Computers and Electronics in Agriculture 156 (2019) 119–128

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Wheat leaf rust detection at canopy scale under different LAI levels using machine learning techniques

Mohsen Azadbakht[a,*], Davoud Ashourloo[a], Hossein Aghighi[a], Soheil Radiom[b], Abbas Alimohammadi[c]

[a] Remote Sensing and GIS Research Center, Faculty of Earth Sciences, Shahid Beheshti University, Tehran 653641255, Iran
[b] Applied Remote Sensing Department, Iranian Space Research Center, Tehran, Iran
[c] GIS Engineering Department, Faculty of Geodesy and Geomatic Engineering, K.N. Toosi University of Technology, Tehran 19967-15433, Iran

ARTICLE INFO

Keywords:
Disease detection
Wheat leaf rust
Hyperspectral signature
Canopy scale
LAI levels
Machine learning techniques

ABSTRACT

Accurate diagnosis of wheat leaf rust is of high interest for precision farming. Spectral data have been increasingly employed to detect this disease at leaf or canopy scales; however, less attention has been paid to the variations of leaf area index (LAI). Therefore, in this study, identification of wheat leaf rust was investigated at canopy scale and under different LAI levels, namely high, medium and low. Four machine learning (ML) methods including ν-support vector regression (ν-SVR), boosted regression trees (BRT), random forests regression (RFR) and Gaussian process regression (GPR) were built to estimate disease severity (DS) levels at canopy scale, where the reflectance data were measured in field situ by using a spectroradiometer, in which records spectra from 350 to 2500 nm. Results showed that ν-SVR outperformed the other ML methods at all three LAI levels with the R^2 measures all being around 0.99. The results, particularly, showed that the performances of the ML methods were improved with increasing LAI value, where RFR reported the worst R^2 value of 0.79 (RMSE = 8.5%) at low LAI level. The variable importance obtained using BRT showed three distinct regions of wavelengths that were appropriate across different LAI levels. The results of this research confirmed that hyperspectral signature can be reliably considered to identify wheat leaf rust disease at different LAI levels. Moreover, performances of several spectral vegetation indices (SVIs) were compared with those of the ML techniques. The results showed that the SVIs were consistently outperformed by the ML methods, particularly at low LAI level in which the SVIs were adversely affected. Nevertheless, all the SVIs, except for the RVSI, performed moderately well at high and medium LAI levels.

1. Introduction

Wheat leaf rust caused by *P. recondita f. sp. tritici (Prt)* is one of the most common and destructive fungal diseases of wheat which threats global food security. Plants infected by this disease exhibit various symptoms at different stages of development which can be simultaneously observed in various parts of the infected leaves and leaves can be in a variety of colors such as yellow, orange, gray or brown (Bolton et al., 2008).

Wheat leaf rust has been reported frequently in many wheat growing areas worldwide and it reduces both wheat yield quantity and quality. Hence, detection and monitoring of this disease is of paramount importance for precision farming (PF). Traditional disease detection methods often observe crops in situ which is costly and time consuming in large-scale studies. Therefore, an efficient method for rapid identification of plant diseases in large-scale wheat growing regions is of utmost necessity. Spectral data acquired from *in situ*, airborne and space borne sensors provide reliable information for early plant disease detection, being utilized in a plethora of studies (West et al., 2003; Mewes et al., 2011) at different scales including leaf-level (Yuan et al., 2013; Zhang et al., 2012), canopy level (Wang et al., 2015) and aerial/space remote sensing (Franke and Menz, 2007). Leaf-scale studies only consider optical properties of an infected leaf at different disease severity levels (Devadas et al., 2009). Disease detection at canopy scale, however, reflects not only the leaf properties but also the background soil, leaf area index and plant geometry (Franke and Menz, 2007). Finally, atmospheric effects on path radiance are investigated in aerial and space remote sensing (Kaufman and Tanre, 1992). In plant disease

* Corresponding author.
E-mail addresses: m_azadbakht@sbu.ac.ir (M. Azadbakht), d_ashourloo@sbu.ac.ir (D. Ashourloo), h_aghighi@sbu.ac.ir (H. Aghighi), soheil.radiom@gmail.com (S. Radiom), alimoh_abb@kntu.ac.ir (A. Alimohammadi).

https://doi.org/10.1016/j.compag.2018.11.016
Received 8 September 2018; Accepted 10 November 2018

detection studies, canopy scale studies are the basis for airborne and space borne disease detection methods (Zhang et al., 2012).

Hyperspectral signature are exceptionally prevalent in plant disease detection applications because they provide valuable information about disease development at early stages. Recorded reflectance data at the visible and near infrared (VNIR) wavelengths have shown high potential for plant disease detection compared to the short wave infrared (SWIR) and thermal infrared (TIR) wavelengths. This is because most diseases influence biochemical material content and structure of the leaves, and thus infected areas are more sensitive to the VNIR wavelengths. Hence, narrow spectral bands in the VNIR regions have been investigated in many studies to detect plant diseases such as wheat leaf rust (Ashourloo et al., 2014; Mahlein et al., 2013).

Leaf rust, yellow rust and stem rust are the three types of wheat rust diseases. Among them, leaf rust has the highest frequency of occurrence in the world, making it an important disease for future studies. However, it has been studied in few studies (Franke and Menz, 2007; Devadas et al., 2009; Ashourloo et al., 2014; Pryzant et al., 2017; Wang et al., 2015). A variety of techniques including classification (Franke and Menz, 2007), spectral vegetation indices (SVIs) (Thenkabail et al., 2000), spectral disease indices (SDIs) (Mahlein et al., 2013) and machine learning (ML) methods (Ashourloo et al., 2016) have been successfully adopted for plant disease detection. SVIs and SDIs are popular due to their simplicity and interpretability in disease detection. Most studies have utilized prevalent SVIs for disease detection, and few have developed SDIs for identification of a specific disease. However, SVIs and SDIs usually use few spectral bands in disease detection and not the full spectrum. On the other hand, ML methods have shown promising results due to their capabilities in handling complex associations between dependent and independent variables (Friedman and Meulman, 2003; Leathwick et al., 2006; Elith et al., 2008). Moreover, ML methods are known to handle data with high dimensionality. Plant disease detection is usually conducted in narrow regions of the spectrum. These methods are appearing superior to vegetation indices and they have shown satisfactory outcomes in detecting plant diseases (Ashourloo et al., 2014). In a study, conducted by Ashourloo et al. (2016), three ML techniques were investigated for wheat leaf rust detection at canopy level. Among the ML methods, GPR was superior, where all of the ML methods showed higher accuracy than the SVIs. In addition, various symptoms of the disease did not affect the ML methods as much as they did the SVIs (Ashourloo et al., 2016).

Spectral reflectance of vegetation is affected by several factors, including the background soil, leaf area index (LAI), leaf angle, leaf thickness, and chlorophyll content (Verhoef, 1984). Of these parameters, LAI plays an important role in affecting vegetation spectral signature (Clevers and Verhoef, 1993) and it changes across different wheat cultivars and under rainfed or irrigated wheat farms. So far, most of the disease detection studies have been conducted in field situ or in the lab without specifying variations of LAI. Therefore, it is necessary to evaluate the potential of different ML methods in disease detection at various levels of LAI. To the best of the authors' knowledge, identification of leaf rust disease using hyperspectral remote sensing data at canopy level has not been investigated under different LAI levels. Therefore, this paper is aimed at determination of leaf rust disease severity levels at canopy scale and under three different LAI levels including low, medium and high levels through adoption of boosted regression trees (BRT), GPR, SVR and random forest regression (RFR) as state-of-the-art ML techniques which have successfully been applied in several applications (Aghighi et al., 2018; Azadbakht et al., 2018; Mutanga et al., 2012; Elith et al., 2008).

The remainder of this manuscript is organized as follows. In Section 2, experimental setup, the ML methods and the selected SVIs are presented. Section 3 provides the results of our experiments in detail and under different conditions, and the concluding remarks are given in Section 4.

