

Cálculo da curva número para bacia hidrográfica urbana utilizando diferentes abordagens de classificação para imagem orbital RapidEye: estudo de caso para o arroio Pepino (Pelotas, RS)

Gustavo Willy NAGEL¹, Fabrício da Silva TERRA^{2,4,5}, Jade Silva de OLIVEIRA³,
Ingrid HORÁK-TERRA^{2,5} & Samuel BESKOW^{3,4}

¹ Programa de Pós-Graduação em Sensoriamento Remoto, Divisão de Sensoriamento Remoto, Instituto Nacional de Pesquisas Espaciais. Av. dos Astronautas, 1.758, CEP 12.227-010, São José dos Campos, SP, Brasil (gustavo.nagel@inpe.br).

² Instituto de Ciências Agrárias, Universidade Federal dos Vales do Jequitinhonha e Mucuri. Av. Universitária, 1.000, CEP 38.610-000, Unaí, MG, Brasil (fabricio.terra@ufvjm.edu.br, ingrid.horak@ufvjm.edu.br).

³ Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas. Rua Gomes Carneiro, 1, CEP 96.010-610, Pelotas, RS, Brasil (jade.sdeoliveira@gmail.com, samuel.beskow@ufpel.edu.br)

⁴ Programa de Pós-Graduação em Recursos Hídricos, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas. Rua Gomes Carneiro, 1, CEP 96.010-610, Pelotas, RS, Brasil.

⁵ Programa de Pós-Graduação em Produção Vegetal, Instituto de Ciências Agrárias, Universidade Federal dos Vales do Jequitinhonha e Mucuri. Av. Universitária, 1.000, CEP 38.610-000, Unaí, MG, Brasil.

Resumo. O valor da curva-número (CN) é um parâmetro empírico usado na determinação do escoamento superficial direto a partir dos excessos de precipitações, sendo dependente das mudanças de uso e cobertura da superfície. Imagens de alta resolução espacial são importantes para identificar tais mudanças em bacias hidrográficas urbanas. O objetivo deste trabalho foi comparar os efeitos de diferentes mapas de uso e cobertura, produzidos a partir de classificações não-supervisionada (K-médias) e supervisionadas (MaxVer, SAM e SVM) em uma imagem orbital de alta resolução espacial, no cálculo do valor CN da bacia hidrográfica urbana do Arroio Pepino (Pelotas, RS). A hipótese é de que diferentes algoritmos de classificação produzem diferentes mapas de superfície que por sua vez afetam o valor CN final. As classificações foram realizadas em uma imagem RapidEye e 10 classes foram identificadas: água, asfalto, estrada de terra, vegetação (3 tipos) e coberturas (4 tipos). O valor CN de cada classe foi obtido pela comparação com valores tabulados, e o valor CN total foi calculado pela média ponderada considerando a área proporcional de cada classe. O SVM foi o algoritmo de melhor desempenho (acurácia global de 70,36% e índice kappa de 0,66). Os valores finais de CN apresentaram distintas intensidades: $CN_{total} = 88,96$ para SAM, $CN_{total} = 89,66$ para K-médias, $CN_{total} = 89,94$ para SVM e $CN_{total} = 90,71$ para MaxVer. A proximidade entre estes valores foi influenciada pela baixa capacidade de drenagem da bacia estudada mesmo em áreas vegetadas. Diferenças nas proporções das classes afetam o valor do CN final da bacia, e sua qualidade é altamente dependente da acurácia da imagem classificada.

Palavras-chave. máxima verosimilhança gaussiana, mapeador de ângulo espectral, k-médias, máquina de vetor suporte, hidrologia urbana, escoamento superficial.

Abstract. CURVE NUMBER CALCULATION FOR URBAN WATERSHED USING DIFFERENT CLASSIFICATION APPROACHES FOR RAPIDEYE IMAGE: CASE STUDY FOR THE PEPINO STREAM (PELOTAS, RS). The curve-number (CN) value is an empirical parameter used for predicting direct runoff from rainfall excess, and it depends on land use and cover changes. High spatial resolution images have been important to identify these changes in urban watersheds. This research aimed to compare the effects of different surface maps, produced by unsupervised (K-means) and supervised classifications (MaxVer, SAM, and SVM) of a high spatial resolution orbital image, on calculation of CN value in the urban watershed of the Pepino stream (Pelotas/RS). The hypothesis is that different classification algorithms produce divergent maps which in turn affect the final CN value. A RapidEye image was classified in order to map the surface, and ten classes were identified: water, asphalt, dirt road, vegetation (3 types) and roofs (4 types). The CN value of each class was obtained by comparing to corresponding tabulated values, and the total CN value was calculated by weighted average taking into account the proportional area of each class. The SVM was the best performed classifier (global accuracy: 70.36% and kappa index: 0.66). The final CN values

showed different intensities: $CN_{total} = 88.96$ for SAM, $CN_{total} = 89.66$ for K-means, $CN_{total} = 89.94$ for SVM and $CN_{total} = 90.71$ for MaxVer. The proximity among these values was influenced by the poor drainage capacity of the watershed even in vegetation areas. Differences on proportions of the classes affect the final CN value of the watershed, and its quality is highly dependent on accuracy of the classified image.

Keywords. maximum gaussian likelihood, spectral angle mapper, k-means, support vector machine, urban hidrology, runoff.

1 Introdução

A urbanização traz consigo mudanças significativas nas propriedades físicas da superfície territorial, principalmente no que se refere à redução da relação infiltração/deflúvio da água devido à impermeabilização. Neste contexto, resultados obtidos a partir de estudos hidrológicos urbanos baseados em coleta de dados a campo e modelagens matemáticas tem sido uma necessidade fundamental para o manejo da água objetivando alagamentos, inundações, abastecimento de reservatórios, qualidade da água, etc. (Du *et al.*, 2015; Xu & Zhao, 2016; Hameed, 2017; Li *et al.*, 2018; Ohana-Levi *et al.*, 2017). Na hidrologia urbana torna-se importante a aplicação de imagens orbitais de alta resolução espacial para que seja possível identificar, classificar e mapear com qualidade e acurácia os diferentes usos e coberturas da superfície (Jiménez *et al.*, 2018; Rimal *et al.*, 2018). Esse mapeamento tem sido feito a partir do reconhecimento de padrões espectrais de alvos urbanos e pode ser empregado na identificação das mudanças da superfície e de fatores que afetem a sustentabilidade urbana (Kadhim *et al.*, 2016; Şatir, 2016).

Para o mapeamento do uso e cobertura do solo é indispensável a classificação de imagens orbitais, supervisionada ou não, baseada nas energias armazenadas pelos pixels que as compõem, pois, assim, é possível a identificação e o agrupamento de alvos comuns presentes dentro do perímetro investigado. Estas energias armazenadas no formato de imagem são refletidas pelos diferentes alvos que fazem parte da superfície terrestre e representam os seus comportamentos espectrais. Segundo Meneses & Almeida (2012), a classificação de imagens multispectrais é o processo de associar pixels de uma imagem a um número adequado de classes que representem alvos do mundo real. No entanto,

segundo Hosseini & Kandovan (2017), a acurácia dessas classificações depende de fatores como a resolução espectral da imagem, a quantidade de amostras de pixels conhecidos (para a classificação supervisionada) e do número de alvos na imagem. Segundo os mesmos autores, imagens com grande número de classes, como em ambientes urbanos, aumentam a complexidade computacional e, portanto, as confusões de classificação. Dentre os diversos algoritmos de classificação baseados no reconhecimento de padrões espectrais de imagens e clusterização, destacam-se: Mapeador de Ângulo Espectral (*Spectral Angle Mapper* - SAM) (Kruse *et al.*, 1993), Máxima Verossimilhança Gaussiana (*Maximum Gaussian Likelihood* - MaxVer), K-médias (*K-means*) (Richards & Jia, 2006) e Máquina de Vetor Suporte (*Support Vector Machine* - SVM) (Mountrakis *et al.*, 2011).

O método curva-número (CN) visa o cálculo do deflúvio total gerado em uma bacia hidrográfica a partir da chuva efetiva, segundo o Departamento de Agricultura dos Estados Unidos (USDA, 1986), e seus valores variam em uma escala de 1 a 100, que se relaciona ao volume de água escoada superficialmente (deflúvio). Esses valores estão diretamente relacionados ao uso e cobertura da bacia, assim como à capacidade de infiltração do solo (Mello & Silva, 2013; Saviddou *et al.*, 2018). O cálculo do deflúvio é importante para o planejamento de estruturas hidráulicas de grande ou pequeno porte e para o manejo de descargas de uma determinada bacia, com o emprego de metodologias que utilizam o CN, como a do Hidrograma Triangular (Banasik *et al.*, 2014; Kowalik & Walega, 2015; Maria Filho *et al.*, 2016; Mekonnen *et al.*, 2016; Zelelew, 2017). A aplicação do método CN é recomendada somente para bacias hidrográficas, urbanas ou não, relativamente pequenas, cujas áreas não excedam 500 ha (Smedema & Rycroft, 1983). É possível a utilização de imagens orbitais de alta

resolução espacial no cálculo do método CN em bacias hidrográficas urbanas (BHU), uma vez o processo de urbanização modifica as coberturas de pequenas áreas da superfície diminuindo a sua capacidade de infiltração.

De acordo com Zope *et al.* (2015), Pradhan-Salike & Pokharel (2017), Shanableh *et al.* (2018), Webber *et al.* (2018), Wu *et al.* (2018) e Yao *et al.* (2018), inundações devido aos excessos de escoamento superficial e à insuficiente capacidade de sistemas de drenagem são problemas frequentes em regiões urbanizadas no mundo. Ainda, estudos identificaram que a conversão do uso de terra agrícola para urbana tem contribuído significativamente para o aumento de risco à inundações (Tripathi *et al.*, 2014; Bajracharya *et al.*, 2015; Miller & Hess, 2017). Como consequência, eventos de inundação afetam milhões de pessoas e causam danos econômicos para cidades e países (Brown *et al.*, 2016; Zimmermann *et al.*, 2016; Spekkers *et al.*, 2017; Luo *et al.*, 2018; Zhou *et al.*, 2018). A BHU do Arroio Pepino, localizada no município de Pelotas (RS), vem sofrendo intenso processo de urbanização ocasionando maior escoamento superficial em períodos chuvosos e, conseqüentemente, inundações na respectiva região de abrangência (Cardoso, 2012).

