

USO DO ALGORITMO EXPECTATION-MAXIMIZATION (EM) PARA A ANÁLISE DA VARIABILIDADE PLUVIOMÉTRICA TEMPORAL NA BACIA DO RIO PIQUIRI



USO DO ALGORITMO EXPECTATION-MAXIMIZATION (EM) PARA A ANÁLISE DA VARIABILIDADE PLUVIOMÉTRICA TEMPORAL NA BACIA DO RIO PIQUIRI

Analysis of the temporary pluviometric variability of the piquiri river basin under the extreme events using data mining techniques

Ana Claudia Guedes Silva
Universidade Estadual do Oeste do Paraná - UNIOESTE
anaa_guedess@hotmail.com

Gabriel de Menezes Trevisan
Universidade Federal de Santa Maria - UFSM
gabryelmt@hotmail.com

Recebido: 24/04/2019

Aceito: 29/10/2019

RESUMO: O presente trabalho teve como objetivo identificar a distribuição espacial da precipitação pluviométrica e a variabilidade temporal na Bacia do Rio Piquiri, no Estado do Paraná, identificando eventos extremos ocorridos no período de 1982 a 2018. As séries históricas foram adquiridas no sistema de informações hidrológicas Hidroweb. A metodologia utilizada teve como base o processo Knowledge Discovery in Database - KDD. Foram definidas áreas pluviometricamente homogêneas para os períodos de 1982-1994 (Período 1), 1995-2006 (Período 2) e 2007-2018 (Período 3). Constatou-se associação entre a existência de eventos extremos e as características físicas da bacia através da distribuição espacial da precipitação com distintas interações entre os conjuntos de dados analisados ao longo da variabilidade temporal. No sul e sudeste da bacia estudada se concentrou altos volumes de precipitação no decorrer de todos os períodos estudados. Com relação aos menores volumes de precipitação, o Período 1 apresentou o menor valor, porém o Período 3 se destacou dos demais com a maior quantidade de estações classificadas como baixa, e para as máximas altas, o Período 2 ganhou destaque. Diante disso, comprova-se a existência de eventos extremos de precipitação, apontando a alta variabilidade temporal e espacial na bacia ao longo da série, podendo interferir nas atividades econômicas locais.

Palavras-chave: Precipitação; Algoritmo EM; Análise fatorial; Gestão ambiental; Agrupamento.

ABSTRACT: The present work aimed at a rainfall spatial distribution and its temporal variability in the Piquiri River Basin in the State of Paraná, identify extreme events occurring in the period from 1982 to 2018. The historical series were acquired in the Hidroweb hydrological information system. The methodology used has based on the Knowledge Discovery in Database - KDD. Areas considered pluviometrically homogenous for the periods of 1982-1994 (Period 1), 1995-2006 (Period 2) and 2007-2018 (Period 3). The association between experience of events and physical characteristics of Basin was verified through of the spatial distribution of occurrences, with different interactions occurring between data sets along of the temporal variability. In southern South America, the concentration was concentrated in high volumes of precipitation, without running all the periods studied. With the same volumes of precipitation, Period 1 presented the lowest value, but Period 3 stood out from the rest with the highest number of plants classified as low, and for the high highs Period 2 called the highlight. Therefore, extreme precipitation events are evidenced, indicating a high temporal and spatial variability in the Basin throughout the series, which may interfere with local vital activities.

Keywords: Precipitation; EM Algorithm; Factorial analysis; Environmental management; Clustering.

INTRODUÇÃO

As condições climáticas e meteorológicas impactam de forma direta os diversos setores da economia, especialmente a agricultura, por isso é de grande interesse o acompanhamento de eventos naturais ao longo do tempo (IPARDES, 2018).

É importante também conhecer a incidência dos eventos extremos do clima, como secas severas e enchentes, identificando as regiões mais suscetíveis à ocorrência desses eventos (CEBDS, 2017).

Tal conhecimento é desejado pela conjuntura agrícola atual, com o propósito de auxiliar o produtor na redução dos riscos inerentes aos investimentos, visando a maximização dos resultados oriundos da produção agrícola (ROMANI et al., 2016).

Com o propósito de auxiliar na análise da frequência de eventos climáticos extremos, as técnicas de mineração de dados surgem como alternativa para transformar informações contidas em grandes volumes de informações em conhecimento passível de ser aplicado no dia-a-dia do empresário rural. Dentre os métodos mais utilizados está o *Knowledge Discovery in Databases* (KDD), ou Descoberta de Conhecimento em Bancos de Dados (FAYYAD et al., 1996; SKUPIEN, CARVALHO, 2017).

O KDD é um processo não trivial de extração de informações implícitas de um conjunto de dados contidos em uma série histórica de estações pluviométricas, sendo executado em etapas: Seleção, Pré-Processamento, Transformação, Mineração de Dados (Data Mining) e Interpretação e Validação (GOMES et al., 2014).

Segundo Santos e Sousa (2018) a escolha da técnica de mineração exige análise detalhada do problema em questão e a decisão sobre o tipo de padrão que será encontrado nos dados. Em particular, a análise de agrupamento é uma técnica capaz de transformar séries históricas de dados pluviométricos em zonas pluviométricas homogêneas.

Ressalta-se que muitos trabalhos nas áreas da hidrologia enfatizam a aplicação de análise de agrupamento para a definição de regiões climáticas homogêneas no do Brasil (SANTOS; SOUSA, 2018; BOSCHI et al., 2017; DOURADO et al., 2013; FECHINE, GALVINCIO, 2008), no presente trabalho, essa técnica será aplicada para analisar a distribuição da precipitação pluviométrica na Bacia do Rio Piquiri, Estado do Paraná, bem como a ocorrência de eventos extremos de precipitação (mínima e máxima absolutas) no período compreendido entre 1982 a 2018.

METODOLOGIA

Área de estudo

A Bacia Hidrográfica do Rio Piquiri localiza-se no Estado do Paraná, com 485 Km de extensão e área total de 24.171,70 Km², o que representa cerca de 12% da área do Estado (SEMA, 2013). O Rio Piquiri nasce no Terceiro Planalto paranaense, região situada no Centro-Sul, e apresenta sua foz no Rio Paraná, no setor oeste do Estado, conforme Figura 1 (IAP, 2008).

