remote sensing

MDPI

Article

Predicting Wheat Leaf Nitrogen Content by Combining Deep Multitask Learning and a Mechanistic Model Using UAV Hyperspectral Images

Xiao Ma [1,2], Pengfei Chen [1,3,*] and Xiuliang Jin [4]

[1] State Key Laboratory of Resources and Environment Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
[2] University of Chinese Academy of Sciences, Beijing 100049, China
[3] National Earth System Science Data Center, National Science and Technology Infrastructure of China, Beijing 100101, China
[4] Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China
* Correspondence: pengfeichen@igsnrr.ac.cn; Tel.: +86-10-64889895

Abstract: Predicting leaf nitrogen content (LNC) using unmanned aerial vehicle (UAV) images is of great significance. Traditional LNC prediction methods based on empirical and mechanistic models have limitations. This study aimed to propose a new LNC prediction method based on combining deep learning methods and mechanistic models. Wheat field experiments were conducted to make plants with different LNC values. The LNC and UAV hyperspectral images were collected during the critical growth stages of wheat. Based on these data, a method combining the deep multitask learning method and the N-based PROSAIL model was proposed and compared with traditional LNC prediction methods, including spectral index (SI), partial least squares regression (PLSR) and artificial neural network (ANN) methods. The results show that the new proposed method obtained the best LNC prediction results, with R^2, *RMSE* and *RMSE*% values of 0.79, 20.86 μg/cm^2 and 18.63%, respectively, during calibration and 0.82, 18.40 μg/cm^2 and 16.92%, respectively, during validation. The other methods obtained R^2, *RMSE* and *RMSE*% values between 0.29 and 0.68, 25.71 and 38.52 μg/cm^2 and 22.95 and 34.39%, respectively, during calibration and between 0.43 and 0.74, 22.79 and 33.55 μg/cm^2 and 20.96 and 30.86%, respectively, during validation. Thus, this study provides an accurate LNC prediction tool for precise nitrogen (N) management in the field.

Keywords: leaf nitrogen content; hybrid method; UAV; hyperspectral image

Citation: Ma, X.; Chen, P.; Jin, X. Predicting Wheat Leaf Nitrogen Content by Combining Deep Multitask Learning and a Mechanistic Model Using UAV Hyperspectral Images. *Remote Sens.* **2022**, *14*, 6334. https://doi.org/10.3390/rs14246334

Academic Editor: Prasad S. Thenkabail

Received: 1 November 2022
Accepted: 10 December 2022
Published: 14 December 2022

1. Introduction

Nitrogen (N) is an important nutrient element that is necessary for wheat growth and development. N plays an important role in improving crop photosynthetic capacity and assimilation products. Due to the large spatial variations in soil N content in fields, farmers tend to use excess N fertilizer to ensure wheat yield. However, excessive N in fields not only reduces wheat yield and quality [1], but also negatively affects the global ecosystem [2]. Therefore, obtaining the N nutrition status of wheat over time and applying N fertilizer reasonably according to plant needs is of great significance.

Leaf nitrogen content (LNC) is an important indicator for determining wheat N nutrition status. The traditional LNC detection methods depend on destructive sampling and labor-intensive analyses in the laboratory, which are time-consuming and expensive [3]. Previous studies have shown that crop biophysical and biochemical parameters can be detected by remote sensing technology [4,5]. In recent years, the rapid development of unmanned aerial vehicle (UAV) remote sensing technology has provided an opportunity to obtain high spatial and time-resolution images in a flexible manner [6], which is extremely suitable for precisely managing fields.

At present, the prediction methods for crop biophysical and biochemical parameter inversion include empirical methods, mechanistic methods and hybrid methods. Considering LNC prediction, empirical methods apply quantitative regression models between the canopy spectral reflectance and LNC using calibration datasets to predict the LNC in the target area. Zhang et al. used UAV hyperspectral images to compare the performance of partial least square regression (PLSR), generic algorithm integrated with the PLSR (GA–PLSR), random forest (RF) and extreme gradient boosting (XGBoost) methods for LNC prediction in alpine meadow ecosystems [7]. Based on field measured hyperspectral data, Jay et al. used the spectral index (SI) method to predict the LNC of sugar beet [8]. Empirical methods are easy to apply but lack mechanistic explanations and require large quantities of data for model training. The designed models may have errors when applied in places with no training data [9]. Using vegetation biophysical and biochemical parameters as input variables, mechanistic models describe the radiative transmission process of electromagnetic waves in the canopy. The reverse process of the mechanistic models can be used to predict the desired input parameters [10]. According to the relationship between chlorophyll and nitrogen, Yang et al. used the equivalent N absorption coefficient to replace the chlorophyll-specific absorption coefficient in the PROSPECT model and designed a nitrogen-based PROSPECT (N-PROSPECT) model, which can be used to combine the canopy scale model named SAIL (scattering by arbitrarily inclined leaves) to invert LNC values using canopy spectra data [11]. According to previous studies, the model that combined the N-PROSPECT model and the SAIL model can be called N-PROSAIL model for short. Using the N-PROSAIL model, Li et al. successively predicted wheat LNC by using the look-up table method [12]. Mechanistic methods have clear mechanistic explanations. However, they require many input parameters that are not easy to obtain in practice [13]. When the input parameters are uncertain, problems may be encountered during the model inversion process [14]. Additionally, due to the complexity of the electromagnetic transmission process, existing mechanism models cannot fully describe the transmission process, which causes the simulated data generated by the mechanism models to not be exactly consistent with the actual measured data, resulting in prediction errors. Hybrid methods first generate a large quantity of simulated data based on the mechanistic models, then use the simulated data to help train the empirical model, and finally apply the trained model to predict the biophysical and biochemical parameters in the target area [15]. The purpose of hybrid methods is to try to combine the advantages of empirical and mechanistic methods. Until recently, few studies have been conducted using hybrid wheat LNC prediction methods and UAV images. However, related work has been performed on the inversion of chlorophyll, which can provide a reference for LNC inversion based on the hybrid methods. Xu et al. simulated a rice dataset based on the PROSAIL model, and then used the dataset to train a Bayesian network model, and finally applied the trained model to predict rice canopy chlorophyll content using multispectral UAV images obtained in the target area [16]. Zhang et al. used the dataset generated by PROSAIL to pretrain a deep neural network (DNN), and then fine-tuned the parameters of the pretrained model using field-measured data, and finally used the model to predict wheat leaf chlorophyll content [17]. The above studies represent existing strategies used in hybrid methods. However, as stated before, the simulated spectra may differ from the actual measured spectra; thus, these approaches are not always efficient. Consistent with existing research [18,19], our initial study showed that both models trained using only simulated data and models pretrained using simulated data and fine-tuned with measured data do not achieve good inversion results in LNC prediction, and these models face overfitting problems.

Deep multitask learning is an inductive transfer learning method. The model trains the network based on multiple tasks. By sharing information between interrelated tasks, the robustness of the model is effectively improved by taking advantage of the similarities and differences between different tasks [20–22]. Due to the complexity of the vegetation canopy structure, the simulated spectra data may differ from the actual measured spectra data, but they have the common features that follow the absorption properties of biochemical parameters. In other words, when the measured data are limited, the simulated data can help to train the model for LNC prediction, but the simulated data cannot be used as the measured data directly during model training. Thus, based on the idea of multitask deep learning, we can set two tasks: one is to train the model based on the simulated data and use it to predict simulated LNC, and another is to train the model based on actual measured data and use it to predict measured LNC. The two tasks can share part of the network structure to learn their common features, while using their separate network structures to learn their unique features. The model learns information from the simulated and measured data simultaneously and transfers the information learned from the simulated data to improve the training process of the measured data, ensuring that the model focuses on important features. In addition, the noise in the simulated data with different patterns helps to improve the model robustness, reducing the overfitting problem [23]. Therefore, we can combine deep multitask learning methods and mechanistic models to develop a hybrid method for LNC prediction; however, no such approach has been established.

In this study, a field experiment was conducted to obtain LNC and hyperspectral image data of winter wheat under different growth conditions. To provide technical support for high-precision LNC prediction, based on these data, the following objectives were set: (i) to propose a new robust hybrid method for LNC prediction by combining a deep multitask learning method and a mechanistic model and (ii) to compare the proposed method with traditional methods (SI, PLSR, artificial neural network (ANN)) to recommend the best model for LNC prediction.

