Using different retrieval methods for evaluating retrieval performance based on UAV- hyperspectral data

Asmaa Abdelbaki [1,4], Martin Schlerf [2], Jochem Verrelst [3] and Thomas. Udelhoven[1]

1 Trier University, Environmental Remote Sensing and Geoinformatics, Trier, Germany
2 Luxembourg Institute of Science and Technology, Remote Sensing and Eco-hydrological Modelling Group, L-4422 Belvaux, Luxembourg
3 University of Valencia, Image Processing Laboratory (IPL), Parc Científic,, 46980 Paterna, València, Spain
4 Soils and Water Science Department, Faculty of Agriculture, Fayoum university, Fayoum 63514 , Egypt

EGU General Assembly 2020

Universität Trier
Environmental Remote Sensing & Geoinformatics

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(e.g., 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., PLSR, LSLR) and non-linear regressions (e.g., RF, SVR, GPR, CCF).

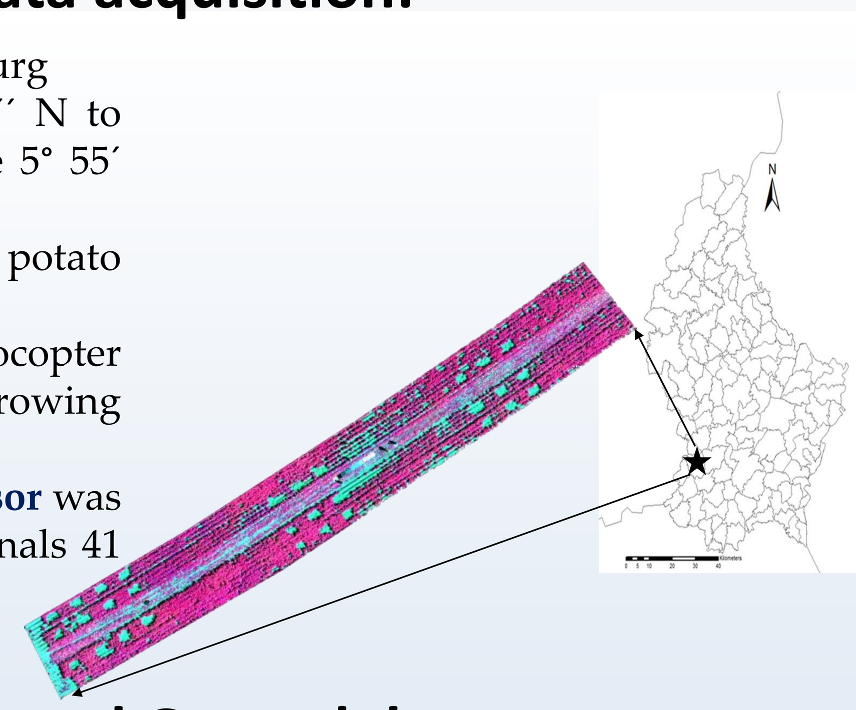
❑ Research questions

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

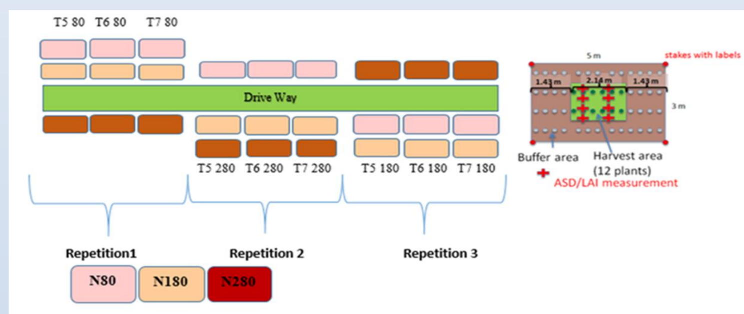
❑ UAV-hyperspectral data acquisition:

- **Region:** south west of Luxembourg
- **Location:** latitude 49° 36´ 47.13´´ N to 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 specral signals 41 bands ranging from 474-925nm.