2. Materials and methods

2.1. Experimental setup

Field experiments were conducted at a site located in the Moghan fertilized plain, in the North West of Iran, in which approximately 7000 hectares of this plain is usually cultivated by wheat annually. This area is expanded from latitudes 39.465° N to 39.615° N and longitudes 47.548° E to 48.009° E.

To examine the potential of the ML methods in disease detection, hyperspectral reflectance data of healthy and diseased leaves were collected at canopy scale under natural environmental conditions. The spectra were collected using an ASD FieldSpec spectroradiometer (Analytical Spectral Devices, Inc., Boulder, CO, USA) with 5° field of view which operates at 350–2500 nm. The wavelength range of 450–1000 nm was only extracted from the spectra in order to avoid water absorption effects at longer wavelengths and also to acquire acceptable signal-to-noise ratio (SNR) (Mahlein et al., 2013). All hyperspectral spectra were recorded on clear and sunny days between 10:00 h and 14:00 h local time on 3–6 May 2017, where white board correction was conducted precedent to recording each individual hyperspectral data.

At 147.3 cm above ground and with the spectrometer's footprint of 3350 cm^2, a total of 284 samples were collected from different

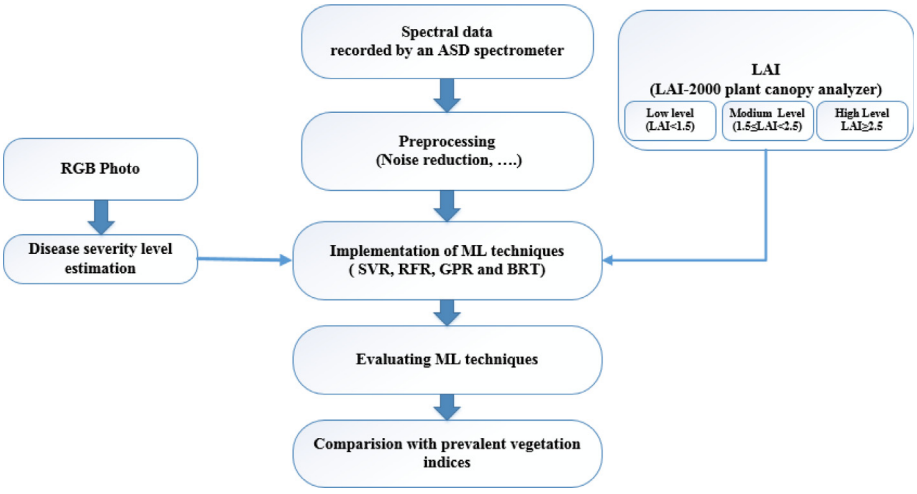

Fig. 1. Flowchart of the experimental setup.

arrangements of healthy and diseased leaves within the canopy, where the LAI values were measured using the LAI-2000 plant canopy analyzer (Li-Cor, Inc., Lincoln, NE, USA). Flowchart of the research procedure in this study is given in Fig. 1. In order to examine whether different LAI levels influence accuracy of the wheat leaf rust inversion models, the recorded spectra samples were divided into three distinct categories of low, medium and high LAI levels characterized with LAI $< 1.5\,m^2/m^2$, $1.5\,m^2/m^2 \leqslant$ LAI $< 2.5\,m^2/m^2$, and LAI $\geqslant 2.5\,m^2/m^2$, respectively. The collected spectra were comprised of 104, 92 and 88 samples at high, medium and low LAI levels, respectively. The spectra were then smoothed using a moving average filter of length 5 (Savitzky and Golay, 1964) and then quantized at the sample rate of 1:10 to generate spectra with the spectral resolution of 10 nm, comparable to common hyperspectral imaging sensors.

Subsequently, in order to determine disease severities at canopy scale, the process that is illustrated in Algorithm 1 was followed. Once the digital photos were taken from above the canopy, an RGB to HIS spectral transformation was then applied to make images device-independent, where the I component (intensity) was ignored from the next process (Al-Hiary et al., 2011). This was followed by performing the k-means clustering method on the images, and an appropriate mask was then employed to exclude edge pixels of the infected areas as well as pixels related to the background soil. Different levels of disease convey various colors owing to the different effects on wheat pigments (Fig. 2a). The H and S components were employed to extract the GLCM matrices in order to eliminate the light intensity effects during the process of detecting diseased areas. The maximum likelihood classifier was then employed and disease levels were extracted accordingly (Fig. 2c). Overall accuracy and Kappa coefficient for most of the images were above 0.95 and 0.9, respectively.

Algorithm 1. Schematic illustration of the disease severity detection.

1. Taking RGB photos.
2. RGB to HIS transformation.
3. k-Means clustering.
4. Masking the background soil.
5. Extracting the texture (from Step 2).
6. Maximum likelihood classification.

2.2. Wheat leaf rust disease inversion based on canopy hyperspectral data

In order to reduce any bias regarding the order of observations the data samples are initially shuffled, as a pre-processing step, before running the ML methods for 10 times. Prior to each run, samples are splitted into training, validation and test data, through adopting 5-fold cross-validation; where at a time, the dataset is divided into 5 equally sized folds and the test fold is left aside. Of the four remained folds, three are used for training at a time and the remained is used for validation of the model. This step is repeated four times to select a model which is applied to the related test fold. In the following subsections, the applied regression models, the performance evaluation measures, and relative feature importance measurement are presented.

2.2.1. Support vector regression

Unlike least squares regression, Support Vector (SV) machines pursue finding coefficients, in which only residuals larger than a constant threshold are of interest to minimize the loss function (James et al., 2013). SVM regression approach initially maps the input space onto a high-dimensional feature space, followed by performing linear regression in the latter feature space using ϵ-insensitive loss (Vapnik, 2013; Cherkassky and Mulier, 2007; Schölkopf et al., 1999). In order to adjust a nonlinear boundary between the attributes in the feature space, kernels are applied (James et al., 2013).

In this study, the recorded spectral signatures in the selected wavelengths and the disease severity level of the observations are representatives of the input and response variables, respectively. From derivatives of SVR, we implement ν-SVR, where the number of support vectors and training errors are restrained by using $\nu \in [0,1]$ (Chang and Lin, 2002). Given $x_i \in R^n$ $(i = 1,\ldots,M)$ as a set of M input observations comprised of n attributes and $y_i \in R$ as a target response, the following objective function is minimized using ν-SVR (Chang and Lin, 2002; Schölkopf et al., 2000; Vapnik, 2013).

$$\frac{1}{2}\|w\|^2 + C\left(\upsilon\varepsilon + \frac{1}{M}\sum_{i=1}^{M}\left(\xi_i + \xi_i^*\right)\right) \tag{1}$$

$$(w^T\Phi(x_i) + b) - y_i \leqslant \varepsilon + \xi_i$$

$$y_i - (w^T\Phi(x_i) + b) \leqslant \varepsilon + \xi_i^*$$

$$\xi_i, \xi_i^* \geqslant 0, \quad \varepsilon \geqslant 0$$

where ε is an error term, b is the constant offset, C is the regularization parameter, $\xi_i^{(*)}$ are the slack variables, and the function $\Phi(x_i)$ projects the input vectors x_i into a higher dimensional space, where a linear separator can be readily used to differentiate between them.

Of the existing kernels, Radial basis function (RBF) is used here because of its reliability and satisfactory accuracy level reported in the literature (James et al., 2013). To optimize the hyperparameters ν and C of Eq. (1) and $\gamma > 0$ of the RBF kernel, we performed grid search on the relevant range of values that these parameters can take from, using the e1071 package version 1.6–8 (Meyer et al., 2017) in R (R Core Team, 2017) version 3.4.2. The range $2^{-8} - 2^{+8}$ was considered for both the regularization parameter C and the RBF kernel parameter γ (Wang et al., 2015), while ν was searched within the interval $[0,1]$.

