

15º SIMPÓSIO NACIONAL DE

GEOMORFOLOGIA

INTEGRAÇÃO DE LiDAR E APRENDIZAGEM DE MÁQUINA PARA ESPACIALIZAÇÃO DE TEORES DISPONÍVEIS DE SÓDIO E POTÁSSIO EM SOLOS DO SEMIÁRIDO POTIGUAR

Fabiano Souza de Macedo ¹
Rone José dos Santos Costa ²
Ana Livia Araújo de Azevedo ³
Renato Batista Cavalcante ⁴
Francisco das Chagas Araújo de Paiva ⁵
João Santiago Reis ⁶

RESUMO

A utilização de tecnologias avançadas como LiDAR, associada a técnicas de aprendizado de máquina, oferece uma nova abordagem para modelar a variabilidade espacial de atributos das paisagens. O presente trabalho tem como objetivo modelar espacialmente os teores de sódio (Na) e potássio (K) disponíveis nos solos de um assentamento rural no semiárido do Rio Grande do Norte, com técnicas de aprendizado de máquina utilizando como variáveis preditoras produtos derivados de levantamento realizado com LiDAR. Para a determinação dos pontos de amostragem e predição dos teores de Na e K disponíveis, os algoritmos foram alimentados com produtos geomorfométricos produzidos após processamento de modelo digital de terreno (MDT). Para predição dos teores de Na e K disponíveis, foram testados modelos gerados por quatro algoritmos. Para avaliação do desempenho dos modelos, foram utilizados o coeficiente de determinação (R^2), erro médio absoluto (MAE) e raiz do erro quadrático médio (RMSE). Foram elaborados mapas de distribuição espacial de teores de Na e K disponíveis, assim como seus respectivos mapas de incerteza das predições. Para K disponível, a melhor performance observada foi do algoritmo GBM com $R^2 = 0.04$, MAE = 0.56 e RMSE = 1.46, sendo as variáveis mais importantes para a predição: Índice de Balanço de Massa, MDT, Orientação das Vertentes, Índice de Umidade Topográfica (IUT) e Índice de Convexidade. Para Na disponível, o melhor desempenho foi obtido pelo SVM com $R^2 = 0.24$, MAE = 0.33 e RMSE = 1.19, que indicou como variáveis mais importantes: Perfil de Curvatura, Índice de Rugosidade do Terreno, Índice Multi-resolução de Planicidade de Topos de Vale e IUT. Observou-se que utilizar apenas variáveis de relevo gerou resultados parcialmente satisfatórios para modelar a distribuição de atributos dos solos. A metodologia mostrou-se promissora, desde que hajam variáveis preditoras relacionadas aos demais fatores de formação de solos.

INTRODUÇÃO

Compreender como os nutrientes estão distribuídos no solo, a sua dinâmica e distribuição espacial, tem sua importância para vários grupos de interessados: agricultura em larga escala ou familiar, instituições públicas e privadas, e a ciência do solo em geral,

¹ Graduando de Licenciatura em Geografia da Universidade Federal do Rio Grande do Norte - UFRN, fabiano.souza.103@ufrn.edu.br;

² Graduando de Licenciatura em Geografia da Universidade Federal do Rio Grande do Norte - UFRN, rone.jose.costa.018@ufrn.edu.br;

³ Mestranda do Programa de Pós-Graduação em Geografia (GEOCERES) da Universidade Federal do Rio Grande do Norte - UFRN, analivia.ob@hotmail.com;

⁴ Graduando de Bacharelado em Geografia da Universidade Federal do Rio Grande do Norte - UFRN, renatocavalcantetreze@hotmail.com;

⁵ Mestrando do Programa de Pós-Graduação em Manejo do Solo e Água da Universidade Federal Rural do Semiárido - UFERSA, franciscochagas621@gmail.com;

⁶ Programa de Pós-Graduação em Geografia (CERES) da Universidade Federal do Rio Grande do Norte - UFRN, joao.reis@ufrn.com.



com o intuito de classificar e caracterizar os solos, além de subsidiar demandas específicas.

Em ambientes semiáridos pelas próprias características dos solos na região, em que há uma lixiviação parcial pela baixa pluviosidade, e pela influência do tipo de material de origem (Chaves, 1983), acontece a concentração de sódio (Na) e potássio (K) em diferentes posições da paisagem, justificando a importância do entendimento das relações solo-relevo.