2. Materials and Methods

2.1. Field Experiment

This study was conducted at the Yucheng Integrated Experiment Station (116°34′13″E, 36°50′00″N) of the Chinese Academy of Sciences during 2020–2021. The winter wheat cultivar "Jimai 22" was sown with 20 cm row spacing. The experiment included two irrigation treatments and five N treatments in 32 plots. Each plot was 10 m × 5 m in size. The two irrigation levels were 60% and 80% of field water capacity and arranged with a split-plot design. The five N levels were 0 kg N/ha (N1), 70 kg N/ha (N2), 140 kg N/ha (N3), 210 kg N/ha (N4) and 280 kg N/ha (N5) and were arranged in a randomized block design, with N1-N4 having three replicates and N5 having four replicates. Except for the N fertilizer, the other management measures were the same in each plot, and the experimental design is shown in Figure 1.

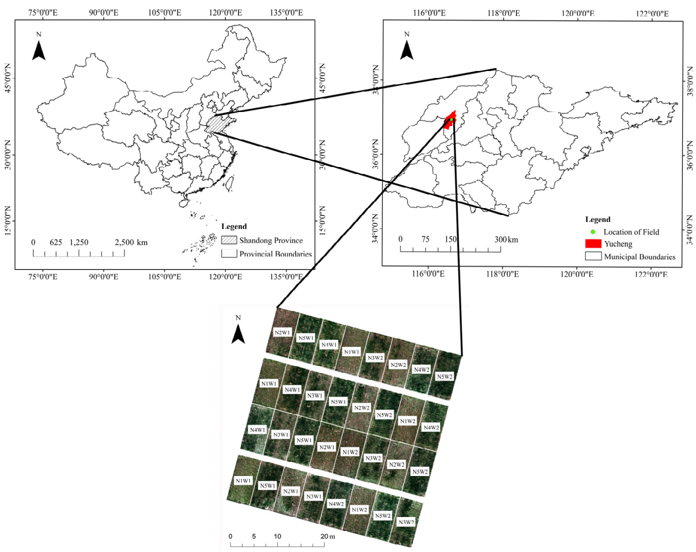

Figure 1. A schematic diagram of the water and nitrogen coupling experiment (W1: 80% field capacity; W2: 60% field capacity; N1: no fertilizer; N2: 70 kg N/ha; N3: 140 kg N/ha; N4: 210 kg N/ha; N5: 280 kg N/ha).

2.2. Field Data Acquisition

Field campaigns were conducted to obtain UAV hyperspectral images and ground-measured LNC data at critical wheat growth stages. These wheat growth stages are Feekes 4–5 (14 April 2021), Feekes 10.2 (7 May 2021) and Feekes 11.1 (17 May 2021).

2.2.1. UAV Image Data

The hyperspectral images were acquired by an M600 pro six-rotor UAV (SZ DJI Technology Co., Guangzhou, China) equipped with an S185 imager (Cubert, Ulm, Germany). S185 is a new type of snapshot hyperspectral sensor characterized by short exposure and integration times. The sensor can capture wavelengths from the visible to near-infrared spectra (450–950 nm) at a 4 nm spectral resolution. In one shot, it can obtain a panchromatic image with a pixel resolution of 1000 × 1000 and a hyperspectral image with a pixel resolution of 50 × 50. Based on the software provided by the sensor manufacturer, these two image types can be fused to make hyperspectral images having same resolution of panchromatic image. To minimize the influence of changes in the incident angle of the sun on the images, the flight time was selected between 10:00 and 12:00. The flight height was set to 30 m, and the forward and side overlaps were both set to 80%. A white panel image was obtained prior to UAV take-off and used to convert the image data from digital number (DN) values to reflection values in the subsequent processing step. In addition, to geometrically correct the obtained UAV images, high-accuracy ground control points (GCPs) were measured with a global navigation satellite system (GNSS) receiver GEO7X handheld global positioning system (GPS) device (Trimble, CA, USA) in network real-time kinematic (NRTK) mode. The NRTK service provided by Qianxun Company (Shanghai, China) and the GCPs measured under these conditions had errors below 1 cm. In the UAV image preprocessing step, first, Curbert Utils Touch (Cubert, Ulm, Germany) software was used to fuse hyperspectral and panchromatic images and export the fused images in .tiff format with approximately 0.84 cm of spatial resolution. Second, the Agisoft Photoscan (Agisoft LLC, St. Petersburg, Russia) software program was used to mosaic the fused images and convert the image DN values to reflectance values. Third, the sampled high-accuracy GCPs were used to geometrically correct the mosaicked images taken at the Feekes 4–5 growth stage, and the mosaicked images from other growth stages were then geo-rectified based on the above geo-corrected image. Finally, the images were resized and masked to retain only the experimental zones.

2.2.2. Field Sampled Data

Field sampling was carried out immediately after obtaining the UAV images. For each plot, representative areas with relatively uniform wheat growth status were selected as sampling sites. First, LAI-2200 (LI-COR Inc., Lincoln, NE, USA) was used to measure the leaf area index (LAI). Then, two rows of wheat plants with lengths of 30 cm were sampled at each selected site and taken to the laboratory, where the plants were divided into leaves and stems (including spikes) and dried in an oven until a constant weight was obtained. The N concentration values in each plant part were measured using the Dumas combustion method with a vario-MACRO cube analyzer (Elementar, Hanau, Germany). Finally, according to the LAI, leaf dry weight and leaf nitrogen concentration, LNC ($\mu g/cm^2$) values were calculated for each sampling site using Formula (1). Notably, the distances between the sampling points and the boundaries were recorded to allow for each sampling point to be accurately located in the geo-corrected images.

$$\mathrm{LNC} = \frac{\mathrm{W} \times \mathrm{N}\%}{\mathrm{LAI}} \tag{1}$$

where LNC represents the leaf nitrogen content ($\mu g/cm^2$) based on leaf area, W represents the leaf dry weight per unit ground area (g/m^2), LAI represents the leaf area index value and N% represents the leaf nitrogen concentration (%).

2.3. A New LNC Prediction Method

According to the logic framework stated in the Introduction Section, this study proposes a new LNC prediction method that combines the deep multitask learning method and the N-PROSAIL mechanism model. It is called the multitask learning-based hybrid model (ML-HM) in this study. As simulated spectrum and actual measured spectrum have both similarities and differences, the model adopts hard parameter sharing to construct a deep multitask learning network [24]. It includes two subtasks: training the network model based on the simulated data of the N-PROSAIL model to predict simulated LNC and training the network model based on the measured data to predict measured LNC. By sharing part of the network parameters among different tasks, information can be shared among multiple tasks. Therefore, the task of inverting the simulated LNC based on the model trained by the simulated data is used as an auxiliary task to improve the inversion accuracy of inverting the actual LNC based on the measured data. The training flowchart of the model is shown in Figure 2. The model proposed in this study has three key modules: a shared layer network, a subtask layer network and multitask optimization. The fundamental goal and implementation details of each module are described below.

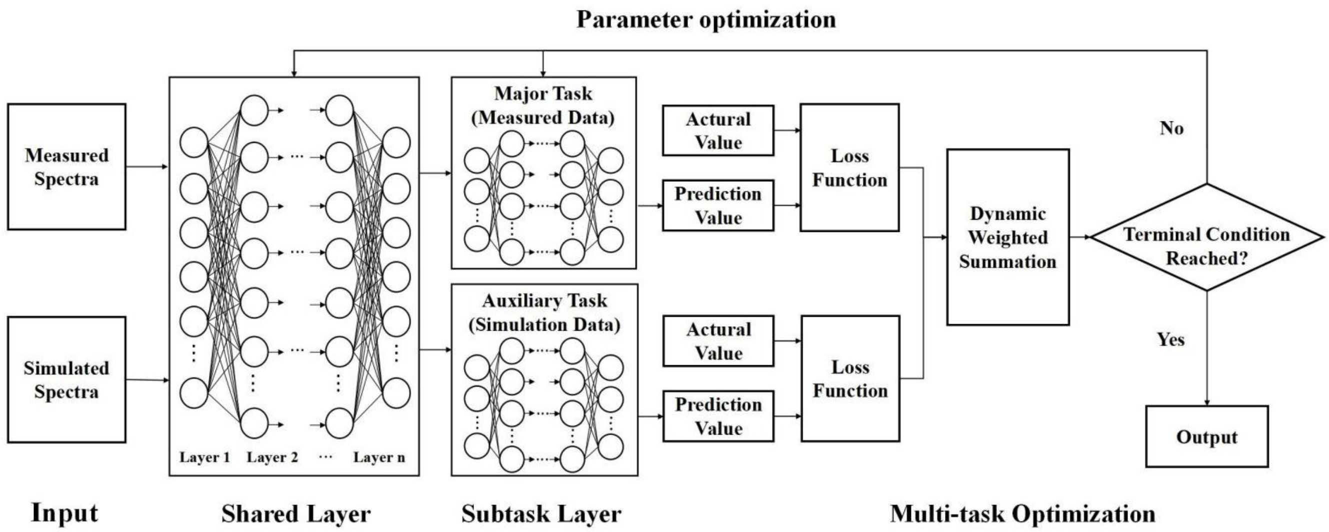

Figure 2. Main structure of the proposed model.

