Quantificação da qualidade da interpolação de dados de métodos potenciais por fontes equivalentes utilizando validação cruzada em blocos

Ellen Fernandes Marcos e Leonardo Uieda, Universidade de São Paulo

Copyright 2024, SBGf - Sociedade Brasileira de Geofísica
Este texto foi preparado para a apresentação no X Simpósio Brasileiro de Geofísica, Salvador, 8 a 10 de outubro de 2024. Seu conteúdo foi revisado pelo Comitê Técnico do X SimBGf, mas não necessariamente representa a opinião da SBGf ou de seus associados. É proibida a reprodução total ou parcial deste material para propósitos comerciais sem prévia autorização da SBGf.

Resumo

Como parte do seu processamento, dados de gravimetria e magnetometria são rotineiramente interpolados em malhas regulares a altitudes constantes. O método das fontes equivalentes é ideal para realizar esse processamento por levar em consideração a física do problema, como a altitude variável dos pontos de observação e o fato dos dados representarem funções harmônicas. A qualidade do resultado obtido varia de acordo com dois parâmetros: a profundidade das fontes e o parâmetro de regularização. Ambos controlam o grau de suavidade da solução e a capacidade do método de prever valores plausíveis onde existem lacunas nos dados observados. Neste trabalho, utilizamos o método da validação cruzada em blocos (Roberts et al. 2017; doi:10.1111/ecog.02881) para estimar a capacidade das fontes equivalentes de prever dados onde existem lacunas nos levantamentos. Na validação cruzada, os dados são divididos aleatoriamente em dois subconjuntos, um para ajuste (treinamento) do modelo de fontes equivalentes e outro para testar as previsões do modelo. Para dados de métodos potenciais, a divisão aleatória dos dados resulta em um conjunto de dados para teste que estão espacialmente próximos dos dados para treinamento. Isso provoca a superestimação da qualidade do ajuste aos dados de teste, pois dados de métodos potenciais são autocorrelacionados, ou seja, dados espacialmente próximos tendem a ter valores semelhantes. A separação dos dados por blocos ao invés de aleatoriamente foi introduzida para sanar esse problema. Porém, sua viabilidade para a técnica de fontes equivalentes não foi explorada. Utilizamos dados sintéticos e reais para avaliar o uso das metodologias "block k-fold" e "leave-one-block-out" para estimar a qualidade da interpolação de dados gravimétricos e magnetométricos por fontes equivalentes. Nossa avaliação inclui um teste sistemático do efeito da variação do tamanho dos blocos na estimativa da qualidade, tanto para levantamentos terrestres quanto aerolevatamentos. Para conjuntos de dados grandes (> 10,000), testamos também a aplicação da validação cruzada em uma parcela dos dados e a viabilidade da extrapolação dos resultados para o resto do levantamento. Por fim, os resultados nos informam como a utilização destes métodos impactam na produção de malhas que combinam diversos levantamentos para grandes territórios, como Brasil, Antártica e Austrália. As fontes equivalentes serão as de Soler & Uieda (2021; doi:10.31223/X58G7C) e a validação cruzada por blocos será a implementação do pacote Verde (Uieda, 2018; doi:10.21105/joss.00957).