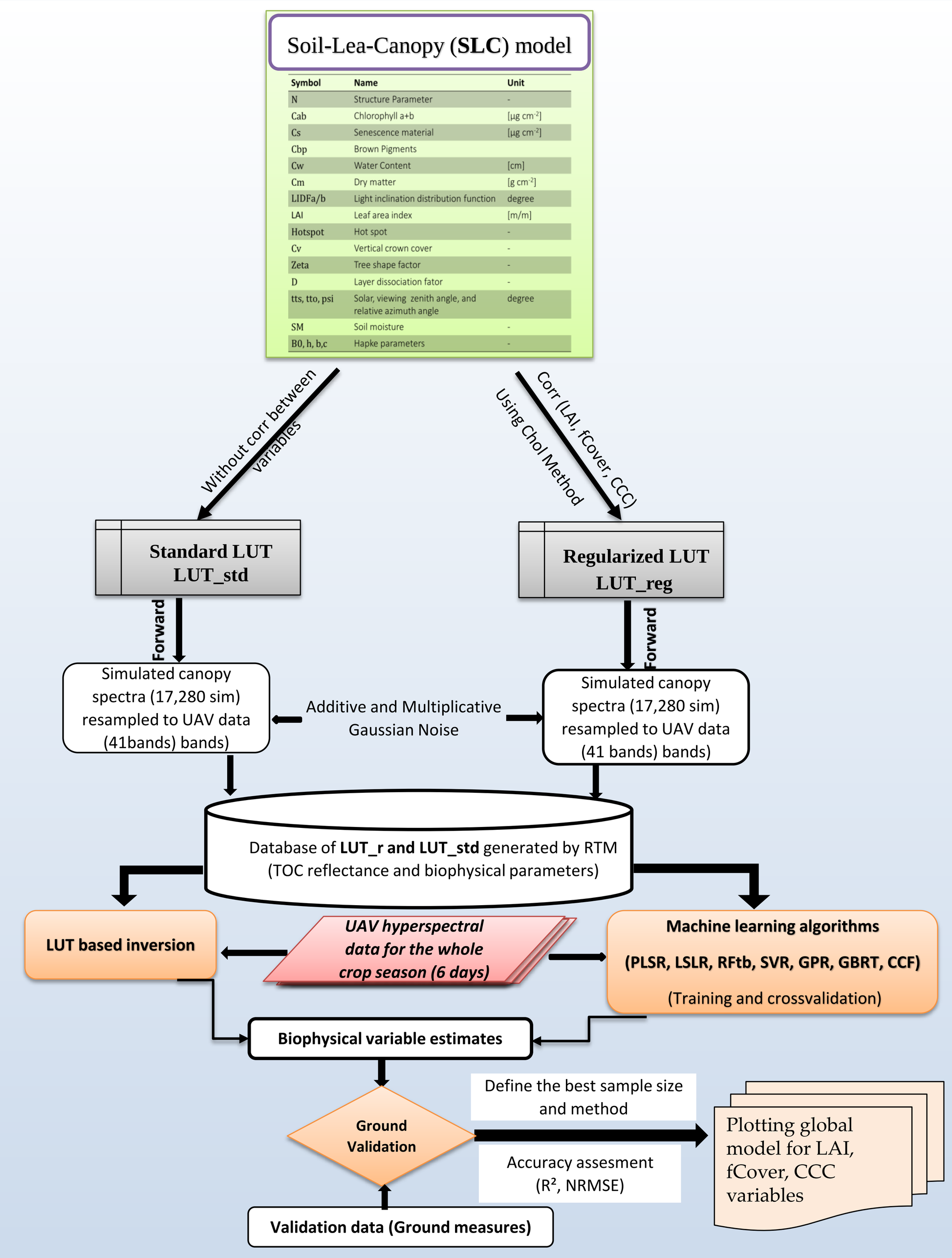


❑ Experimental design and Ground data:

- 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.



