Remote Sensing and Machine Learning for Prediction of Wheat Growth in Precision Agriculture Applications

Yuxi Fang

In the fulfilment of the requirement for the degree of

Master of Philosophy

Centre for excellence in Signal and Image Processing

Department of Electronic and Electrical Engineering

University of Strathclyde, Glasgow

Supervised by

Doctor Jinchang Ren

Doctor Hong Yue

March 6, 2020

Declaration of Authorship

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Yuxi Fang

March 6, 2020

Acknowledgements

As time goes by, my master's study life is coming to an end. In the beautiful city of Glasgow, I will never forget my study experience at the famous University of Strathclyde. Walking in the beautiful campus, watching the teachers and students of different colors walking by, I feel infinite emotion in my heart. In retrospect, I felt very confused in the face of the pressure of language, the discomfort in life, and the loneliness in a foreign country. Fortunately, I was very lucky to have my supervisors and colleagues, who helped me successfully carry out my work and finish my studies.

First of all, I would like to thank my First supervisor, Dr. Jinchang Ren, who is knowledgeable, gentle and considerate like my father, for providing me with valuable learning opportunities, providing me careful guidance and help in my studies, and providing me meticulous care in my life. I was touched by his rigorous academic attitude and integrity. I would like to express my heartfelt thanks and sincere respect. I'd also like to thank my second Supervisor, Dr. H. Yue, for her help in my studies.

Thanks to Professor Daming Dong in Beijing, who provided me with research data for my experiment and gave me valuable opinions on my thesis. Thank you, Mr. He Sun, who was in the same laboratory with me, for discussing with me in the process of research and experiment, which gives me a lot of inspiration and great help for the smooth completion of the thesis. Miss Xiaoquan Li, Dr. Yijun Yan, Mr. Guoliang Xie and other colleagues' care in daily study and life made me deeply feel the warmth of the big family. It's nice to meet you!

Thank my parents in particular for providing me with such a good learning opportunity. Your full love and selfless dedication are the eternal driving force and strong backing on my learning path.

Finally, I would like to express my heartfelt thanks and sincere blessing to my teachers, classmates, family and friends who have helped and cared for me on the way of growing up!

Abstract

This thesis focuses on remote sensing and machine learning for prediction of wheat growth in precision agriculture applications.

Agriculture is the primary productive force, which plays an important role in human activities. Wheat, as one of the essential sources of food, is also a widely planted crop. The impact of weather and climate and some other uncertain factors on wheat production is crucial. Therefore, it is necessary to use reliable and statistically reasonable models for crop growth and yield prediction based on vegetation index variables and other factors, so as to obtain reliable prediction for efficient production. Applying certain artificial intelligence algorithms to the precision agriculture can significantly improve the efficiency of traditional agriculture in crop planting and reduce the consumption of human and natural resources. Remote sensing can objectively, accurately and timely provide a large amount of information for ecological environment and crop growth in agriculture applications. By combining the image and spectral data obtained by remote sensing technology with machine learning, information about wheat growth, yield and insect pests can be learned in time.

This thesis focuses on its applications in agriculture, particularly using effective prediction models such as the back propagation neural network and some optimisation algorithms for predicting wheat growth, yield and aphid. The work presented in this thesis address the issues of wheat growth prediction, yield assessment and aphid validation by model building and machine learning algorithm optimisation by means of remote sensing data. Specifically, the following objectives are defined: 1. Analyse multiple vegetation indexes based on the TM 1-4 band data of Landsat satellite and use regression algorithms to train the models and predict wheat growth; 2. Analyse and compare multiple vegetation indexes models by means of spectral data and use regression algorithms to predict wheat yield; 3. Combine spectral vegetation indexes and multiple regression algorithms to predict wheat aphid; 4. Use accurate evaluation criteria for validating the efficacy of the various algorithms.

In this thesis, the remote sensing data from the satellite has been applied instead of the airborne-based remote sensing data. Based on the TM 1-4 band image data of Landsat satellite, multiple vegetation indexes were used as the input of regression algorithms. After that, four kinds of regression algorithms such as the multiple linear regression (MR) algorithm, back propagation network (BPNN) algorithm, genetic algorithm (GA) optimised BPNN algorithm and particle swarm optimisation (PSO) optimised BPNN algorithm were used to train the model and predict the LAI and SPAD. The prediction results of each algorithm were compared with the ground truth information collected by hand held instruments on the ground.