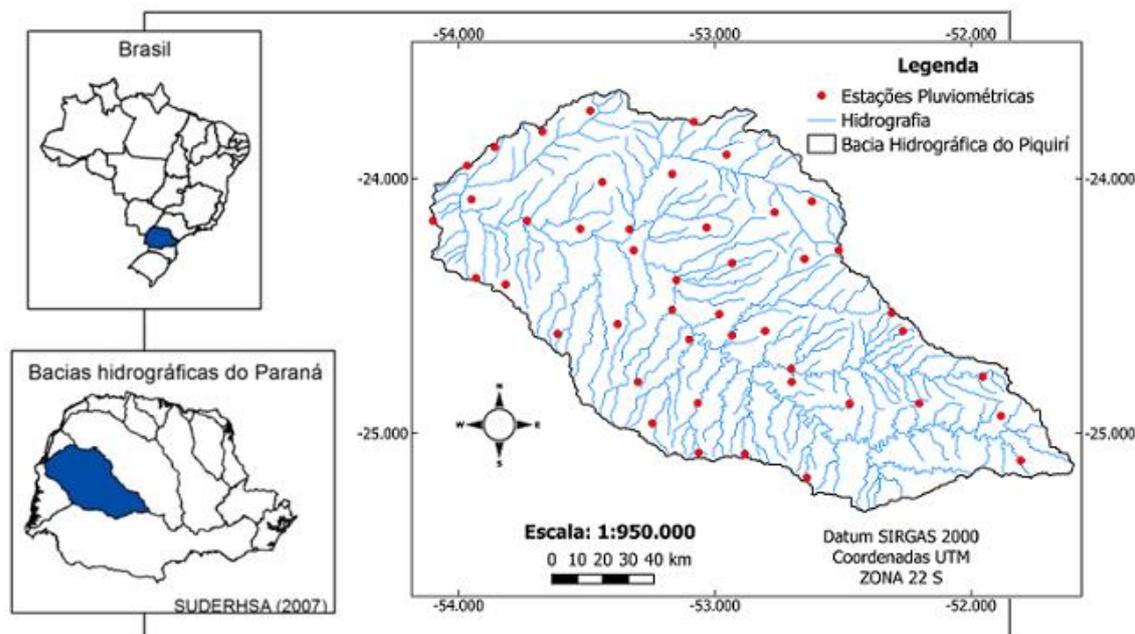


Figura 1. Localização geográfica da bacia hidrográfica e estações pluviométricas do Rio Piquiri no estado do Paraná - SUDERHSA (2007).

Em geral a bacia em estudo situa-se em uma região com variação das temperaturas entre 17,5 e 22,9°C (APARECIDO et al., 2016); e precipitação média anual de 1762,5 mm (CORREA; GALVANI, 2017a). Contudo, a região apresenta peculiaridades espaciais na distribuição espacial das chuvas em função do controle do relevo orografia e a situação geográfica entre os dois grandes tipos climáticos do território nacional, o clima subtropical do Brasil Meridional dos tipos “Cfa” e “Cfb”, com predomínio do tipo climático C2rA’a’ segundo a classificação climática de Camargo (1991) (CAVAGLIONE et al., 2000; APARECIDO et al., 2016).

Essa condição favorece um regime de precipitação quase equitativa o longo do ano, e o clima tropical do Brasil Central, com a concentração de chuvas no período de primavera e verão e a redução significativa nos meses de outono e inverno (NIMER, 1989; NERY, 2006; FRITZSONS et al., 2011).

Terassi, Correa e Galvani (2017) e Correa e Galvani (2017b) descreveram uma menor pluviosidade média anual e redução das chuvas para o setor norte da bacia em estudo e, inversamente, maiores totais pluviométricos para o setor sul, que segundo Waltrick et al. (2015) se caracterizam por uma das mais elevadas médias de erosividade anual do estado do Paraná, superior a 11.000 MJ.mm.ha⁻¹.h⁻¹.ano⁻¹.

Outro fato de destaque na caracterização é sobre distribuição populacional na Bacia do Rio Piquiri, em que Correa (2017) aborda em seu estudo, que as regiões sul e sudeste se destacam por haver maior número de pessoas vivendo nas zonas rurais, em que a principal atividade econômica é a agropecuária.

Seleção dos dados

Inicialmente foram obtidos dados diários de 72 estações pluviométricas da Agência Nacional das Águas, disponível no Sistema de Informações Hidrológicas (HidroWeb) (ANA, 2016), inseridos na bacia hidrográfica do Rio Piquiri, dessas realizou-se uma filtragem dos dados mais consistentes no qual resultou uma quantia de 48 estações pluviométricas com uma série histórica de coleta de dados de 36 anos compreendida entre o período de 1982 a 2018 (Tabela 1).

Tabela 1. Coordenadas geográficas dos postos pluviométricos utilizados no estudo.

ESTAÇÃO	LONGITUDE	LATITUDE	ESTAÇÃO	LONGITUDE	LATITUDE
2453009	-53,3333	-24,2000	2453052	-53,9500	-24,0831
2453050	-53,5256	-24,1983	2453043	-53,1500	-24,4000
2452033	-52,9331	-24,3331	2353003	-53,8831	-23,8500
2454006	-54,1000	-24,1667	2353047	-53,9667	-23,9500
2452040	-52,6500	-24,3167	2353006	-53,1667	-23,9831
2452029	-52,6214	-24,0906	2453008	-53,4397	-24,0147
2452010	-52,7667	-24,1333	2453047	-53,8167	-24,4167
2453037	-53,3800	-24,5731	2452001	-52,9333	-24,6167
2452000	-52,7019	-24,7489	2352026	-52,9547	-23,9067
2453001	-53,1667	-24,5167	2453048	-53,9325	-24,3919
2452011	-52,8042	-24,5997	2453030	-53,6142	-24,6111
2552047	-52,6517	-25,2114	2453010	-53,3167	-24,2831
2451010	-51,8833	-24,9333	2453016	-53,0319	-24,1933
2451021	-51,9544	-24,7792	2453013	-53,1000	-24,6333
2552006	-52,8833	-25,0833	2551009	-51,8064	-25,1092
2452016	-52,2028	-24,8844	2452014	-52,5167	-24,2831
2553019	-53,0667	-25,1000	2452041	-52,2436	-24,5006
2552008	-52,2583	-25,1122	2452015	-52,2667	-24,6000
2552019	-52,4311	-25,2225	2452009	-52,9833	-24,5333
2452035	-52,5500	-24,4331	2453056	-53,2439	-24,9628
2353016	-53,6756	-23,7969	2453014	-53,0667	-24,8831
2353004	-53,0822	-23,7778	2452019	-52,4739	-24,8858
2353005	-53,4864	-23,7342	2452012	-52,7000	-24,8000
2453000	-53,7333	-24,1667	2453012	-53,3000	-24,8000

Fonte: autores, 2019.

O critério adotado neste estudo segue a análise de alguns autores que dividiram o período de estudo em intervalos decadais (DOURADO *et al.*, 2013; BOSCHI *et al.*, 2017). Essa abordagem possibilita um exame mais detalhado e amplo do período escolhido, sendo uma série histórica de dados pluviométricos de 1982 a 2018 dividido em: Período 1: 1982 – 1994, Período 2: 1995 – 2006 e Período 3: 2007 – 2018.

A série temporal delimitada segue proposição da *World Climate Organization* (WMO) (1989) que salienta que o critério de delimitação de uma série histórica de dados

meteorológicos deve conter, pelo menos, 30 anos, respaldando desta forma as inferências relacionadas às análises realizadas.

Após a separação da série em períodos, foi realizado a média de cada estação por período, e o valor encontrado, foi interpretado e avaliado na classificação dos quartis, ou seja, na ocorrência de eventos extremos.

Pré-processamento e transformação

Dentro desse período de 36 anos, as falhas foram identificadas e seguindo as recomendações da ANA (2011), eliminaram-se os meses que apresentaram mais de 5% de registros sem informação, bem como as falhas não foram preenchidas.

