EGU2020-18406
https://doi.org/10.5194/egusphere-egu2020-18406
EGU General Assembly 2020

Optimizing hybrid retrieval strategies for crop attribute retrieval using hyperspectral UAV data

Asmaa Abdelbaki, Martin schlerf, Jochem Verrelst, and Thomas Udelhoven

Unmanned aerial vehicle-based (UAV) hyperspectral imagery is of great significance to estimate crop attributes at a landscape scale, which is required for many environmental and agricultural applications. Multiple methods have been proposed such as empirical regression, radiative transfer, and hybrid models to derive target information (leaf area index (LAI), canopy chlorophyll content (CCC), and fractional vegetation cover (fCover)). Yet, it remains a challenge to select the most suitable method, since each method has its respective advantages and disadvantages. In this study, a hybrid strategy is proposed, as it combines the flexibility of regression with the universality of radiative transfer models (RTM) compared to other retrieval methods concerning model accuracy, computational efficiency under varying sample sizes and different levels of artificial noise. Two datasets of canopy spectra were simulated from two types of Look-up-tables (LUTs) for simulating a range of canopy reflectance-based on a set of input parameters from a Soil-Leaf-canopy RTM. The first type (LUTstd) was derived from a set of independent input parameters, while the other type (LUTreg) relied on the variable correlations by using the Cholesky algorithm. The LUTs were used for training linear and non-linear nonparametric regression algorithms for estimating the relevant parameters for characterizing 27 potato plots. Subsequently, the best approach of non-parametric regression methods was applied to UAV-based hyperspectral data for mapping of crop properties.

Results showed for LAI and fCover estimates that the principal component regression, partial least square regression, and least squares regression line (PCR, PLSR and LSLR) outperformed any of machine learning regression algorithms (MLRAs) and LUT inversion approaches. Besides, analysis of multiple LUT sizes ranging from 1000 to 17280 revealed that the 1000 simulations were sufficient for training LUTs. Also, adding 1% of noise to the simulations was adequate to imitate the uncertainty of UAV data. By using the independent ground data for validation, the PCR and PLSR methods yielded the lowest errors (R^2= 0.81, NRMSE=11.47% for LUTreg than LUTstd (R^2= 0.51, NRMSE= 22.61%). Regarding fCover, the accuracy of linear non-parametric and LUT-inversion approaches in LUTreg (R^2=0.75 and NRMSE=14.53% for PLSR and R^2= 0.78 and NRMSE=14.37% for LUT inversion based) was increased slightly rather than the results obtained from MLRAs (R^2= 0.76 and NRMSE= 14.74% for kernel ridge regression (KRR)). Regarding CCC, the best result was obtained using Random forest of tree bagger (RFTB) and fit ensemble (RFFE) for both LUTs. The accuracy of LUTreg did not improve as much as LUTstd through changing sample sizes (R^2= 0.80, NRMSE= 14.71% for LUTreg and R^2= 0.81, NRMSE= 13.93% for LUTstd). In terms of processing speed, the linear non-parametric methods were the fastest one as compared to MLRAs

(PCR=0.0097 and PLSR=0.013 seconds). In conclusion, compared to the two analyzed hybrid strategies (Linear and non-linear non-parametric regression), the use of LUT-inversion is not recommended for large images because of low prediction accuracy and slow processing speed.

Keywords: SLC model, LUT inversion based, linear non-parametric regression, machine learning, hybrid model, Leaf area index, fractional vegetation cover, leaf and canopy chlorophyll content.