2.2.2. Gaussian process regression

A Gaussian process (GP) is known as a generalization of the Gaussian probability distribution (Rasmussen and Williams, 2006). GPR is known as a kernel-based statistical learning method to deal with nonlinear regression models (Lázaro-Gredilla et al., 2014; Verrelst et al., 2012). A GP model is a non-parametric model (Rasmussen and Williams, 2006), where a prior GPR is generated through considering the training dataset, which is then used in generating a posterior GPR (Ashourloo et al., 2016). Given the training inputs $\{x_i^{(b)}\}_{i=1}^N$ and target variables $\{y_i\}_{i=1}^N$, GPR captures various relations between these two in order to describe how they are correlated (Williams, 1998).

$$\widehat{y} = f(X) = \sum_{i=1}^{N} \alpha_i K\left(x_i.\, x_t\right) + \omega_0 \tag{2}$$

Here, ω_0 is a bias term, $\alpha_i \in \Re$ is the weight of the training inputs, and function K evaluates the similarities between the training inputs $x_i^{(b)}$ and the test set x_t.

According to the literature, changes in initial parameters does not affect the final results obtained from GPR (Bu and Pan, 2014). In this

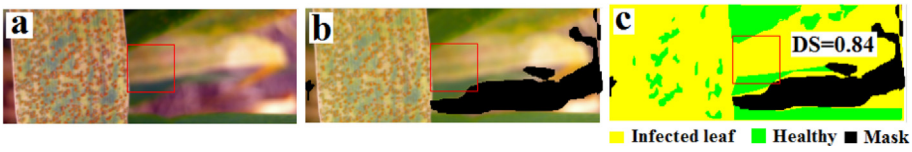

Fig. 2. Detection and classification of the rust-infected leaf regions (a: primary leaf, b: employing the mask, c: classified image with disease severity (DS) level of 0.84).

study, the essential hyperparameters of the squared-exponential covariance function; namely the length-scale ℓ, the signal variance σ_f and the noise variance σ_n are optimized using a freely available software at http://www.gaussianprocess.org/gpml/ (Rasmussen and Williams, 2006).

2.2.3. Boosted regression trees

Boosting is known as a compelling statistical learning concept, where the rationale behind it is that a combination of weak learners is used to create a single prevailing predictive model (Hastie et al., 2013). The integration of boosting (Schapire, 2003) with regression trees (Breiman et al., 1984) leads to BRT (Elith et al., 2008). BRT is a ML method which has shown statistical interpretability (Friedman et al., 2000). In addition to deriving complex nonlinearities between input vectors, provision of feature importance and handling high correlation between input predictors, it does not demand excluding outliers or to transform data before performing this method (Friedman and Meulman, 2003; Leathwick et al., 2006; Elith et al., 2008). BRT has been adopted for prediction of other plant diseases such as Fusarium head blight (FHB) (Landschoot et al., 2012; Shah et al., 2014).

At each step of boosting, a new best tree that provides the largest shrink of the loss function (residuals) is added to update the model, where the previous trees remain unchanged and a linear combination of all these trees build the final BRT model (Elith et al., 2008). In contrast to random forest which only reduces the model variance through averaging over all weak learners (trees), forward sequential fitting of BRT makes bias reduction of the model possible (Elith et al., 2008; James et al., 2013). However, boosting may overfit if large number of trees are used (James et al., 2013; Azadbakht et al., 2018). Therefore, regularization is essential for BRT to mutually optimize three hyperparameters: interaction depth, learning rate and number of trees. In this study, Bayesian optimization algorithm (Snoek et al., 2012) is employed to select the best hyperparameters.

A BRT model is as follows (Hastie et al., 2013):

$$\hat{f}_{BRT}^{A} = \sum_{m=1}^{M} T\left(x;\theta_m\right) \tag{3}$$

where $T(x;\theta_m)$ is m-th tree, and at each step of the procedure the following cost function should be solved (Hastie et al., 2013):

$$\hat{\theta}_m = \underset{\theta_m}{\operatorname{argmin}} \sum_{i=1}^{N} L(y_i, f_{m-1}(x_i) + T(x;\theta_m)) \tag{4}$$

where $f_{m-1}(x)$ is the current model, $\theta_m = \{R_j m, \gamma_j m\}_1^{J_m}$ is a set of the region set R_{jm} and constants γ_{jm} in m-th tree, J is a meta-parameter, R_{jm} refers to j-th disjoint non-overlapping region in m-th tree, y_i is the response associated with x_i, and L is the loss function.

2.2.4. Random forest regression

Random forests (RF) (Breiman, 2001) is a modification of bootstrap aggregation (bagging), where it attempts to further reduce the variance of the final prediction model through reduction of the between-trees correlation (Hastie et al., 2013; Azadbakht et al., 2018). To this end, a random subset of input variables is used to create splits for each tree. The RF prediction model $\hat{f}_{rf}^{B}$ for regression applications is finally made as an average of the total trees grown (Breiman, 2001; Hastie et al., 2013):

$$\hat{f}_{rf}^{B} = \frac{1}{B}\sum_{b=1}^{B} T\left(x;\theta_b\right) \tag{5}$$

where B is the total number of trees, $T(x;\theta_b)$ represents b-th tree, x is the target point that requires making prediction for, and θ_b is a random vector generated for b-th tree independent of previous vectors.

To construct a tree, a random bootstrap sample of n features is considered at each split. The left-out observations that are not used to build the predictive model are referred to as the out-of-bag (OOB) observations and an estimation of the test error can be obtained using the overall out-of-bag MSE of the model. The Bayesian optimization algorithm (Snoek et al., 2012) is applied to select the optimum parameters of RFR, including minimum leaf size and the number of predictors to use at each node. Through training a random forest of 500 regression trees, minimization of the quantile regression error delivers the best parameters.

2.3. Performance evaluation and ANOVA analysis

Performance of the methods are evaluated in terms of the coefficient of determination (R^2) and the root mean square error (RMSE), as common measures in assessing regression models (Chai and Draxler, 2014). To measure how consistant the ML methods perform across 10 runs, Bowley's measure of skewness (Bowley, 1917) is applied to the distributions of the assessment metrics (Eq. (6)).

$$SK_{Bowley} = (Q_3 + Q_1 - 2Q_2)/(Q_3 - Q_1) \tag{6}$$

where Q_i represents the i-th quartile of the distribution of the measures. In this case, $SK_{Bowley} = 0$ refers to a symmetrical distribution, while $SK_{Bowley} < 0$ and $SK_{Bowley} > 0$ refer to negatively and positively skewed distributions, respectively. Given the R^2 measure, a negatively skewed distribution is more preferred, indicating that the median (Q_2) is located closer to the third quartile (Q_3) than the first (Q_1) and thus the method is likely to perform favorably.

The results are tested for statistical significance at the $\alpha = 5\%$ level using a one-factor analysis of variance (ANOVA) (Berenson et al., 1983; Devore and Berk, 2011), by testing the hypothesis that the average performance measures of the presented methods are equal (null hypothesis) against the alternative hypothesis that at least one is different (see Eq. (7)). In the latter case, a pairwise comparison of the regression models is implemented to identify which models are significantly different using Tukey's honestly significant difference (HSD) and t-test statistics.

$$\begin{cases} H_0\text{: } \mu_{GPR} = \mu_{BRT} = \mu_{RFR} = \mu_{SVR} \\ H_1\text{:at least one is different} \end{cases} \tag{7}$$

where μ_X represents the mean performance measure, either R^2 or RMSE, of the X method.

2.4. Relative variable importance estimation

For BRT, the relative importance of the wavelengths, as input variables, is measured via the number of times that a given variable is used for splitting, where the sum of squared improvements of the model over all internal nodes averaged over all trees as follows (Friedman and Meulman, 2003; Hastie et al., 2013):

$$I_\ell^2 = \frac{1}{M}\sum_{m=1}^{M} \iota_\ell^2(T_m) \tag{8}$$

where T_m refers to the m-th tree, M is the total number of trees, $\iota_\ell^2(T_m)$ is the relative relevance for the predictor variable X_ℓ.