Já os demais dados faltantes foram preenchidos por meio da técnica de ponderação regional baseado nas correlações com as estações vizinhas. Para a aplicação do método foi adotado como critério mínimo a obtenção de um coeficiente de determinação superior a 0,7, segundo recomendações de Barbosa et al. (2005) e Pruski et al. (2004) e a proximidade das estações.

Em seguida, foi realizada a análise da consistência de cada estação, dentro de uma visão regional, com o objetivo de comprovar o grau de homogeneidade dos dados disponíveis num posto com relação às observações registradas em postos vizinhos. Para tanto, foi utilizado a técnica da Dupla Massa (BERTONI & TUCCI, 2007).

Nesta técnica, selecionam-se, por vez, a estação de interesse e as mais próximas, cujos totais anuais acumulados de um posto de observação que se pretende avaliar a consistência dos dados, são plotados nas ordenadas de um gráfico e nas abscissas os totais médios anuais das demais estações.

Haverá consistência dos totais anuais da estação analisada quando houver uma tendência linear em relação às estações vizinhas. A verificação da linearidade entre os totais anuais da estação analisada com relação às demais é avaliada pelo ajuste da equação da reta e do coeficiente de determinação, obtidos pela técnica da minimização da soma dos quadrados dos desvios (OLIVEIRA et al., 2010).

Uma etapa essencial na aplicação das técnicas de clusterização quando se utiliza atributos com diferentes ordens de grandeza é a normalização dos dados. No entanto, para o presente trabalho esta etapa não foi necessária, pois segundo Dourado et al. (2013) e Boschi et al. (2017) quando se estuda apenas uma variável, não há necessidade de aplicar essa técnica, uma vez que a única informação utilizada nesse trabalho foi chuva anual, o conjunto de dados se encontrava na mesma unidade, não necessitando, portanto, da normalização.

Mineração de dados

O programa computacional utilizado para as análises dos dados foi o *Waikato Environment for Knowledge Analysis* (WEKA), versão 3.8. Um software livre, disponível sob licença da GNU (*General Public License*) composto de uma coleção de

algoritmos nas áreas de aprendizado de máquina e mineração de dados (HALL et al., 2009; WITTEN et al., 2011).

A técnica de modelagem escolhida para seleção das zonas homogêneas de precipitação pluviométrica foi a clusterização, por meio do algoritmo *Expectation-Maximization* (EM).

Para Johann et al. (2013), este algoritmo baseia-se nas estatísticas de máxima verossimilhança para estimar os parâmetros da distribuição normal, realizando uma mistura de várias distribuições normais univariadas de mesma variância, estimando as médias de cada distribuição normal e, através de um processo iterativo, formam-se os grupos.

O usuário pode escolher o número de *clusters* (grupos) desejados, porém o algoritmo EM também pode encontrar o número adequado de grupos a serem utilizados pelo modelo sem a predefinição inicial, sendo essa última, a opção escolhida no estudo.

Para a avaliação da qualidade dos modelos gerados e o número de clusters adequados para cada intervalo, utilizou-se o log da verossimilhança (*log-likelihood*).

Para a elaboração dos mapas, utilizou-se o Modelo Digital de Elevação (MDE) a qual foi obtido através do projeto TOPODATA (folhas 24S54; 24S525; 25S54 e 25S525), elaborado a partir da missão *Shuttle Radar Topography Mission* (SRTM), cujos dados são disponibilizados pelo *United States Geological Survey* (USGS).

Os dados originais do projeto SRTM tem resolução espacial de 90m, entretanto, no projeto Topodata, os modelos foram reamostrados para uma resolução de 30m, melhorando a qualidade dos dados brutos, como as folhas estão, originalmente, na projeção geográfica do elipsóide de referência WGS 1984, durante a preparação dos dados para o processamento realizou-se o mosaico das folhas, o qual foi reprojetoado para a projeção UTM do mesmo elipsoide SIRGAS 2000, para a zona 22S.

Os processos de delimitação automática da bacia e extração da hidrografia foram realizados no software ArcGis versão 2.16.1. Inicialmente buscou-se executar o preenchimento de depressões espúrias, corrigindo possíveis ruídos atmosféricos minimizando erros de coleta de dados (*r.fill.dir*), obtendo-se assim o Modelo Digital de Elevação Hidrologicamente Consistente (MDEHC).

Após obter a delimitação e a rede de drenagem da bacia hidrográfica do Piquiri, realizou-se os agrupamentos por períodos, especializando as estações no ambiente GIS, através das coordenadas geográficas, e assim, elaborando os mapas de distribuição das zonas pluviométricas homogêneas da Bacia do Piquiri.

Interpretação e avaliação

Após a geração e espacialização dos *clusters* avaliaram-se os agrupamentos buscando identificar o comportamento da precipitação para cada zona homogênea através da média e desvio padrão.

Esse procedimento possibilitou conhecer a tendência de precipitação em cada cluster no período estudado, permitindo estimar a configuração geral e a distribuição espaciais.

Aplicou-se a técnica dos quartis para estabelecer os limiares do índice pluviométrico, com o intuito de classificar a ocorrência de eventos extremos na escala temporal anual, baseando-se nos quartis inferior (Q_{25}), mediano (Q_{50}) e superior (Q_{75}), recebendo a classificação conforme exposto na Tabela 2.

Tabela 2. Valores para a Classificação - Quartis.

Quartis	Limites de Precipitação (mm)	Classificação
Q25	685,6 - 1474,5	Baixa
Q50	1474,6 - 1813,2	Média
Q75	1813,3 - 2104,8	Alta
Valor máximo	2104,9 - 3339,4	Máxima Alta

Fonte: autores, 2019.

RESULTADOS E DISCUSSÃO

Para os anos de 1982-1994 e 1995-2006 foram encontrados 2 agrupamentos em cada período ($K=2$), ou seja, regiões pluviométricamente homogêneas, onde o Período 1 apresentou 58% dos dados concentrados na região 0, ficando a maioria agrupado nos setores nordeste e sudeste da bacia, sendo a média dessa região pluviométricamente homogênea menor do que da região 1 (1660,85 mm) (Tabela 3).

Tabela 3. Quantidade de estações, média e desvio padrão em cada região pluviométricamente homogêneas, respectivamente ao período analisado.

	Anos	Regiões pluviométricamente homogêneas	
		0	1
Estações		28	20
Média (mm)	Período 1 1982-1994	1660,85	2012,64
Desvio		202,7	196,71
Estações		25	23
	Período 2 1995-2006		
Média (mm)		2052,32	1666,84
Desvio		189,36	205,66

Fonte: Autores, 2019.

No Período 2 ocorre o oposto ao período anteriormente analisado. A menor média de precipitação pluviométrica encontrada foi na região pluviométricamente homogênea 1 (1666,84 mm), porém, apenas 48% dos dados foram agrupados nesse grupo, correspondendo a 23 estações pluviométricas, ficando a maioria agrupadas nos setores central e noroeste da bacia.

Em relação ao Período 3, este foi o único a obter 3 regiões pluviométricamente homogêneas (K=3). Fato este ocorrido em função da grande diferença de variância entre os grupos formados.