A hipótese deste estudo é que diferentes algoritmos de classificação de imagens produzem mapas temáticos de uso e cobertura da superfície divergentes que, por sua vez, afetam o valor CN final de uma dada BHU. Sendo assim, os nossos objetivos foram: a) comparar o efeito de diferentes mapas produzidos através de classificações não-supervisionada (K-médias) e supervisionadas (SAM, MaxVer e SVM), em uma imagem orbital de alta resolução espacial, no cálculo do método CN para a BHU do Arroio Pepino; e b) propor uma abordagem para o cálculo do método CN em BHU baseada em mapas de uso e cobertura da superfície terrestre, obtidos a partir de imagens orbitais de alta resolução espacial.

2 Material e métodos

O local de estudo corresponde à BHU do Arroio Pepino que possui uma área de

19,89 km² e está situada no perímetro urbano do município de Pelotas (RS) (Fig. 1). A BHU investigada encontra-se sobre antigas áreas de várzeas (planossolos, gleissolos e organossolos) que foram aterradas e que apresentam pouca diferença de cota topográfica. A delimitação da bacia foi obtida através dos processos de digitalização, georreferenciamento e vetorização do mapa disponível em Cardoso (2012). De acordo com Saraiva & Polidori (2008), os bairros presentes na BHU do Arroio Pepino são 4: os bairros Areal (com 14% da sua área dentro da bacia), Centro (37%), São Gonçalo (65%) e Três Vendas (8%). O divisor de águas da bacia passa pela zona central (região de maior densidade populacional), enquanto o curso d'água foi considerado como limite para a delimitação dos bairros mencionados.

Como informação orbital, utilizou-se uma imagem do sensor RapidEye (Alves Júnior *et al.*, 2015) adquirida em 09/11/2014, com resolução espacial de 5 m após ortorretificação, resolução radiométrica de 12 bits e resolução espectral como segue: banda 1 (azul: 440-510 nm), banda 2 (verde: 520-590 nm), banda 3 (vermelho: 630-685 nm), banda 4 (vermelho limítrofe: 690-730 nm) e banda 5 (infravermelho próximo: 760-850 nm). Esta imagem foi adquirida junto ao Geo Catálogo do Ministério do Meio Ambiente (MMA, 2014) através do Programa de Regularização Ambiental (instituído pelo Decreto Presidencial 7.830).

Para as classificações da imagem foi utilizado software especializado em processamento digital de imagens. Na classificação não-supervisionada, onde não há conhecimento prévio sobre as classes existentes na imagem, utilizou-se o algoritmo K-médias, que objetiva diminuir a variabilidade interna das classes geradas a partir da soma da distância ao quadrado (erros) entre cada pixel e o centro da classe ao qual foi alocado (Richards & Jia, 2006; Abbas *et al.*, 2016; Sowan & Qattous, 2017; Santos *et al.*, 2018). Para o funcionamento deste algoritmo, a priori apenas são informadas a quantidade de classes que se deseja criar, o número de iterações e o limiar da classificação, ou seja, a porcentagem máxima de pixels que não mudarão de classes entre as iterações

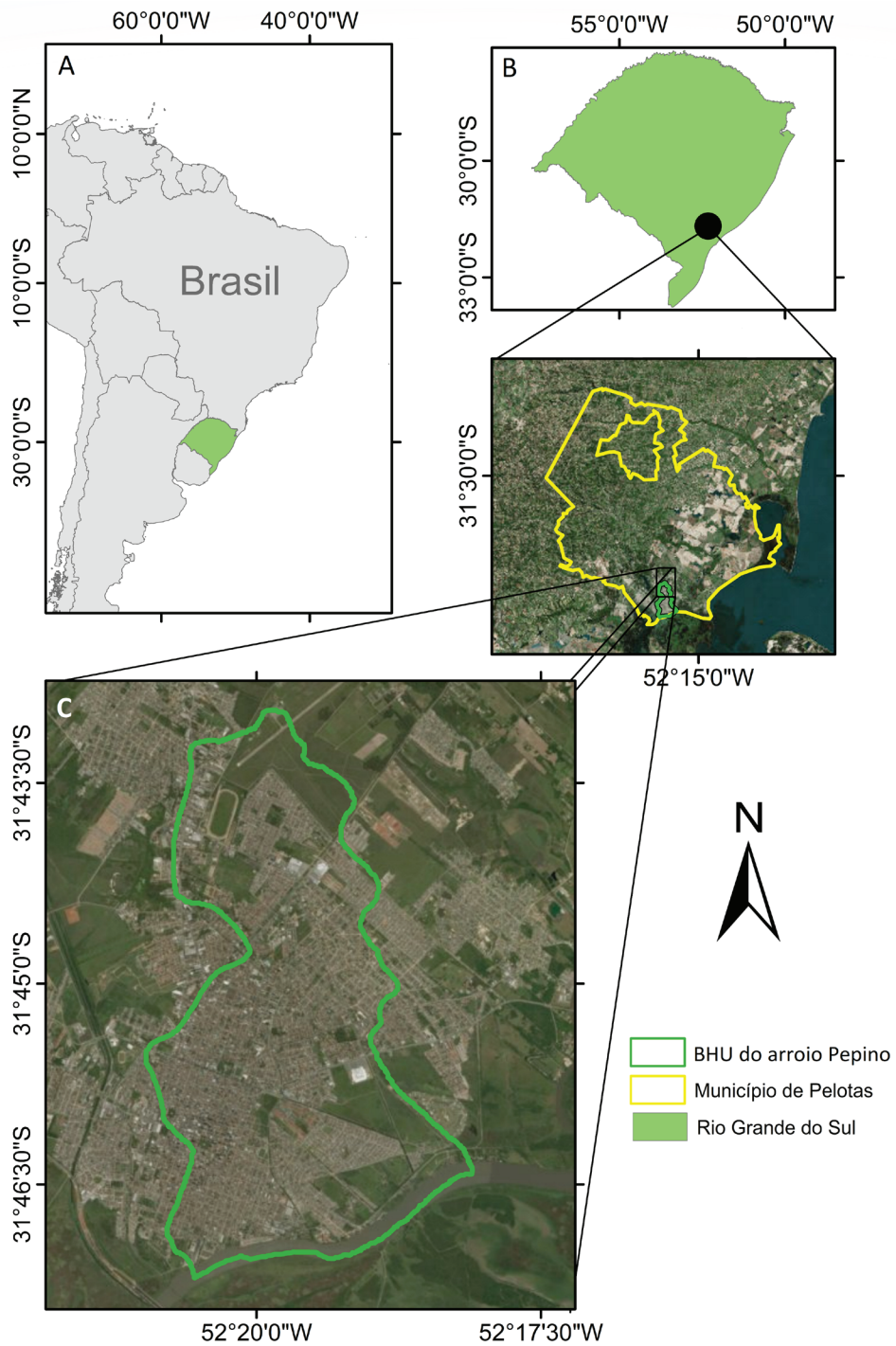


Figura 1. Mapa de localização da área de estudo. A) Estado do Rio Grande do Sul; B) Município de Pelotas; C) BHU do Arroio Pepino (baseada em imagem do Google Earth Pro).
Figure 1. Location map of the study area. A) State of Rio Grande do Sul; B) Municipality of Pelotas; C) Pepino stream BHU (based on image from the Google Earth Pro).

(Meneses & Almeida, 2012). O algoritmo divide os n pixels da imagem em k diferentes grupos (classes), baseados nas suas dissimilaridades espectrais, entre os pixels e os centróides das classes, conforme equação 1 (Naldi *et al.*, 2011):

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2 \quad (1)$$

Na equação 1, J é a função objetiva a ser minimizada, j é classe em questão analisada (variando de 1 a k), k é o número total de classes definido pelo analista (k ésima classe), i é o pixel em questão analisado (variando de 1 a n), n é o número total de pixels da imagem (n ésimo pixel), x_i é o vetor composto por todas as bandas do pixel i , c_j é o centróide do classe j e $\|x_i^{(j)} - c_j\|^2$ é a função distância (distância euclidiana). As informações das bandas estão contidas em cada pixel analisado e no centróide dos grupos. Foram geradas 10 classes a partir de 10 interações com um limiar de 5%.

Na classificação supervisionada, o objetivo é classificar pixels de identidade desconhecida a partir de pixels amostrais de identidade conhecida, os quais foram selecionados a partir de amostras coletadas visualmente na imagem RapidEye. Dessa forma, os algoritmos precisaram ser previamente treinados pelo analista utilizando pixels amostrais representativos das classes que se desejou criar. Para cada classe, foram utilizados aproximadamente 500 pixels para o conjunto de treinamento e outros 500 pixels para o teste dos classificadores avaliados. Os algoritmos utilizados para este fim foram: MaxVer, SAM e SVM. Por meio da interpretação visual da imagem foram estabelecidas 10 classes: água, asfalto, estrada de terra, vegetações (arbórea, arbustiva e herbácea), coberturas (telhados de barro, de zinco, de concreto e escuro). Os pixels utilizados como amostras de treinamento e teste para cada classe pré-definida foram os mesmos para os três algoritmos de classificação investigados, a fim de possibilitar a comparação entre resultados.