The relationship between wheat yield and spectral data has been studied. Based on the BPNN algorithm, four kinds of models such as visible hyperspectral index (VHI) model, hyperspectral vegetation index (HVI) model, difference hyperspectral index (DHI) model and normalized hyperspectral index (NHI) model have been utilized to predict wheat yield. For the optimal NHI model, three regression algorithms such as back propagation network (BPNN) algorithm, genetic algorithm (GA) optimised BPNN algorithm and particle swarm optimisation (PSO) optimised BPNN algorithm, were compared to predict wheat yield, and RMSE and R-square of the three algorithms were compared and analysed.

Finally, the relationship between wheat aphid and spectral data has been investigated. Nine vegetation indexes related to aphid have been estimated from spectral data as the input of regression algorithms. Five kinds of regression algorithms such as back propagation network (BPNN) algorithm, genetic algorithm (GA) optimised BPNN algorithm, particle swarm optimisation (PSO) optimised BPNN algorithm, ant colony (ACO) optimisation algorithm optimised BPNN algorithm and cuckoo search (CS) optimised BPNN algorithm have been implemented to predict wheat aphid, which was validated with the ground truth information measured by hand-held instruments on the ground. The prediction results of each algorithm have been analysed.

The major original contributions of this thesis are as follows:

1. A variety of optimisation algorithms are used to improve the regression analysis of the BPNN algorithm, so that the prediction results of each model for wheat growth, yield and aphid are more accurate.

2. The spectral characteristics of winter wheat canopy have been analysed. The correlation between the absorption band and the associated physical and chemical properties of crops, specially the red edge slope, with the crop yield and wheat aphid damage is established.

3. Adjusted MSE and un-centered R-square, as accurate evaluation criteria for practical applications, are used to compare the prediction results of the models under different dimensions of the observed data.

4. Improve algorithm training by using the cross-validation method to obtain reliable and stable models for the prediction of wheat growth, yield, and aphid. Through repeated cross-validation, a better model can be obtained in the last.

Key word： Precision agriculture; BP network, wheat growth assessment; wheat yield prediction, wheat aphid validation

Content

List of Figures

List of Tables

Acronyms

AI	Artificial Intelligence
API	Aphid Index
ANN	Artificial Neural Networks
BPNN	Back Propagation Neural Network
CNNs	Convolutional Neural Networks
DHI	Difference Hyperspectral Index
DSSI1	Damage Sensitive Spectral Index1
DVI	Difference Vegetation Index
GA	Genetic Algorithm
GNDVI	Green Normalized Difference Vegetation Index
GVI	Green Vegetation Index
HVI	Hyperspectral Vegetation Index
LAI	Leaf Area Index
MCARI	Modified ChlorophyII Absorption Reflectance Index
ML	Machine Learning
MR	Multiple-linear Regression
MSE	Mean Squared Error
NBNDVI	Narrow-Band Normalized Difference Vegetation Index
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NHI	Normalized Hyperspectral Index
PRI	Photochemical Reflectance Index
PSO	Particle Swarm Optimisation
PVI	Perpendicular Vegetation Index
RMSE	Root Mean Squared Error
RVI	Ratio Vegetation Index

RVSI	Red-edge Vegetation Stress Index
SAVI	Soil Adjusted Vegetation Index
SIPI	Structure Insensitive Pigment Index
SPAD	Soil and Plant Analyser Development
SSE	Sum of Squares due to Error
SSR	Sum of Squares of the Regression
SST	Total Sum of Squares
TM	Thematic Mapper
VHI	Visible Spectral Index

Chapter 1.

Introduction

1.1 Motivation and objectives

Agriculture plays a crucial role for the global economy. As the population increases, the agricultural system is under increasing pressure to increase the productivity of crops and grow more crops. The precision agriculture, also known as digital agriculture, has become a new trend in scientific fields. Various data intensive methods have been employed to improve the agricultural productivity and minimise the impact from environment.