2.3.1. Shared Layer

This part of the network is shared by the major and auxiliary tasks. This component extracts the common features among the simulated and measured data. The input variables of the major and auxiliary tasks are the measured spectra data and simulated spectra data. The output variables are the common features used in the subsequent analyses for each task. Using this network structure reduces the risk of overfitting the network when training the major tasks based on limited measured data. The shared layer consists of six layers, with 20, 20, 20, 10, 10 and 10 neurons in sequence. A rectified linear unit (ReLU) is used as the activation function in each layer.

2.3.2. Subtask Layer

The subtask layers of the major and auxiliary tasks learn the unique characteristics of the simulated and measured data, respectively. They use the output variables of the shared layer as input variables, and the LNCs from the simulated and measured data are the output variables. Both subtask layers have three layers, with 10 neurons in each layer. The first layer uses ReLU as the activation function, and the latter two layers use traditional hyperbolic tangent (tanh) as the activation function. The subtask layers allow for the network to focus not only on the common features of the two tasks but also on the unique features of each task.

2.3.3. Multitask Optimization

This module integrates the LNC prediction error of the two tasks to determine the model parameters [25]. As each task impacts the network and different tasks have various convergence rates during network training, if each task has the same weight, the model is dominated by certain tasks. Thus, according to Liu et al. [26], we calculate the dynamic weighting error (DWA) according to the two tasks. The cost function is calculated using Formulas (2)–(5). To measure the prediction error, the mean square error (MSE) is calculated according to the predicted LNC and actual LNC for each task. For the entire process of this module, the deep multitask learning model runs according to the input data, and the corresponding cost function value is calculated; then, with the goal of minimizing the cost function, the backpropagation algorithm is run to optimize the model parameters. This process loops until the terminal condition is reached.

$$MSE_{all}(t) = \alpha_k(t) \times MSE_k(t) \tag{2}$$

$$\alpha_k(t) = \frac{W(t)}{\sum_{i=1}^{N} W_i(t)} \tag{3}$$

$$W_k(\mathrm{t}) = \frac{N \exp(r_k(t-1)/T)}{\sum_{i=1}^{N} \exp(r_i(t-1)/T)} \tag{4}$$

$$r_k(t-1) = \frac{MSE_k(t-1)}{MSE_k(t-2)} \tag{5}$$

In the above formulas, $MSE_{all}(t)$ represents the total LNC prediction error at training time t, $MSE_k(t)$ represents the mean square error of task k at training time t, $\alpha_k(t)$ represents the normalized weight of task k at training time t, $W_k(t)$ represents the unnormalized weight of task k at training time t, N represents the total number of tasks and is set to 2 in this study, $r_k(t-1)$ represents the training speed of task k at training time $t-1$, T is a constant and is set to 0.5 in this study and $MSE_k(t-1)$ and $MSE_k\ (t-2)$ represent the mean square error of task k at training times $t-1$ and $t-2$, respectively.

2.4. Data Analysis Method

To verify the performance of the ML-HM method, we compared the proposed approach with traditional methods (SI, PLSR and ANN) for LNC prediction. During this process, using collected UAV images, the averaged spectrum of all the pixels in the circle with the measured position of sampling points as center and two-times row spacing (40 cm) as the radius was used to correspond LNC values in the sampling point. It should be noted that as the image of S185 imager is the fused data from a high-resolution pan image and coarse resolution hyperspectral image, the pixels of the image are still the mixed pixels corresponding to the resolution of original coarse image and removing soil pixels cannot obtain a high crop N prediction accuracy [27]. Thus, we did not remove soil pixels. A total of 96 pairs of sampling data points were collected in this study. Three of these data points were found to have errors and were discarded. Thus, a total of 93 sampling points were retained. Based on the method of systematic random sampling, 75% of the samples were selected as the calibration dataset, and the remaining 25% were used as the validation dataset. The calibration dataset was used for model training, and the validation dataset was used for model validation. The R^2, $RMSE$ and $RMSE\%$ ($RMSE$/mean $\times$ 100%) values were used to evaluate the performance of each model. It should be noted that a large amount of data was also simulated using the N-PROSAIL model, and these data were used as the simulated data in the training of the ML-HM method. The main flow chart is shown in Figure 3 and the detailed description of each method is as follows.

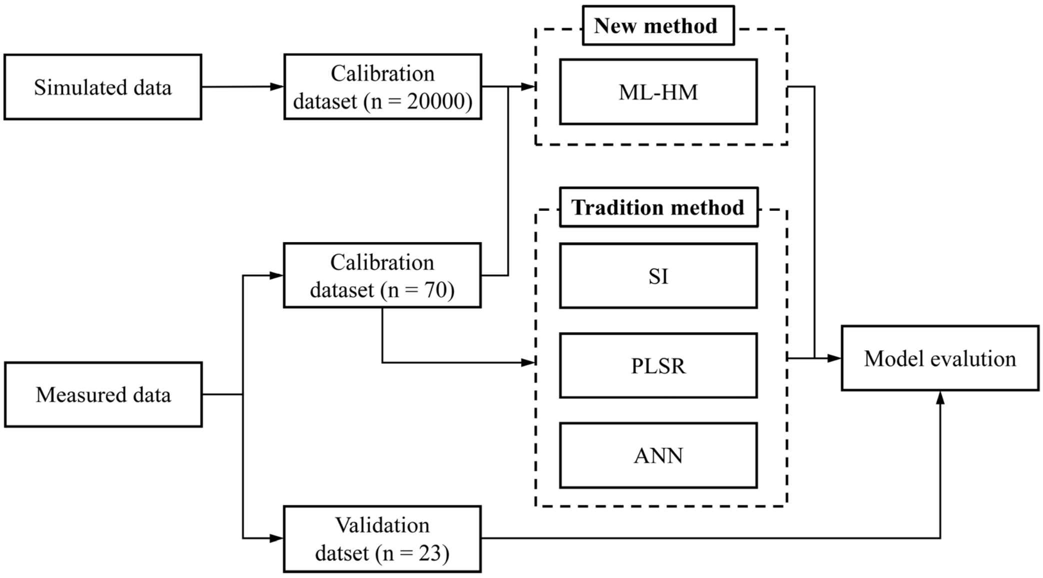

Figure 3. Main flow chart of the LNC prediction models designed based on SI, PLSR, ANN and ML-HM methods.

2.4.1. SI Method

The spectral index uses the band combination method to eliminate background noise and improves the sensitivity to target parameters. It is the most commonly used method for LNC prediction. In this study, based on previous studies, the commonly used spectral indices for estimating LNC were selected and are shown in Table 1. For each index, four types of models (linear, exponential, power and logarithmic models) were used to design LNC prediction models based on calibration dataset, and the best model with the highest R^2 and lowest *RMSE* and *RMSE*% values was selected and validated using validation dataset.

Table 1. Spectral indices used in this study.