Using the OOB samples, variable importance is also reported for RFR, where the prediction strength of the variables is measured. In this case, the decrease in the model accuracy for each predictor variable while it is permuted against the OOB observations is averaged over all trees (Hastie et al., 2013). Given the j-th variable $X^{(j)}$, the OOB dataset of the ℓ-th tree as $\mathscr{D}_{\ell,n}$, $\mathscr{D}_{\ell,n}^{j}$ as the OOB dataset with $X^{(j)}$ being randomly permuted, and $m_n(.;\theta_\ell)$ as the ℓ-th tree estimate (Biau and Scornet, 2016), then:

$$\widehat{MDA}(X^{(j)}) = \frac{1}{M}\sum_{\ell=1}^{M} \left(R_n(m_n(.;\theta_\ell), \mathscr{D}_{\ell,n}^{j}) - R_n(m_n(.;\theta_\ell), \mathscr{D}_{\ell,n})\right) \tag{9}$$

Considering Y_i the response vector, R_n in Eq. (9) for either $\mathscr{D}_{\ell,n}$ or

Table 1
SVIs used for wheat leaf rust severity detection.

SVI	Description	Reference
NBNDVI	$(\rho_{850} - \rho_{680})/(\rho_{850} + \rho_{680})$	(Thenkabail et al., 2000)
PRI	$(\rho_{570} - \rho_{531})/(\rho_{570} + \rho_{531})$	(Gamon et al., 1992)
GI	ρ_{554}/ρ_{677}	(Broge and Leblanc, 2001)
RVSI	$((\rho_{570} - \rho_{531})/2) - \rho_{732}$	(Merton and Huntington, 1999)

$\mathscr{D}^{j}_{\ell,n}$ is defined as follows (Biau and Scornet, 2016):

$$R_n(m_n(.;\theta_\ell), \mathscr{D}) = \frac{1}{|\mathscr{D}|} \sum_{i:(X_i,Y_i)\in\mathscr{D}}^{M} (Y_i - m_n(X_i;\theta_\ell))^2 \quad (10)$$

2.5. Comparison of the ML methods with the SVIs

For further evaluation of the ML methods, their performances were compared against the common SVIs in detection of wheat leaf rust disease. The selected indices are the narrow-band normalized vegetation index (NBNDVI), photochemical reflectance index (PRI), green index (GI) and red-edge vegetation stress index (RVSI). Table 1 describes these indices, with ρ_λ being the spectral reflectance at the given wavelength λ (Ashourloo et al., 2014).

3. Results and discussion

3.1. Performance evaluation of the ML methods

In order to reduce the bias of random splitting, the ML methods were implemented 10 times on three categories of observations, representing different LAI levels. Boxplots in Fig. 3 represent the calculated R^2 values of the actual disease severity levels against those predicted from each model under 10 runs, indicating how well each model approximates the real disease levels. As seen, larger R^2 values are reported for all the regression models at high LAI level, followed by marginally deficient performance of the methods at medium LAI level. At low LAI level, however, the worst performances of the models are evident, with less deterioration of the ν-SVR performance. As shown in Fig. 3, ν-SVR was continuously superior at all LAI levels, showing the shortest inter-quartile ranges ($Q_3 - Q_1$) together with the largest values of the first, second and third quartiles at all three LAI levels. On the other hand, the longest interquartile ranges were emerged for GPR at both high and medium LAI levels, and for RFR at low LAI level (Fig. 3b). According to Eq. (6) and Table 2, the largest negative Bowley skewness values at both high and low LAI levels were for GPR (followed by RFR) and BRT, respectively. This indicates that these methods performed more favorably, indeed after ν-SVR, in terms of the distribution of the R^2 values. In these cases, the differences between the 3rd and 2nd quartiles ($Q_3 - Q_2$) were smaller than those between the 2nd and 1st quartiles ($Q_2 - Q_1$). In other words, most of the R^2 values of these methods were concentrated on the upper bound of the boxplots. At medium LAI level, however, all methods show small to moderate positive skewness values.

Table 2
Differences between the median and the 1st and 3rd quartiles of the R^2 values of the regression models.

Regression model	High LAI		Medium LAI		Low LAI	
	$Q_2 - Q_1$	$Q_3 - Q_2$	$Q_2 - Q_1$	$Q_3 - Q_2$	$Q_2 - Q_1$	$Q_3 - Q_2$
GPR	**0.02**	**0.01**	0.02	0.03	0.04	0.08
BRT	0.00	0.00	0.01	0.01	**0.06**	**0.01**
RFR	**0.01**	**0.00**	0.02	0.02	0.06	0.12
ν-SVR	0.00	0.00	0.00	0.00	0.00	0.00

As seen in Table 2 and Fig. 3b, the inter-quartile ranges of the methods varied from 0.0 to 0.02, 0.0 to 0.04 and 0.01 to 0.18 at high, medium and low LAI levels, respectively. ν-SVR was followed by GPR at all three LAI levels in terms of the three quartiles (1st, 2nd and 3rd), while the inter-quartile range of BRT was less than those of all other methods, except ν-SVR, as an indication of its less varied performance. At low LAI level, comparatively tall boxplots of RFR and GPR, together with their long bar charts indicate that these methods performed rather differently across several runs.

Fig. 4 shows the boxplot comparison of the four regression models at three different LAI levels in terms of the RMSE measure. Similar to Fig. 3a, the best overall performance was reported for ν-SVR in terms of the 1st, 2nd and 3rd quartiles. As it is shown in Fig. 4b, the shortest inter-quartile ranges were also resulted for ν-SVR at the three LAI levels, followed by BRT. Higher variabilities of GPR and RFR were concluded as a result of long inter-quartile values of the RMSE values, particularly at low and high LAI levels. The positive Bowley skewness values (Eq. (6)) of both BRT and ν-SVR were much larger at low LAI level, with $Q_2 - Q_1 < Q_3 - Q_2$ as in Table 3. RFR and GPR showed the highest skewness values at high LAI level; while similar to the results of the R^2 values in Table 2 and Fig. 3, no positive skewness values were emerged at medium LAI level.

An ANOVA analysis on the R^2 measure was performed to evaluate the regression models, where the p-values and the F-statistics of all three circumstances (LAI levels) were, respectively, less than the significance level 0.05 and larger than the critical F-values. These obviously indicate that we reject the null hypothesis and, thus, at least one of the methods performed significantly different at the 95% confidence level. The results of Tukey's HSD test are presented in Table 4 for pairwise comparisons between the ML methods, where the adjusted p-values of smaller than 0.05 indicate significant differences between the pairs. At medium LAI level, for example, all the methods performed significantly different on average, with the adjusted p-values of smaller than 0.05. At high LAI level, however, we failed to reject the null hypothesis ($\mu_1 = \mu_2$) between some pairs and the cases with no significant differences in average performances emerged between ν-SVR and GPR on one hand, and between BRT and RFR on the other hand. In fact, the

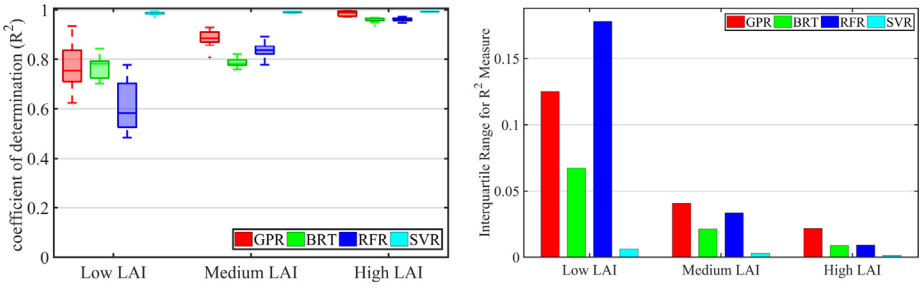

Fig. 3. (a) Boxplots, and (b) inter-quartile ranges of the methods for coefficient of determination (R^2).

