financeiro no projeto nº 424915/2016-2 - Edital Universal MCTI/CNPq nº 01/2016.

6. REFERÊNCIAS BIBLIOGRÁFICAS

ALVES, E. D. L. Ilha de calor urbana em cidade de pequeno porte e a influência de variáveis geourbanas. Revista Brasileira de Climatologia. v. 20, p. 97-116, 2017.

ASTROM, K. On the choice of sampling rates in parametric identification of time series. Information Sciences, v. 1, n. 3, p. 273 – 278, 1969.

CHUNG F.L., F. T. L. R. N. V. Flexible time series pattern matching based on perceptually important points. In: International Joint Conference on Artificial Intelligence (IJCAI) Workshop on Learning from Temporal and Spatial Data. 2001.

ESLING, P.; AGON, C. Time-series data mining. ACM Comput. Surv., ACM,New York, NY, USA, v. 45, n. 1, p. 12:1–12:34, dez. 2012.

FU, T. C. A review on time series data mining. Engineering Applications of Artificial Intelligence, v. 24, n. 1, p. 164 – 181, 2011.

GUO, X. Application of meteorological big data. 16th International Symposium on Communications and Information Technologies (ISCIT), Qingdao, p. 273-279, 2016.

KAISER, Ronald Andreu; VIVACQUA, Adriana Santarosa. Aplicação do algoritmo Perceptually Important Points em séries temporais de datacenters. 2010.

KEOGH, E.; CHAKRABARTI, K.; PAZZANI, M.; MEHROTRA, S. Dimensionality reduction for fast similarity search in large time series databases. Knowledge and Information Systems, Springer-Verlag London Limited, v. 3, n. 3, p. 263–286, 2001.

LEE, S.; KWON, D.; LEE, S. Dimensionality reduction for indexing time series based on the minimum distance. Journal of Information Science and Engineering, v. 19, n. 4, p. 697-711, 2003.

OSCAR JÚNIOR, A. C. S. Extremos atmosféricos e desastres hidrometeorológicos em Duque de Caxias (RJ). Revista Brasileira de Climatologia, v. 17, p. 189-205, 2015.

QUERINO, K. A. S.; BIUDES, M. S.; MACHADO, N. G.; QUERINO, J. K. A. S.; SANTOS NETO, L. A.; SILVA, M. J. G.; ARRUDA, P. H. Z.; NOGUEIRA, J. S. Balanço de ondas curtas sobre floresta sazonalmente alagável do Pantanal Mato-Grossense. Revista Brasileira de Climatologia. v. 20, p. 252-266, 2017.

RODRIGUES, M.; FRANCISCO, A. D.; REBOITA, M. S. Temperatura aparente de superfície na bacia hidrográfica do rio Tavares, Florianópolis. Revista Brasileira de Climatologia. v. 21, p. 1-19, 2017.

RODRIGUES, T. R.; VOURLITIS, G. L.; LOBO, F. A.; SANTANNA, F. B.; ARRUDA, T. R.; NOGUEIRA, J. S. Modeling canopy conductance under contrasting seasonal conditions for a tropical savanna ecosystem of south central Mato Grosso, Brazil. Agricultural and Forest Meteorology, v. 218–219, p. 218-229, 2016.

VENTURA, T. M.; SANTANA, L. L. R.; MARTINS, C. A.; FIGUEIREDO, J. M. Análise da aplicabilidade de métodos estatísticos para preenchimento de falhas em dados meteorológicos. Revista Brasileira de Climatologia, v. 19, p. 168-177, 2016.

WILLMOTT, C. J.; MATSUURA, K.; ROBESON, S. M. Ambiguities inherent in sums-of-squares-based error statistics. Atmospheric Environment, v. 43, n. 3, p. 749–752, 2009.

WOOLDRIDGE, J. M. Introductory econometrics: a Modern Approach. South–Western College Publishing, a division of Thomson Learning, 2000