O algoritmo MaxVer avalia a variância e covariância das categorias dos padrões de resposta espectral para a classificação de

pixels desconhecidos, assumindo-se que a distribuição dos valores dos pixels em cada banda nas classes é Gaussiana (distribuição normal). A partir desta suposição, calcula-se a probabilidade de um dado pixel pertencer a uma classe específica (Omo-Irabor, 2016; Wang *et al.*, 2018). O pixel será classificado de acordo com a classe que apresentar maior probabilidade de pertencimento. O limiar (*threshold*) nesse algoritmo é a mínima probabilidade aceitável para que o pixel seja classificado em uma determinada classe (Meneses & Almeida, 2012). O algoritmo MaxVer é baseado no teorema de Bayes (equação 2):

$$D = \ln(a_c) - [0,5 \ln(|Cov_c|)] - [0,5(X - M_c)T(Cov_c^{-1})(X - M_c)] \quad (2)$$

onde a_c é a probabilidade de que cada pixel é membro de uma classe c (o padrão é 1.0), Cov_c é a covariância da matriz de pixels da classe c , X é o vetor de um pixel candidato, M_c é o vetor médio de uma classe c e T é a função de transposição. D é o peso da distância, que se resume na probabilidade do pixel pertencer ao vetor da classe M_c , sendo a variável c uma determinada classe. Os vetores, seja do pixel candidato ou de uma determinada classe, são calculados com base nos valores das bandas da imagem. O vetor candidato irá pertencer à classe a que possui a maior probabilidade, ou seja, o maior valor de D . O limiar de 0 % foi utilizado para que todos os pixels fossem classificados, a fim de possibilitar a comparação com o algoritmo K-médias.

O algoritmo SAM determina a similaridade espectral entre os pixels da imagem e os pixels utilizados como amostra de referência, a partir do cálculo do ângulo formado entre os espectros (Kruse *et al.*, 1993; Csendes & Mucsi, 2016). O classificador trata os espectros como se fossem vetores em um ambiente onde o número de dimensões é o mesmo do número de bandas da imagem. Pequenos ângulos formados entre pixels da imagem e referência significam maior correlação entre esses, enquanto grandes ângulos significam menor correlação. Pixels de um mesmo alvo localizados em regiões com pouca intensidade de iluminação serão

localizados mais próximos da origem (pixels negros), quando comparados aos pixels com alta exposição à iluminação. Porém, o ângulo formado será o mesmo para ambos os pixels independentemente da condição de iluminação (Kruse *et al.*, 1993). O limiar neste algoritmo é o máximo ângulo permitido, onde pixels com ângulo maior do que o limiar não serão classificados. Neste trabalho, foi utilizado o limiar máximo de 1 radiano, para que todos os pixels pudessem ser classificados. Para um dado pixel, o algoritmo SAM determina a similaridade entre o seu valor de reflectância (t) e a reflectância de referência (r), para uma dada banda (i) até o número total de bandas (nb), através do arccosseno (\cos^{-1}) do produto das reflectâncias utilizadas na classificação apresentada na equação 3 (Kruse *et al.*, 1993; Csendes & Mucsi, 2016):

$$\cos^{-1} \left(\frac{\sum_{i=1}^{nb} t_i r_i}{\left(\sum_{i=1}^{nb} t^2 \right)^{1/2} \times \left(\sum_{i=1}^{nb} r^2 \right)^{1/2}} \right) \quad (3)$$

O algoritmo SVM tem seu método baseado na teoria da aprendizagem estatística, onde o objetivo é determinar um hiperplano que separa um conjunto de dados em diferentes classes, a partir de um processo iterativo com pixels de treinamento (Ustener *et al.*, 2015; Mustafa *et al.*, 2018). O hiperplano de melhor ajuste é o que possui a maior distância em relação aos pixels vizinhos de ambas as classes. Os dados analisados no algoritmo são transformados em vetores, que possuem características numéricas atribuídas para cada banda (Mountrakis *et al.*, 2011; Hosseini & Kandovan, 2017; Varma & Nynalasetti, 2017). O hiperplano do algoritmo pode ser definido a partir da equação ($w \times x + b = 0$), onde x é o ponto do eixo x do hiperplano, w é a normal do hiperplano e b é a distância do hiperplano à origem. A variável ($\|w\|$) corresponde à normal euclidiana de w . A separação linear de um hiperplano é definida a partir da comparação entre duas classes (y) para uma banda (i) $y_i \in \{-1, 1\}$, em que $w \times x + b \geq +1$ para $y_i = +1$ e $w \times x + b \leq -1$ para $y_i = -1$. As equações podem ser combinadas na equação 4 (Foody & Mathur, 2004):

$$y_i(w \times x + b) - 1 \geq 0 \quad (4)$$

Os dados de treinamento são chamados de vetores suporte, e os vetores de duas classes são paralelos ao hiperplano que os separa. A margem entre esses planos é $(2)/(\|w\|)$, e o hiperplano que melhor se ajusta é determinado a partir da solução da equação 5:

$$\min \left\{ \frac{1}{2} \|w\|^2 \right\} \quad (5)$$

Para avaliar o desempenho dos classificadores, foram utilizados a Exatidão Global (EG , Eq. 6), e o Índice Kappa (K , Eq. 7), calculados a partir da matriz confusão (Cohen, 1960):

$$EG = \left(\frac{A}{N} \right) \times 100 \quad (6)$$

$$K = \frac{\left[N \times \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_i \times x_{i+1}) \right]}{\left[N^2 - \sum_{i=1}^r (x_i \times x_{i+1}) \right]} \quad (7)$$

onde A é o total de pixels classificados corretamente, N é o total de pixels correspondentes às amostras de teste do classificador, r é o número de linhas da matriz, x_{ii} é o número de pixels na linha i e coluna i e x_{i+} e x_{+i} são respectivamente os totais marginais da linha i e coluna i .

Uma vez que a imagem RapidEye utilizada é de 2014, para geração da matriz confusão, as amostras para teste dos algoritmos de classificação foram obtidas através da interpretação visual da imagem do *Google Earth (Digital Globe Quickbird)*, da mesma época (08/11/2014), e da delimitação manual (adição de polígonos) de algumas áreas complementares referentes às 10 classes investigadas e diferentes das usadas para treinamento. Os polígonos correspondentes às amostras teste foram então individualmente exportados no formato *Keyhole Markup Language - Kml (*.kml)*, para serem posteriormente convertidos para o formato vetorial *Shapefile - Shp (*.shp)* e agrupados conforme as suas respectivas classes. Por fim, o

arquivo vetorial gerado foi sobreposto à imagem RapidEye no software de processamento digital de imagens para que os pixels envolvidos pelos polígonos e correspondentes às classes pudessem ser identificados como as amostras teste.

As imagens classificadas foram então transportadas para o software de geoprocessamento para que as demais etapas pudessem ser realizadas. Os pixels pertencentes à delimitação da BHU foram extraídos das imagens classificadas utilizando o vetor da área de estudo como máscara. Posteriormente, os conjuntos de pixels pertencentes às mesmas classes e próximos entre si foram agrupados e transformados em polígonos. Os polígonos pertencentes à mesma classe e separados espacialmente foram agrupados em uma única entidade. As nuvens existentes na cena foram removidas posteriormente, pois o efeito de suas sombras causa confusão nas classificações. Após estes pré-processamentos, para cada imagem classificada, calcularam-se as áreas referentes às classes e suas proporções (em %) em relação à área total da BHU.

O cálculo do método CN foi adaptado e obtido de forma ponderada considerando as porcentagens da área de cada classe a fim de obter-se um único valor representativo para toda a BHU (Eq. 8):

$$CN_{total} = \frac{\sum_{i=1}^n (CN_i \times A_i)}{\sum_{i=1}^n (A_i)} \quad (8)$$

onde CN é o valor curva-número (adimensional), A é a área de cada classe (i) e n é o número total de classes.

Segundo Sartori *et al.* (2005a, 2005b), pelo fato da bacia estar, predominantemente, em área de várzea e sobre solos pouco resistentes à erosão, rasos, com baixa taxa de infiltração e de drenagem, bem como com razão textural variável (planossolos, gleissolos e organossolos), a BHU enquadrou-se no grupo hidrológico D proposto pelo método CN. A figura 2 mostra a metodologia do trabalho em formato de fluxograma.

Algumas adaptações e considerações foram feitas para associar as classes de uso e cobertura obtidas da superfície com as categorias tabeladas do valor CN, propostas por Mello & Silva (2013). Para os alvos correspondentes à cobertura vegetal, se considerou três categorias: herbáceas (gramíneas de poucos centímetros), arbustivas (arbustos e vegetação mais desenvolvida de poucos metros de altura) e arbóreas (árvores bem desenvolvidas). As considerações estão expressas no quadro 1.

3 Resultados e discussão

3.1 Desempenho da classificação não-supervisionada

Tomando como base as amostras de teste, classificações errôneas acima de 10% foram consideradas as principais confusões. Na classificação K-médias, os principais erros observados ocorreram entre as classes: a) asfalto e telhado de zinco (22,07%); b) telhado de concreto e estrada de terra (21,85%), seguida de telhado de zinco (10,57%); c) telhado escuro e asfalto (11,93%); d) telhado de barro e de zinco (33,33%), seguido por estrada de terra (15,02%) e asfalto (10,30%); e) telhado de zinco e asfalto (33,36%), seguido por estrada de terra (20,76%); f) vegetação arbórea e herbácea (41%), seguida por arbustiva (29,39%); e g) vegetação arbustiva e herbácea (47,84%), seguida por arbórea (19,88%) (Tab. 1).

De forma geral, a similaridade no comportamento espectral entre alguns dos alvos estudados, devido à proximidade das suas constituições físico-químicas, associada à baixa resolução espectral do sensor RapiEye ajudam a explicar os erros de classificação. Tais confusões foram intensificadas, pois não houve treinamento prévio com pixels dos alvos conhecidos para este algoritmo.