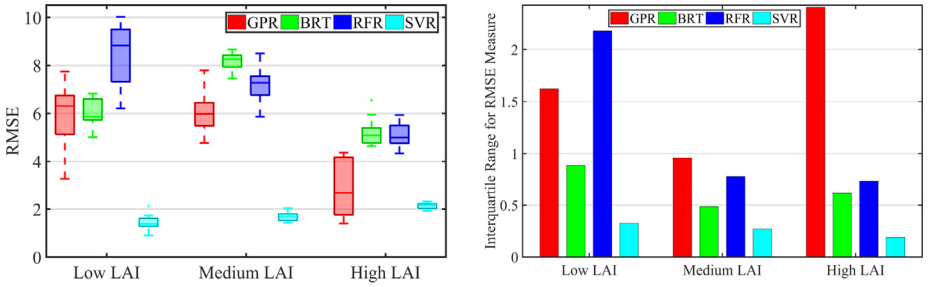

Fig. 4. (a) Boxplots, and (b) interquartile ranges of the methods for RMSE measure.

Table 3
Differences between the median and the 1st and 3rd quartiles of the RMSE values of the regression models.

Regression model	High LAI		Medium LAI		Low LAI	
	$Q_2 - Q_1$	$Q_3 - Q_2$	$Q_2 - Q_1$	$Q_3 - Q_2$	$Q_2 - Q_1$	$Q_3 - Q_2$
GPR	**0.92**	**1.49**	0.50	0.46	1.19	0.44
BRT	0.31	0.30	0.32	0.17	**0.13**	**0.76**
RFR	**0.23**	**0.50**	0.51	0.27	1.49	0.69
ν-SVR	0.12	0.07	0.14	0.13	**0.10**	**0.23**

former pair performed on average significantly different from the latter. Similar to high LAI level, we failed to reject the null hypothesis of the equality of the two means between BRT and GPR, at low LAI level. All other pairs of methods, however, performed significantly different at low LAI level; with the best performance of ν-SVR, followed by BRT, GPR and RFR, respectively (see Fig. 3a). A two-sample t-test pairwise comparison of the methods was also calculated by the R^2 performance metric with a 95% confidence interval. As shown in Table 4, the results were entirely similar to those obtained from Tukey's HSD test.

Similar to the results of R^2, the ANOVA analysis of the RMSE values showed that the average performance of at least one ML method, at all LAI levels, is significantly different at the 95% confidence level. Smaller p-values than the significance level 0.05 and larger F-values than the critical F-values are strong evidences of this. Therefore, the Tukey's HSD test was performed between the pairs of models, with the adjusted p-values being shown in Table 5. As seen, the only pairs with non-significant differences between the average performances were GPR-BRT at low LAI level, and ν-SVR-GPR and RFR-BRT at high LAI level. We reject the null hypothesis of the equality of means ($\mu_1 = \mu_2$) of all other pairs under other situations. Pairwise comparison of the methods using a two-sample t-test over the RMSE values with a 95% confidence interval was also conducted. As shown in Table 5, analogous to the results of the Tuckey's HSD test, we failed to reject the null hypothesis only when considering ν-SVR-GPR and BRT-RFR pairs at high LAI level, and BRT-GPR at low LAI level.

Fig. 5 shows the actual wheat leaf rust severity levels plotted against

Table 5
Pairwise comparison (p-value) of the regression models using Tuckey's HSD and t-test statistics on RMSE values.

Pairs	High LAI		Medium LAI		Low LAI	
	Tuckey's HSD	t-test	Tuckey's HSD	t-test	Tuckey's HSD	t-test
BRT-GPR	0.00	0.00	0.00	0.00	**1.00**	**1.00**
RFR-GPR	0.00	0.00	0.00	0.00	0.00	0.00
ν-SVR-GPR	**0.17**	**0.24**	0.00	0.00	0.00	0.00
RFR-BRT	**0.95**	**1.00**	0.02	0.02	0.00	0.00
ν-SVR-BRT	0.00	0.00	0.00	0.00	0.00	0.00
ν-SVR-RFR	0.00	0.00	0.00	0.00	0.00	0.00

the average predicted values across 10 runs of the regression models. As seen in this figure, all regression models performed satisfactorily on average at high LAI level (Fig. 5a–d), where GPR and ν-SVR showed marginally better R^2 values. At medium and low LAI levels, predicted values using ν-SVR, on average, provided the least deviations from the actual severity levels, with a goodness of fit similar to that obtained at high LAI level (Fig. 5h and l). The R^2 values indicated that the second least worst performance across different LAI levels was for GPR; whereas, RFR showed the worst performance at low LAI level ($R^2 = 0.79$). BRT, on the other hand, performed similarly at both medium ($R^2 = 0.82$) and low ($R^2 = 0.83$) LAI levels, performing marginally better than RFR at low LAI level ($R^2 = 0.79$).

Fig. 6 shows the relative relevance of each single wavelength for wheat leaf rust disease inversion model at the three LAI levels, averaged across 10 runs and normalized into the [0, 1] interval. The wavelengths of high relevance are highlighted with high relative importance levels, using BRT and RFR at each LAI level. The figures provided by BRT, as the mean decreases in Ginni index, indicated that almost distinct wavelength ranges play roles at different LAI levels. For example, the wavelengths of 450–550 nm contributed significantly at high LAI level, and the wavelength intervals of approximately 550–700 nm were also emerged as important, though with markedly lower importance level compared to the former interval. At medium LAI level, the highly relevant wavelengths were expanded from around 600 to 700 nm, while the wavelength ranges of 700–800 nm were highly relevant at low LAI

Table 4
Pairwise comparison (p-value) of the regression models using Tuckey's HSD and t-test statistics on R^2 values.

Pairs	High LAI		Medium LAI		Low LAI	
	Tuckey's HSD	t-Test	Tuckey's HSD	t-Test	Tuckey's HSD	t-Test
BRT-GPR	0.00	0.00	0.00	0.00	**0.99**	**1.00**
RFR-GPR	0.00	0.00	0.00	0.00	0.00	0.00
ν-SVR-GPR	**0.30**	**0.50**	0.00	0.00	0.00	0.00
RFR-BRT	**0.85**	**1.00**	0.00	0.00	0.00	0.00
ν-SVR-BRT	0.00	0.00	0.00	0.00	0.00	0.00
ν-SVR-RFR	0.00	0.00	0.00	0.00	0.00	0.00