Para ter precisão e qualidade no mapeamento e distribuição de nutrientes, a tecnologia tem fundamental importância. A ferramenta LiDAR - *Light Detection and Ranging* - funciona com pulsos de laser que, com precisão, mapeia e identifica objetos, podendo cobrir grandes áreas, com nível de detalhamento e grau de precisão extremamente alto (Reddy et. al. 2015). Aliado a essa ferramenta, técnicas de aprendizagem de máquina (*machine learning*) são ferramentas promissoras como método para modelagem espacial de nutrientes, treinando algoritmos a prever a espacialização dos teores de Na e K disponíveis.

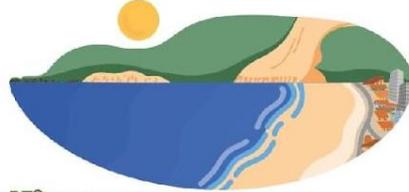
Considerando a contribuição do LiDAR para a obtenção de uma resolução detalhada e um alto grau de precisão em relação ao terreno, e a crescente eficácia das técnicas de aprendizagem de máquina e dos modelos de predição para a identificação de padrões dos atributos dos solos, este trabalho integra ambas as técnicas visando fornecer uma maior segurança na compreensão do comportamento da espacialização dos teores de Na e K disponível, o que pode orientar as atividades de campo de maneira otimizada e precisa.

Portanto, compreendendo a relevância desse estudo no semiárido do Seridó potiguar, o objetivo deste trabalho é realizar a modelagem espacial dos teores disponíveis de Na e K em solos de um assentamento rural do Rio Grande do Norte, com técnicas de aprendizagem de máquina, utilizando como variáveis preditoras de produtos derivados de levantamento realizado com técnica LiDAR.

MATERIAIS E METODOS

A área de estudo localiza-se no município de Jucurutu/RN, inserido dentro da região imediata de Caicó/RN, e possui área total de 189 ha (Figura 1).

A partir de levantamento de LiDAR, foi obtido um Modelo Digital do Terreno (MDT) com resolução de 25 centímetros, através do sensor DJI Zenmuse L1 acoplado a

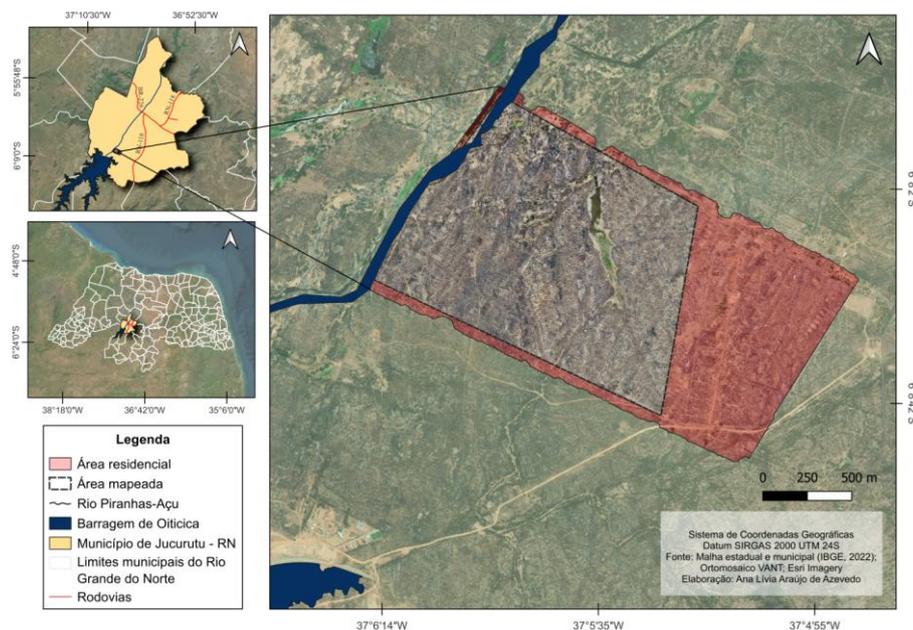


15º SIMPÓSIO NACIONAL DE GEOMORFOLOGIA

veículo aéreo não tripulado (VANT). O MDT foi reamostrado para 15 metros, e foram extraídas covariáveis do terreno em ambiente SAGA (*System for Automated Geoscientific Analyses*) (Conrad et al., 2015).

