An effective approach for filling large gaps in remotely sensed time-series

Zhipeng Tang (1,2), Janne Heiskanen (1,2), Hari Adhikari (1,2), Petri Pellikka (1,2)

(1) Department of Geosciences and Geography, University of Helsinki, Helsinki, Finland (zhipeng.tang@helsinki.fi), (2) Institute for Atmospheric and Earth System Research, University of Helsinki, Helsinki, Finland (zhipeng.tang@helsinki.fi)

Gaps of missing data due to clouds and sensor malfunctioning limit the use of optical satellite images for land surface characterization and monitoring. A large number of methods have been developed to fill gaps in the satellite images. However, there is still no method that can produce high-quality gap-filled images, especially in the presence of large gaps caused by clouds. In this study, we propose a novel machine learning based approach to fill the large gaps in Landsat time series. In the method, spectral-temporal metrics based on the time series, such as mean, median, and standard deviation, are extracted from the time-series of satellite images, and Random Forest algorithm is used to predict reflectance in the gaps in one target image of the time series. Furthermore, a moving neighborhood window with a specific size (for example, 8 pixels × 8 pixels) is used to calculate the values in the gaps. However, not all values can be filled by this method as the neighborhood is usually smaller than the gaps. We finally combine the values filled by random forest and those filled by the neighborhood window so as to integrate spectral-temporal and spatial context information. The results based on one year of Landsat 8 OLI images demonstrate that the proposed method can restore values in the large gaps accurately in the regions with high land cover heterogeneity.