Index	Name	Formula	Developed by
NDVI	Normalized difference vegetation index	$(R_{800} - R_{670})/(R_{800} + R_{670})$	[28]
GNDVI	Green normalized difference vegetation index	$(R_{800} - R_{550})/(R_{800} + R_{550})$	[29]
MSAVI	Modified soil adjusted vegetation index	$(2R_{800} + 1 - \text{sqrt}((2R_{800} + 1) - 8(R_{800} - R_{670})))/2$	[30]
OSAVI	Optimized adjusted vegetation index	$1.16(R_{800} - R_{670})/(R_{800} + R_{670} + 0.16)$	[31]
EVI	Enhanced vegetation index	$2.5(R_{800} - R_{670})/(R_{800} + 6R_{670} - 7.5R_{490} + 1)$	[32]
TVI	Triangular vegetation index	$0.5(120(R_{750} - R_{550}) - 200(R_{670} - R_{550}))$	[33]
MTVI2	Modified triangular vegetation index 2	$1.5(1.2(R_{800} - R_{550}) - 2.5(R_{670} - R_{550}))/\text{sqrt}((2R_{800} + 1)^2 - (6R_{800} - 5\text{sqrt}(R_{670})) - 0.5)$	[34]
RVI	Ratio vegetation index	R_{810}/R_{560}	[35]
NDRE	Normalized difference red-edge index	$(R_{790} - R_{720})/(R_{790} + R_{720})$	[36]
VI_{opt}	Optimal vegetation index	$(1 + 0.45)((R800)^2 + 1)/(R670 + 0.45)$	[37]
DNCI	Double peak canopy nitrogen index	$(R720 - R700)/(R700 - R670)/(R720 - R670 + 0.03)$	[38]
MCARI/MTVI2	Combined index I †	MCARI/MTVI2 MCARI: $(R700 - R670 - 0.2(R700 - R550))(R700/R670)$ MTVI2: $1.5(1.2(R800 - R550) - 2.5(R670 - R550))/\text{sqrt}((2R800 + 1)^2 - (6R800 - 5\text{sqrt}(R670)) - 0.5)$	[39]
MTCI	MERIS terrestrial chlorophyll index	$(R750 - R710)/(R710 - R680)$	[40]
TCARI/OSAVI	Combined index II †	TCARI: $3((R_{700} - R_{670}) - 0.2(R_{700} - R_{550})(R_{700}/R_{670}))$ OSAVI: $1.16(R_{800} - R_{670})/(R_{800} + R_{670} + 0.16)$	[41]
REP	Red-edge position	$700 + 40(R_{re} - R_{700})/(R_{740} - R_{700})$ R_{re}: $(R_{670} + R_{780})/2$	[42]
R-M	Red model	$R_{750}/R_{720} - 1$	[43]
RTVI	Red-edge triangular vegetation index	$(100(R_{750} - R_{730}) - 10(R_{750} - R_{550}))\text{sqrt}(R_{700}/R_{670})$	[44]

† indicates named in this study.

2.4.2. PLSR Method

PLSR is a commonly used method for addressing variable collinearity [45]. When establishing the PLSR-based LNC prediction models, this study selected the first five principal components (PCs) as model input data, as using more PCs did not improve model performance. Notably, as the spectral bands above 850 nm in the images obtained by the hyperspectral sensor S185 have poor quality according to Chen et al. [46], only the spectral bands between 450 nm and 850 nm were used as input variables. In this study, the PLSR models were calibrated using a calibration dataset and validated using a validation dataset. MATLAB 2020a code was written and used to perform the above process.

2.4.3. ANN Method

The ANN is a mathematical model inspired by the biological brain that simulates the complex information processing of the brain–nervous system [47] and has been widely used in many fields [48]. Compared with other ANN models, backpropagation (BP) ANN is one of the most widely used neural network models due to its good performance and ability to handle any linear or nonlinear relationship between the input and output variables [49]. In addition, Homik [50] et al. documented that a three-layer BP network could perform all N-dimensional to M-dimensional transformations. Therefore, a three-layer BP neural network with an input layer, a hidden layer and an output layer was used in this study. The hidden layer had three neurons, the activation function from the input layer to the hidden layer was the log-sigmoid function, and the activation function from the hidden layer to the output layer was the tan-sigmoid function. The maximum number of model iterations was set to 10,000, and the learning rate was set to 0.01. To reduce the number of input variables, principal component analysis (PCA) was performed on the band reflectance values, and the five first PCs were selected as the model input data, similar to the PLSR methods. The ANN model was trained based on all calibration samples and validated based on all validation samples. The above process was also performed using MATLAB 2020a code.

2.4.4. ML-HM Method

As stated previously, when the measured data are limited due to time and cost constraints, the training of deep learning models can be assisted by simulated data. The ML-HM method needs simulated data to help train the model. In this study, the N-PROSPECT model proposed by Li et al. was used to generate simulated data [51]. Based on the distribution characteristics of the field-measured data and related references, the ranges and sampling methods for the input parameters were determined and are shown in Table 2. These parameters were used to generate representative datasets under different conditions. Finally, a total of 20,000 data were generated.

Similar to the ANN method, when designing the LNC prediction model based on the ML-HM method, the bands between 450 and 850 nm were selected, PCA was used to reduce the number of input variables and the five first principal components were taken as input variables. As mentioned in Section 2.3, the two tasks of the ML-HM model included an auxiliary task, namely, training a network based on the simulated data, and a major task, namely, training a network based on the measured data. The two tasks shared part of the network. During the model calibration process, all 20,000 simulated data and all 70 calibration samples of the measured data were used to train the network. For major task model validation, all 23 validation samples of the measured data were used to validate the model performance. Considering the training process, the adaptive moment estimation (Adam) algorithm was used as the optimizer, as it requires less memory and converges more rapidly than other optimizers. In addition, the learning rate was set to 0.001, the weight decay coefficient was set to 0.0016, the learning rate decay strategy was set to reduce the learning rate to 0.2 times the current value every 20 cycles and the number of training cycles was set to 100. Due to the fact that the quantity of training data for the two tasks significantly differed, to avoid the shared network being mainly influenced by simulated data, the method used in Narayannan's study was used [52] and described as follows. Each training cycle contained 1000 iterations to continuously train the network. In each iteration, 20 samples from the simulated dataset and 8 samples from the measured data were randomly selected to train the model. The above process was called as mini-batch gradient descent training process [53]. All the above was performed using Python 3.8 code. To note that, as our objective is to obtain a model for predicting actual LNC, we did not using simulated data to validate the performance of auxiliary task for simulated LNC prediction.

Table 2. Ranges and sampling methods of the input parameters for the N-PROSAIL model.

Variable	Min	Max	AVG	SD	Sampling Method	Reference
LNC (Leaf nitrogen content, $\mu g/cm^2$)	20	220	110	45	Gauss	Measured dataset
C_{brown} (Brown pigment content, $\mu g/cm^2$)	0	0	-	-	Fixed	[54]
C_w (Equivalent water thickness, cm)	0.004	0.04	-	-	Uniform	[12]
C_{dm} (Dry matter content, g/m^2)	0.001	0.02	-	-	Uniform	[55]
$N_{structer}$ (Leaf structure)	1.2	1.8	1.5	0.3	Gauss	[55]
LID (Leaf inclination distribution, deg)	30	80	60	30	Gauss	[56]
LAI (Leaf area index, m^2/m^2)	0.1	9	3.9	1.6	Gauss	Measured dataset
S_L (Hot spot parameter)	0.1	0.5	0.2	0.5	Gauss	[57]
θ_s (Solar zenith angle, deg)	20	45	-	-	Uniform	Measured dataset
Rsoil (Soil brightness parameter)	0.2	0.9	0.4	0.4	Gauss	[58]

3. Results

3.1. LNCs in the Field

Analysis of variance (ANOVA) and Duncan's new multiple range test (MRT) analyses were performed on the LNC values of wheat under different irrigation and N treatment levels during various growth stages and are shown in Table 3. There were significant differences ($p < 0.05$) in LNC among the different N levels in all wheat growth stages. The LNC values in the experimental year varied between 20.45 and 216.38 $\mu g/cm^2$, covering a wide LNC range. Thus, our dataset is a good dataset for LNC model design.

Table 3. Mean LNC ($\mu g/cm^2$) in wheat under different water and N treatments during different growth stages.

Growth Stage	Irrigation Treatment	N Fertilizer Treatment *				
		N1	N2	N3	N4	N5
Feekes 4–5	W1	28.65 ^{a}	81.09 ^{a,b}	116.40 ^{b}	138.01 ^{b}	137.66 ^{b}
	W2	45.26 ^{a}	98.47 ^{b}	131.78 ^{b,c}	122.22 ^{b,c}	144.31 ^{c}
Feekes 10.2	W1	78.86 ^{a}	114.45 ^{a,b}	164.65 ^{b,c}	142.07 ^{b,c}	173.68 ^{c}
	W2	51.54 ^{a}	70.66 ^{a}	124.96 ^{b}	140.28 ^{b}	154.72 ^{b}
Feekes 11.1	W1	59.05 ^{a}	64.85 ^{a,b}	115.29 ^{a,b}	153.08 ^{b}	119.18 ^{a,b}
	W2	28.21 ^{a}	63.49 ^{b}	114.81 ^{c}	141.14 ^{c}	118.35 ^{c}

*: Numbers indicate the average LNC in the corresponding N and irrigation treatments. Within each row, the different letters indicate significant differences at the 0.05 level ($p < 0.05$); W1: 80% field capacity; W2: 60% field capacity; N1: 0 kg N/ha^{-1}; N2: 70 kg N/ha^{-1}; N3: 140 kg N/ha^{-1}; N4: 210 kg N/ha^{-1}; N5: 280 kg N/ha^{-1}.