No caso do asfalto, a semelhança espectral desse alvo com o telhado de zinco corrobora a explicação anterior. A classe telhado de barro foi confundida com estrada de terra em consequência da similaridade espectral entre ambas, o que pode ser explicada pela mesma natureza constituicional dos dois alvos.

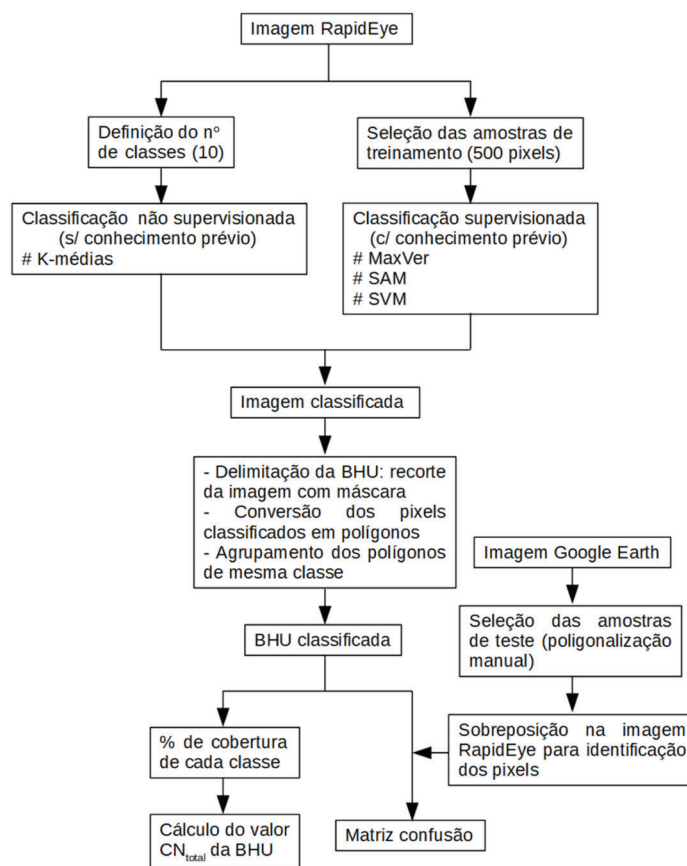


Figura 2. Fluxograma da metodologia para a obtenção do valor de CN e para a determinação do melhor classificador.
 Figure 2. Methodological flow chart for obtaining the CN value and for determination of the best classifier.

Quadro 1. Classes de uso e cobertura da superfície e suas correspondentes categorias tabeladas para cada valor de CN (adaptado de Melo & Silva, 2013).
 Chart 1. Classes of surface use and coverage and their corresponding categories tabulated for each CN value (adapted from Melo & Silva, 2013).

Classes	Categorias tabeladas	CN
Asfalto	Ruas e estradas ¹	93
Estrada de terra	Ruas e estradas ²	89
Telhado de barro	Lotes urbanos	94
Telhado de zinco	Lotes urbanos	94
Telhado escuro	Lotes urbanos	94
Telhado de concreto	Lotes urbanos	94
Vegetação arbórea	Florestas normais	76
Vegetação herbácea	Campos permanentes ³	83
Vegetação arbustiva	Campos permanentes ⁴	79
Água	-	-

¹Pavimentadas e coletores abertos; ²Estradas de terra; ³Esparsos de baixa transpiração; ⁴Normais.

Telhado de barro confundiu-se com a classe asfalto em regiões onde o pavimento possui uma película de solo, principalmente em locais próximos a estradas de terra. Porém, a confusão entre telhados de barro e de zinco é devida, principalmente, à proximidade dos dois alvos nos bairros da cidade, o que faz com que ocorra influência das duas classes em um mesmo pixel. Quanto às classes de vegetação, a maior parte dos pixels de vegetação arbórea e arbustiva foi atribuída à vegetação herbácea neste algoritmo, provavelmente, pelas mesmas explicações mencionadas anteriormente.

3.2 Desempenho das classificações supervisionadas

As maiores confusões (> 10 %) do algoritmo MaxVer foram relacionadas às classes: a) asfalto e telhado de zinco (26,60%); b) telhado escuro e asfalto (17,82 %), seguido por telhado de zinco (13,44%); c) telhado de zinco e asfalto (13,74%), seguido por telhado de concreto (10,76%); d) vegetação herbácea e estrada de terra (14,67%), seguida por vegetação arbustiva (12,07%); e e) vegetações arbórea e arbustiva (69,10%), seguida pela herbácea (19,82%), conforme tabela 2.

As confusões entre vegetações arbórea e arbustiva ficaram bastante evidentes no algoritmo Maxver, que classificou corretamente apenas 0,57% dos pixels correspondentes à classe de vegetação arbórea. Vale ressaltar que

as confusões entre as classes de vegetação (herbácea, arbórea e arbustiva) e estrada de terra ocorreram, principalmente, devido as suas proximidades em regiões residenciais da cidade. Nota-se que o algoritmo classificou erroneamente pixels correspondentes às amostras teste de vegetação herbácea como estrada de terra e telhado de zinco, mas pixels correspondentes às amostras teste de estrada e telhado não foram classificados como vegetação (Tab. 2). Isso mostrou certa tendência do algoritmo em favorecer classes que não apresentaram vegetação em pixels com mistura espectral.

Para o algoritmo SAM, as maiores confusões (> 10%) ocorreram entre as classes: a) asfalto e telhado de zinco (15,92%); b) estrada de terra e telhado de zinco (12,47%), seguido pelo telhado de barro (11,60%); c) telhado escuro e asfalto (14,64%), seguido pelo telhado de zinco (12,15%); d) telhado de barro e estrada de terra (34,44%), seguida pelo telhado de zinco (14,60%); e) telhado de zinco e asfalto (38,27%); f) vegetações arbórea e herbácea (30,21%); e g) vegetações arbustiva e arbórea (80,31%), seguida pela herbácea (16,13%), conforme tabela 3. Este algoritmo obteve excelente desempenho para a classificação da vegetação herbácea, com 87,83% de acerto. No entanto, o SAM produziu a confusão mais proeminente, para classes de mesma natureza, entre todos os algoritmos de classificação (Tab. 1, 2, 3 e 4).

Tabela 1. Matriz confusão do algoritmo K-médias.
Table 1. Confusion matrix of the K-means algorithm.

Classes reais (amostras teste)	Classes previstas									
	1	2	3	4	5	6	7	8	9	10
(1) Asfalto	71,74	1,74	0,06	1,16	1,44	22,07	1,80	0,00	0,00	0,00
(2) Estrada de terra	0,10	83,54	5,42	0,00	5,32	5,62	0,00	0,00	0,00	0,00
(3) Telhado de concreto	0,82	21,85	66,61	0,06	0,09	10,57	0,00	0,00	0,00	0,00
(4) Telhado escuro	11,93	0,00	0,00	78,85	0,00	3,63	1,81	0,00	0,00	3,78
(5) Telhado de barro	10,30	15,02	0,61	2,08	35,89	33,33	2,77	0,00	0,00	0,00
(6) Telhado de zinco	33,36	20,76	0,46	0,84	4,43	39,16	0,99	0,00	0,00	0,00
(7) Vegetação herbácea	0,00	0,00	1,75	0,00	0,44	0,88	95,22	0,00	0,00	0,00
(8) Água	0,00	0,68	0,00	0,07	3,45	0,43	0,00	96,94	0,05	0,10
(9) Vegetação arbórea	0,00	0,53	0,11	0,38	0,98	1,06	41,00	0,00	26,55	29,39
(10) Vegetação arbustiva	0,09	0,00	0,00	0,00	0,51	0,32	47,84	0,00	19,88	31,36

Tabela 2. Matriz confusão do algoritmo MaxVer.
 Table 2. Confusion matrix of the MaxVer algorithm.

Classes reais (amostras teste)	Classes preditas									
	1	2	3	4	5	6	7	8	9	10
(1) Asfalto	66,69	1,52	2,96	0,14	2,10	26,60	0,00	0,00	0,00	0,00
(2) Estrada de terra	0,10	85,58	1,02	0,00	4,81	8,49	0,00	0,00	0,00	0,00
(3) Telhado de concreto	0,64	1,19	94,45	0,00	0,00	3,72	0,00	0,00	0,00	0,00
(4) Telhado escuro	17,82	0,00	1,21	62,99	3,47	13,44	1,06	0,00	0,00	0,00
(5) Telhado de barro	2,86	5,67	1,30	0,09	81,77	7,97	0,35	0,00	0,00	0,00
(6) Telhado de zinco	13,74	5,57	10,76	0,08	1,07	68,78	0,00	0,00	0,00	0,00
(7) Vegetação herbácea	0,00	14,67	0,02	0,00	0,05	7,52	65,67	0,00	0,00	12,07
(8) Água	0,00	0,18	2,19	0,00	0,00	0,70	0,00	96,94	0,00	0,00
(9) Vegetação arbórea	0,00	3,63	0,08	0,00	0,08	6,73	19,82	0,00	0,57	69,10
(10) Vegetação arbustiva	0,00	9,55	0,00	0,00	0,00	1,93	6,84	0,00	0,00	81,63

Por meio da matriz confusão (Tab. 3), observou-se que uma pequena porcentagem de amostras teste correspondentes à vegetação herbácea foi classificada como alvos urbanos (estrada de terra e telhados de barro e zinco), classes normalmente muito próximas em bairros residenciais da BHU. No entanto, uma maior quantidade de amostras teste de asfalto, estrada de terra e telhados (escuro, de barro e de zinco) foi atribuída principalmente à vegetação herbácea. Desta forma, percebeu-se que o algoritmo SAM apresentou uma tendência em classificar pixels com mistura espectral como alvos de vegetação. Pixels correspondentes à água, assim como nos outros algoritmos, foram corretamente classificados. No entanto, este bom resultado restringiu-se as grandes massas de água, como o canal São Gonçalo, na parte sul da BHU (Fig. 1). Nos canais menores, como o Arroio Pepino, houve confusões com telhado escuro e asfalto, devido à coloração escura do canal em regiões de sombra, e com vegetação herbácea, em razão da grande eutrofização dos canais da cidade com predomínio de algas que se assemelham às classes de vegetação definidas.