As variáveis preditoras utilizadas foram produtos geomorfométricas gerados a partir do MDT, e foram utilizadas para orientar a amostragem de solos pelo método do Hipercubo Latino Condicionado (HLC) e para a predição dos atributos químicos. Foram utilizados os seguinte produtos: Declividade (DECL), Altura da Declividade (ADECL), Índice de Posição Topográfica (IPT), Perfil da Curvatura (PC) e Classificação do relevo Baseada no Índice de Posição Topográfica (CR), propostas por Gallant & Wilson (2000); Posição de meia encosta (PME), Altura Padronizada (AP) e Elevação Normalizada (EN), segundo Conrad et al. (2015); o Índice de Multirresolução de Planicidade do Fundo do Vale (IMPFV) e o Índice de Multirresolução de Planicidade do Topo do Vale (IMPTV), baseados em Gallant e Dowling (2003) e, por fim, Profundidade do Vale (PV) (Rodríguez, F. et. al., 2002), Índice de Rugosidade do Terreno (IRT) (Rylei, De Gloria & Elliot, 1999), Índice de Balanço de Massa (IBM) (Friedrich, 1996), Índice de Convexidade (CONV) (Koethe, R & Lehmeier, F., 1996) e o Modelo Digital do Terreno (MDT).

Figura 1: Localização da área de estudo.



Fonte: Os autores (2025)

As amostras de solos foram coletadas em profundidade de 0 a 20 cm, em pontos orientados pelo método do Hipercubo Latino Condicionado (HLC). O HLC consiste em um procedimento para espacializar pontos de amostragem que sejam representativos da



área estudada (Carvalho Filho et al, 2014). Sendo assim, o algoritmo foi alimentado com os produtos geomorfométricos derivados do LiDAR para geração de 100 pontos representativos para a área de estudo, considerando as características do relevo local.

Após trabalhos de campo para coleta, as amostras de solo foram submetidas a secagem ao ar, e peneiramento em malha de 2 mm no Laboratório Didático de Geociências (LADGEO-UFRN). Todos os procedimentos de laboratório seguiram metodologias do Manual de Métodos de Análise de Solos (Teixeira et al., 2017). Para determinar os teores de sódio (Na) e potássio (K) disponíveis nos solos foi utilizada extração com a solução Mehlich-1 ($\text{HCl } 0.05 \text{ mol L}^{-1} + \text{H}_2\text{SO}_4 \text{ } 0.125 \text{ mol L}^{-1}$) e dosagem em fotometro de chama.

A etapa da modelagem espacial e predição dos teores de Na e K disponíveis consistiu nas seguintes etapas: i) empilhamento da variáveis preditoras (geomorfométricas) e extração de valores nos pontos de coleta/análise de solos; ii) treinamento dos modelos em 4 algoritmos diferentes; iii) execução da predição em 100 repetições; iv) validação e avaliação da performance dos algoritmos para seleção do melhor modelo; vi) elaboração de mapa médio e mapa de incerteza. Os algoritmos utilizados foram o Support Vector Machine (SVM), Random Forest (RF), Gradient Boosting Machines (GBM) e Cubist, e todo fluxo foi realizado no software R através do pacote *caret* (Kuhn, 2008).

RESULTADOS E DISCUSSÃO

A partir da estatística descritiva das análises em laboratório dos teores de ambos os elementos estudados, foi possível perceber os baixos teores encontrados nos solos (Tabela 1). Porém, percebe-se que o K possui um menor coeficiente de variação em comparação com o Na, indicando distribuição de valores mais homogênea.

Tabela 1: Estatística descritiva dos teores de sódio e potássio nas amostras utilizadas na modelagem.

Variável	Parâmetros					Quartis		
	Média	Mediana	DP	CV (%)	Min.	Máx.	Q1	Q3
Na (cmolc/kg)	0,057	0	0,240	421	0,000	1,52	0	0,027
K (cmolc/kg)	0,245	0,194	0,162	66,1	0,050	1,24	0,015	0,304

DP - Desvio-padrão; CV - Coeficiente de variação; Min - Mínimo; Máx - Máximo; Q1 - primeiro quartil; Q3 - terceiro quartil.