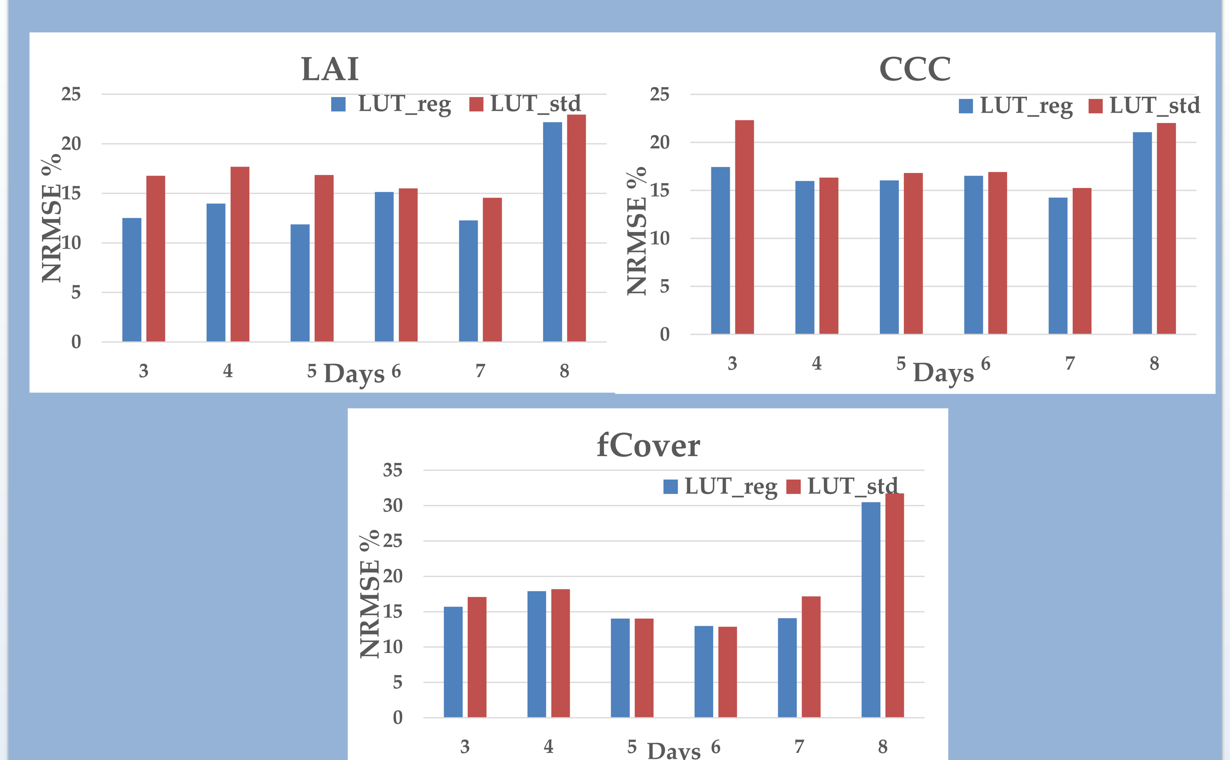
Methodology



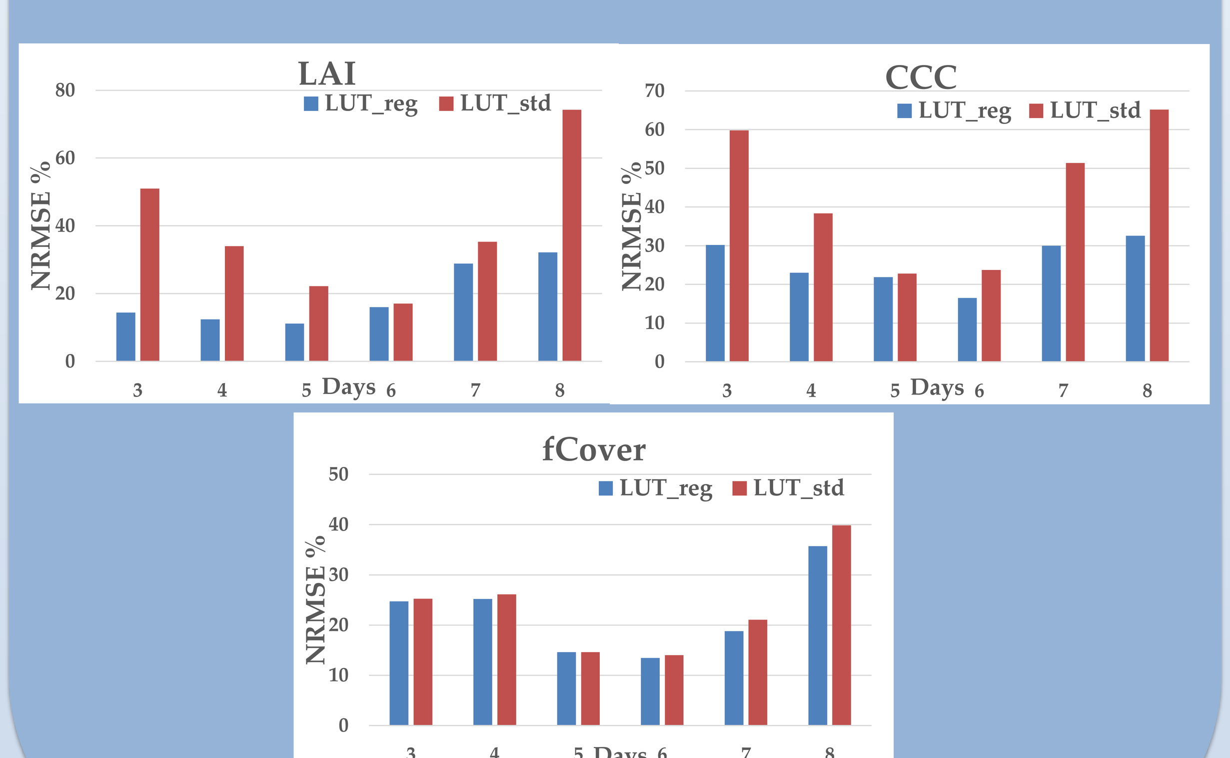
Results