Currently, the agriculture has become a field of high input and high cost without wise use of fertilizer and plant protection measures. Due to uncertain factors such as weather, production, policy, price, etc., agriculture is no longer as profitable as before, which often brings losses to farmers. In today's changing situation, it is crucial to predict all aspects related to agriculture.

Despite the strong need for reliable and timely forecasts, the current situation is far from satisfactory. The impact of weather and climate on food production is crucial. Weather variables have different effects on crops at different stages of development. Therefore, it is necessary to use reliable and statistically reasonable models for crop yield prediction based on vegetation index variables and other factors, so as to obtain reliable prediction for efficient production.

Smart agriculture [1] is essential to address the challenges of agricultural production in terms of productivity, environmental impact, food security and Sustainability [2]. With the sustained growth of the global population, it is necessary to achieve a substantial increase in food production [3], while maintaining global food availability and high nutrition levels, and protecting natural ecosystem through sustainable agricultural processes.

To tackle these issues, it is necessary to better understand complex, diverse and

unpredictable agricultural ecosystems through continuous monitoring, measurement and analysis of various physical aspects and phenomena. This means the need to analyse big agricultural data [4] and the use of new information and communication technologies (ICT) [5], both for short-term crop / farm management and for large-scale ecosystem observation to enhance existing management and decision / policy making tasks.

Remote sensing technology [6] provides a large-scale field information of agricultural environment, which is carried out by satellites, aircraft and unmanned aerial vehicles (UAV). Remote sensing technology is a well-known non-destructive method, which can collect information about the earth characteristics and obtain data in a large geographical area. An illustration of remote sensing technology for agriculture is shown in Figure 1.1.

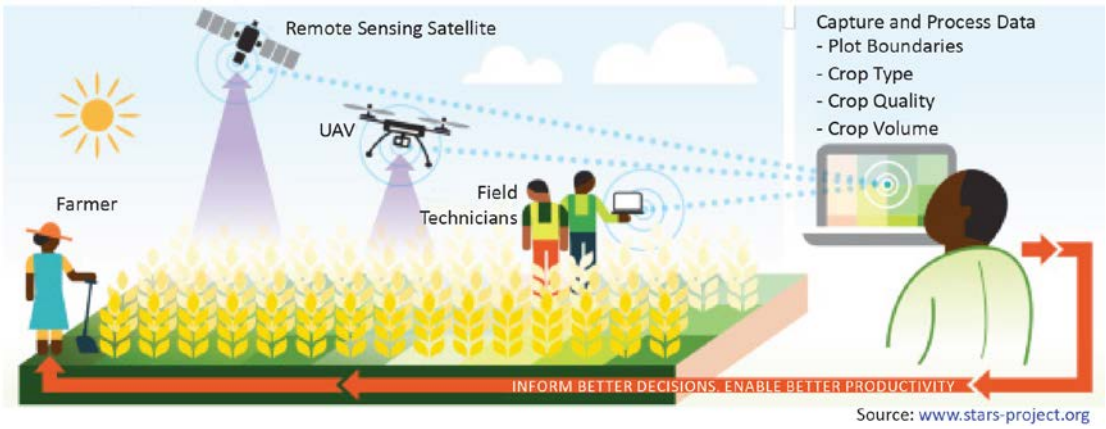


Figure 1.1 :Sketch map of remote sensing technology for agriculture [7]

One of the main challenges of food security is yield estimation, which can accurately predict crop yield before harvest. Agricultural monitoring, especially in developing countries, can improve food production and support humanitarian efforts based on climate change and drought [8].

Yield prediction is one of the most important topics in precision agriculture, which is of great significance to yield mapping, yield estimation, crop supply and demand matching and crop management to improve productivity.

It is estimated that 795 million people still do not have enough food to live on

(FAO, 2015) [9], and by 2050, 2 billion will need to be fed [8]. Eliminating hunger and improving food security are the main goals of UN agenda for sustainable development in the 2030 (UN, 2015) [10].

