

Hierarchical spatio-temporal model for interpolation of missing data in satellite image series: The R-INLA approach

Vinicius Soares Martins Alves ¹
Chang Chiann ²

Abstract

Increasingly advanced satellites and sensors produce large volumes of data, and usual statistical models may be unfeasible for analysing these sets. In addition, missing data occurs due to clouds and other environmental artifacts. This work evaluates dynamic models with spatio-temporal random effects for interpolation in data sets with observations arranged in a grid, commonly offered through remote sensing systems. The evaluated model imposes a temporal structure with an autoregressive dynamic of order 1 with spatially correlated residuals. An approach using Stochastic Partial Differential Equations allows for representing a continuous Gaussian Field with Matérn covariance function as a discretely indexed field, providing great computational benefits by inducing sparse spatial precision matrices. The R-INLA package provides these calculations by representing the continuous field through basis functions defined on a triangulation of the spatial domain. Additionally, the package estimates parameters and obtains predictions using nested Laplace approximations. Simulations compare the mixed models evaluated with traditional methods, such as kriging and the conditional autoregressive model, showing reduced in execution time, adequate mean square prediction errors, and good coverage for the prediction confidence intervals. The model is evaluated on a set of images obtained by a satellite for an area of the Metropolitan Region of São Paulo.

Keywords: Kriging; Remote sensing; Spatial statistics; Spatio-temporal data; Time series. Stochastic Partial Differential Equation.

¹Institute of Mathematics and Statistics, University of São Paulo, São Paulo - vsmalves@ime.usp.br

²Institute of Mathematics and Statistics, University of São Paulo, São Paulo - chang@ime.usp.br

Modelo espaço-temporal hierárquico para a interpolação de dados faltantes em séries de imagens de satélite usando R-INLA

Vinicius Soares Martins Alves ¹
Chang Chiann ²

Resumo

Satélites e sensores cada vez mais avançados produzem grande volume de dados, e modelos estatísticos usuais podem ser inviáveis para o tratamento desses conjuntos. Ademais, dados faltantes ocorrem devido à presença de nuvens e outros artefatos atmosféricos. Este trabalho avalia o uso de modelos mistos dinâmicos com efeitos espaço-temporais aleatórios para a interpolação de dados faltantes em conjuntos de dados espaço-temporais dispostos em grade, comumente oferecidos através de sistemas de sensoriamento remoto. O modelo avaliado impõe uma estrutura temporal com dinâmica autorregressiva de ordem 1 com resíduos espacialmente correlacionados. Uma abordagem que usa Equações Diferenciais Parciais Estocásticas permite representar um Campo Gaussiano contínuo com função de covariância de Matérn em uma forma discretamente indexada, provendo grande vantagem computacional ao induzir matrizes de precisão espaciais esparsas. O pacote INLA do R provê esses cálculos representando o campo contínuo através de funções-base definidas numa triangulação do domínio espacial. Adicionalmente, o pacote estima parâmetros e obtém previsões usando aproximações aninhadas de Laplace. Simulações comparam o modelo avaliado com métodos tradicionais como a krigagem e modelo autorregressivo condicional, apresentando redução em tempo de execução, erro quadrático médio de predição satisfatório e boa cobertura para os intervalos de confiança de predição. O método também é aplicado em um conjunto de imagens obtidas via satélite para uma área da Região Metropolitana de São Paulo.

Palavras-chave: Sensoriamento remoto. Krigagem. Dados espaço-temporais. Estatística espacial. Séries temporais. Equações Diferenciais Parciais Estocásticas.

¹Instituto de Matemática e Estatística da Universidade de São Paulo, São Paulo - vsmalves@ime.usp.br

²Instituto de Matemática e Estatística da Universidade de São Paulo, São Paulo - chang@ime.usp.br