As principais confusões (> 10%) geradas pelo algoritmo SVM foram entre as classes: a) asfalto e telhado de zinco (18,70%); b) estrada de terra e telhado de zinco (12,27%); c) telhado de concreto e asfalto (14,59%); d) telhado escuro e asfalto (22,81%); e) telhado de barro e de zinco (12,55%); f) telhado de zinco e asfalto (18,02%); g) vegetação herbácea e arbustiva (17,12%); h) vegetação arbórea e herbácea (34,04%), seguida pela arbustiva (25,23%); e i) vegetação arbustiva

e arbórea (36,91%), seguida pela herbácea (15,56%), conforme se verifica na tabela 4.

Como já mencionado, parte destes erros de classificação deveu-se, principalmente, à similaridade espectral entre os alvos em questão. As confusões entre as classes estrada de terra e telhado de zinco, bem como entre as classes telhado de barro e de zinco, provavelmente ocorreram devido à proximidade destes alvos em regiões residenciais (pixels de 5 metros com resposta espectral de diferentes classes). O algoritmo SVM obteve os melhores resultados para o alvo asfalto, quando comparado aos demais algoritmos (Tab. 1, 2, 3 e 4). Os maiores equívocos deste classificador, observado também pelos demais, foi em relação às classes de vegetação. Houve confusões entre os alvos de vegetação arbórea, arbustiva e herbácea. A baixa resolução espectral da imagem contribuiu para este erro de classificação.

3.3 Comparação entre algoritmos de classificação

Para cada classe investigada, os algoritmos de melhor desempenho e que obtiveram as maiores taxas de acerto foram: K-médias para as classes telhado escuro (78,85%) e vegetação herbácea (95,22%), tabela 1, Maxver para as classes estrada de terra (85,58%), telhados de concreto (94,45%), barro (81,77%), zinco (68,78%) e vegetação arbustiva (81,63%), tabela 2, SAM para as classes vegetação arbórea (67,88%) e água (98,95%), tabela 3, e SVM para a classe asfalto (76,93%), tabela 4.

Tabela 3. Matriz confusão do algoritmo SAM.
 Table 3. Confusion matrix of the SAM algorithm.

Classes reais (amostras teste)	Classes preditas									
	1	2	3	4	5	6	7	8	9	10
(1) Asfalto	68,42	1,47	3,24	7,00	0,62	15,92	2,12	1,21	0,00	0,00
(2) Estrada de terra	0,44	73,96	0,22	0,00	11,60	12,47	1,31	0,00	0,00	0,00
(3) Telhado de concreto	6,05	0,03	91,99	0,24	0,00	1,04	0,03	0,61	0,00	0,00
(4) Telhado escuro	14,64	0,00	0,00	71,96	0,00	12,15	1,25	0,00	0,00	0,00
(5) Telhado de barro	4,77	34,44	0,66	0,09	41,62	14,60	3,78	0,00	0,05	0,00
(6) Telhado de zinco	38,27	2,57	4,45	1,63	0,00	51,20	1,88	0,00	0,00	0,00
(7) Vegetação herbácea	0,00	0,40	0,00	0,00	0,05	0,07	87,83	0,00	9,70	1,95
(8) Água	0,35	0,00	0,09	0,35	0,00	0,09	0,18	98,95	0,00	0,00
(9) Vegetação arbórea	0,08	0,08	0,00	0,00	0,04	0,16	30,21	0,00	67,88	1,56
(10) Vegetação arbustiva	0,00	0,00	0,00	0,05	0,00	0,00	16,13	0,00	80,31	3,51

Tabela 4. Matriz confusão do algoritmo SVM.
 Table 4. Confusion matrix of the SVM algorithm.

Classes reais (amostras teste)	Classes preditas									
	1	2	3	4	5	6	7	8	9	10
(1) Asfalto	76,93	1,10	0,19	0,14	1,69	18,70	1,24	0,00	0,00	0,00
(2) Estrada de terra	0,10	80,78	0,31	0,00	6,54	12,27	0,00	0,00	0,00	0,00
(3) Telhado de concreto	14,59	1,10	82,63	0,00	0,00	1,68	0,00	0,00	0,00	0,00
(4) Telhado escuro	22,81	0,00	0,00	70,39	1,06	1,81	3,93	0,00	0,00	0,00
(5) Telhado de barro	4,46	3,98	0,22	0,30	76,10	12,55	2,29	0,00	0,00	0,09
(6) Telhado de zinco	18,02	5,56	6,41	0,00	0,31	68,63	0,84	0,00	0,00	0,15
(7) Vegetação herbácea	0,00	0,78	0,00	0,00	0,00	7,42	74,68	0,00	0,00	17,12
(8) Água	0,18	0,00	1,66	0,00	0,00	1,23	0,00	96,94	0,00	0,00
(9) Vegetação arbórea	0,00	0,49	0,08	0,00	0,00	2,76	34,04	0,00	37,41	25,23
(10) Vegetação arbustiva	0,00	0,00	0,00	0,00	0,00	1,84	15,56	0,00	36,91	45,68

No geral, o algoritmo SVM obteve o melhor desempenho de classificação com os maiores valores de EG (70,36%) e K (0,66), apesar de ter apresentado a maior taxa de acerto apenas para a classe asfalto, quando comparado aos demais algoritmos, como mencionado anteriormente. O melhor desempenho do algoritmo SVM explicou-se pela homogeneidade das taxas de acerto entre as classes investigadas, onde mais 70 % dessas apresentaram taxas superiores a 70 % (Tab. 4). Para os demais classificadores, apenas 50 % das classes apresentaram taxas de acerto superiores a 70 % (Tab. 1, 2 e 3). O algoritmo SAM apresentou a segunda melhor performance de classificação (EG: 68,95% e K: 0,64), seguido dos classificadores MaxVer (EG: 67,98% e K: 0,63) e K-médias (EG: 64,79% e K: 0,59).

Todos os classificadores obtiveram bons resultados para as classes telhado de concreto ($66,71\% \leq \text{taxa de acerto} \leq 94,45\%$) e água ($96,94\% \leq \text{taxa de acerto} \leq 98,95\%$). No

entanto, para as classes de vegetação, houve muita divergência dos resultados obtidos entre os diferentes algoritmos de classificação. Para todos os classificadores, as confusões mais proeminentes ocorreram entre as vegetações herbáceas, arbustivas e arbóreas. Uma das razões para tais confusões foi a falta de amostras florestais na imagem, utilizadas na etapa de treinamento dos classificadores, e a moderada resolução espectral da imagem que, somadas a pouca diferença de comportamento espectral entre esses alvos de vegetação, ampliou os erros de classificação.

Os resultados obtidos corroboram com Pires *et al.* (2014), os quais encontraram que, para uma imagem rural da série Landsat, o classificador Maxver teve melhor desempenho que o algoritmo K-médias, e com Simões *et al.* (2015), que obteve melhor desempenho do SAM em comparação ao K-médias. Moughal (2013) identificou o algoritmo SVM como o de melhor

desempenho para uma região urbana, seguido dos classificadores SAM e MaxVer, exatamente como os resultados observados neste estudo. Esta mesma sequência (SVM, SAM e MaxVer) de melhores resultados de classificação também foi encontrada por Moura (2016) para uma imagem RapidEye. Para Hedge *et al.* (2014), os resultados da classificação também foram melhores com o algoritmo SVM em relação ao SAM para uma imagem hiperespectral de área urbana. No entanto, Cattani *et al.* (2013) obtiveram resultados diferentes para classificação utilizando imagem RapidEye de uma zona rural, onde o melhor desempenho foi encontrado para o algoritmo MaxVer em comparação ao SAM.

Considerando o índice Kappa, os resultados das classificações foram fortes ($0,61 < K < 0,8$) para os algoritmos SVM, MaxVer e SAM, e moderados ($0,41 < K < 0,6$) para o algoritmo K-médias, segundo Landis & Koch (1977). Uma possível justificativa para os desempenhos encontrados pelos algoritmos testados pode estar na moderada resolução espectral da imagem utilizada, uma vez que os algoritmos em questão utilizam os valores de reflectância dos alvos em cada uma das bandas empregadas para realizar as suas respectivas comparações (Richards & Jia, 2006; Mountrakis *et al.*, 2011; Meneses & Almeida, 2012). Portanto, quanto maior a resolução espectral da imagem, maiores serão as chances de se distinguirem os alvos. No entanto, a alta resolução espacial da imagem pode ter contribuído na melhor distinção entre os pixels, melhorando o desempenho dos classificadores, uma vez que alvos puderam ser mais bem discriminados.

3.4 Distribuição espacial e proporção das classes

Através das figuras 3 e 4, onde são respectivamente mostrados os mapas obtidos por cada um dos classificadores estudados e as proporções das classes, pode-se observar que o algoritmo SAM produziu o mapa com a maior porcentagem de área verde, de maior permeabilidade/infiltração, atribuída às classes de vegetação arbórea, herbácea e arbustiva, ocupando 34,26% da área de estudo. O algoritmo

K-médias apresentou a segunda maior proporção com 29,70% de ocupação da BHU, seguido pelo SVM com 28,18% e por último o classificador MaxVer com apenas 19,94% do mapa atribuído às classes de vegetação. É visualmente perceptível que o mapa gerado pelo classificador SAM apresenta muito mais áreas de vegetação (coloração verde e amarela), ao contrário do algoritmo MaxVer, com mais áreas com cores correspondentes às classes urbanas (Fig. 3 e 4).