3.2. Simulated and Measured Spectral Reflectance

All the measured and some simulated spectra data are shown in Figure 4a. It can be seen that the change range of the measured spectrum is within the change range of the simulated spectrum and that they have a similar spectral structure, with absorption valleys at the red and blue bands and peaks at the near infrared band. However, when considering the correlation coefficients between each band and LNC, they not only have similarities between bands 450 nm and 734 nm, both having significant ($p < 0.05$) correlation with LNC, but also have differences between bands 738 nm and 766 nm, with simulated spectra having a significant correlation with LNC and measured data having no significant correlation with LNC. In summary, the data demonstrated our theory according to which the simulated data has similarities with the measured data, but it also has differences with the measured data and cannot be used to train the model directly for actual LNC prediction.

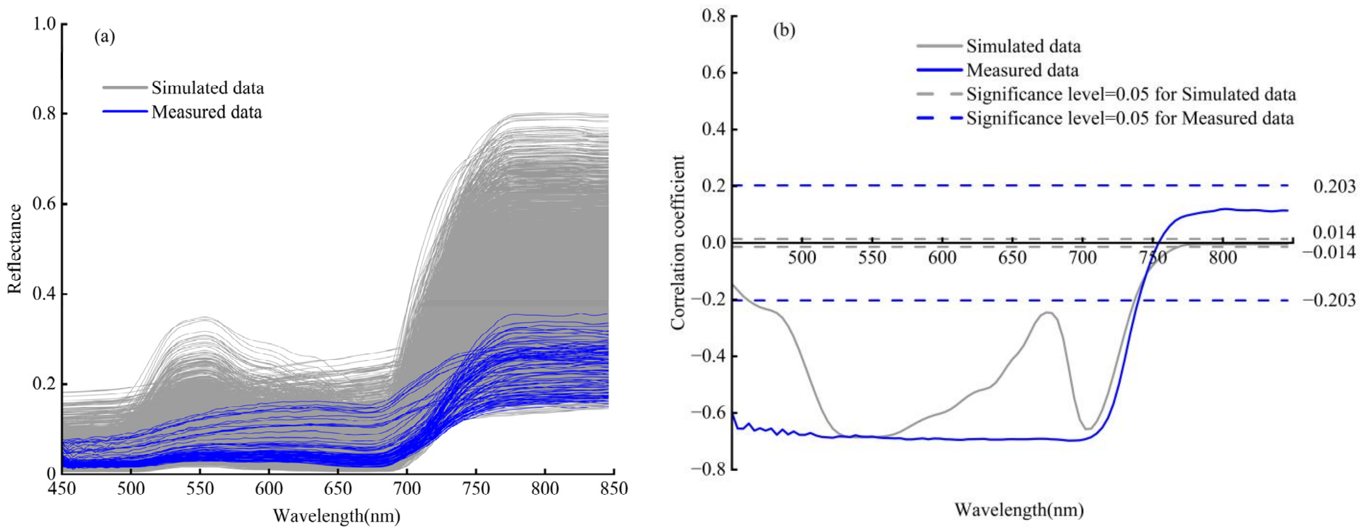

Figure 4. Simulated and measured spectral reflectance curves (**a**) and the correlation coefficients between each band in simulated spectrum and LNC and between each band in measured spectrum and LNC (**b**).

3.3. LNC Prediction Results by the SI Method

The LNC prediction results for the models designed using different spectral indices are shown in Table 4 and ranked from best to worst performance according to the calibration and validation results. Among them, MCART/MTVI2 and TCARI/OSAVI achieved the best results, with R^2 values of 0.68 and 0.65, *RMSE* values of 29.86 and 31.26 μg/cm^2, and *RMSE*% values of 26.66% and 27.91%, respectively, during calibration and R^2 values of 0.68 and 0.68, *RMSE* values of 25.25 and 25.69 μg/cm^2 and *RMSE*% values of 23.23% and 23.63%, respectively, during validation. Compared with the other vegetation indices, MSAVI and TVI had the worst accuracy, with R^2 values of 0.33 and 0.29, *RMSE* values of 37.41 and 38.52 μg/cm^2 and *RMSE*% values of 33.40% and 34.39%, respectively, during calibration and R^2 values of 0.48 and 0.43, *RMSE* values of 32.20 and 33.55 μg/cm^2 and *RMSE*% values of 29.62% and 30.86%, respectively, during validation. The remaining spectral indices achieved moderate LNC prediction results. During the calibration stage, the R^2 values varied between 0.35 and 0.57, the *RMSE* values varied between 29.98 and 36.85 μg/cm^2 and the *RMSE*% values varied between 26.77% and 32.90%. During the validation stage, the R^2 values varied between 0.49 and 0.70, the *RMSE* values varied between 23.86 and 31.60 μg/cm^2 and the *RMSE*% values varied between 21.95% and 29.07%.

Table 4. LNC prediction results for the models designed using different spectral indices. Linear, logarithmic, exponential and power models were used for fitting each index. The model type with the best results is expressed in the table.

SI	Model Type	Calibration			Validation		
		R^2	*RMSE* (μg/cm^2)	*RMSE%*	R^2	*RMSE* (μg/cm^2)	*RMSE%*
MCARI/MTVI2	Exponential	0.68	29.86	26.66%	0.68	25.25	23.23%
TCARI/OSAVI	Exponential	0.65	31.26	27.91%	0.68	25.69	23.63%
MTCI	Logarithmic	0.56	30.20	26.96%	0.70	23.86	21.95%
REP	Linear	0.57	29.98	26.77%	0.68	25.06	23.05%
RM	Logarithmic	0.52	31.53	28.15%	0.66	25.37	23.34%
NDRE	Logarithmic	0.53	31.44	28.07%	0.65	25.75	23.68%
GNDVI	Logarithmic	0.51	31.91	28.49%	0.65	26.02	23.93%
DCNI	Logarithmic	0.51	31.94	28.52%	0.64	26.49	24.37%
NDVI	Logarithmic	0.48	33.09	29.54%	0.61	27.50	25.29%
VI_{opt}	Logarithmic	0.46	33.64	30.03%	0.60	28.09	25.84%
RTVI	Power	0.54	35.26	31.48%	0.49	28.33	26.06%
OSAVI	Logarithmic	0.43	34.35	30.67%	0.56	29.15	26.81%
RVI	Logarithmic	0.40	35.45	31.65%	0.57	29.10	26.77%
MTVI2	Logarithmic	0.37	36.22	32.34%	0.51	30.90	28.42%
EVI	Logarithmic	0.35	36.85	32.90%	0.50	31.60	29.07%
MSAVI	Logarithmic	0.33	37.41	33.40%	0.48	32.20	29.62%
TVI	Logarithmic	0.29	38.52	34.39%	0.43	33.55	30.86%

3.4. LNC Prediction Results by the PLSR Method

The LNC prediction results of the PLSR method are shown in Figure 5. During calibration, the PLSR method achieved an R^2 value of 0.68, an *RMSE* value of 25.71 μg/cm^2 and an *RMSE%* value of 22.95%. Moreover, during the validation stage, the model had an R^2 value of 0.74, an *RMSE* value of 22.79 μg/cm^2 and an *RMSE%* value of 20.96%. In addition, although most of the points are clustered near the 1:1 line, some points deviate from the 1:1 line. The PLSR method may underestimate samples with higher LNC values and overestimate samples with lower LNC values.

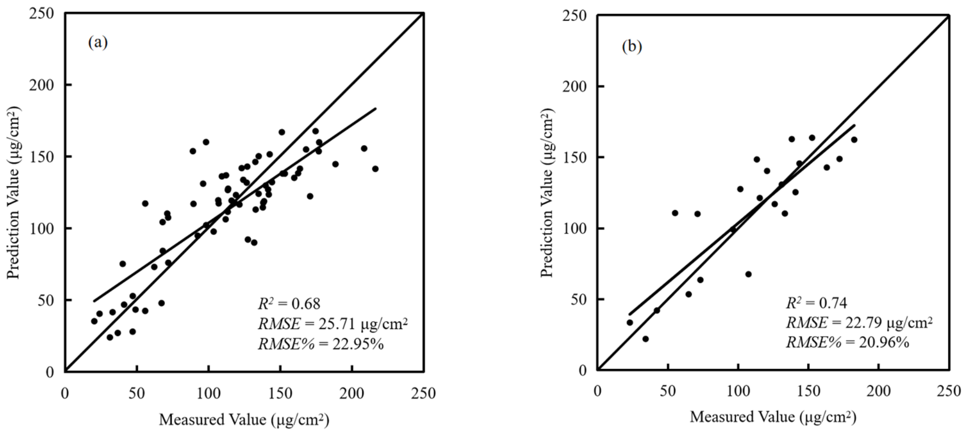

Figure 5. PLSR model results. (**a**) Calibration; (**b**) Validation.