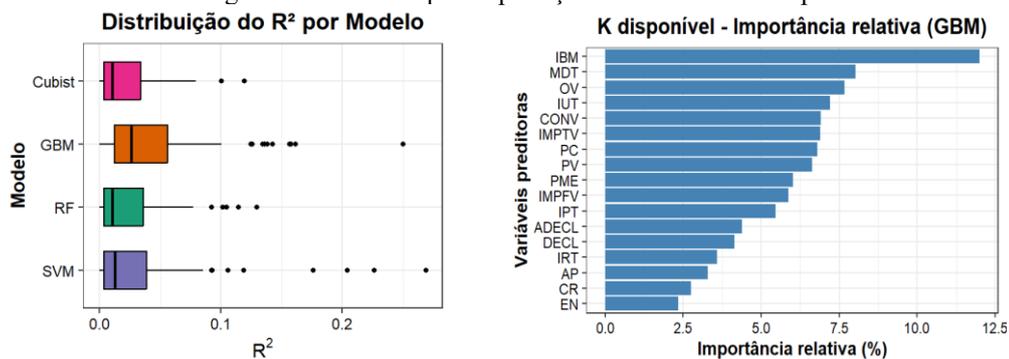
Analisando os resultados, observamos que para K disponível, o algoritmo que apresentou melhor coeficiente de determinação (R^2) na predição dos teores foi o GBM (Figura 2), que teve média com valor de 0.044 e mediana de 0.026 (Tabela 2).

A Figura 2 apresenta a importância relativa das variáveis utilizadas para elaboração do modelo pelo GBM. Foram ranqueadas como as cinco variáveis preditoras de maior importância índice de balanço de massa (IBM), MDT, orientação das vertentes (OV), índice de umidade topográfica (IUT) e índice de convexidade (CONV).

Tabela 2: Estatística descritiva dos parâmetros de avaliação da performance dos modelos preditivos de K disponível.

Parâmetros		Modelos			
		SVM	RF	GBM	Cubist
R^2	Média	0.023	0.024	0.044	0.030
	Mediana	0.011	0.011	0.026	0.013
	Desvio padrão	0.027	0.029	0.046	0.047
MAE	Média	0.109	0.121	0.125	0.113
	Mediana	0.106	0.119	0.124	0.108
	Desvio padrão	0.018	0.016	0.019	0.018
RMSE	Média	0.165	0.172	0.176	0.168
	Mediana	0.149	0.157	0.159	0.150
	Desvio padrão	0.050	0.047	0.049	0.051

Figura 2: Coeficiente de determinação (R^2) considerando todas as repetições, e importância relativa das variáveis geomorfométricas para a predição dos teores de K disponível.



Para o K disponível, analisando os mapas gerados por todos os algoritmos, que a sua distribuição (Figura 3) apresentou maiores valores nas áreas mais elevadas da paisagem, coincidindo com os locais de maior concentração de afloramentos rochosos e solos muito pedregosos, observados em campo, corroborando com estudos de Martins et