Gadal & Ouerghemmi (2016) encontraram resultado semelhante em uma região urbana da Lituânia, onde o algoritmo SAM também superestimou regiões de vegetação. Esta superestimativa confirmou-se pela matriz de confusão do algoritmo SAM (Tab. 3), na qual se verifica que este foi o classificador supervisionado que mais atribuiu erroneamente a vegetação para áreas de telhado de zinco, de barro, estrada de terra e asfalto. Essas confusões são explicadas, pois as áreas próximas às zonas verdes sofrem influência espectral das vegetações e dos alvos urbanos que as rodeiam, como o asfalto e as coberturas residenciais. Enquanto que o algoritmo SAM classificou essas zonas como vegetação, por outro lado, o classificador MaxVer atribuiu essas regiões às classes urbanas, o que fez desse algoritmo o com menor área de vegetação (Fig. 3 e 4).

Os mapas gerados pelos algoritmos SVM e K-médias foram os mais similares entre si e apresentaram áreas de classes urbanas e vegetais de forma mais igualmente distribuídas (Fig. 3). Os polígonos brancos no mapa correspondem às áreas de nuvens que foram recortadas para possibilitar as análises. Em relação à área total, o algoritmo SVM produziu áreas com vegetação muito parecidas com o algoritmo K-médias, com porcentagens de 28,18% e 29,70%, respectivamente (Fig. 4). Esta similaridade contrastou com o desempenho de ambos os algoritmos, onde, segundo os indicadores de performance (EG e K), o SVM foi considerado o melhor classificador e o K-médias o pior. Este fato pode ser explicado, pois os tipos de vegetação são distintos e, desta forma, confusões entre alvos de vegetação também são considerados erros, o que prejudicou o desempenho do K-médias. Soma-se a esta explicação, o tipo

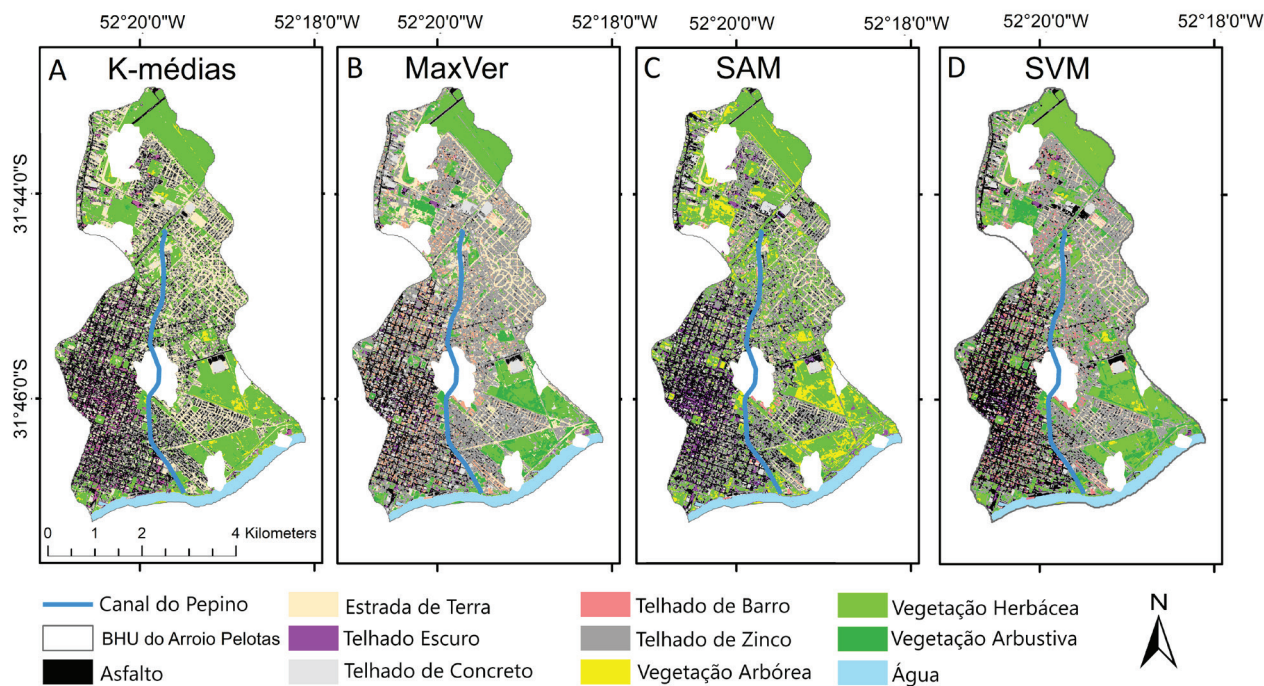


Figura 3. Distribuição espacial das classes obtidas a partir dos seguintes algoritmos. A) K-médias; B) MaxVer; C) SAM; D) SVM.
 Figure 3. Spatial distribution of the classes obtained by the following algorithms. A) K-means; B) MaxVer; C) SAM; D) SVM.

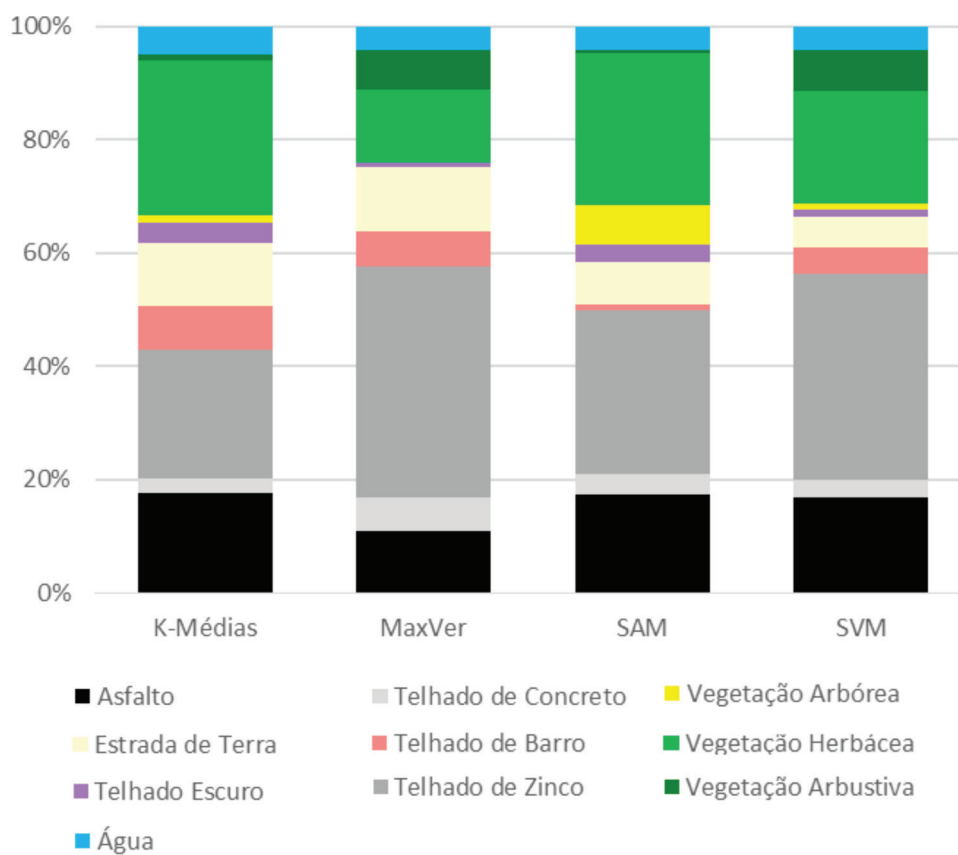


Figura 4. Proporções das classes obtidas em relação à área total estudada para os algoritmos K-médias, MaxVer, SAM e SVM.
 Figure 4. Proportions of the obtained classes in relation to the total studied area for the SVM, K-means, MaxVer and SAM algorithms.

de funcionamento não-supervisionado do classificador K-médias que não utilizou amostras de treinamento de vegetação, neste caso, para o processo de clusterização (Richards & Jia, 2006). Quando analisadas outras classes, o algoritmo SVM também foi superior na quase totalidade dos alvos, quando comparado ao K-médias (Fig. 4).

Na figura 4, observou-se que as principais discrepâncias entre proporções das classes geradas pelos algoritmos de classificação testados ocorreram para: a) telhado de zinco: de 22,67% a 40,65%; b) vegetação herbácea: de 12,83% a 27,40%; c) vegetação arbórea: de 0,01% a 6,88%; d) telhado de barro: 1,06% a 7,83%; e e) asfalto: de 10,97% a 17,56%. A representação gráfica das proporções das demais classes obtidas em relação à área total estudada pode ser visualizada na figura 4.

3.4 Cálculo do CN_{total}

Tendo em vista que os valores de CN_{total} (Eq. 8), sofrem grande influência da urbanização (Li *et al.*, 2018), os mapas que geraram estes maiores valores foram justamente os que apresentaram menores porcentagens de vegetação e, conseqüentemente, maiores proporções de classes consideradas urbanas. Neste sentido, o mapa de uso e cobertura produzido pelo algoritmo MaxVer resultou no maior valor CN com 90,71, seguido do SVM com 89,94, K-médias com 89,66 e, por último, SAM com 88,97. Os valores obtidos foram considerados altos, em que quase a totalidade da chuva que chega à bacia é transformada em escoamento superficial, o que pode ajudar a explicar a grande frequência de inundações na área. Para fins de comparação, a BHU do Arroio Dilúvio, situada na cidade de Porto Alegre, que é densamente povoada, possui valor estimado de CN igual a 79 (Tassi *et al.*, 2006).