Para Batista (2009), a hipótese com maior verossimilhança terá menor log-verossimilhança negativa, mostrando-se assim, que o conjunto de dados desse período obteve grande dispersão estatística, já que seu log-likelihood foi o menor dos demais períodos (-64,55271), justificativa para que ocorresse esse agrupamento diferenciado (Tabela 4).

Tabela 4. Valores de log-likelihood de cada respectivo período.

Período	Número de agrupamento (K)	log-likelihood
1	2	-67,63627
2	2	-67,58028
3	3	-64,55271

Fonte: autores, 2019.

Para esse mesmo período, a menor média de precipitação pluviométrica foi encontrada na região 0 (1495,40 mm), para um total de 18 estações (37%). Nas regiões 1 e 2 foram agrupadas 15 estações em cada grupo (31,5%), encontrando a maior média na região 2, situando-se no setor central da bacia, conforme Tabela 5.

Tabela 5. Quantidade de estações, a média e desvio padrão no Período 3.

Anos	Regiões pluviométricamente homogêneas			
	0	1	2	
Estações	18	15	15	
Média (mm)	Período 3 2007-2018	1495,4	1660,54	21039,83
Desvio		145,71	143,81	141,19

Fonte: autores, 2019.

Assim, por meio da geração das regiões pluviométricamente homogêneas, os resultados obtidos para a série história do período de 1982 a 2018 foram representados por mapas com a espacialização das estações na respectiva bacia hidrográfica, cada um com o respectivo período analisado (Figura 2):

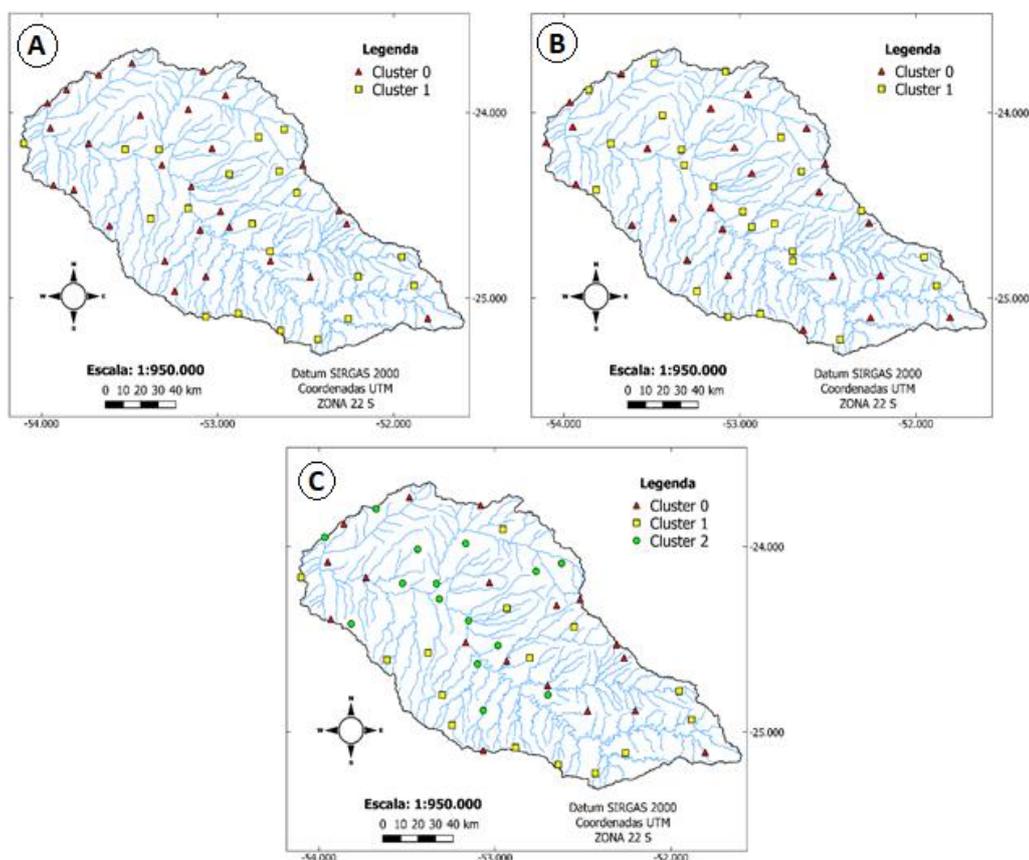


Figura 2. Espacialização das regiões pluviométricamente homogêneas (*cluster*) - Período 1 (A), Período 2 (B) e Período 3 (C). Org: autores, 2019.

Análise de precipitações e eventos extremos

Em relação à análise da precipitação e a existência de eventos extremos, houve distintas interações entre os conjuntos de dados ao longo dos períodos estudados na Bacia do Piquiri (Tabela 6).

Tabela 6. Estatística descritiva da precipitação pluviométrica realizada para os Períodos 1 (1982 – 1994), Período 2 (1995 – 2006) e Período 3 (2007 – 2018).

	Regiões pluviométricamente homogêneas	Média Mínima (mm)	Média (mm)	Média Máxima (mm)	Mediana (mm)	Desvio Padrão	Coefficiente de Variação (%)
Período 1	0	1465,1	1741,4	2083,4	1718,9	±175,8	10,1
	1	1491,0	1898,5	2216,0	1894,9	±207,8	10,9
Período 2	0	1471,8	1841,9	2142,1	1811,5	±190,4	10,3
	1	1275,5	1842,5	2272,9	1848,4	±253,3	13,7
Período 3	0	1337,9	1681,6	2009,8	1666,4	±194,5	11,6
	1	1503,1	1809,5	2015,2	1864,5	±172,1	9,5
	2	1454,5	1649,8	1907,0	1675,7	±112,0	6,8

Fonte: autores, 2019.

A análise de todos os dados de precipitação pluviométrica anual do Período 1, a mínima encontrada foi de 685,6 mm e a máxima de 3339,4. Já para o mesmo período, a região 0 apresentou a menor média de volume de chuvas se comparado com a região 1, respectivamente. Isso se confirma com a associação da técnica dos quartis (Figura 3), onde a região 0 foi o único em que apareceu a classificação baixa, com o mínimo de 1465,1 mm.