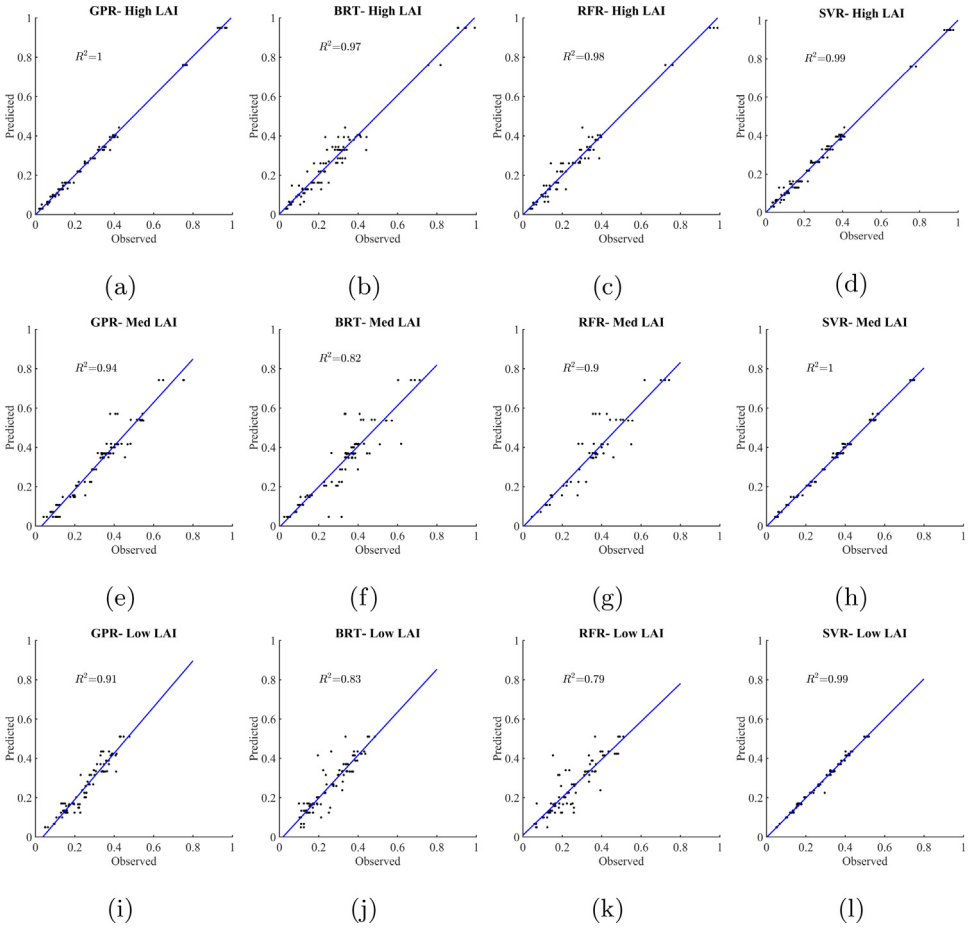

Fig. 5. Scatter plots of the actual and predicted disease severity levels of the regression models at high LAI level (a–d), medium LAI level (e–h), and low LAI level (i–l).

level. The important wavelengths highlighted using RFR, however, did not highlight distinct ranges at the three LAI levels. At high and low LAI levels, the wavelengths of 450–750 nm were highlighted with high contributions, showing a wider range of wavelengths in comparison with those selected by BRT, while a narrow wavelength range from 700 to 750 nm was highly emphasized in the case of medium LAI level. These differences can be explained in part by the fact that trees in these two ensemble regression methods are built using different algorithms, and also different variable importance measures are used in RFR and BRT methods (see Sections 2.2.3, 2.2.4, 2.4).

In this study, wheat leaf rust disease severity levels were investigated at three different LAI levels. We assumed that at a constant disease severity level, spectral reflectance values of the infected plants vary with the changes of LAI levels. As shown in Figs. 3–5, the ML methods, except ν-SVR, performed with different accuracies at the three LAI levels, indicating substantial effects of LAI on leaf rust disease detection. Therefore, the spectral signatures of random samples at two different disease severity levels at three LAI levels are illustrated in Fig. 7. As expected, the spectral responses differed between the samples of approximately similar leaf rust disease severity levels of different LAI values. In Fig. 7a, the canopy spectral reflectance of the sample with medium LAI at wavelengths beyond 700 nm was marginally higher than that of the equivalent sample with low LAI value, while the spectral reflectance of the sample of high LAI at a similar wavelength range was far higher than the spectral signatures of these two. In Fig. 7b, with the disease severity level of approximately 34%, the low LAI canopy showed higher spectral reflectance than both the canopy samples of high and medium LAI values at wavelengths from 450 to 700 nm. The

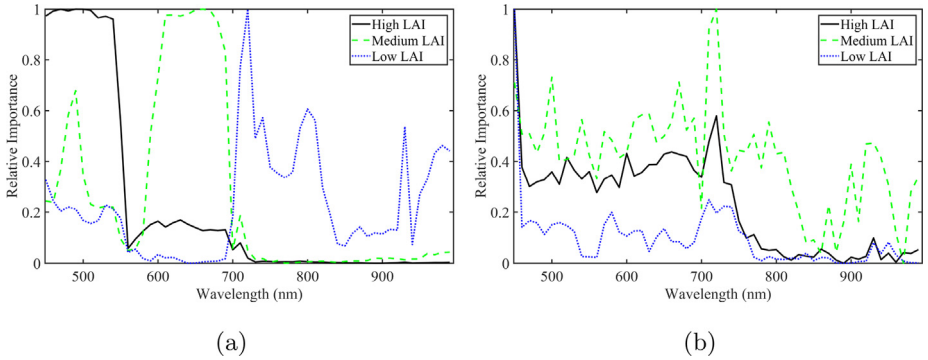

Fig. 6. Relative importance of the wavelengths for wheat leaf rust severity inversion at high, medium and low LAI levels for (a) BRT and (b) RFR regression models.

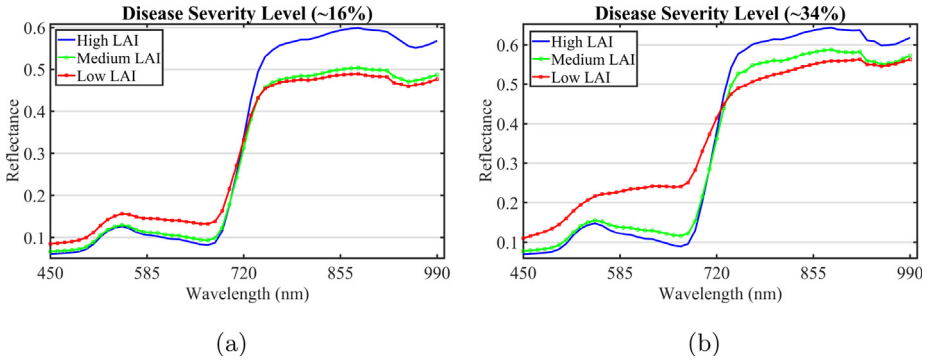

Fig. 7. Reflectance spectra of two samples of approximately (a) 16% and (b) 34% WLR disease severity levels.

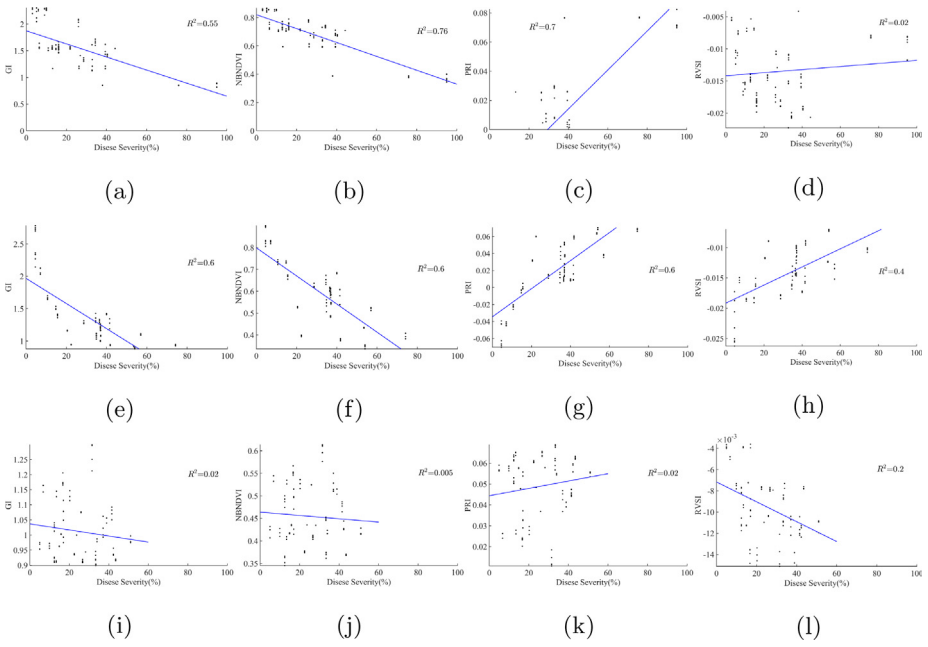

Fig. 8. Relations between the SVIs and disease severity levels (in %) at high LAI level (a–d), medium LAI level (e–h), and low LAI level (i–l).

canopy spectral reflectance of the high LAI sample, however, overtook the reflectance values of the other two corresponding samples at wavelengths beyond 700 nm. In Fig. 7b, the spectral signatures of the low LAI sample showed a steady increasing trend of the reflectance values within the wavelength interval of 550–700 nm, together with a significantly decent slope in the red edge area of the spectrum as an indication of the background soil contributing in the reflected spectra in both cases. These examples clearly show the significance of exploring wheat leaf rust disease severity levels under different canopy densities as well as to understand under what canopy density circumstances estimation of disease severity gets challenging.