Tomando como base o resultado de menor valor do CN_{total} obtido com o mapa produzido pelo algoritmo SAM, os incrementos de valores finais foram de 0,77 para o algoritmo K-médias, de 1,09 para o algoritmo SVM e de 1,95 para o algoritmo MaxVer. Estas pequenas diferenças nos valores finais de CN_{total} foram atribuídas às baixas variações dos valores de CN por categoria

(Quadro 1), uma vez que esses foram obtidos para ambientes classificados hidrologicamente no grupo D, ou seja, regiões sobre solos de várzea com pouca infiltração (Sartori *et al.*, 2005a, 2005b). Este é o caso da BHU do Arroio Pepino, a qual possui grande capacidade de gerar escoamento mesmo em áreas com cobertura vegetal, superficial. Diferenças nos valores de CN_{total} seriam mais evidentes se os classificadores fossem empregados em regiões urbanas sobre solos arenosos profundos e com baixo teor de argila, pertencentes ao grupo hidrológico A (Mello & Silva, 2013).

Quanto maior o valor do CN de uma bacia hidrográfica, maior será o escoamento superficial estimado para uma chuva considerada em estudos hidrológicos (Mello & Silva, 2013; Saviddou *et al.*, 2018). Isso pode fazer com que obras hidráulicas para contenção de inundação ou para sistemas de drenagem sejam superdimensionadas, acarretando desperdícios de capital econômico. Por outro lado, se o valor CN for subdimensionado, além de prejuízos econômicos, danos ambientais e humanos também poderão ocorrer por conta de enchentes, inundações e alagamentos devido ao mau funcionamento destas mesmas obras hidráulicas.

4 Conclusões

Com base nos resultados apresentados, o cálculo da curva-número da BHU do Arroio Pepino é afetado pelas diferenças nas proporções das classes de uso e cobertura da superfície, obtidas através de mapas temáticos produzidos por diferentes algoritmos de classificação supervisionada (SVM, MaxVer e SAM) e não-supervisionada (K-médias). Desta forma, verifica-se que a hipótese deste trabalho foi comprovada, onde a qualidade do valor de CN_{total} é altamente dependente da acurácia do produto final da classificação. Portanto, o algoritmo que apresentar o melhor desempenho na classificação fornecerá o mapa de melhor qualidade para o cálculo do CN_{total} mais confiável.

Sendo assim, o algoritmo de classificação SVM é considerado o melhor estimador para o mapeamento da superfície da BHU em estudo,

tendo em vista seu desempenho satisfatório indicado pelos índices EG e K. O resultado do valor de CN_{total} obtido pelo produto deste classificador prediz que, aproximadamente, apenas 10,06% da água proveniente de uma chuva irá infiltrar, o que ajuda a explicar as frequentes enchentes, inundações e alagamentos que a região sofre em períodos chuvosos. É importante destacar que, para BHUs sobre solos hidrologicamente classificados no grupo D (solos com elevado potencial de escoamento, baixa taxa de infiltração, rasos ou de camada impermeável), as diferenças entre os valores de CN_{total} podem ser muito próximas independentemente do algoritmo de classificação utilizado para gerar os mapas de uso e cobertura da superfície.

O método proposto para o cálculo do CN_{total} de BHUs baseado em mapas de uso e cobertura da superfície obtidos a partir de imagens orbitais de alta resolução espacial é viável e pode ser utilizado para outras bacias urbanas. Porém, os resultados de CN_{total} serão cada vez mais divergentes entre os algoritmos de classificação testados à medida que os solos da bacia estudada possuem maiores capacidades de infiltração, pois as áreas vegetadas terão maior influência sobre o cálculo final. Neste caso, super ou subestimativas de classes de vegetação impactarão significativamente na modelagem do escoamento superficial direto de uma bacia hidrográfica urbana.

A imagem de satélite RapidEye, assim como outras imagens orbitais de alta resolução espacial, apresenta-se como um produto importante e necessário para o cálculo do valor CN em BHUs, onde as proporções correspondentes à cada classe de uso e cobertura da superfície podem ser satisfatoriamente mensuradas e consideradas no cálculo total da bacia. No entanto, ressalta-se que a falta de resolução espectral deste sensor pode afetar o desempenho de classificadores que se baseiam no comportamento espectral dos pixels para discriminar as classes da imagem, assim como os testados neste trabalho. Uma alternativa a este problema pode ser o uso de algoritmos de classificação baseados na forma dos objetos, tal como a segmentação por contornos, adequada para imagens de alta resolução espacial. A segmentação por

contornos baseia-se em algoritmos de detecção de borda, tais como Canny, Rothwell e Nalwa.

Agradecimentos. Os autores agradecem ao Ministério do Meio Ambiente (MMA) pela disponibilização da imagem RapidEye através do Programa de Regularização Ambiental (instituído pelo Decreto Presidencial 7.830).

Referências

- Abbas, A.W., Minallh, N., Ahmad, N., Ahmad, S.A.R. & Khan, M.A.A. 2016. K-Means and ISODATA clustering algorithms for landcover classification using remote sensing. *Sindh University Research Journal*, 48(2): 315-318.
- Alves Júnior, J.B., Cruz Júnior, N. & Chuerubim, M.L. 2015. Tecnologia de Satélite: RapidEye. In: XII CONGRESSO NACIONAL DE ENGENHARIA DE AGRIMENSURA, 1., 2015, Salvador. Anais... Salvador, CONEA, v. 1, p. 1-10.
- Bajracharya, A.R., Rai, R.R. & Rana, S. 2015. Effects of urbanization on storm water run-off: a case study of Kathmandu metropolitan city, Nepal. *Journal of the Institute of Engineering*, 11(1): 36-49.
- Banasik, K., Krajewski, A., Sikorska, A. & Hejduk, L. 2014. Curve number estimation for a small urban catchment from recorded rainfall-runoff events. *Archives of Environmental Protection*, 40(3): 75-86.
- Brown, P., Daigneault, A. & Gawith, D. 2016. Climate change and the economic impacts of flooding on Fiji. *Climate and Development*, 9(6): 493-504.
- Cardoso, E.C.A. 2012. *Mapeamento das transformações socioambientais da bacia hidrográfica do arroio Pepino, 1916 – 2011 / Pelotas (RS)*. Rio Grande, 86p. Dissertação de Mestrado, Programa de Pós-Graduação em Geografia, Instituto de Ciências Humanas e da Informação, Universidade Federal do Rio Grande.
- Cattani, C.E.V., Mercante, E., Souza, C.H.W. & Wrublack, S.C. 2013. Desempenho de algoritmos de classificação supervisionada dos Satélites RapidEye. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 16., 2013, Foz do Iguaçu. Anais... São José dos

- Campos, SBSR, v. 1, p. 8005-8010.
- Cohen, J.A. 1960. Coeficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1): 37-46.
- Csendes, B. & Mucsi, L. 2016. Identification and spectral evaluation of agricultural crops on hyperspectral airborne data. *Journal of Environmental Geography*, 9(3): 49-53.
- Du, S., Rompaey, A.V., Shi, P. & Wang, J. 2015. A dual effect of urban expansion on flood risk in the Pearl River Delta (China) revealed by land-use scenarios and direct runoff simulation. *Natural Hazard*, 77(1): 111-128.
- Foody, G.M. & Mathur, A. 2004. Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification. *Remote Sensing of Environment*, 93(1): 107-117, 2004.
- Gadal, S. & Ouerghemmi, W. 2016. Urban objects recognition feasibilities by airborne hyperspectral and multispectral remote sensing. In: PROCEEDINGS OF SPATIAL ACCURACY, 8., 2016, Montpellier. Anais... Montpellier, PSA, v. 1, p. 41-66.
- Hameed, H.M. 2017. Estimating the effect of urban growth on annual runoff volume using gis in the erbil sub-basin of the Kurdistan region of Iraq. *Hydrology*, 4(12): 1-13.
- Hedge, G., Ahemed, M., Hebbar, R. & Raj, U. 2014. Urban land cover classification using hyperespectral data. In: ISPRS TECHNICAL COMMISSION VIII SYMPOSIUM, 12., 2014, Hyderabad. Anais... Hyderabad, ISPRS, v. 1, p. 1-10.
- Hosseini, L. & Kandovan, R.S. 2017. Hyperspectral image classification based on hierarchical SVM algorithm for improving overall accuracy. *Advances in Remote Sensing*, 6(1): 66-75.
- Jiménez, A.A., Vilchez, F.F., González, O.N. & Flores, S.M.L.M. 2018. Analysis of the land use and cover changes in the metropolitan area of Tepic-Xalisco (1973–2015) through Landsat images. *Sustainability*, 10(1): 1-15.
- Kadhim, N., Mourshed, M. & Bray, M. 2016. Advances in remote sensing applications for urban sustainability. *Euro-Mediterranean Journal for Environmental Integration*, 1(7): 1-22.
- Kowalik, T. & Walega, A. 2015. Estimation of CN parameter for small agricultural watersheds using asymptotic functions. *Water*, 7(1): 940-955.
- Kruse, F.A., Lefkoff, A.B., Bordman, J.W., Heidebrecht, K.B., Shapiro, A.T., Barloon, P.J. & Goetz, A.F.H. 1993. The spectral image processing system (SIPS) - interative visualization and analysis of image spectrometer data. *Remote Sensing of Environment*, 44(1): 145-163.
- Landis, J.R. & Koch, G.G. 1977. The measurement of observer agreement for categorical data. *International Biometric Society*, 33(1): 159-174.
- Li, C., Liu, M., Hu, Y., Shi, T., Zong, M. & Walter, M.T. 2018. Assessing the impact of urbanization on direct runoff using improved composite CN method in a large urban area. *International Journal of Environment Research and Public Health*, 15(775): 1-14.
- Luo, P., Mu, D., Xue, H., Ngo-Duc., Dang-Dinh, K., Takara, K., Nover, D. & Schladow, G. 2018. Flood inundation assessment for the Hanoi Central Area, Vietnam under historical and extreme rainfall conditions. *Scientific Reports*, 8(1): 1-11.
- Maria Filho, P.J., Targa, M.C., Santos, P.S. & Mostarda Neto, N. 2016. Barragens de detenção em bacia hidrográfica com uso de curvas cota-volume e hidrogramas triangulares sintéticos. *Ambiente & Água - An Interdisciplinary Journal of Applied Science*, 11(1): 1120-1135.
- Mekonnen, M., Melesse, A.M. & Keesstra, S.D. 2016. Spatial runoff estimation and mapping of potential water harvesting sites: A GIS and remote sensing perspective, Northwest Ethiopia. In: Melesse, A.M. & abteu, W. (Ed.). *Landscape dynamics, soils and hydrological processes in varied climates*, London, Springer, p. 565-584.
- Mello, C.R. & Silva, A.M. 2013. *Hidrologia: princípios e aplicações em sistemas agrícolas*. Lavras, UFLA, 455p.
- Meneses, P.R. & Almeida, T. 2012. *Introdução ao processamento de imagens de sensoriamento remoto*. Brasília, UnB, 266p.
- Miller, J.D. & Hess, T. 2017. Urbanisation impacts on storm runoff along a rural-urban gradient.