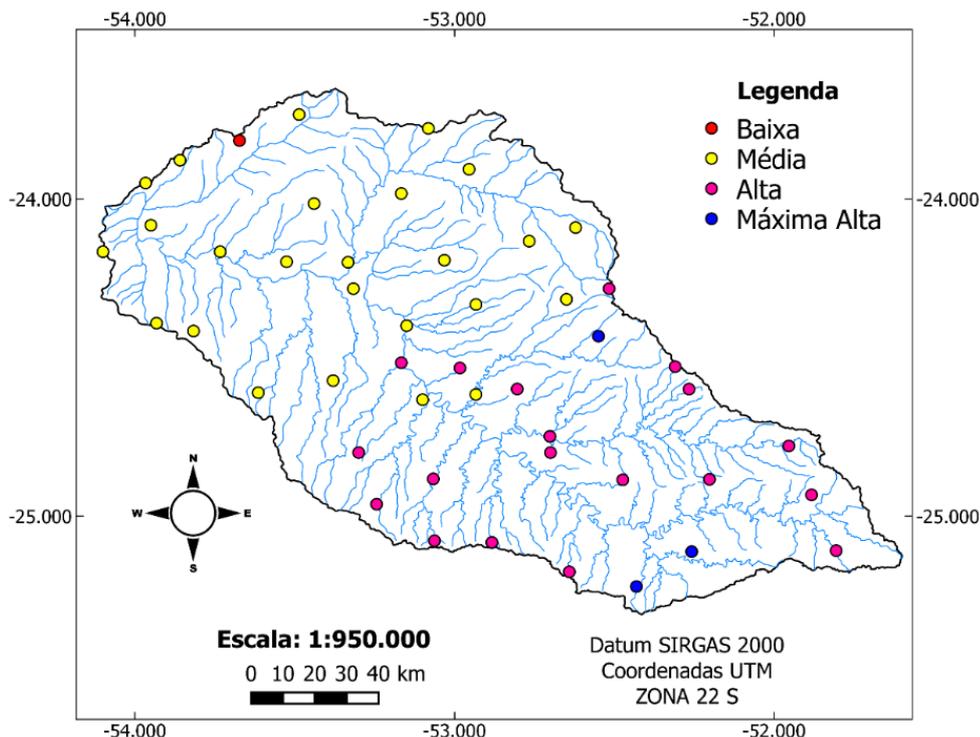


Figura 3. Espacialização dos eventos extremos (classificados no quartis) para o Período 1 nas regiões pluviométricamente homogêneas 0 e 1. Fonte: autores, 2019.

Ainda na região 0, também foi possível identificar a existência de 17 estações classificadas como média e 10 como alta, sendo a máxima de 2083,4 mm. Diferente da região 1 que foram identificadas 8 estações na classificação média e 9 estações como alta, situadas ao sudeste da bacia.

Já a classificação máxima alta obteve 3 dados nesse item, sendo o de maior média com volumes de chuva de 2216 mm, ou seja, mais estações chuvosas na região 1, reforçando a ideia inicial essa região possui maiores volumes de precipitação do que a região 0.

Evidenciando os dados do Período 1, os menores volumes de precipitação foram encontrados nos setores noroeste e nordeste com 685,6 mm e 1474,5 mm. Fato assegurado pelo predomínio do tipo climático C2rA'a' segundo a classificação climática de Camargo (1991), caracterizado por ser um clima úmido subúmido com

pequena deficiência hídrica (APARECIDO et al., 2016), corroborando o encontrado nesses setores de baixos volumes de precipitação.

Com relação à espacialização do Período 2, esta apresentou a mínima de 773,7 mm e a máxima de 2900 mm, valores superiores se comparado com o Período 1.

Com relação as regiões pluviométricamente homogêneas, as menores e maiores médias de volumes de chuvas do Período 2 se concentraram na região 1, com mínima de 1275,5 mm e máxima de 2272,9 mm.

Além disso, para a região 0 encontrou-se apenas 1 estação classificada como baixa, 10 como alta e 2 como máxima alta. Já para a região 1 encontrou-se o mesmo valor para baixa, mas para alta e máxima alta se obteve 11 e 3 estações nessas classificações, reforçando a característica de precipitação volumétrica encontrada anteriormente.

Com relação ao Período 2 visualizaram-se maiores aglomerações de estações de classificação alta e máxima alta nas regiões central, sudeste e sudoeste do sistema fluvial, além de haver maiores valores de precipitação média máxima, e menor valor de precipitação média mínima.

Contudo, em relação a classificação da espacialização dos eventos extremos, não houve diferenciação expressiva em relação ao Período 1 (Figura 4).

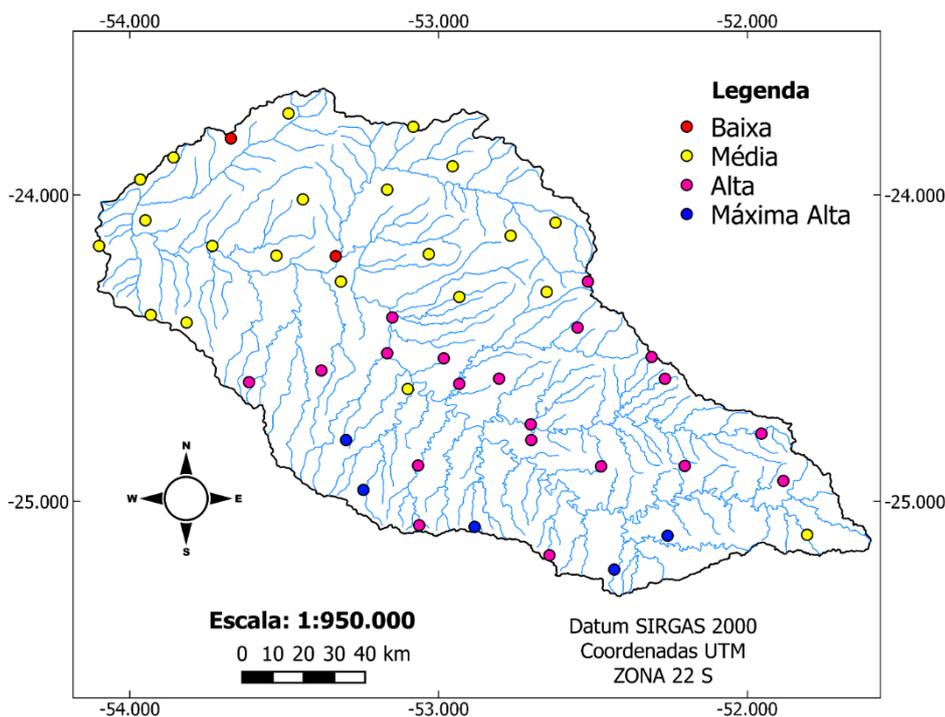


Figura 4. Espacialização dos eventos extremos (classificados) para o Período 2 nas regiões 0 e 1. Fonte: autores, 2019.

Segundo Araújo *et al.* (2015), as declividades nessa região são superiores a 45%, podendo chegar a mais de 75% no sentido sul da bacia, havendo as maiores altitudes

da bacia, variando de 700 – 1200 m, isto influencia na ocorrência de precipitação pois quando uma massa de ar se aproxima de uma montanha ela é forçada a deslocar-se para cima acompanhado o relevo do terreno, ocasionando a precipitação (CARVALHO; ASSAD; PINTO, 2012).

Devido também à declividade dessa região, os solos não são espessos e com elevado grau de instabilidade pedológica, predominando Neossolos Litólicos, que, segundo o Sistema Brasileiro de Classificação de Solos – SiBCS/Embrapa (2006), são solos que possuem pouca profundidade e presença de rochas, dificultando a penetração de raízes e de água no interior do mesmo.

Outro aspecto importante no Período 2 é que se centra a aglomeração de estações de classificação dos eventos médios e baixos nos setores norte e nordeste, as quais possuem menores volumes de precipitação se comparado as demais regiões da bacia em estudo, situando-se no compartimento do clima tipo Cfa, caracterizado por possui mais de 30 mm de chuva no mês mais seco, segundo a classificação climática de Köppen.

Devido a essa característica e as baixas declividades, a região norte e nordeste tornam-se suscetíveis a possíveis inundações e alagamentos caso ocorra eventos com grandes volumes de escoamento superficial, não havendo a necessidade de ser eventos de longa duração, mas sim, chuvas de maior intensidade.