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

A) LUT inversion:

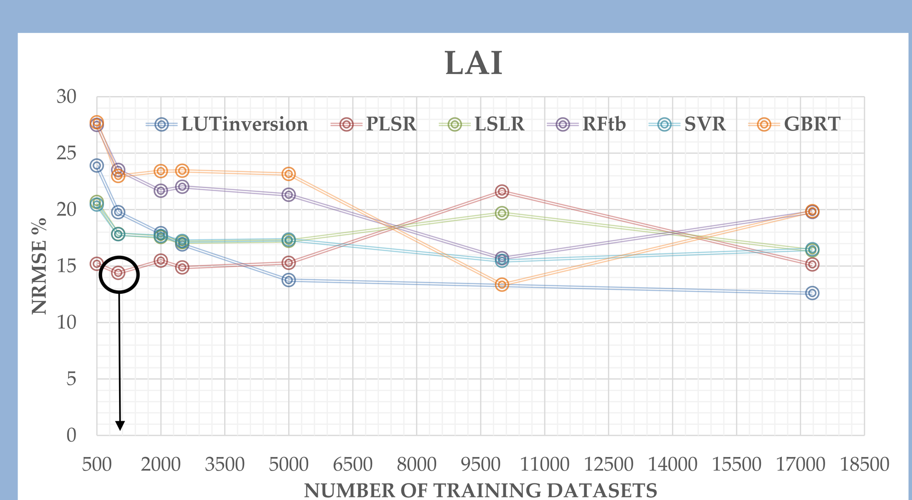


B) MLRAs (PLSR)