Remote sensing technology of various scales is usually proved to be a suitable tool for agricultural crop monitoring. Satellite images can monitor the crop growth in different stages, collect images in time series from the same place, and provide complete multi-dimensional space-time information. After acquiring such valuable information, we can build an effective model for the crop growth, which can reflect the growth and development of crops, so as to achieve continuous monitoring of crop growth. Through this constructed model, the yield evaluation can be also achieved. In particular, the remote sensing technology supported by UAV can capture more accurate information through high-resolution data. In recent years, the development of spectral recording system based on UAV has made great progress. Compared with the man-machine based system, the sensors for the UAV is smaller, lighter, and cheaper in the process of data acquisition and processing. The great potential of this technology has been demonstrated. The data generated in modern agricultural operations are provided by a variety of different sensors, which can better understand the operating environment (dynamic crop, interaction of soil and weather conditions) and the operation itself (mechanical data), so as to make the results more accurate and faster decisions [11].

The process-based crop growth model has been successfully used to predict the physiological development, growth and yield of crops in the field based on the interaction between vegetation index variables and plant physiological processes. These models require a large number of model parameters and inputs related to climate conditions, soil characteristics, management practices and crop variables, which are the main limiting factors of regional yield application. Remote sensing data provides spatiotemporal information of land surface in various scales. The remote sensing data is widely used to evaluate the size and change of crop condition parameters. The integration of remote sensing data and crop growth model represents an important research direction of precision agriculture.

Wheat is one of the most common and important economic crops in the world, which benefits from its versatility. Reliable and timely prediction of wheat yield provides important and useful input for correct, farsighted and wise planning.

Wheat yield prediction model has become an important research area due to its potential contribution to food security, because it may be regarded as a good indicator of global food supply. Wheat is one of the important food crops, and it is also an essential staple food for people. By increasing the yield of wheat and reducing the occurrence of diseases and insect pests, the loss can be greatly reduced. Early detection is essential for all kinds of wheat diseases. The traditional method of plant disease detection depends on the visual inspection. The common problem for the tradition method is that it is inefficient. For trained computers, the diagnosis of plant diseases is essentially a process of pattern recognition. After thousands of images of diseased plants are classified, machine learning algorithm can determine the category and severity of disease, and even give suggestions in the future to reduce the loss of disease.

Machine learning (ML) is a trend technology, which can be used in modern agriculture. The use of ML in agriculture helps to create more healthy seeds. The principle that Arthur Samuel [12] used in machine learning experiments has been applied in modern agriculture. Artificial machine learning in agriculture is one of the fastest growing fields. The agricultural sector is using artificial technology to improve accuracy and find solutions.

Artificial intelligence (AI) has grown steadily as part of the technological development of the industry. As early as last century, artificial intelligence technologies have begun to exploit in the field of agriculture, but it has not made much progress. After entering the 21st century, the great efficiency of artificial intelligence in the industrial field has brought more opportunities for precise agricultural.

By combining the intelligent information technology and traditional agriculture, the production mode of modern precision agriculture is deduced. In this mode, we can obtain high-precision spectral data of crops through satellites and unmanned aerial

vehicles. Also, we can obtain crops characteristics in different growth stages, analyse and predict the growth of crops, and make decisions in advance.

With the help of machine learning, plant breeding becomes more and more accurate and efficient, and can evaluate and predict a wider range of variables.

The work presented in this thesis will address the issues of wheat growth prediction, yield assessment and aphid validation by model building and machine learning algorithm optimisation by means of remote sensing data. Specifically, the following objectives are defined.

1. Analyse multiple vegetation indexes based on the TM 1-4 band data of Landsat satellite and use regression algorithms to train the models and predict wheat growth.

2. Analyse and compare multiple vegetation indexes models by means of spectral data and use regression algorithms to predict wheat yield.

3. Combine spectral vegetation indexes and multiple regression algorithms to predict wheat aphid.

4. Use accurate evaluation criteria for validating the efficacy of the various algorithms.

1.2 Original contributions

In this thesis, several different methods have been introduced to address the main objectives highlighted in Section 1 of Chapter 1. The major contributions of this thesis are summarised as follows:

1. A variety of optimisation algorithms are used to improve the regression analysis of the BPNN algorithm, so that the prediction results of each model for wheat growth, yield and aphid are more accurate.

2. The spectral characteristics of winter wheat canopy have been analysed. The correlation between the absorption band and the associated physical and chemical properties of crops, specially the red edge slope, with the crop yield and wheat aphid damage is established.

3. Adjusted MSE and un-centered R-square, as accurate evaluation criteria for

practical applications, are used to compare the prediction results of the models under different dimensions of the observed data.