Já para o Período 3 a máxima foi de 2813,3 mm e a mínima de 1019,3 mm, sendo que esse período se destacou dentre os demais pois foi o único que apresentou a formação de 3 regiões pluviométricamente homogêneas quando aplicado o algoritmo EM, bem como não apresentou classificação máxima alta como os demais períodos analisados (Figura 5).

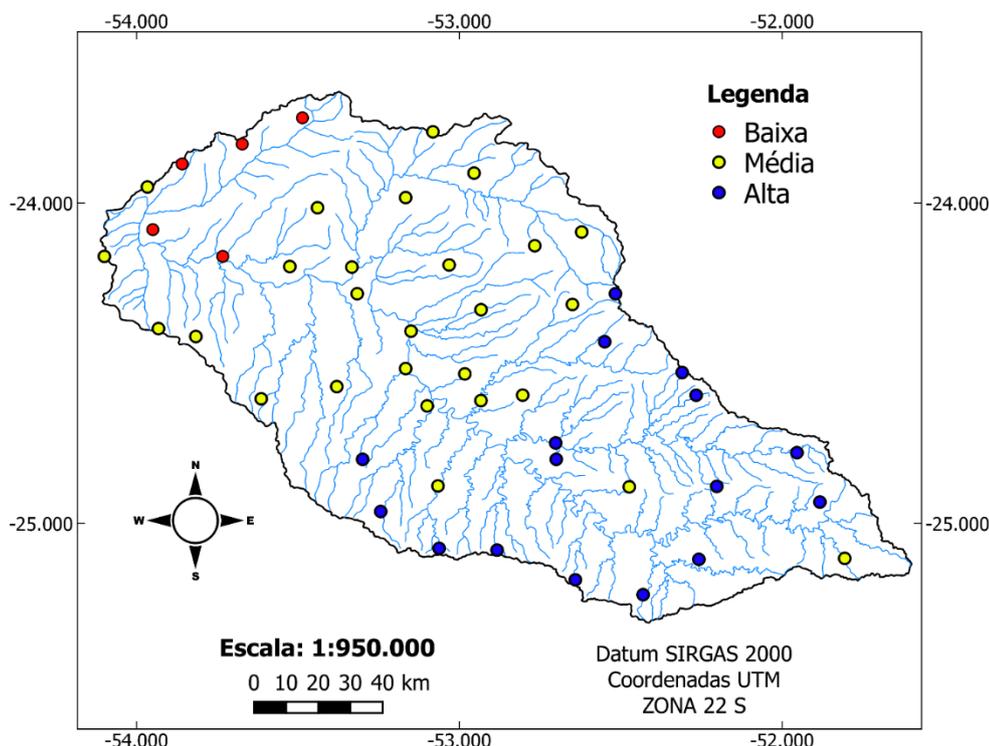


Figura 5 - Espacialização dos eventos extremos (classificados) para o Período 3 nas regiões pluviométricamente homogêneas 0, 1 e 2. Fonte: autores, 2019.

Na região 0 a mínima foi de 1337,9 mm, máxima 1681,6 mm e a média 1681,6 mm, além de terem sido encontradas 4 estações baixas, 8 médias e 6 altas.

Para a região 1 a mínima foi de 1503,1 mm, máxima de 1809,5 e a média de 2015,2 mm. Não houve estações classificadas como baixa ou máxima alta, mas apenas 9 altas e 6 médias.

Na região 2 encontrou-se apenas 1 de estação como alta e baixa, para cada classificação. Para a média foram encontradas 13 estações, onde a máxima dessa região foi de 1907 mm e a mínima de 1454,5 mm, com média de 1649,8 mm.

Com a espacialização das classificações das estações é possível visualizar a aglomeração das altas nos setores sul e sudeste, as médias no setor central e as baixas nos setores norte e nordeste.

Visualizou-se que 4 estações que estavam classificadas como média no Período 2, passaram a ser classificadas como baixas nesse período. O mesmo comportamento se deu para 8 estações que estavam classificadas como alta no período anterior, as quais passaram a serem médias no Período 3.

Já as estações classificadas como Máxima Alta no período anterior, passaram a ser Alta no Período 3.

Como nesse período houve aumento do número das estações baixas, de 1 para 5 unidades, isso demonstra que nesse período, caracterizou-se com os menores

valores de precipitação, fato reforçado pela ausência de estações máxima alta bem como os menores valores de precipitações máximas.

CONCLUSÃO

Através da análise da série histórica de precipitação pluviométrica e sua espacialidade na Bacia do Rio Piquiri foi possível aliar tecnologias computacionais, para identificar as regiões com maior incidência de chuvas, bem como, a ocorrência de classes de valores extremos no período de 1982 a 2018.

Para tanto, utilizou-se a mineração de dados, através da técnica de agrupamento por meio do algoritmo EM que delimitou com coerência as zonas pluviométricas da bacia em questão, mostrando-se eficiente para essa proposição.

Em relação à análise da precipitação e a existência de eventos extremos, percebeu-se distintas interações entre os conjuntos de dados na Bacia do Piquiri. A série destaca os altos volumes de chuvas que se concentram no sul e sudeste da bacia hidrográfica, o qual pode vir a influenciar futuramente na dinâmica ambiental do local em estudo.

A análise dos valores extremos demonstrou que no primeiro período alterações na distribuição espacial das precipitações, com grandes diferenças entre as mínimas e máximas (média com aproximadamente 1836,8 mm), mas no último intervalo, Período 3, foi visto uma regularização e normalização dos dados, não havendo grandes alterações, principalmente em termos de eventos extremos, fato comprovado no último período que não houve a existência de máximas altas.

Já em oposição ao cenário anterior, o Período 1 apresentou o menor volume de chuva, bem como, foi o único intervalo que apresentou apenas uma estação com a classificação baixa (2353016), dando ênfase para as regiões noroeste e nordeste.

Em relação às máximas altas ficaram em destaque no Período 2, com o maior número de estações nessa classificação, sendo situadas principalmente na região sudeste.

Assim, por meio da espacialização dos dados foi possível visualizar que ao longo da série estudada houve a existência de estações classificadas como máxima alta e baixa, apresentando, portanto, um panorama da variabilidade de precipitação em relação a existência de eventos extremos na bacia do rio Piquiri, ficando evidente a influência da precipitação orografia para a caracterização da área estudada, em que as áreas a jusante são as que apresentam os menores totais pluviométricos e a porção a montante aquela em que foram percebidos os maiores totais.

REFERÊNCIAS

AGÊNCIA NACIONAL DE ÁGUAS - ANA. *Inventário Estações Pluviométricas*. Brasília, DF, 2013. Disponível em: <<http://arquivos.ana.gov.br/inf hidrologicas/InventariodasEstacoesPluviometricas.pdf>>. Acessado em: 12 dez. 2017.

AGÊNCIA NACIONAL DE ÁGUAS - ANA. *Sistema de Informações Hidrológicas - HidroWeb*. Brasília, DF, 2017. Disponível em: <<http://hidroweb.ana.gov.br/HidroWeb.asp?TocItem=1080&TipoReg=7&MostraCon=false&CriaArq=false&TipoArq=1&SerieHist=true>>. Acessado em: 30 jan. 2018.