3.5. LNC Prediction Results by the ANN Method

The LNC prediction results of the ANN method are shown in Figure 6. The ANN model shows a moderate LNC prediction performance. During calibration, the ANN method obtained an R^2 value of 0.56, an *RMSE* value of 33.62 μg/cm^2 and an *RMSE%* value of 30.01%. During validation, the model had an R^2 value of 0.62, an *RMSE* value of 30.42 μg/cm^2 and an *RMSE%* value of 27.98%. In addition, many points deviate from the 1:1 line. The ANN method underestimates samples with higher LNC values and overestimates samples with lower LNC values.

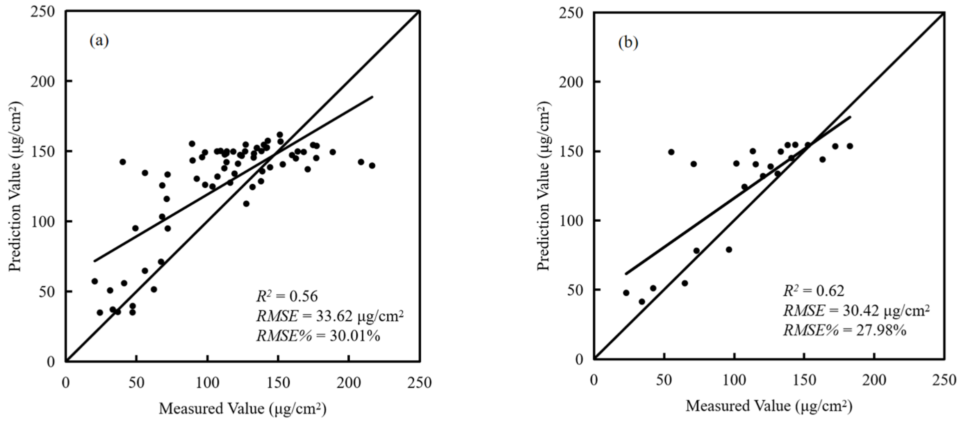

Figure 6. ANN model results. (**a**) Calibration; (**b**) Validation.

3.6. LNC Prediction Results by the ML-HM Method

For the ML-HM method, based on measured data, it obtained an R^2 value of 0.79, an *RMSE* value of 20.86 μg/cm^2 and an *RMSE*% value of 18.63% during calibration as well as an R^2 value of 0.82, an *RMSE* value of 18.40 μg/cm^2 and an *RMSE*% value of 16.92% during validation (Figure 7). In addition, all points were near the 1:1 line. Thus, the ML-HM method showed very good LNC prediction performance.

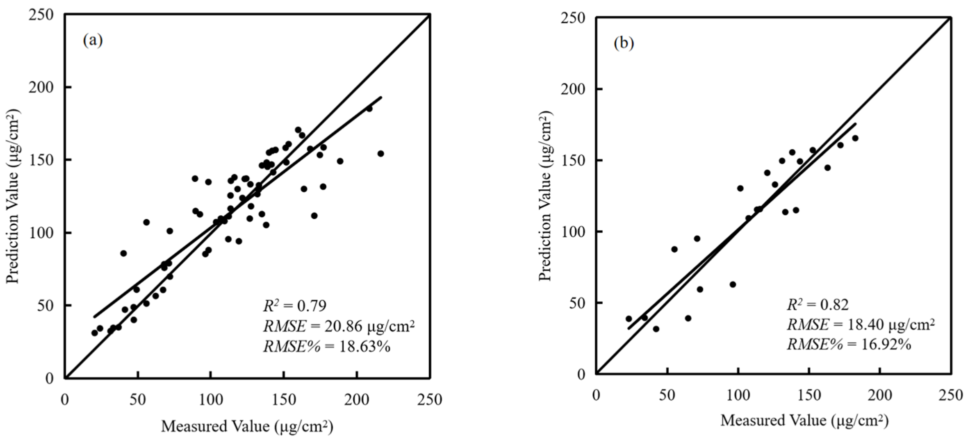

Figure 7. ML-HM model results. (**a**) Calibration; (**b**) Validation.

4. Discussion

4.1. Comparison with Previous Studies

In this study, the newly proposed ML-HM method achieved the best LNC prediction results, with an R^2 value of 0.82 during validation. The corresponding R^2 values of the SI, PLSR and ANN methods were between 0.43 and 0.74. Using field-measured hyperspectral data, Jia et al. designed various spectral indices for wheat LNC prediction and obtained the maximum R^2 value of 0.66 [59]. Moreover, based on UAV hyperspectral images, Zhang et al. used different methods (PLSR, GA-PLSR, RF and XGBoost) for LNC prediction, and the maximum R^2 value of the four methods is 0.55 [7]. Compared with previous studies, the results of the traditional methods (SI, PLSR and ANN methods) in our study are within reasonable ranges, and the novel ML-HM method performed better than the traditional methods.

4.2. Best Structure for the Hybrid Method

Previous studies on hybrid methods that combined deep learning methods and mechanistic models used single-task schemes to design the model. To verify the effectiveness of the multitask learning scheme proposed in this study, several single-task schemes were designed according to previous studies and compared with the multitask scheme proposed in this study. These single-task schemes include the following: (a) the simulated data generated in Section 2.4.4 and the calibration dataset of the measured data were used to train the same network for LNC prediction, and the validation dataset of the measured data was used to validate the model. For comparison, the network used the same structure as the major task network in the ML-HM method and is denoted as the single-task method I. (b) Only the simulated data were used to train the network for LNC prediction, and the validation dataset of the measured data was used to validate the model. The network used the same structure as the major task network in the ML-HM method and is denoted as the single-task method II. (c) Only the calibration dataset of the measured data was used to train the network, and the validation dataset of the measured data was used to validate the model. The network used the same structure as the major task network in the ML-HM method and is denoted as the single-task method III.

The results of the different LNC prediction schemes are shown in Table 5. As the ML-HM method, the single-task method I used the simulated data and part of the measured data to train the network. However, this scheme considered the simulated data to be the same as the measured data. In fact, due to the complexity of the electromagnetic transmission process, existing mechanistic models cannot fully describe the transmission process; thus, the simulated data differ from the measured data. Thus, although single-task method I performs better during model calibration, its performance decreases significantly during model validation. Single-task method II used only the simulated data to train the model. Similar to the prior discussion, this scheme also has an overfitting problem, with good performance during calibration and poor performance during validation. Single-task method III used part of the measured data to train the network and the remaining data to validate the network. Compared with the above two single-task schemes, this scheme has a stable performance during model calibration and validation. However, its LNC prediction results are worse than those of the multitask method proposed in this study. This result may have occurred due to the following reasons: i) the measured data have a relatively small number of calibration samples and thus cannot represent data in different situations, and ii) although the simulated data differ from the measured data, they represent data in different situations. In our ML-HM model, we effectively extracted the common features shared by the simulated and measured data for LNC prediction through the network in the shared layer and used these shared features to assist in training the network on the measured data, which effectively improved the LNC prediction accuracy.

Table 5. LNC prediction results using different structures to design the hybrid model.

Model Type	Calibration			Validation		
	R^2	*RMSE* (μg/cm^2)	*RMSE%*	R^2	*RMSE* (μg/cm^2)	*RMSE%*
Single-task method I	0.89	13.26	11.90%	0.68	24.56	22.59%
Single-task method II	0.88	14.34	12.88%	0.22	38.58	35.49%
Single-task method III	0.70	24.93	18.65%	0.71	23.77	21.86%
ML-HM method (this study)	0.79	20.86	18.63%	0.82	18.40	16.92%

4.3. Optimal LNC Prediction Method

Many commonly used methods were selected and compared with the ML-HM method proposed in this study. The ML-HM method performed the best among the comparison methods. This good performance may be due to the following reasons: (i) multiple factors (e.g., LAI, chlorophyll, leaf mesophyll structure, leaf water content, leaf angle distribution function, background) influence canopy reflectance, and the combined influence of these factors on the reflectance spectrum is not purely linear. The PLSR method is a linear regression method that cannot effectively determine the nonlinear relationship between the spectral reflectance and LNC. (ii) The SI model is based on multispectral information and utilizes only limited spectral features. (iii) Due to the limited quantity of measured data, the

ANN model training was not sufficient. Thus, it could not express the relationship between spectral reflectance and LNC. In the ML-HM method, we used useful information in the simulated data to help train the network on the measured data, resulting in an improved LNC prediction accuracy.