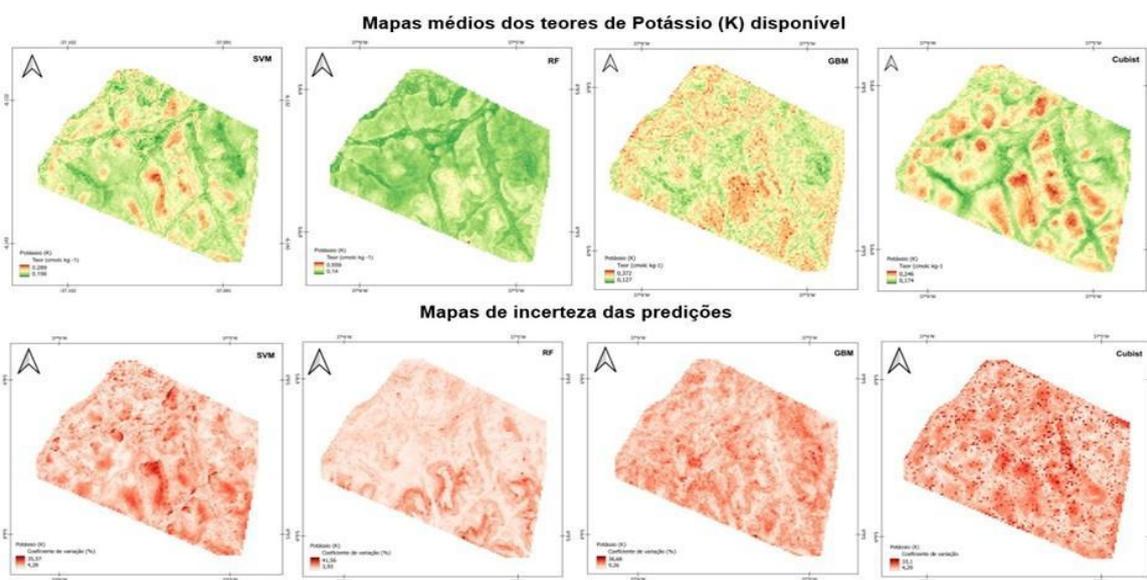


15º SIMPÓSIO NACIONAL DE GEOMORFOLOGIA

al. (2004) que afirmam que a reserva de K disponível nos solos está relacionada ao material de origem e seu estágio de desenvolvimento.

No caso de o Na disponível, observamos que o R^2 teve uma considerável variabilidade, tendo o SVM como o modelo de melhor desempenho, apresentando resultado de média de 0.242 e mediana de 0.307 (Tabela 3). A Figura 4 mostra a distribuição desses valores e percebe-se como RF, GBM e Cubist estão mais próximos, indicando desempenho muito inferior.

Figura 3: Distribuição espacial da predição dos teores de K disponível para cada algoritmo.



Na Figura 4, são apresentadas quais as variáveis foram elencadas com maior importância relativa pelo SVM. Para elaboração do modelo, as cinco variáveis com maior importância relativa foram o perfil de curvatura (PC), índice de rugosidade do terreno (IRT), índice multirresolução de planicidade de fundo de vale (IMPFV), índice de umidade topográfica (IUT) e índice de convexidade (CONV).

É observada uma distribuição de maiores teores de Na disponível nas posições mais baixas da paisagem, quando analisados os mapas médios gerados por todos os modelos (Figura 5). O mapa gerado pelo SVM destaca teores mais elevados em locais da área de estudo sazonalmente inundáveis. Este resultado corrobora o comportamento esperado para o Na em solos de ambientes semiáridos, em que a lixiviação ocorrida em solos de partes superiores das encostas tende a acumular sais em partes baixas da paisagem, em razão de condições de drenagem reduzida e evapoconcentração (Hailu; Mehari, 2021).



15º SIMPÓSIO NACIONAL DE
GEOMORFOLOGIA

Tabela 3: Estatística descritiva dos parâmetros de avaliação da performance dos modelos preditivos de Na disponível.

Parâmetros		Modelos			
		SVM	RF	GBM	Cubist
R²	Média	0.242	0.019	0.028	0.005
	Mediana	0.307	0.005	0.011	0.002
	Desvio padrão	0.202	0.038	0.047	0.009
MAE	Média	0.338	0.624	0.570	0.469
	Mediana	0.201	0.470	0.525	0.116
	Desvio padrão	0.268	0.476	0.162	0.584
RMS E	Média	1.196	1.748	1.463	1.776
	Mediana	0.275	0.755	0.670	0.379
	Desvio padrão	1.567	1.616	1.403	1.852

Figura 4: Coeficiente de determinação (R²) considerando todas as repetições, e importância relativa das variáveis geomorfométricas para a predição dos teores de Na disponível.

