Introduction

Recently, the hybrid model, that combines elements of physically-based and statistical regression methods, has been integrated to overcome the limitation of the parametric and physical methods.

Practically, the machine learning models (MLRAs) are trained on a simulated radiative transfer model (RTM) or SLC database to establish complex linear and non-linear non-parametric models linking the biophysical and biochemical variables and spectral reflectance.

The MLR toolbox within the ARTMO software package was used in this study to implement non-parametric modelling algorithms. These approaches were classified into linear (e.g., GLS, LSR, LSLR) and non-linear regressions (e.g., RF, SVR, GPR, CCF).

Research questions

1. To what extent does integrating the correlation structure of selected variables into the LUT approach using the SLC model improve the interested variables (LAI, fCover, CCC)?

2. Which non-parametric algorithm provides the best estimates toward the accuracy compared to LUT-inversion for LAI, fCover, and CCC retrievals?

3. Which non-parametric algorithm provides the best estimates toward the accuracy compared to LUT-inversion for LAI, fCover, and CCC retrievals?

4. How does the number of training sample size influence the performance of LUT inversion and MLRAs?

Data and Experimental design

UAV-hyperspectral data acquisition:

- **Region:** south west of Luxembourg
- **Location:** latitude 49°36’ 50.06” N and longitude 5°55’ 06.73” E to 5°55’12.52” E
- **Vegetation:** Victoria Variety of potato crop.
- **Six UAV flights** with a DJI octocopter were performed during the growing season 2016.
- **The hyperspectral Gamaya sensor** was capable of collecting spectral signals 41 bands ranging from 474–925 nm.

Experimental design and Ground data:

- The experimental field was subjected to three nitrogen fertilisation levels of 80, 180, and 280 kg/ha nitrogen for 9 replications.
- LAI was measured by Licor LAI-2000 instrument.
- Fractional vegetation cover measured visually.
- The SPAD-502 Konica Minolta instrument used to measure leaf chlorophyll.

Methodology

1. Comparison between LUT_reg and LUT_std through the whole crop season:

A) LUT inversion:

B) MLRAs (PLSR)

2. The sensitivity of training different sample size (500, 1000, 2000, 5000, 10000) from original dataset (17280 points):

A. Day 3 (in the early stage, cloudy cover 60%)

B. Day 5 (Tuber bulking and flowering season under sunny condition)

C. Day 7 (in the late stage, cloud condition 80%)

3. Evaluating different MLRAs of day 5 for LAI, fCover and CCC estimations:

Discussion and Conclusion

- using LUT inversion and MLRAs the Cholesky Decomposition algorithm in LUT approach of SLC-RTM (LUT_reg) has been improved the interested variables (LAI, fCover and CCC) through the crop growing season of potato compared to LUT_std which it did not take into account the correlation between variables.
- The findings of LAI revealed that 1000 of training datasets was sufficient for training MLRAs to get better accuracy rather than other subset of samples (500, 2000, 5000, 10000).
- In contrary, with LUT inversion the best accuracy was achieved when the original dataset (17,280 simulations) was used for estimations.
- Among the 7 non-parametric modelling algorithms evaluated here, PLSR performed best for LAI except the last two dates which were under the cloudy conditions, although the non-linear non-parametric regression methods were the best for estimating CCC for all dates, especially RF(TB).
- For fCover, the accuracy of LSLR and SVR predictions were the best and both methods derived similar results in term of NRMSE % compared to others in the whole dates of potato experiment.

Literature