4. Improve algorithm training by using the cross-validation method to obtain reliable and stable models for the prediction of wheat growth, yield, and aphid. Through repeated cross-validation, a better model can be obtained in the last.

1.3 Thesis organization

The remainder of the thesis is organised as follows:

Chapter 2 provides the relevant literature review focusing on recent developments in the field of growth prediction, yield assessment and aphid validation.

Chapter 3 introduces the fundamental theories and background required to understand the work done in this thesis.

Chapter 4 estimates multiple vegetation indexes related to wheat growth based on the TM 1-4 band data of Landsat satellite and uses four kinds of regression algorithms to train the models and predict the LAI and SPAD.

Chapter 5 calculates four kinds of vegetation indexes models related to wheat yield by means of spectral data and uses three kinds of regression algorithms to train the models and predict the yield; investigates the relationship between spectral data and wheat aphid, constructs nine vegetation indexes related to wheat aphid and uses five kinds of regression algorithms to train and predict wheat aphid.

Chapter 6 gives some concluding remarks about the work in this thesis and detailed plans to further improve the introduced results in this thesis.

Chapter 2.

Literature Review

The related work in machine learning and the application of spectral data in precision agriculture are briefly introduced in this chapter. The aim of this literature review is to highlight the contributions in the following two chapters of this thesis in context with the relevant research in wheat growth prediction, yield assessment and aphid validation. Firstly, Artificial neural network (ANN) for forecasting is introduced in section 2.1 of Chapter 2. Prediction of crop growth by remote sensing data such as Leaf Area Index (LAI) predictions, crop yield prediction and harmfulness and wheat aphid prediction are presented in sections 2.2 of Chapter 2. Finally, a brief chapter summary is provided in section 2.3 of Chapter 2.

2.1 Artificial neural network for forecasting

Compared with the model-based non-linear method, artificial neural network (ANN) was a non-linear data-driven method that can perform non-linear modelling without knowing in advance the relationship between input and output variables. Therefore, it was a more versatile and flexible modelling tool for prediction.

The idea of using artificial neural networks to make predictions was not introduced recently. The first application of ANN dates back to 1964. Hu [13] used Widrow's adaptive linear network for weather forecasting in his thesis. After the back-propagation algorithm was introduced [14], the application of ANNs in predictions had greatly developed due to the strong ability of back-propagation. Werbos [15] proposed back propagation and found that neural networks trained with back propagation have better performance than traditional statistical methods such as regression and Box-Jenkins methods [16].

Tang et al. [17], Tang and Fishwick[18], et al. reported several prediction comparison results between Box-Jenkins and ANN models. Zhang Guoqiang et al.,

[19] introduced the latest situation of the application of artificial neural network in prediction. The purpose was to find an effective neural network or a set of adaptive neural networks for prediction.

KOSCAK et al., [20] compared common weather forecasting methods with artificial neural networks and found that the performance of artificial neural networks had high accuracy. Mahdi Pakdaman Naeini et al., [21] used two neural networks, Forward Multilayer Perception and Elman Recursive Network, to predict a company's stock value based on its stock value history. The experimental results showed that compared with Elman regression network and linear regression method, the application of Forward Multilayer Perception neural network in stock value change prediction was more promising.

Ji et al., [22] investigated whether artificial neural network (ANN) models could effectively predict the yield of rice for typical climatic conditions of the mountainous region and compared the effectiveness of multiple linear regression models with ANN models. A. Irmak et al., [23] developed a back-propagation neural network (BPNN) model to predict the spatial distribution of soybean yields and to understand the causes of yield variability.

An intelligent system for effectively predicting Thrips Tabaci Linde (Thrips) population in cotton fields was proposed. In the design of the intelligent system, a forward multilayer perception neural network and a back propagation training algorithm were used. The neural network was trained and tested, and data was obtained. The experimental results showed that the system was effective for predicting the population dynamics of Trips pests in cotton fields. In addition, a comparative analysis was performed between the proposed system and the two existing models.

Artificial intelligence (AI) technology has been applied to soil, plants, rainfall and yield. Through the study of two artificial neural network technologies, the application of artificial neural network to improve production forecasting and nitrogen management in Western Australia was developed [24]. William W. Guo et al., [25] introduced a method for combining crop and yield analysis and forecasting. Then

statistical analysis and neural network inference were performed.