4.4. Application Potential and Limitations of This Study

In this study, an ML-HM LNC prediction method was designed. In this method, the simulated data from the mechanistic model were used as an auxiliary task to help train the LNC prediction model, thereby improving the robustness of the model and reducing the number of measured samples required by the model. For field applications, UAV images can be obtained; then, the ML-HM method can be used to predict LNC in the field. Next, an N application map can be produced by considering the LNC in the field and the law of wheat N demand during its life cycle. Finally, precise N management can be conducted according to the N application map.

Although the method proposed in this paper achieved good results, some limitations remain. In this study, during mini-batch gradient descent training process, the quantity of simulated and measured data was determined by experience with no theoretical support. An excessive amount of simulated data can significantly increase the calculation time. However, a small amount of simulated data are insufficient for helping the network learn enough information to assist the training of the main task network. Therefore, determining a reasonable ratio of simulated to measured data can improve model training efficiency and model accuracy, which will be investigated in future work.

5. Conclusions

In this study, wheat field experiments under different irrigation and N treatments were conducted, and LNC and UAV hyperspectral images were collected during critical growth stages. Based on these data, a new hybrid method combining the deep multitask learning method and the N-PROSAIL model was proposed and compared with traditional LNC prediction methods (SI, PLSR and ANN methods). Additionally, we compared the proposed approach with traditional hybrid methods using single-task schemes. The results show that the proposed ML-HM method has better LNC prediction performance than the SI, PLSR and ANN methods. Moreover, compared with the hybrid method using a single-task scheme, the ML-HM method more effectively extracts useful information from the simulated data to help train the network on the measured data and discard confusing information, increasing the robustness of the model. Thus, this study provides an accurate LNC prediction tool for precise N management in the field.

Author Contributions: Data analysis, writing original draft, X.M.; Conceptualization, review and editing, supervision, P.C.; Review and editing, X.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by The National Natural Science Foundation of China (Grant Numbers 41871344), the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant Number XDA28040502).

Data Availability Statement: The data presented in this study are available upon request from the corresponding author. The data are not publicly available due to their use in subsequent studies.

Acknowledgments: We thank the Yucheng Integrated Experiment Station of the Chinese Academy of Sciences and Jing Li for providing the experimental field used for this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Albornoz, F. Crop responses to nitrogen overfertilization: A review. *Sci. Hortic.* **2016**, *205*, 79–83. [CrossRef]
2. Gruber, N.; Galloway, J.N. An earth-system perspective of the global nitrogen cycle. *Nature* **2008**, *451*, 293–296. [CrossRef] [PubMed]

3. Panda, S.S.; Hoogenboom, G.; Paz, J.O. Remote sensing and geospatial technological applications for sits-specific management of fruit and nut crop: A review. *Remote Sens.* **2010**, *2*, 1973–1997. [CrossRef]
4. Mondal, P.; Basu, M. Adoption of precision agriculture technologies in India and in some developing countries: Scope, present status and strategies. *Prog. Nat. Sci.* **2009**, *19*, 659–666. [CrossRef]
5. Zhang, L.; Niu, Y.; Zhang, H.; Han, W.; Li, G.; Tang, J.; Peng, X. Maize canopy temperature extracted from UAV thermaland RGB imagery and its application in water stress monitoring. *Front. Plant Sci.* **2019**, *10*, 1270. [CrossRef]
6. Shafian, S.; Rajan, N.; Schnell, R.; Bagavathiannan, M.; Valasek, J.; Shi, Y.; Olsenholler, J. Unmanned aerial systems-based remote sensing for monitoring sorghum growth and development. *PLoS ONE* **2018**, *13*, e0196605. [CrossRef]
7. Zhang, Y.; Wang, T.; Guo, Y.; Skidmore, A.; Zhang, Z.; Tang, R.; Song, S.; Tang, Z. Estimating community-level plant functional traits in a species-rich alpine meadow using UAV image spectroscopy. *Remote Sens.* **2022**, *14*, 3399. [CrossRef]
8. Jay, S.; Maupas, F.; Bendoula, R.; Gorretta, N. Retrieving LAI, chlorophyll and nitrogen contents in sugar beet crops from multi-angular optical remote sensing: Comparison of vegetation indices and PROSAIL inversion for field phenotyping. *Field Crop. Res.* **2017**, *210*, 33–46. [CrossRef]
9. Croft, H.; Arbian, J.; Chen, J.M.; Shang, J.; Liu, J. Mapping within-field leaf chlorophyll content in agricultural crops for nitrogen management using Landsat-8 imagery. *Precis. Agric.* **2020**, *21*, 856–880. [CrossRef]
10. Li, Z.; Jin, X.; Wang, J.; Yang, G.; Nie, C.; Xu, X.; Feng, H. Estimating winter wheat (Triticum aestivum) LAI and leaf chlorophyll content from canopy reflectance data by integrating agronomic prior knowledge with the PROSAIL model. *Int. J. Remote Sens.* **2015**, *36*, 2634–2653. [CrossRef]
11. Yang, G.; Zhao, C.; Pu, R.; Feng, H.; Li, Z.; Li, H.L.; Sun, C. Leaf nitrogen spectral reflectance model of winter wheat (Triticum aestivum) based on PROSPECT: Simulation and inversion. *J. Appl. Remote Sens.* **2015**, *9*, 095976. [CrossRef]
12. Li, Z.; Jin, X.; Yang, G.; Drummond, J.; Yang, H.; Clark, B.; Li, Z.; Zhao, C. Remote sensing of leaf and canopy nitrogen status in winter wheat (Triticum aestivum L.) based on N-PROSAIL model. *Remote Sens.* **2018**, *10*, 1463. [CrossRef]
13. Darvishzadeh, R.; Atzberger, C.; Skidmore, A.; Schlerf, M. Mapping grassland leaf area index with airborne hyperspectral imagery: A comparison study of statistical approaches and inversion of radiative transfer models. *ISPRS-J. Photogramm. Remote Sens.* **2011**, *66*, 894–906. [CrossRef]
14. Durbha, S.S.; King, R.L.; Younan, N.H. Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer. *Remote Sens. Environ.* **2007**, *107*, 348–361. [CrossRef]
15. Annala, L.; Honkavarra, E.; Tuominen, S.; Pölönen, I. Chlorophyll concentration retrieval by training convolutional neural network for stochastic model of leaf optical properties (SLOP) inversion. *Remote Sens.* **2020**, *12*, 283. [CrossRef]
16. Xu, X.; Lu, J.; Zhang, N.; Yang, T.; He, J.; Yao, J.; Cheng, T.; Zhu, Y.; Cao, W.; Tian, Y. Inversion of rice canopy chlorophyll content and leaf area index based on coupling of radiative transfer and Bayesian network models. *ISPRS-J. Photogramm. Remote Sens.* **2019**, *150*, 185–196. [CrossRef]
17. Zhang, Y.; Hui, J.; Qin, Q.; Sun, Y.; Zhang, T.; Hong, S.; Li, M. Transfer-learning-based approach for leaf chlorophyll content estimation of winter wheat from hyperspectral data. *Remote Sens. Environ.* **2021**, *267*, 112724. [CrossRef]
18. Jiao, Q.; Sun, Q.; Zhang, B.; Huang, W.; Ye, H.; Zhang, Z.; Zhang, X.; Qiang, B. A random forest algorithm for retrieving canopy chlorophyll content of wheat and soybean trained with PROSAIL simulation using adjusted average leaf angle. *Remote Sens.* **2022**, *14*, 98. [CrossRef]
19. Svendsen, D.H.; Martino, L.; Campos-Taberner, M.; Garcia-Haro, F.J.; Camps-Valls, G. Joint gaussian processes for biophysical parameter retrieval. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 1718–1727. [CrossRef]
20. Baxter, J. A model of inductive bias learning. *J. Artif. Intell. Res.* **2000**, *12*, 149–198. [CrossRef]
21. Thrun, S. Is Learning the n-th Thing Any Easier than Learning the First? In Proceedings of the Advances in Neural Information Processing Systems 8, NIPS, Denver, CO, USA, 27–30 November 1995; Touretzky, D.S., Mozer, M., Hasselmo, M.E., Eds.; MIT Press: Cambridge, MA, USA, 1995; pp. 640–646.
22. Caruana, R. Multi task learning. *Mach. Learn.* **1997**, *28*, 41–75. [CrossRef]
23. Zhang, Y.; Yang, Q. An overview of multi-task learning. *Natl. Sci. Rev.* **2018**, *5*, 30–43. [CrossRef]
24. Baxter, J. A bayesian/information theoretic model of learning to learn via multiple task sampling. *Mach. Learn.* **1997**, *28*, 7–39. [CrossRef]
25. Zhang, Q.; Wu, S.; Wang, X.; Sun, B.; Liu, H.M. A PM2.5 concentration prediction model based on multi-task deep learning for intensive air quality monitoring stations. *J. Clean. Prod.* **2020**, *275*, 122722. [CrossRef]
26. Liu, S.; Johns, E.; Davison, A.J. End-to-End multi-task learning with attention. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 1871–1880.
27. Chen, P.; Wang, F. Effect of crop spectra purification on plant nitrogen concentration estimations performed using high-spatial-resolution images obtained with unmanned aerial vehicles. *Field Crop Res.* **2022**, *288*, 108708. [CrossRef]
28. Rouse, J.; Haas, R.H.; Deering, D.; Schell, J.A.; Harlan, J. Monitoring the vernal advancement and retrogradation (Green wave effect) of natural vegetation. *Environ. Sci.* **1974**, *1974*, 1–37.
29. Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* **1996**, *58*, 289–298. [CrossRef]
30. Qi, J.; Chehbouni, A.; Huete, A.R.; Kerr, Y.H.; Sorooshian, S. A modified soil adjusted vegetation index. *Remote Sens. Environ.* **1994**, *48*, 119–126. [CrossRef]