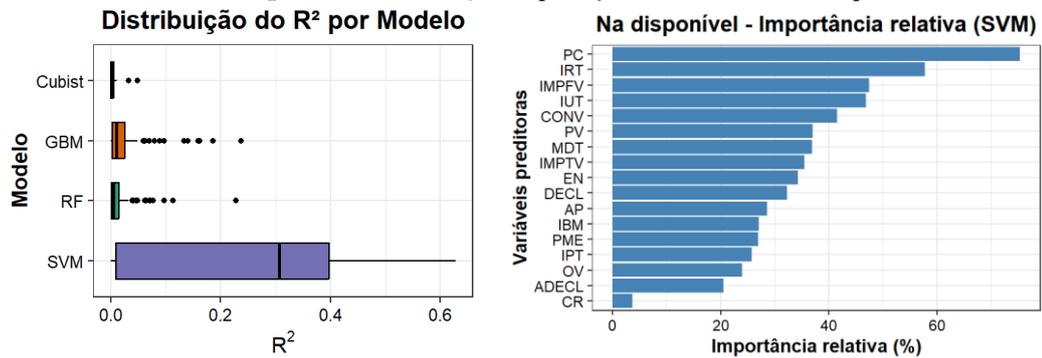
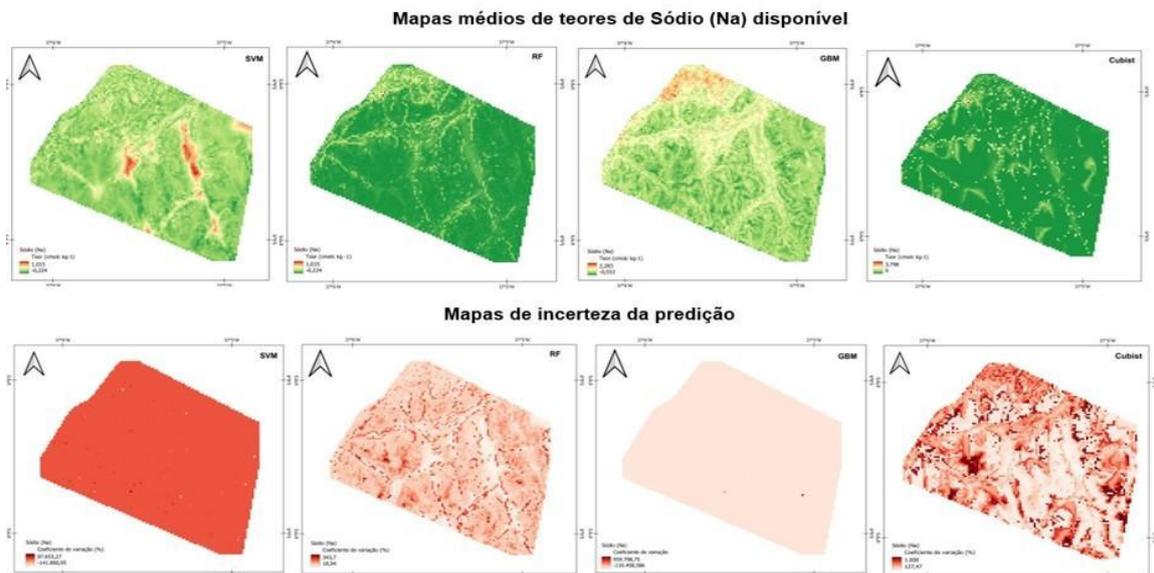


Figura 5: Distribuição espacial da predição dos teores de K disponível para cada algoritmo.





CONSIDERAÇÕES FINAIS

Considerado os resultados encontrados neste estudo, conclui-se a utilização de apenas variáveis preditoras relacionadas ao relevo, mesmo utilizando dados de alta resolução derivados de LiDAR, gerou resultados parcialmente satisfatórios. Testando diferentes algoritmos de aprendizado de máquina, foi evidenciado que apesar dos diferentes modelos gerados conseguirem representar a tendência espacial da distribuição dos teores de Na disponível de forma satisfatória, os moderados valores de R^2 , MAE e RMSE para o melhor modelo (SVM) indicam que os teores podem estar ligeiramente superestimados ou subestimados. No caso de K disponível a modelagem não foi considerada satisfatória, pelos baixos valores de R^2 , e altos valores de MAE e RMSE, apesar de haver sentido na distribuição espacial gerada pela predição.

Foi demonstrado que para cada conjunto de dados, e tipo de atributo do solo a ser predito, os algoritmos testados apresentaram performances diferentes, assim como diferentes variáveis preditoras foram elencadas como mais importantes para elaboração dos modelos. Portanto, conclui-se que não há um melhor algoritmo para todos os casos, assim como não há sequência de importância de variáveis preditoras fixa para utilização em predição de atributos de solos em geral. É importante sempre testar diferentes tipos de algoritmos para selecionar, através da avaliação da performance da etapa de validação, aquele que melhor se adequa a sua realidade de estudo.