AGÊNCIA NACIONAL DE ÁGUAS – ANA. *Outorga de direito de uso de recursos hídricos. Cadernos de capacitação em recursos hídricos*. v. 6, 2011.

APARECIDO, L. E. O.; ROLIM, G. S.; RICHETTI, J.; SOUZA, P. S.; JOHANN, J. A. Köppen, Thornthwaite and Camargo climate classifications for climatic zoning in the State of Paraná, Brazil. *Ciência e Agrotecnologia* 40(4):405-417, Jul/Aug. 2016.

ARAÚJO, F. C.; MELLO, E. L.; SILVA, B. B.; MERCANTE, E.; GOLLIN, G. M. Utilização de dados de sensoriamento remoto para obtenção das características físicas da bacia hidrográfica do rio Piquiri - PR. In: *XVII Simpósio Brasileiro de Sensoriamento Remoto*, João Pessoa - PB, 2015. Disponível em: <<http://www.dsr.inpe.br/sbsr2015/files/p1084.pdf>>. Acessado em: 12 dez. 2017.

BARBOSA, S. E. da S.; BARBOSA JÚNIOR, A. R.; SILVA, G. QUEIROZ da.; CAMPOS, E. N. B.; RODRIGUES, V. de C. Geração de modelos de regionalização de vazões máximas, médias de longo período e mínimas de sete dias para a Bacia do Rio do Carmo, Minas Gerais. *Engenharia Sanitária e Ambiental*, v.10, n.1, p.64-71, 2005.

BATISTA, J. L. F. Verossimilhança e Máxima Verossimilhança. *Material de apoio*. Universidade de São Paulo - USP. Campus Piracicaba. 2009. Disponível em: <<http://cmq.esalq.usp.br/BIE5781/lib/exe/fetch.php?media=leituras:verossim.pdf>>. Acessado em: 12 dez. 2017.

BERTONI, J. C.; TUCCI, C. E. M. Precipitação. In: TUCCI, C. E. M. *Hidrologia: ciência e aplicação*. 2.ed. Porto Alegre: ABRH/Editora, UFRGS, 2007.

BOSCHI, R. S.; OLIVEIRA, S. R. M.; ASSAD, E. D. Técnicas de mineração de dados para análise da precipitação pluvial no Rio Grande do Sul. *Engenharia Agrícola* (Impresso). v.31, p.1189 - 1201, 2017. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-69162011000600016>. Acessado em: 10 dez 2017.

CAMARGO, A. P. de. Classificação climática para zoneamento de aptidão agroclimática. *Revista Brasileira de Agrometeorologia*. 8:126-131, 1991.

CARDOSO, A. O.; DIAS, P. L. S. Identificação de Trimestres Extremos no Regime Pluviométrico no Sul e Sudeste do Brasil e Relação com Anomalias com a TSM. *Revista Brasileira de Meteorologia*, v.19, nº1, 149-162p. 2004. Disponível em: <www.rbmet.org.br/port/revista/revista_dl.php?id_artigo=64&id_arquivo=155>.

Acessado em: 12 dez. 2017.

CARVALHO, J. R. P.; ASSAD, E.D.; PINTO, H.S. *Interpoladores geoestatísticos na análise da distribuição espacial da precipitação anual e de sua relação com altitude*.

Pesquisa Agropecuária Brasileira. 47:1235-1242, 2012. Disponível em: <<https://seer.sct.embrapa.br/index.php/pab/article/view/10819/7983>>. Acessado em: 24 abr. 2017.

CONSELHO EMPRESARIAL BRASILEIRO PARA O DESENVOLVIMENTO SUSTENTÁVEL – CEBDS. *Riscos Climáticos: Como o setor empresarial está se adaptando?* Câmara Temática de Energia e Mudança do Clima (CTClima), Rio de Janeiro, 2017. Disponível em: <http://cebds.org/wp-content/uploads/2017/11/Gerenciamento-Clima-Completo_Final.pdf>.

Acessado em: 30 mar. 2017.

CORREA, M. G. G. *Distribuição espacial e variabilidade da precipitação pluviométrica na bacia do rio Piquiri-PR*. Dissertação (Mestre em Geografia Física) -Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, São Paulo, 2017.

Disponível em: <www.teses.usp.br/teses/disponiveis/8/8135/tde-28082013-122143/>. Acessado em: 12 dez. 2017.

CORREA, M. G. G.; GALVANI, E. Análise estatística da variabilidade da precipitação pluviométrica para a bacia hidrográfica do rio Piquiri-PR. *Boletim Paulista de Geografia*, n. 96, p. 1-10, 2017a.

CORREA, M. G. G.; GALVANI, E. Variabilidade Espacial e Sazonal da Precipitação Pluviométrica na Bacia Hidrográfica no Rio Piquiri-PR. *Revista do Departamento de Geografia*, v. 34, 2017b.

DOURADO, C. da S.; OLIVEIRA, S. R. de M.; AVILA, A. M. H. Análise de zonas homogêneas em séries temporais de precipitação no Estado da Bahia. *Revista Bragantia*, Campinas, v. 72, n. 2, p.192-198, 2013. Disponível em: <<http://www.redalyc.org/pdf/908/90827918006.pdf>>.

Acessado em: 30 mar. 2017.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA – EMBRAPA. *Sistema Brasileiro de Classificação de Solos – SiBCS*. 2. ed. Rio de Janeiro. 2006. Disponível em: <www.agrolink.com.br/downloads/sistema-brasileiro-de-classificacao-dos-solos2006.pdf>.

Acessado em: 12 dez. 2017.

FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. From data mining to knowledge discovery: An overview. In: *Advances in Knowledge Discovery and Data Mining*, AAAI Press/The MIT Press, England, 1996, p.1-34.

FECHINE, J.A.L.; GALVINCIO, J.D. Agrupamento da precipitação mensal da bacia hidrográfica do rio Brígida–PE, através da multivariada. *Revista Brasileira de Geografia Física*, v.1, p.39-46, 2008.

FRITZSONS, E.; MANTOVANI, L.E.; WREGE, M.S.; CHAVES NETO, A. Análise da pluviometria para definição de zonas homogêneas no Estado do Paraná. *RA'E GA: o Espaço Geográfico em Análise*, Curitiba, v.23, n.1, p.555-572, 2011.

GOMES, D.C.; CARVALHO, D.R.; CUBAS, M.R.; SHMEIL, M.A.H. Mineração de Dados no Serviço de Atendimento de Urgências. *J. Health Inform*, v. 6, n. 4, p. 141-6, 2014.

HALL, M.; FRANK, W.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P.; WITTEN, I. H. The Weka data mining software: An update. *SIGKDD Explorations*. 11(1), 10-18. 2009. Disponível em: <<http://dl.acm.org/citation.cfm?id=1656278>>. Acessado em: 10 dez 2016.

INSTITUTO AMBIENTAL DO PARANÁ - IAP. *Avaliação Ambiental Integrada - Bacia do Rio Piquiri*. 2008. Disponível em: <www.iap.pr.gov.br>. Acesso em: 11 dez. 2017.