The leaf rust disease inversion models utilized to provide estimation of the disease severities at canopy scale were GPR, BRT, RFR and ν-SVR. The regression models built here worked acceptably on the recorded canopy spectra, predicting the disease severity level based only on the spectral signatures of the samples. Nevertheless, better performance of the regression models was obtained at high LAI level, and the loss of LAI negatively influenced the performance of the models. These results deepened our understanding through providing further support for the hypothesis that the reduction of LAI can make detection of wheat leaf rust disease at canopy level difficult, in which contribution of the background soil in the recorded spectra increases.

Wheat leaf rust disease can cause both reduction of chlorophyll content of leaves and alterations of cellular structure, as well as affecting leaf water content (Ashourloo et al., 2014; Zhang et al., 2014; Line, 2002; Kuckenberg et al., 2009). A combination of these situations can influence canopy spectral response of the samples, which leads to various spectral signatures depending on the dominant phenomenon within the sensor's field of view such as background soil, low water or chlorophyll content of the leaves. This indicates that in order for leaf rust disease identification using airborne and spaceborne remote sensing sensors, *in situ* study of the canopy spectra from samples of different LAI levels should be carefully considered as a prerequisite.

3.2. Performance evaluation of the SVIs

The selected SVIs are plotted against the estimated disease severity levels under different LAI levels in Fig. 8. Although none of the SVIs provided acceptable R^2 values at low LAI level, they all performed moderately well at medium LAI level, with the RVSI showing the smallest R^2 value (0.4). Small R^2 values of all SVIs at low LAI level (Fig. 8i–l) can be linked to the contribution of background soil in the received spectrum as an explanation of insensitivity of the SVIs to model wheat leaf rust disease severity at low LAI level. At high LAI level, the NBNDVI and PRI performed better than the GI and RVSI, with the latter representing the worst performance (R^2 = 0.02). In general, the best performances were emerged at high LAI level for the NBNDVI (R^2 = 0.76) and PRI (R^2 = 0.7), while the RVSI did not perform well at all three LAI levels.

Fig. 8 reveals that the SVIs cannot successfully model this disease at

early stages, when the LAI level is low. Comparison of the results in Fig. 8 with those of the ML methods in Section 3.1 verifies robustness of the ML methods in modeling the wheat leaf rust disease severity under various LAI conditions, less afected by the contribution of background soil. Apart from the significant differences in performance of SVIs at different LAI levels, even the best performed SVIs (NBNDVI and PRI at high LAI level) were significantly outperformed by all the ML methods. However, performance evaluation of SVIs in wheat leaf rust modeling at different LAI conditions demands further investigation, using airborne and spaceborne sensors.

4. Conclusions

In this study, wheat leaf rust disease inversion models were built using ML methods and SVIs at canopy scale and under three distinct LAI levels. Hyperspectral signature data collected in field situ by using a spectroradiometer were served as inputs of the models, where all the ML inversion models performed acceptably at high and medium LAI levels. Of the implemented methods, however, ν-SVR was less affected by variations of LAI, followed by GPR. More research is required to determine the efficacy of the wheat leaf rust disease inversion models built in this study, on different remote sensing sensors, at different LAI levels and using different wheat cultivars. Moreover, feature selection conducted by BRT showed separate wavelength ranges with the highest contributions in the inversion model at the three LAI levels. The lower the LAI level, the higher the wavelength range highlighted as the most useful range that can result in superior outcomes. Future research should therefore concentrate on the investigation of this issue, using similar experimental set up in order to extract the most relevant wavelengths at each LAI level. Further modeling work will also have to be conducted in order to determine the performance of different SVIs in estimation of wheat leaf rust severity at different LAI conditions.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.compag.2018.11.016.

References

Aghighi, H., Azadbakht, M., Ashourloo, D., Shahrabi, H.S., Radiom, S., 2018. Machine learning regression techniques for the silage maize yield prediction using time-series images of landsat 8 oli. IEEE J. Sel. Top. Appl. Earth Observ. Rem. Sens. 1–15. https://doi.org/10.1109/JSTARS.2018.2823361.

Al-Hiary, H., Bani-Ahmad, S., Reyalat, M., Braik, M., ALRahamneh, Z., 2011. Fast and accurate detection and classification of plant diseases. Mach. Learn. 14. https://doi.org/10.5120/2183-2754.

Ashourloo, D., Aghighi, H., Matkan, A.A., Mobasheri, M.R., Rad, A.M., 2016. An investigation into machine learning regression techniques for the leaf rust disease detection using hyperspectral measurement. IEEE J. Sel. Top. Appl. Earth Observ. Rem. Sens. 9, 4344–4351. https://doi.org/10.1109/JSTARS.2016.2575360.

Ashourloo, D., Mobasheri, M.R., Huete, A., 2014. Developing two spectral disease indices for detection of wheat leaf rust (pucciniatriticina). Rem. Sens. 6, 4723–4740. https://doi.org/10.3390/rs6064723.

Azadbakht, M., Fraser, C.S., Khoshelham, K., 2018. Synergy of sampling techniques and ensemble classifiers for classification of urban environments using full-waveform lidar data. Int. J. Appl. Earth Observ. Geoinform. 73, 277–291.

Berenson, M., Levine, D., Goldstein, M., 1983. Intermediate Statistical Methods and Applications: A Computer Package Approach. Prentice-Hall. <https://books.google.com/books?id=87NvQgAACAAJ> .

Biau, G., Scornet, E., 2016. A random forest guided tour. Test 25, 197–227. https://doi.org/10.1007/s11749-016-0481-7.

Bolton, M.D., Kolmer, J.A., Garvin, D.F., 2008. Wheat leaf rust caused by puccinia triticina. Mol. Plant Pathol. 9, 563–575.

Bowley, A., 1917. Elements of Statistics. Studies in Economics and Political Science: Monographs. P.S. King & son. <https://books.google.com/books?id=M4ZDAAAAIAAJ> .

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Breiman, L., Friedman, J., Stone, C., Olshen, R., 1984. Classification and Regression Trees. The Wadsworth and Brooks-Cole Statistics-Probability Series. Taylor & Francis. <https://books.google.com/books?id=JwQx-WOmSyQC> .

Broge, N.H., Leblanc, E., 2001. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. Rem. Sens. Environ. 76, 156–172.

Bu, Y., Pan, J., 2014. Stellar atmospheric parameter estimation using gaussian process regression. Monthly Not. Roy. Astron. Soc. 447, 256–265. https://doi.org/10.1093/mnras/stu2063.

Chai, T., Draxler, R.R., 2014. Root mean square error (rmse) or mean absolute error (mae)? Geosci. Model Dev. Discuss. 7, 1525–1534.

Chang, C., Lin, C., 2002. Training nu-support vector regression: theory and algorithms. Neural Comput. 14, 1959–1978. https://doi.org/10.1162/089976602760128081.

Cherkassky, V., Mulier, F., 2007. Learning from Data: Concepts, Theory, and Methods. Wiley-IEEE, Wiley. <https://books.google.com/books?id=IMGzP-IIaKAC> .

Clevers, J., Verhoef, W., 1993. Lai estimation by means of the wdvi: a sensitivity analysis with a combined prospect-sail model. Rem. Sens. Rev. 7, 43–64.

Devadas, R., Lamb, D., Simpfendorfer, S., Backhouse, D., 2009. Evaluating ten spectral vegetation indices for identifying rust infection in individual wheat leaves. Precis. Agric. 10, 459–470. https://doi.org/10.1007/s11119-008-9100-2.

Devore, J., Berk, K., 2011. Modern Mathematical Statistics with Applications. Springer Texts in Statistics. Springer. <https://books.google.com/books?id=cv3pcEJ7amMC> .

Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. J. Anim. Ecol. 77, 802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.x.

Franke, J., Menz, G., 2007. Multi-temporal wheat disease detection by multi-spectral remote sensing. Precis. Agric. 8, 161–172. https://doi.org/10.1007/s11119-007-9036-y.