31. Rondeaux, G.; Steven, M.; Baret, F. Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* **1996**, *55*, 95–107. [CrossRef]
32. Huete, A.; Justice, C.; Liu, H. Development of vegetation and soil indices for MODIS-EOS. *Remote Sens. Environ.* **1994**, *49*, 224–234. [CrossRef]
33. Broge, N.H.; Leblanc, E. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sens. Environ.* **2002**, *76*, 156–172. [CrossRef]
34. Haboudans, D.; Miller, J.R.; Pattey, E.; Zarco-Tejada, P.J.; Strachan, I.B. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sens. Environ.* **2004**, *90*, 337–352. [CrossRef]
35. Xue, L.; Cao, W.; Luo, W.; Dai, T.; Zhu, Y. Monitoring leaf nitrogen status in rice with canopy spectral reflectance. *Agron. J.* **2004**, *96*, 135–142. [CrossRef]
36. Fitzgerald, G.J.; Rodriguez, D.; Christensen, L.K.; Belford, R.; Sadras, V.O.; Clarke, T.R. Spectral and thermal sensing for nitrogen and water status in rainfed and irrigated wheat environments. *Precis. Agric.* **2006**, *7*, 233–248. [CrossRef]
37. Reyniers, M.; Walvoort, D.J.J.; Baardemaaker, D.J. A linear model to predict with a multi-spectral radiometer the amount of nitrogen in winter wheat. *Int. J. Remote Sens.* **2006**, *27*, 4159–4179. [CrossRef]
38. Chen, P.; Haboudance, D.; Tremblay, N.; Wang, J.; Vigneault, P.; Li, B. New spectral indicator assessing the efficiency of crop nitrogen treatment in corn and wheat. *Remote Sens. Environ.* **2010**, *114*, 1987–1997. [CrossRef]
39. Eitel, J.U.H.; Long, D.S.; Gessler, P.E.; Smith, A.M.S. Using in-situ measurements to evaluate the new RapideEyeTM satellite series for prediction of wheat nitrogen status. *Int. J. Remote Sens.* **2007**, *28*, 4183–4190. [CrossRef]
40. Dash, J.; Curran, P.J. The MERIS terrestrial chlorophyll index. *Int. J. Remote Sens.* **2004**, *25*, 5403–5413. [CrossRef]
41. Haboudane, D.; Miller, J.R.; Tremblay, N.; Zaro-Tejada, P.J.; Dextraze, L. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sens. Environ.* **2002**, *81*, 416–426. [CrossRef]
42. Cho, M.A.; Skidmore, A.K. A new technique for extracting the red edge position from hyperspectral data: The linear extrapolation method. *Remote Sens. Environ.* **2006**, *101*, 181–193. [CrossRef]
43. Gitrlson, A.A.; Viña, A.; Cogand, V.; Rundquist, D.C.; Arkebauer, T.J. Remote estimation of canopy chlorophyll content in crops. *Geophys. Res. Lett.* **2005**, *32*, L08403.
44. Chen, P.; Tremblay, N.; Wang, J.; Vigneault, P.; Huang, W.; Li, B. New index of crop canopy fresh biomass estimation. *Spectrosc. Spectr. Anal.* **2010**, *30*, 512–517. (In Chinese)
45. Yang, B.; Wang, M.; Sha, Z.; Wang, B.; Chen, J.; Yao, X.; Cheng, T.; Cao, W. Evaluation of aboveground nitrogen content of winter wheat using digital imagery of unmanned aerial vehicles. *Sensors* **2019**, *19*, 4416. [CrossRef] [PubMed]
46. Chen, P.; Li, G.; Shi, Y.; Xu, Z.; Yang, F.; Cao, Q. Validation of an unmanned aerial vehicle hyperspectral sensor and its application in maize leaf area index estimation. *Sci. Agric. Sin.* **2018**, *51*, 1464–1474. (In Chinese)
47. McCulloch, W.S.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* **1943**, *5*, 115–133. [CrossRef]
48. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *512*, 436–444. [CrossRef]
49. Chen, P.; Wang, J.; Huang, W.; Tremblay, N.; Ou, Y.; Zhang, Q. Critical nitrogen curve and remote detection of nitrogen nutrition index for corn in the northwestern plain of shandong province, China. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 682–689. [CrossRef]
50. Hornik, K.; Stinchcombe, M.; White, H. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. *Neural Netw.* **1990**, *3*, 551–560. [CrossRef]
51. Li, Z.; Li, Z.; Fairbairn, D.; Li, N.; Xu, B.; Feng, H.; Yang, G. Multi-LUTs method for canopy nitrogen density estimation in winter wheat by field and UAV hyperspectral. *Comput. Electron. Agric.* **2019**, *162*, 174–182. [CrossRef]
52. Narayanan, S.; Mannam, K.; Achan, P.; Ramesh, M.V.; Rangan, V.; Rajan, S. A contextual multi-task neural approach to medication and adverse events identification from clinical test. *J. Biomed. Inform.* **2022**, *125*, 103960. [CrossRef]
53. Senthil, T.; Chinnasamy, R.; Jeevanandham, D. An improved optimization technique using deep neural networks for digit recognition. *Soft Comput.* **2021**, *25*, 1647–1658. [CrossRef]
54. Wang, L.; Chen, S.; Peng, Z.; Huang, J.; Wang, C.; Jiang, H.; Zheng, Q.; Li, D. Phenology effects on physically based estimation of paddy rice canopy traits from UAV hyperspectral imagery. *Remote Sens.* **2021**, *13*, 1792. [CrossRef]
55. Upreti, D.; Huang, W.; Kong, W.; Pascucci, S.; Pignatti, S.; Zhou, X.; Ye, H.; Casa, R. A Comparison of hybrid machine learning algorithms for the retrieval of wheat biophysical variables from Sentinel-2. *Remote Sens.* **2019**, *11*, 481. [CrossRef]
56. Li, H.; Liu, G.; Liu, Q.; Chen, Z.; Huang, C. Retrieval of winter wheat leaf area index from Chinese GF-1 satellite data using the PROSAIL model. *Sensors* **2018**, *18*, 1120. [CrossRef] [PubMed]
57. Delloye, C.; Weiss, M.; Defourny, P. Retrieval of the canopy chlorophyll content from Sentinel-2 spectral bands to estimate nitrogen uptake in intensive winter wheat cropping systems. *Remote Sens. Environ.* **2018**, *216*, 245–261. [CrossRef]
58. Liang, L.; Geng, D.; Yan, J.; Qiu, S.; Di, L.; Wang, S.; Xu, L.; Wang, L.; Kang, J.; Li, L. Estimating crop LAI using spectral feature extraction and the hybrid inversion method. *Remote Sens.* **2020**, *12*, 3534. [CrossRef]
59. Jia, M.; Colombo, R.; Rossini, M.; Celesti, M.; Zhu, J.; Cogliati, S.; Cheng, T.; Tian, Y.; Zhu, Y.; Cao, W.; et al. Estimation of leaf nitrogen content and photosynthetic nitrogen use efficiency in wheat using sun-induced chlorophyll fluorescence at the leaf and canopy scales. *Eur. J. Agron.* **2021**, *122*, 126192. [CrossRef]