INSTITUTO AMBIENTAL DO PARANÁ – IAP. *Caracterização Da Bacia Hidrográfica Do Rio Piquiri*. AAI da Bacia do rio Piquiri. Curitiba. 01 - 10p. 2006. Disponível em: <http://www.iap.pr.gov.br/arquivos/File/EIA_RIMA/baciapiquiri/f_IV_Caracterizacao_da_Bacia.pdf>. Acessado em: 12 dez. 2017.

Instituto de Terras, Cartografia e Geociência do Paraná - ITCG/PR. *Mapa Clima - Estado do Paraná*. 2010. Disponível em: <http://www.itcg.pr.gov.br/arquivos/File/Produtos_DGEO/Mapas_ITCG/PDF/Mapa_Climas_A3.pdf>. Acessado em: 12 dez. 2017.

INSTITUTO PARANAENSE DE DESENVOLVIMENTO ECONÔMICO E SOCIAL – IPARDES. *Cadernos Municipais*. 2018. Disponível em: <www.ipardes.gov.br>. Acesso em: 22 nov. 2018.

JOHANN, J. A.; ROCHA, J. V.; OLIVEIRA, S. R. de M.; RODRIGUES, L. H. A.; LAMPARELLI, R. A. C. Data mining techniques for identification of spectrally homogeneous areas using NDVI temporal profiles of soybean crop. *Eng. Agríc.* vol.33 no.3 Jaboticabal May/June 2013. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-69162013000300008>. Acessado em: 10 dez. 2017.

NERY, J.T. Dinâmica climática da região Sul do Brasil. *Revista Brasileira de Climatologia*, Curitiba, v.1, n.1, p.61-75, 2006.

NIMER, E. *Climatologia do Brasil*. 2ª Edição. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística, 1989. 421p.

OLIVEIRA, L. F. C., FIOREZE, A. P., MEDEIROS, A. M. M., SILVA, M. A. S. Comparação de metodologias de preenchimento de falhas de séries históricas de

precipitação pluvial anual. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 14(11), 1186–1192. 2010.

PRUSKI, F. F.; PEREIRA, S. B.; NOVAES, L. F.; SILVA, D. D.; RAMOS, M. M. Precipitação média anual e vazão específica média de longa duração, na Bacia do São Francisco. *Revista Brasileira de Engenharia Agrícola e Ambiental*, v.8, n.2/3, 2004.

PRUSKI, F.F.; PEREIRA, S. B.; NOVAES, L. F.; SILVA, D. D.; RAMOS, M. M. Precipitação média anual e vazão específica média de longa duração, na Bacia do São Francisco. *Revista Brasileira de Engenharia Agrícola e Ambiental*, v.8, n.2/3, p.247-253, 2004.

ROLIM, G. S.; SENTELHAS, P. C.; BARBIERI, V. Planilhas no ambiente Excel TM para os cálculos de balanços hídricos: normal, sequencial, de cultura e de produtividade real e potencial. *Revista Brasileira de Agrometeorologia*, Santa Maria, v. 6, n. 1, p. 133-137, 1998. Disponível em: <www.ufsm.br/rba/t13313761.html>. Acessado em: 12 dez. 2017.

ROMANI, L. A. S.; BAMBINI, M. D.; COLTRI, P. P.; JUNIOR, A. L.; KOENIGKAN, L. V.; OTAVIAN, A. F.; EVANGELISTA, S. R. M.; JUNIOR ZULLO, J.; PINTO, H. S.; ASSAD, E. D. Sistema de Monitoramento Agrometeorológico - Agritempo: inovação em rede apoiando políticas públicas e a tomada de decisão agrícola. *Revista Agrometeoros*, Passo Fundo, v.24, n.1, p.29-40, out. 2016. Disponível em: <<http://seer.sct.embrapa.br/index.php/agrometeoros/article/view/24879/13561>>. Acessado em: 30 mar. 2017.

SANTOS, E. F. N.; SOUSA, I. F. Análise estatística multivariada da precipitação do estado de sergipe através dos fatores e agrupamentos. *Revista Brasileira de Climatologia*. 14(23). 2018.

SECRETARIA DE ESTADO DO MEIO AMBIENTE E RECURSOS HÍDRICOS – SEMA. *Bacias hidrográficas do Paraná*. Governo do Estado do Paraná. Governo do Estado do Paraná. 2ª ed: Curitiba.

SILVA, S. T.; BRAGA, C. C.; ARAGÃO, M. R. S. Influência do El Nino - Oscilação Sul na distribuição espacial e temporal da precipitação no Estado da Paraíba. In: *Congresso Brasileiro De Meteorologia*, 9, Campos do Jordão-SP. Anais. Sociedade Brasileira de Meteorologia p. 236-239, 1996.

SKUPIEN, S.V.; CARVALHO, D.R. Data Mining on Nursing Consultation for the Postpartum Period. *Iberoamerican Journal of Applied Computing*, v.2, p. 32-42, 2017. Disponível em: <<http://www.revistas2.uepg.br/index.php/ijac/article/view/5362/3457>>. Acessado em: 30 mar. 2017.

SOUSA, W. dos S.; SOUSA, F. de A. S. Rede neural artificial aplicada à previsão de vazão da Bacia Hidrográfica do Rio Piancó. *Revista Brasileira de Engenharia Agrícola*

e *Ambiental* v.14, n.2, p.173–180, 2010. Disponível em: <<http://www.scielo.br/pdf/rbeaa/v14n2/v14n02a08.pdf>>. Acessado em: 12 dez. 2017.

SUPERINTENDÊNCIA DE DESENVOLVIMENTO DE RECURSOS HÍDRICOS E SANEAMENTO AMBIENTAL – SUDERHSA. *Bacias Hidrográficas do Paraná*. Paraná, 2007. Disponível em: <<http://www.aguasparana.pr.gov.br/modules/conteudo/conteudo.php?conteudo=82>>. Acesso em: 27 abr. 2017.

TERASSI, P.M.B.; CORREA, M.G.G.; GALVANI, E. Determinação de regiões pluviométricas homogêneas na bacia hidrográfica do rio Piquiri - Paraná. In: *Os desafios da Geografia Física na fronteira do conhecimento*. Campinas: Instituto de Geociências (UNICAMP), p.1610-1621, 2017.

WALTRICK, P.C.; MACHADO, M.A.M.; DIECKOW, J.; OLIVEIRA, D. Estimativas da erosividade de chuvas no estado do Paraná pelo método da pluviometria: Atualização com dados de 1986 a 2008. *Revista Brasileira de Ciência do Solo*, Viçosa, v.39, n.1, p.256-267, 2015.

WITTEN, I. H.; FRANK, E.; HALL, M. A. *Data mining: practical machine learning tools and techniques*. San Francisco: Morgan Kaufmann. 3, 629. 2011. Disponível em: <<ftp://ftp.ingv.it/pub/manuela.sbarra/Data%20Mining%20Practical%20Machine%20Learning%20Tools%20and%20Techniques%20-%20WEKA.pdf>>. Acessado em: 10 dez. 2017.

WORLD CLIMATE ORGANIZATION – WMO. *Calculation of monthly and anual 30-years standard normals*. Washington, 1989.