

MAPEAMENTO DE SOLOS: CONVENCIONAL VERSUS DIGITAL A PARTIR DE TÉCNICAS DE APRENDIZADO DE MÁQUINA NA ANÁLISE DA BACIA HIDROGRÁFICA DO RIO BONITO, REGIÃO SERRANA DO ESTADO DO RIO DE JANEIRO

João Victor Bueno (joaobueno182@gmail.com)

O objetivo deste trabalho foi avaliar o desempenho de algoritmos no Mapeamento Digital de solos (MDS) — Random Forest (RF) e Support Vector Machine Radial (SVMR) — na discriminação das classes de solos identificadas na bacia hidrográfica do rio Bonito (72,4 km²), em contexto montanhoso do Estado do Rio de Janeiro. Destaca-se a crescente demanda e possibilidades a partir de mapeamentos quantitativos, com o uso de tecnologias de aprendizagem de máquina (Machine Learning), para a predição espacial de classes de solos, que visam auxiliar no planejamento, além da ocupação e do manejo mais eficientes sobre esse recurso natural valioso para a humanidade e que tem sido altamente degradado, o que atinge 33% dos solos do mundo. Nesse sentido, a produção de conhecimento sobre a sua distribuição espacial e propriedades torna-se fundamental, auxiliando na sua conservação, assim como servindo de base para estudos que demandem informações sobre a cobertura pedológica. A metodologia baseou-se exclusivamente em dados legados de solos (dados de perfis e mapa de solos) a partir de Levantamento de Reconhecimento Detalhado de Solos em escala 1:25 000, em índices de sensoriamento remoto obtidos pelo sensor Sentinel-2A e em covariáveis morfométricas derivadas do Modelo Digital de Elevação Copernicus GLO, com

resolução espacial de 30 m. Também foram utilizadas covariáveis categóricas, como litologia, geomorfologia e cobertura e uso da terra. O conjunto de dados legados continha, ao todo, vinte e oito (28) amostras analisadas, variando entre uma (01) e seis (06) amostras por classe, adotando-se um balanceamento a partir de Região de Interesse por Classe (ROIC). As amostras foram avaliadas por validação cruzada em três esquemas: 5-fold, 10-fold e Leave-One-Out (LOOCV). O melhor desempenho foi obtido com o algoritmo Random Forest sob validação cruzada LOOCV, atingindo uma acurácia de 0,92 e índice kappa de 0,91. O RF manteve a média de acurácia de 0,91 e kappa de 0,90, superando o SVM com acurácia média de 0,52 e kappa de 0,46. Para avaliar a incerteza e a consistência da previsão de unidades de solo foi utilizado o levantamento convencional a partir dos dados legados. Constatou-se, mais uma vez, a maior eficiência do Random Forest, quando comparados aos demais modelos. O algoritmo destacou a relevância das seguintes covariáveis: Profundidade do Vale; Elevação; Índice de Umidade Saga; Declividade; e Distância Vertical até o Canal. As características do relevo local, predominantemente montanhoso, desempenharam um papel significativo na compreensão da variação espacial de classes de solo, gerando a maior influência das covariáveis morfométricas na predição das unidades de solos.

Palavras-chave: mapeamento digital de solos; pedometria; covariáveis; relevo montanhoso; dados legados.