Acreditamos que a metodologia de utilização de técnicas de aprendizado de máquina apresenta potencial para utilização em mapeamentos de solos, no que se refere a predição da distribuição espacial de atributos físicos e químicos. Porém, para geração de modelos com melhores performances devem ser utilizadas variáveis preditoras que representem, quando possível, todos os fatores de formação de solos: material de origem, clima, relevo, organismos e tempo.

Palavras-chave: Mapeamento de solos; modelagem espacial, inteligência artificial, solos do semiárido; pedologia.

AGRADECIMENTOS

Agradecemos à PROPESQ-UFRN por todo o apoio e subsídio para a realização deste trabalho através do fomento por meio do programa PIBIC, PROEX-UFRN,



Laboratório Didático de Geociência (LADGEO-UFRN), SEMARH (Secretaria de Meio Ambiente e Recursos Hídricos), e aos moradores da Agrovila Raimundo Nonato.

REFERÊNCIAS

BOEHNER, J., et al. Soil Regionalisation by Means of Terrain Analysis and Process Parameterisation. In: Micheli, E., Nachtergaele, F., Montanarella, L. [Ed.]: Soil Classification 2001. **European Soil Bureau**, Research Report, n. 7, EUR 20398 EN, Luxembourg, p.213-222, 2002.

CARVALHO FILHO, W. et al. Método do Hipercubo Latino Condicionado para a amostragem de solos na presença de covariáveis ambientais visando o mapeamento digital de solos. **Revista Brasileiro de Ciência do Solo**, v. 38, p. 386-396, 2014.

CHAVES, Lúcia Helena Garófalo. **Relação quantidade/intensidade de potássio em solos do trópico semi-árido brasileiro**. 1983. Tese de Doutorado. Universidade de São Paulo.

CONRAD, O., et al. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4, **Geosci. Model Dev.**, v. 8, n. 1991–2007, 2015.

FRIEDRICH, K. Digitale Beliefgliederungsverfahren zur Ableitung bodenkundlich relevanter Flächeneinheiten / von Klaus Friedrich. Hrsg. vom Fachbereich Geowissenschaften der Johann-Wolfgang-Goethe-Universität Frankfurt. - Frankfurt am Main : Inst. für Physische Geographie, 1996.

GALLANT, J.C.; DOWLING, T.I. A multiresolution index of valley bottom flatness for mapping depositional areas. **Water Resources Research**, v. 39, p. 1347-1359, 2003.

GALLANT, J. C.; WILSON, J. P. Primary topographic attributes. In: **Terrain Analysis: Principles and applications**. John Wiley & Sons, New York, 2000.

HAILU, B.; MEHARI, H. Impacts of Soil Salinity/Sodicity on Soil-Water Relations and Plant Growth in Dry Land Areas: A Review. **Journal of Natural Sciences Research**, v. 12, n. 3, p. 1-10, 2021.

KUHN, M. Building predictive models in R using the caret package. **Journal of Statistical Software**, v. 28, n. 5, p. 1–26, 2008.

MARTINS, R.; MELO, V. F. & SERRAT, B.M. Reserva Mineral de potássio em solos dos Campos Gerais, estado do Paraná. **R. Ceres**, v. 51, p. 521-533. 2004.

PEREIRA, L. F. et al. Applying Machine Learning Techniques to Model and Map Soil Surface Texture Using Limited Legacy Data. In: **Pedometrics in Brazil**. Cham: Springer Nature Switzerland, 2024. p. 31-45.

REDDY, A.D.; HAWBAKER, T.J.; WURSTER F; ZHU, Z.; WARD, S; NEWCOMB; D.; MURRAY R. Quantifying soil carbon loss and uncertainty from a peatland wildfire using multi-temporal LiDAR. **Remote Sensing of Environment**, v. 170, p.306–316, 2015.



15º SIMPÓSIO NACIONAL DE
GEOMORFOLOGIA

RILEY, S.J., DE GLORIA, S.D., ELLIOT, R. A Terrain Ruggedness that Quantifies Topographic Heterogeneity. **Intermountain Journal of Science**, v.5, n.1-4, p. 23-27, 1999.

RODRIGUEZ, F., MAIRE, É., COURJAULT-RADÉ, P., DARROZES, J. The Black Top Hat function applied to a DEM: A tool to estimate recent incision in a mountainous watershed (Estibère Watershed, Central Pyrenees). **Geophysical Research Letters**, v. 29, n. 6, 2002.

