A paper from the Proceedings of the 17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture

13-15 July 2026

Porto Alegre, Brazil

Inversion of Potato Chlorophyll Content Based on Radiation Transfer Model and Machine Learning Algorithm

Yuanyuan Ma[1,2], Jie Zhang[1], Di Pan[1], Qiang Wu[1], Xiaoyu Song[1,*], Xingang Xu[1]

（[1] Information Technology Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing,China,100097; [2]KCollege of Informatics, Huazhong Agricultural University, Wuhan 430070, China）

Abstract: *This study proposed an AL-assisted PROSPECT-4 hybrid framework for potato LCC estimation. AL selected informative samples from simulated spectra to train GPR, KRR, and PLSR models. GPR_PROSPECT achieved the best performance, with (R^2=0.742), RMSE= 4.207ug/cm2, and NRMSE = 9.921%, confirming the effectiveness of AL-assisted hybrid inversion.*

Keywords: *Potato, Chlorophyll content; Hyperspectral remote sensing; Hybrid method; Active learning*

1. Introduction

Leaf chlorophyll content (LCC) is a key indicator for crop monitoring and ecological assessment (Cheng et al., 2023; Houborg et al., 2007). Although radiative transfer models (RTMs) are physically interpretable, their inversion is often ill-posed (Lu et al., 2023). Hybrid methods improve RTM inversion by integrating simulations with machine learning (Chakhvashvili et al., 2022). In this study, active learning was used to select informative samples from PROSPECT-4 simulations, and GPR_PROSPECT, KRR_PROSPECT, and PLSR_PROSPECT models were developed for potato LCC estimation.

2. Materials and Methods

2.1. Experimental design

The experiment was carried out in the science Park of Keshan Farm of Qiqihar Branch of Heilongjiang State Administration of Land Reclamation in 2022. The distribution of the experimental plot is as follows.

2.2. Data analysis and methods

2.2.1. PROSPECT model

PROSPECT-4 was used to simulate potato leaf reflectance and generate LUT training data. The input parameters included Cab, Cw, N, and Cm, with ranges defined from measured data and literature. Latin hypercube sampling was applied, and the Cab range was expanded to 0-70 μg/cm^2 to support active learning (Table 1).

Table 1. PROSPECT-4 model input parameters.

Parameter	Unit	Min	Max	Samples
Chlorophyll content(Cab)	ug/cm^2	0	70	2
Equivalent water thickness(Cw)	g/cm^2	0.0001	0.08	2
Blade structure(N)	—	1.5	2.5	2
Dry matter content(Cm)	g/cm^2	0.0001	0.05	2

2.2.2. Active Learning

Active learning (AL) selects informative samples based on uncertainty or diversity criteria to improve regression model performance. It has been shown to enhance inversion accuracy and reduce limitations in hybrid inversion methods. Common AL strategies include PAL, EQB, RSAL, EBD, ABD, and CBD.

3. Results and analysis

3.1. Spectral curve analysis of potato leaves in different data sets

Measured, simulated, and AL-selected spectra showed similar trends and covered the observed spectral range, supporting subsequent model development.

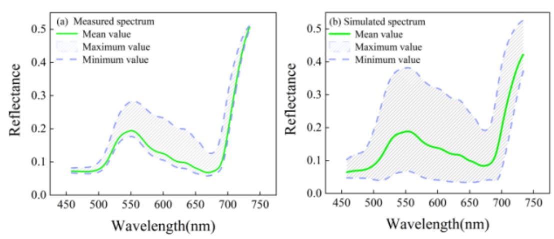

Figure 1. Analysis of Spectral Curves of Potatoes from Different Datasets.

3.2. Validation of inversion model for potato chlorophyll content

Validation using 180 measured samples showed that GPR_PROSPECT performed best, with (R^2=0.742), RMSE=4.207ug/cm^2, and NRMSE = 9.921%, outperforming KRR_PROSPECT and PLSR_PROSPECT. All models achieved the highest accuracy during the tuber growth stage and lower accuracy during the starch accumulation stage.

4. Conclusions and discussion

Active learning improved the efficiency of training sample selection from the PROSPECT-4 simulated data pool. GPR_PROSPECT showed the best validation performance among the three hybrid models, indicating its advantage in nonlinear potato LCC inversion.

References

1. Cheng J, Yang H, Qi J, et al. (2023). Evaluation of the effect of leaf spatial aggregation on chlorophyll content retrieval in open-canopy apple orchards. *International Journal of Applied Earth Observation and Geoinformation*.

2. Houborg R, Soegaard H, Boegh E. (2007). Combining vegetation index and model inversion methods for the extraction of key vegetation biophysical parameters using Terra and Aqua MODIS reflectance data. *Remote Sensing of Environment*, 106, 39-58.

3. Lu Y, Wang L, Zhu C, et al. (2023). Predicting surface solar radiation using a hybrid radiative Transfer‐Machine learning model. *Renewable and Sustainable Energy Reviews*.

4. Chakhvashvili E, Siegmann B, Muller O, et al. (2022). Retrieval of Crop Variables from Proximal Multispectral UAV Image Data Using PROSAIL in Maize Canopy. *Remote sensing*,14.