While designing effective features was challenging, researchers have taken another approach, which was to learn features from the data. Convolutional neural networks (CNNs) were one of the most effective models for learning image features from data. The high representation ability made CNNs at the latest level in various image analysis tasks, such as classification [26], segmentation [27], and object detection [28]), which were superior to traditional methods using manual features [29].

In summary, most researchers only carried out prediction on one aspect to realize the application of artificial neural network in agriculture applications. Few researchers synthetically applied artificial neural network to wheat growth, yield, disease, and pest prediction. What was less common was to carry out the neural network optimisation to improve the efficiency of the algorithm and the accuracy of prediction.

2.2 Prediction of crop growth by remote sensing data

2.2.1 Leaf Area Index (LAI) predictions

In 1947, Watson defined leaf area index (LAI) as the total unilateral area of leaf tissue per unit of ground surface area, giving a dimensionless value usually from 0 (for bare ground) to greater than 7 (for dense vegetation) [30]. LAI was applied to evaporation, transpiration, light absorption, yield estimation, crop growth stage and chemical element cycle in plant and environmental studies [31] .

LAI was an important parameter to indicate photosynthesis and growth state of crops, which was of great significance to the prediction of crop yield [32]. LAI could be used to elucidate the function of plant canopy [33]. LAI might be the most commonly used specific canopy index in crop research. This was related to canopy structure, which was directly related to photosynthesis and biomass accumulation. This index has been used in many ecological models. For example, Knyazikhin, et al.,

[34] used LAI and fraction of photosynthetically effective radiation absorbed by vegetation from atmospheric active multi angle imaging spectrometer data. Wu et al., [35] presented that the effect of integrated LAI and leaf N accumulation (LNA) data was better than that of using each parameter alone to optimise the parameters of crop model. LAI was also used to optimise SPAD trim rates because it was relatively easy to generate large amounts of LAI data.

Traditional LAI measurement methods included sample weighing and blade width to length ratio, both of which were time-consuming, which made it difficult to obtain large amount of LAI data in space and time [35]. Optical instruments were widely used to measure radiation through the canopy, so that LAI could be determined. Among all kinds of commercial optical instruments that could be used for indirect in situ LAI measurement, LAI-2000 plant canopy analyser was the most widely used one . Forest research showed that the LAI measured by LAI-2000 was much smaller than the actual LAI measured directly [44]. Stroppiana et al., [36] also introduced that when LAI > 1, LAI-2000 often underestimated LAI in rice. These studies indicated that the use of LAI-2000 to measure LAI in plants required appropriate validation.

When SPAD (Soil and Plant Analyser Development) value was used to diagnose the nitrogen status of plants, the response of the same leaf to nitrogen at different growth stages should be concerned [37]. Peng et al., [38] have proved that adjusting the SPAD value of specific leaf weight could improve the prediction of dry weight nitrogen status. However, the resulting individual data still did not reflect the rice situation very well. In response, CHL $(CHL = SPAD_{Upper} \bullet LAI_{Green})$ has been proposed to estimate the canopy chlorophyll status [39][40]. Ciganda et al., [41] also demonstrated the relationship between the chlorophyll content in each leaf and the total canopy CHL concentration, which was based on the red edge chlorophyll index: $CHL_{Red-edge} = (R_{NIR} / R_{Red-edge})$.

Multispectral (MS) and hyperspectral (HS) cameras have been widely used to monitor plant growth and biochemical indicators to obtain multiple vegetation index selection [42][43]. In order to measure the coverage and leaf area index (LAI), White,

et al., [44] recommended to optimise long-term monitoring by installing multi-band digital camera on tower platform. From this point of view, there were few practical applications. In addition, the required index for assessing crop chlorophyll status from remote sensing data was not only responsive to chlorophyll concentration, but also insensitive to the influence of background and LAI [45]. Through field radiation measurements, chlorophyll density per unit of land should be easier to estimate than chlorophyll concentration [46]. On the contrary, optical remote sensing of chlorophyll content was more challenging than chlorophyll density. However, Haboudane et al., [47] developed the prediction index of chlorophyll content in crop leaves with LAI value from 0.5 to 6, based on the reflectance data above the canopy. The techniques for collecting the reflectance images of rice canopy in visible red and near-infrared bands have been described, including fixed-point, seasonal continuous observations and their application in rice nitrogen uptake and LAI prediction (Shibayama et al.).