- Journal of Hydrology*, 552(1). 474-489.
- MMA. Ministério do Meio Ambiente. 2014. *Imagens RapidEye - Geocatálogo*. Disponível em: < <http://geocatalogo.ibama.gov.br/>>. Acesso em: 30 jun. 2016.
- Mountrakis, G., Im, J. & Ogole, C. 2011. Support vector machines in remote sensing - a review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(1): 247-259.
- Moughal, T.A. 2013. Hyperspectral image classification using Support Vector Machine. *Journal of Physics: Conference Series*, 439(1): 1-9.
- Moura, N.V.A. 2016. *Comparação entre classificadores supervisionados para fins de análise multitemporal do uso e cobertura da terra no entorno do reservatório da usina hidrelétrica de Corumbá IV/GO*. Brasília, 42p. Dissertação de Mestrado, Programa de Pós-Graduação em Geografia, Universidade de Brasília.
- Mustafa, A., Rienow, A., Saadi, I., Cools, M. & Teller, J. 2018. Comparing support vector machines with logistic regression for calibrating cellular automata land use change models. *European Journal of Remote Sensing*, 51(1): 391-401.
- Naldi, M.C., Campello, R.J.G.B., Hruschka, E.R. & Carvalho, A.C.P.L.F. 2011. Efficiency issues of evolutionary k-means. *Applied Soft Computing*, 11(1): 1938-1952.
- Ohana-Levi, N., Givati, A., Alfasi, N., Peeters, A. & Karnieli, A. 2017. Predicting the effects of urbanization on runoff after frequent rainfall events. *Journal of Land Use Science*, 13(1/2): 1-21.
- Omo-Irabor, O.O.A. 2016. Comparative study of image classification algorithms for landscape assessment of the Niger delta region. *Journal of Geographic Information System*, 8(1): 163-170.
- Pires, L.C., França, A.C.M. & Bacani, V.M. 2014. Análise do desempenho dos classificadores de imagens orbitais no mapeamento do uso da terra e cobertura vegetal da bacia hidrográfica do córrego Indaiá, Aquidauana-MS. In: 5º SIMPÓSIO DE GEOTECNOLOGIAS NO PANTANAL, 1., 2014, Campo Grande. *Anais...* Campo Grande, SGP, v. 1, p. 276-285.
- Pradhan-Salike, I. & Pokharel, R. 2017. Impact of urbanization and climate change on urban flooding: a case of the Kathmandu valley. *Journal of Natural Resources and Development*, 710(1): 56-66.
- Richards, J.A. & Jia, X. 2006. *Remote sensing digital image analysis: an introduction*. New York, Springer-Verlag, 494p.
- Rimal, B., Zhang, L., Keshtkar, H., Haack, B.N., Rijal, S. & Zhang, P. 2018. Land use/land cover dynamics and modeling of urban land expansion by the integration of cellular automata and Markov chain. *International Journal of Geo-Information*, 7(154): 2-21.
- Santos, R.C., Galo, M. & Tachibana, V.M. 2018. Classification of lidar data over building roofs using k-means and principal component analysis. *Bulletin of Geodetic Sciences*, 24(1): 69-84.
- Saraiva, M.V.P. & Polidori, M.C. 2008. Relação bacias urbanas x zoneamento: o caso de Pelotas/RS. *Revista Projectare*, 1(1): 110-119.
- Sartori, A., Lombardi Neto, F. & Genovez, A.M. 2005a. Classificação hidrológica de solos brasileiros para a estimativa da chuva excedente com o método do serviço de conservação do solo dos Estados Unidos. Parte 1: Classificação. *Revista Brasileira de Recursos Hídricos*, 10(4): 5-18.
- Sartori, A., Genovez, A.M. & Lombardi Neto, F. 2005b. Classificação hidrológica de solos brasileiros para a estimativa da chuva excedente com o método do serviço de conservação do solo dos Estados Unidos. Parte 2: Aplicação. *Revista Brasileira de Recursos Hídricos*, 10(4): 19-29.
- Şatir, O. 2016. Mapping the land-use suitability for urban sprawl using remote sensing and GIS under different scenarios. *INTECH Open Science*, 9(1): 205-226.
- Savvidou, E., Efstratiadis, A., Koussis, A.D., Koukouvinos, A. & Skarlatos, D. 2018. The curve number concept as a driver for delineating hydrological response units. *Water*, 10(194): 1-32.
- Shanableh, A., Al-Ruzouq, R., Yilmaz, A. G., Siddique, M., Merabtene, T. & Imteaz, M.A. 2018. Effects of land cover change on urban floods and rainwater harvesting: a case study in Sharjah, UAE. *Water*, 10(5): 1-17.

- Simões, I., Duft, D.G., Miranda, C.G., Souza, C.H.W., Sanches, G.M. & Picoli, M.C.A. 2015. Avaliação do desempenho de classificadores em imagens de veículo aéreo não tripulado (VANT). In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 17., 2015, João Pessoa. *Anais...* São José dos Campos, INPE, v. 1, p. 611-618.
- Smedema, L.K. & Rycroft, D.W. 1983. *Land drainage: planning and design of agricultural drainage systems*. New York, Cornell University Press, 376p.
- Sowan, B. & Qattous, H. 2017. A data mining of supervised learning approach based on kmeans clustering. *International Journal of Computer Science and Network Security*, 17(1): 18-24.
- Spekkers, M., Rozer, V., Thieken, A., Veldhuis, M.C. & Freibich, H. 2017. A comparative survey of the impacts of extreme rainfall in two international case studies. *Natural Hazard and Earth System Science*, 17(1): 1337-1355.
- Tassi, R., Allasia, D., Adalberto, M., Talita, M., Josiane, H. & Silveira, L. 2016. Determinação do parâmetro CN para sub-bacias urbanas do arroio Dilúvio - Porto Alegre/RS. In: SIMPÓSIO DE RECURSOS HÍDRICOS DO SUL-SUDESTE, 1., 2016, Curitiba. *Anais...* Curitiba, SRHSS, v. 1, p. 1-13.
- Tripathi, R., Sengupta, S.K., Patra, A., Chang, H. & Jung, W. 2014. Climate change, urban development, and community perception of an extreme flood: a case study of Vernonia, Oregon, USA. *Applied Geography*, 46(1): 137-146.
- USDA. United States Department of Agriculture. 1986. *Urban Hydrology for Small Watersheds*. Washington, USDA, 164p. (USDA-210-VI-TR-55).
- Ustener, M., Sanli, F.B. & Dixon, B. 2015. Application of support vector machines for land use classification using high-resolution RapidEye images: a sensitivity analysis. *European Journal of Remote Sensing*, 48(1): 403-422.
- Varma, M.K.S. & Nynalasetti, K.K.R. 2017. Satellite image classification using genetic algorithm based on SVM classifier. *International Journal of Control Theory and Applications*, 10(26): 12-20.
- Xu, Z. & Zhao, G. 2016. Impact of urbanization on rainfall-runoff processes: case study in the Liangshui river basin in Beijing, China. *International Association of Hydrological Sciences*, 373(1): 7-12.
- Zezelew, D.G. 2017. Spatial mapping and testing the applicability of the curve number method for ungauged catchments in northern Ethiopia. *International Soil and Water Conservation Research*, 5(4): 293-301.
- Zimmermann, R., Bracalenti, L., Piacentini, R. & Inostroza, L. 2016. Urban flood risk reduction by increasing green areas for adaptation to climate change. *Procedia Engineering*, 161(1): 2241-2246.
- Zope, P.E., Eldho, T.I. & Jothiprakash, V. 2015. Impacts of urbanization on flooding of a coastal urban catchment: a case study of Mumbai City, India. *Natural Hazards*, 75(1): 887-908.
- Zhou, Q., Leng, G. & Huang, M. 2018. Impacts of future climate change on urban flood volumes in Hohhot in northern China: benefits of climate change mitigation and adaptations. *Hydrology and Earth System Science*, 22(1): 305-316.
- Wang, C., Xu, A. & Li, X. 2018. Supervised classification high-resolution remote-sensing image based on interval type-2 fuzzy membership function. *Remote Sensing*, 10(5): 1-22.
- Webber, J.L., Gibson, M.J., Chen, A.S., Savic, D., Fu, G. & Butler, D. 2018. Rapid assessment of surface-water flood-management options in urban catchments. *Urban Water Journal*, 15(3): 2010-2017.
- Wu, J., Yang, R. & Song, J. 2018. Effectiveness of low-impact development for urban inundation risk mitigation under different scenarios: a case study in Shenzhen, China. *Natural Hazard and Earth System Sciences*, 18(1): 2525-2536.
- Yao L., Wei W., Yu Y., Xiao J. & Chen L. 2018. Rainfall-runoff risk characteristics of urban function zones in Beijing using the SCS-CN model. *Journal of Geographical Sciences*, 28(5): 656- 668.