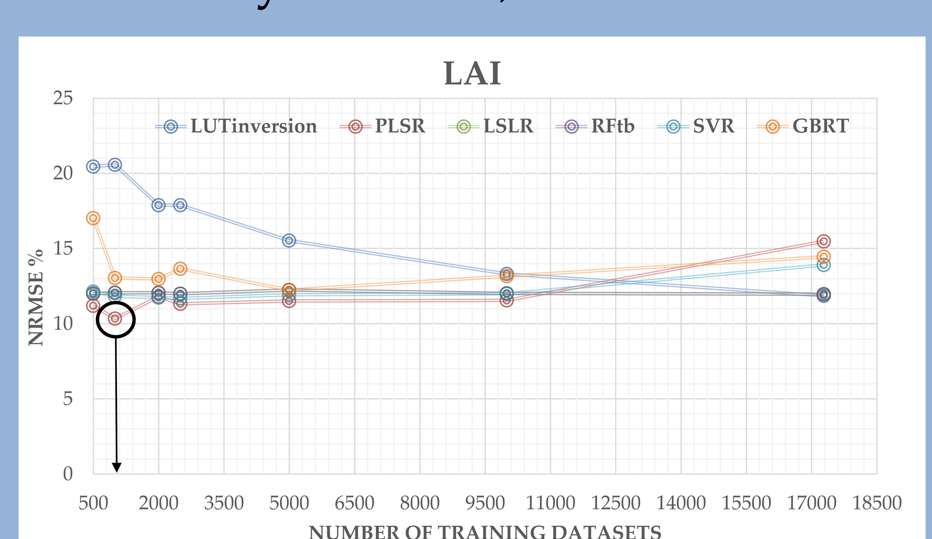


2. The sensitivity of training different sample size (500, 1000, 2000, 5000, 10000) from original dataset (17280sim):

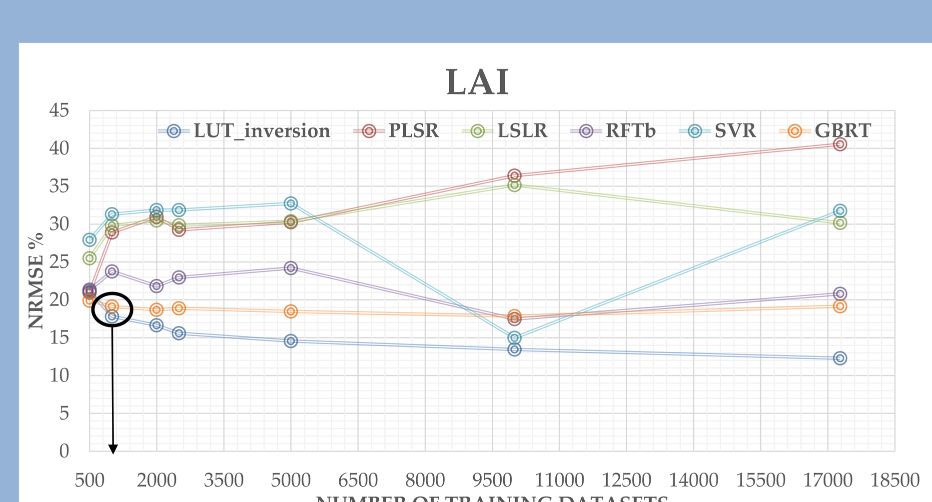
A. Day 3 (in the early stage, cloudy cover 60%)



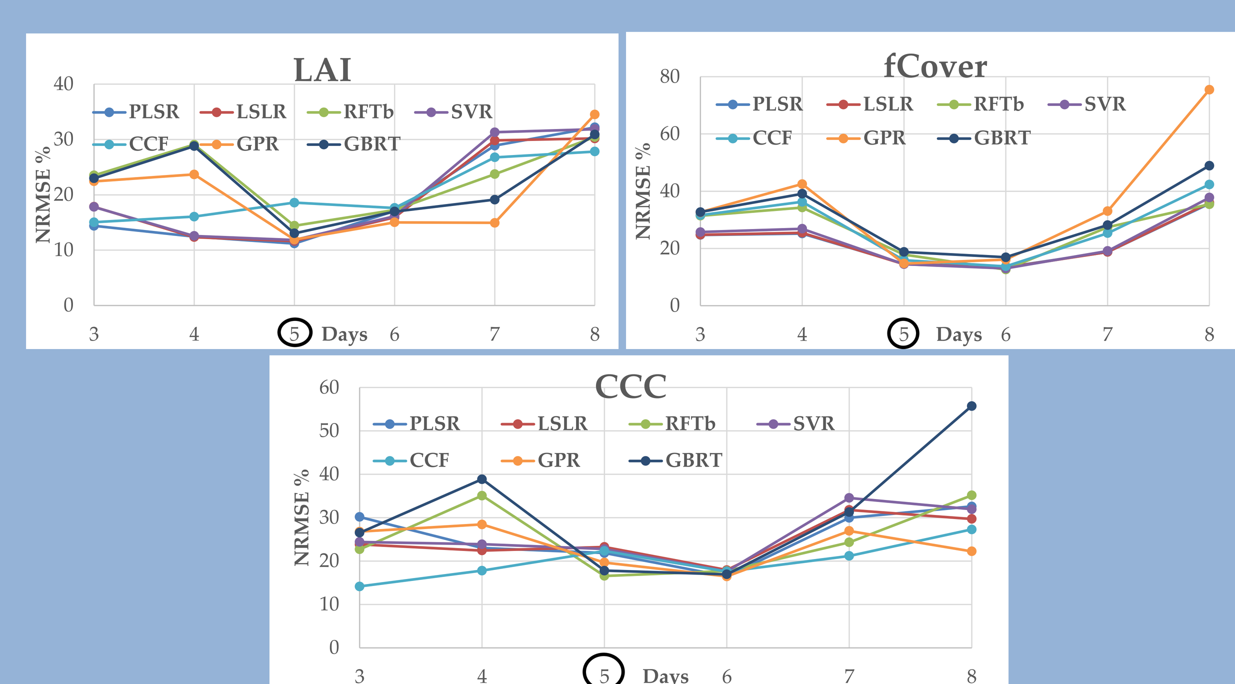
B. Day 5 (Tuber bulking and flowering season under sunny condition)