Friedman, J., Hastie, T., Tibshirani, R., et al., 2000. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). Ann. Stat. 28, 337–407. https://doi.org/10.1214/aos/1016120463.

Friedman, J.H., Meulman, J.J., 2003. Multiple additive regression trees with application in epidemiology. Stat. Med. 22, 1365–1381. https://doi.org/10.1002/sim.1501.

Gamon, J., Penuelas, J., Field, C., 1992. A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. Rem. Sens. Environ. 41, 35–44.

Hastie, T., Tibshirani, R., Friedman, J., 2013. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer Series in Statistics. Springer, New York. <https://books.google.com/books?id=yPfZBwAAQBAJ> .

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An Introduction to Statistical Learning: With Applications in R. Springer Texts in Statistics. Springer. <https://books.google.com/books?id=qcI_AAAAQBAJ> .

Kaufman, Y.J., Tanre, D., 1992. Atmospherically resistant vegetation index (arvi) for eos-modis. IEEE Trans. Geosci. Rem. Sens. 30, 261–270. https://doi.org/10.1109/36.134076.

Kuckenberg, J., Tartachnyk, I., Noga, G., 2009. Detection and differentiation of nitrogen-deficiency, powdery mildew and leaf rust at wheat leaf and canopy level by laser-induced chlorophyll fluorescence. Biosyst. Eng. 103, 121–128. https://doi.org/10.1016/j.biosystemseng.2008.09.018.

Landschoot, S., Waegeman, W., Audenaert, K., Vandepitte, J., Haesaert, G., De Baets, B., 2012. Toward a reliable evaluation of forecasting systems for plant diseases: a case study using fusarium head blight of wheat. Plant Disease 96, 889–896. https://doi.org/10.1094/PDIS-08-11-0665.

Lázaro-Gredilla, M., Titsias, M.K., Verrelst, J., Camps-Valls, G., 2014. Retrieval of biophysical parameters with heteroscedastic gaussian processes. IEEE Geosci. Rem. Sens. Lett. 11, 838–842. https://doi.org/10.1109/LGRS.2013.2279695.

Leathwick, J., Elith, J., Francis, M., Hastie, T., Taylor, P., 2006. Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. Marine Ecol. Prog. Ser. 321, 267–281. https://doi.org/10.3354/meps321267.

Line, R.F., 2002. Stripe rust of wheat and barley in north america: a retrospective historical review. Ann. Rev. Phytopathol. 40, 75–118. https://doi.org/10.1146/annurev.phyto.40.020102.111645.

Mahlein, A.K., Rumpf, T., Welke, P., Dehne, H.W., Plümer, L., Steiner, U., Oerke, E.C., 2013. Development of spectral indices for detecting and identifying plant diseases. Rem. Sens. Environ. 128, 21–30. https://doi.org/10.1016/j.rse.2012.09.019.

Merton, R., Huntington, J., 1999. Early simulation results of the aries-1 satellite sensor for multi-temporal vegetation research derived from aviris. In: Proceedings of the Eighth Annual JPL Airborne Earth Science Workshop, Pasadena, CA, USA, pp. 9–11.

Mewes, T., Franke, J., Menz, G., 2011. Spectral requirements on airborne hyperspectral remote sensing data for wheat disease detection. Precis. Agric. 12, 795. https://doi.org/10.1007/s11119-011-9222-9.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., 2017. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. < https://CRAN.R-project.org/package=e1071. r package version 1.6-8.

Mutanga, O., Adam, E., Cho, M.A., 2012. High density biomass estimation for wetland vegetation using worldview-2 imagery and random forest regression algorithm. Int. J. Appl. Earth Observ. Geoinform. 18, 399–406. https://doi.org/10.1016/j.jag.2012.03.012.

Pryzant, R., Ermon, S., Lobell, D., 2017. Monitoring ethiopian wheat fungus with satellite imagery and deep feature learning. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 1524–1532. https://doi.org/10.1109/CVPRW.2017.196.

R Core Team, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. < https://www.R-project.org/ > .

Rasmussen, C., Williams, C., 2006. Gaussian Processes for Machine Learning. Adaptative computation and machine learning series. University Press Group Limited. <https://books.google.com/books?id=vWtwQgAACAAJ> .

Savitzky, A., Golay, M.J., 1964. Smoothing and differentiation of data by simplified least

squares procedures. Anal. Chem. 36, 1627–1639.
Schapire, R.E., 2003. The boosting approach to machine learning: an overview. In: Nonlinear Estimation and Classification. Springer, pp. 149–171. https://doi.org/10.1007/978-0-387-21579-2_9.
Schölkopf, B., Burges, C., Smola, A., 1999. Advances in Kernel Methods: Support Vector Learning. Philomel Books. <https://books.google.com/books?id=_NYamXKkNM8C> .
Schölkopf, B., Smola, A.J., Williamson, R.C., Bartlett, P.L., 2000. New support vector algorithms. Neural Comput. 12, 1207–1245. https://doi.org/10.1162/089976600300015565.
Shah, D.A., De Wolf, E.D., Paul, P., Madden, L., 2014. Predicting fusarium head blight epidemics with boosted regression trees. Phytopathology 104, 702–714. https://doi.org/10.1094/PHYTO-10-13-0273-R.
Snoek, J., Larochelle, H., Adams, R.P., 2012. Practical bayesian optimization of machine learning algorithms. In: Advances in Neural Information Processing Systems, pp. 2951–2959.
Thenkabail, P.S., Smith, R.B., De Pauw, E., 2000. Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. Rem. Sens. Environ. 71, 158–182. https://doi.org/10.1016/S0034-4257(99)00067-X.
Vapnik, V., 2013. The Nature of Statistical Learning Theory. Springer, New York. <https://books.google.com/books?id=EoDSBwAAQBAJ> .
Verhoef, W., 1984. Light scattering by leaf layers with application to canopy reflectance modeling: the sail model. Rem. Sens. Environ. 16, 125–141.
Verrelst, J., Alonso, L., Camps-Valls, G., Delegido, J., Moreno, J., 2012. Retrieval of vegetation biophysical parameters using gaussian process techniques. IEEE Trans. Geosci. Rem. Sens. 50, 1832–1843. https://doi.org/10.1109/TGRS.2011.2168962.
Wang, H., Qin, F., Liu, Q., Ruan, L., Wang, R., Ma, Z., Li, X., Cheng, P., Wang, H., 2015. Identification and disease index inversion of wheat stripe rust and wheat leaf rust based on hyperspectral data at canopy level. J. Spectrosc. https://doi.org/10.1155/2015/651810.
West, J.S., Bravo, C., Oberti, R., Lemaire, D., Moshou, D., McCartney, H.A., 2003. The potential of optical canopy measurement for targeted control of field crop diseases. Annu. Rev. Phytopathol. 41, 593–614. https://doi.org/10.1146/annurev.phyto.41.121702.103726.
Williams, C.K., 1998. Prediction with gaussian processes: from linear regression to linear prediction and beyond. Nato Ser. Behav. Soc. Sci. 89, 599–621. https://doi.org/10.1007/978-94-011-5014-9_23.
Yuan, L., Zhang, J., Zhao, J., Huang, W., Wang, J., 2013. Differentiation of yellow rust and powdery mildew in winter wheat and retrieving of disease severity based on leaf level spectral analysis. Spectrosc. Spectral Anal. 33, 1608–1614. https://doi.org/10.3964/j.issn.1000-0593(2013)06-1608-07.
Zhang, J., Pu, R., Huang, W., Yuan, L., Luo, J., Wang, J., 2012. Using in-situ hyperspectral data for detecting and discriminating yellow rust disease from nutrient stresses. Field Crops Res. 134, 165–174.
Zhang, J., Pu, R., Loraamm, R.W., Yang, G., Wang, J., et al., 2014. Comparison between wavelet spectral features and conventional spectral features in detecting yellow rust for winter wheat. Comput. Electron. Agric. 100, 79–87. https://doi.org/10.1016/j.compag.2013.11.001.