Effectiveness of gap-filling in tropical tree canopy cover modelling using Landsat time series

Zhipeng Tang[1], Petri K. E. Pellikka[2], and Janne Heiskanen[1]

[1]Department of Geosciences and Geography, University of Helsinki
[2]Department of Geosciences and Geography

November 21, 2022

Abstract

Satellite-based land surface phenology information (e.g. time-series spectral metrics and seasonal composites) can be used to model tropical tree canopy cover (TCC) but is limited by missing observations caused by clouds and cloud shadows and the sensor failure. Gap-filling (i.e. reconstruction of missing data in images) provides a key to solving the limitation, but its effectiveness and necessity for a given application, e.g. TCC modelling have been rarely explained. The main goal of this study was to examine the effectiveness and necessity of gap-filling, applied to the TCC modelling using Landsat time series. We chose Missing Observation Prediction Based on Spectral-Temporal Metrics (MOPSTM) method for its simple tuning, fast computation, and good performance in Landsat missing data reconstruction. MOPSTM models the relationship between valid observations in time series and spectral-temporal metrics in the k-Nearest Neighbor regression to predict the missing observations. We provided a quantitative comparison of TCC modelling using predictor variables (e.g. seasonal composite, spectral-temporal metrics, and harmonic regression coefficients) derived from Landsat time series that included gap-filling versus those that did not include gap-filling. With 1-year Landsat 8 Surface Reflectance time series acquired from a tropical area in Taita Hills, Kenya throughout 2015 and a reference TCC map scaled in $0 - 100$ derived from the Airborne Laser Scanning (ALS) data between 2014 and 2015, we applied random forest regression to model TCC using the predictor variables. The results indicated that TCC modelling using gap-filled predictors yielded smaller root mean square error than that using the non-gap-filled predictors, which proved the effectiveness of gap-filling in tropical TCC modelling. The effects of gap-filling might be reduced when the predictor variables were of high quality, e.g. median composite derived from the time series where sufficient observations exist. We concluded that gap-filling has a positive effect on the accuracy of TCC modelling.

Effectiveness of gap-filling in tropical tree canopy cover modelling using Landsat time series

Zhipeng Tang [a,b], Petri Pellikka [a,b], Janne Heiskanen [a,b]

a. Department of Geosciences and Geography, University of Helsinki, P.O. Box 68, FI-00014, Finland

b. Institute for Atmospheric and Earth System Research, Faculty of Science, University of Helsinki, Finland

Abstract

Satellite-based land surface phenology information (e.g. time-series spectral metrics and seasonal composites) can be used to model tropical tree canopy cover (TCC) but is limited by missing observations caused by clouds and cloud shadows and the sensor failure. Gap-filling (i.e. reconstruction of missing data in images) provides a key to solving the limitation, but its effectiveness and necessity for a given application, e.g. TCC modelling have been rarely explained. The main goal of this study was to examine the effectiveness and necessity of gap-filling, applied to the TCC modelling using Landsat time series. We chose Missing Observation Prediction Based on Spectral-Temporal Metrics (MOPSTM) method for its simple tuning, fast computation, and good performance in Landsat missing data reconstruction. MOPSTM models the relationship between valid observations in time series and spectral-temporal metrics in the *k*-Nearest Neighbor regression to predict the missing observations. We provided a quantitative comparison of TCC modelling using predictor variables (e.g. seasonal composite, spectral-temporal metrics, and harmonic regression coefficients) derived from Landsat time series that included gap-filling versus those that did not include gap-filling. With 1-year Landsat 8 Surface Reflectance time series acquired from a tropical area in Taita Hills, Kenya in 2015 and a reference TCC map scaled in 0 – 100 derived from the Airborne Laser Scanning (ALS) data between 2014 and 2015, we applied random forest regression to model TCC using the predictor variables. The results indicated that TCC modelling using gap-filled predictors yielded smaller root mean square error than that using the non-gap-filled predictors, which proved the effectiveness of gap-filling in tropical TCC modelling. The effects of gap-filling might be reduced when the predictor variables were of high quality, e.g. median composite derived from the time series where sufficient observations exist. We concluded that gap-filling has a positive effect on the accuracy of TCC modelling.

Key words: Tropical tree canopy cover; Landsat; spectral-temporal metrics; seasonal composite; harmonic; machine learning