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



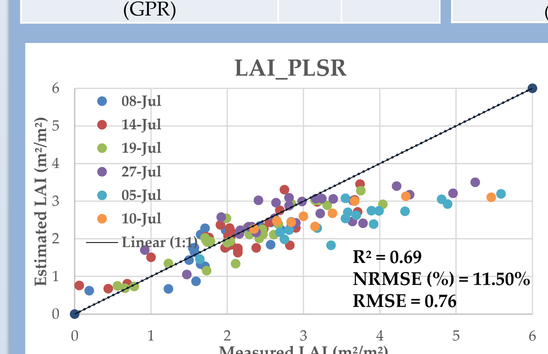
3. Evaluating different MLRAs of day 5 for LAI, fCover and CCC estimations :

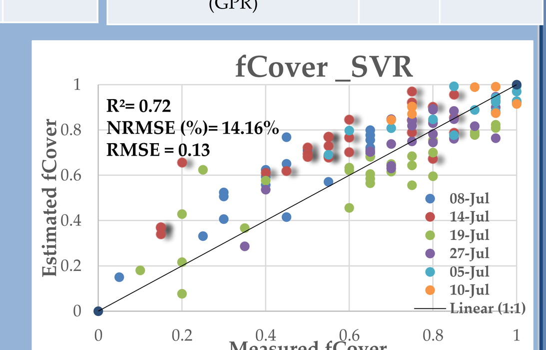


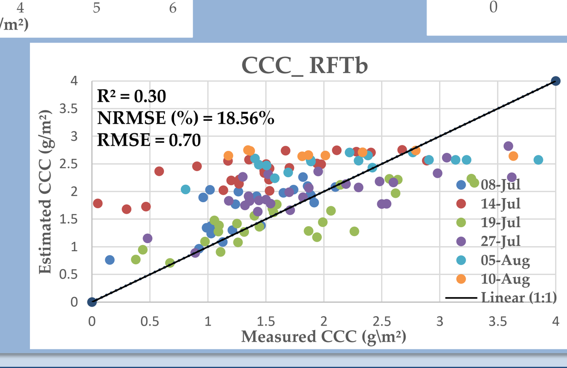
LAI Retrieved Methods	LUT_reg R^2	LUT_reg NRMSE%
LUT inversion (17,280 original dataset)	0.83	11.86
Partial least square regression (PLSR)	**0.82**	**11.19**
Least square linear regression (LSLR)	0.81	11.62
Random Forest (Tree Bagger) (RFTB)	0.82	14.4
Conical Correlation Forests (CCF)	0.62	18.58
Gradient Boosted regression tree (GBRT)	0.79	13.04
Support Vector Regression (SVR)	0.81	11.83
Gaussian Process Regression (GPR)	0.79	11.9

CCC Retrieved Methods	LUT_reg R^2	LUT_reg NRMSE%
LUT inversion (17,280 original dataset)	**0.83**	**11.86**
Partial least square regression (PLSR)	0.78	21.85
Least square linear regression (LSLR)	0.76	23.25
Random Forest (Tree Bagger) (RFTB)	**0.7**	**16.57**
Conical Correlation Forests (CCF)	0.59	22.33
Gradient Boosted regression tree (GBRT)	0.79	17.8
Support Vector Regression (SVR)	0.76	22.77
Gaussian Process Regression (GPR)	0.78	19.66

fCover Retrieved Methods	LUT_reg R^2	LUT_reg NRMSE%
LUT inversion (17,280 original dataset)	**0.83**	**11.86**
Partial least square regression (PLSR)	0.75	14.63
Least square linear regression (LSLR)	**0.76**	**14.52**
Random Forest (Tree Bagger) (RFTB)	0.65	17.86
Conical Correlation Forests (CCF)	0.75	15.93
Gradient Boosted regression tree (GBRT)	0.65	18.79
Support Vector Regression (SVR)	**0.75**	**14.52**
Gaussian Process Regression (GPR)	0.75	14.74







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

- Abdelbaki, A.; Schlerf, M.; Verhoef, W.; Udelhoven, T. Introduction of Variable Correlation for the Improved Retrieval of Crop Traits Using Canopy Reflectance Model Inversion. Remote Sens. 2019, 11, 2681.
- Jochem Verrelst, Juan Pablo Rivera, José Moreno, Gustavo Camps-Valls,Gaussian processes uncertainty estimates in experimental Sentinel-2 LAI and leaf chlorophyll content retrieval,ISPRS Journal of Photogrammetry and Remote Sensing,Volume 86,2013,Pages 157-167,ISSN 0924-2716.