Using two band digital camera images, Takada et al., [48] showed the possibility of monitoring the seasonal variation of SPAD value and stem number of rice leaves, which was roughly estimated as the parameters of plant mountain orientation. Michio et al., [49] conducted a biennial study on fixed-point continuous two-band imaging technology, so as to make a stable prediction of the seasonal change of rice leaf color through SPAD value.

2.2.2 Crop yield prediction

Traditionally, yield prediction relied on ground-based field surveys, which were costly and poor in crop yield assessment (Reynolds et al.)[50]. Therefore, it was an important goal of crop production to develop a low-cost, fast and accurate regional yield prediction method (panda et al.) [51].

It was an effective means for remote sensing technology to monitor crop growth parameters, such as biomass (Fu et al.) [52], leaf area index (LAI) (Haboudane et al.) [53], and chlorophyll content (Haboudane et al.) [54], which could be accurately estimated by vegetation index (VIS). A series of studies have solved the problem of

crop yield prediction (Moran et al.) [55]. Tucker, etc. showed that there was a linear relationship between the normalized difference vegetation index (NDVI) and grain yield. Wang et al. obtained the yield prediction models with canopy reflectance ratio (NIR / red, NIR / GRN) from field measurements. These models have successfully predicted the yield of large-area rice through satellite images. Becker Reshef et al. [57] predicted wheat yield per unit area in Kansas and Ukraine using time series NDVI data from MODIS.

To improve the accuracy of production forecast, a multi template VI [58] was proposed. Wang et al. [56] predicted the yield of wheat from accumulation to the early stage of grain filling with the accumulated VI such as PNDVI (NIR, green) and PRVI (NIR, red), which was more accurate than that of single stage.

Although it was difficult to get enough data from satellites, the accumulated VI was the only form used. Another way to predict yield used agronomic parameters related to yield that could be estimated from remote sensing data. Researchers have estimated the absorbed photosynthetically active radiation (PAR) and LAI from remote sensing data and used them to predict production [59] [60].

Satellite images from Landsat, SPOT5 and QuickBird had high spatial resolution of 30 m, 10 m and 3 m, which overcomed the shortcomings of small survey area and showed high level of crop yield prediction accuracy. However, in the south of China, the application of satellite image was limited by the small plot of rice growing stage, complex terrain, overcast coverage and high cost. The latest technology progress of UAV and the miniaturization of sensors have rapidly expanded its application in precision agriculture. Based on a UAV system, remote sensing data with high spatial-temporal resolution could be obtained in a low-cost and more practical way for crop monitoring [61][62].

A large part of the yield modelling for assessing the impacts of climate change relied on deterministic biophysical crop models [63]. These models were based on detailed representations of plant physiology and were still important, especially for assessing response mechanisms and adaptive selection [64]. However, they were generally superior to statistical models in prediction on a larger spatial scale [65][66].

In particular, many literatures following Schlenker and Roberts [67]have used statistical models to prove a strong link between extreme high temperature and crop failure. These methods depended on classical econometric methods. Recent work has attempted to integrate crop models with statistical models by incorporating crop model output into statistical models [68] and using crop models in parameterization of statistical models [69].

Semi parametric neural network (SNN), has been developed to enhance the statistical models by using deep neural network. As a crop yield modelling framework, the prediction performance of SNN was better than any other published ones. SNN greatly improved the statistical efficiency of typical neural networks by using prior knowledge about important phenomena and functional forms related to the results. By using neural network with the parametric model, it could capture dynamics that did not exist or were not completely specified in parametric models.

2.2.3 Wheat Aphid prediction

In addition, in the current agricultural production practice, the prevention and control of diseases and insect pests were mainly through breeding resistant varieties and spraying insecticides (fungicides). Because of the variety of diseases and insect pests, the adaptability of pathogens and insect pests to variety resistance, and the quality and yield of resistant varieties were difficult to consider, the former cannot eliminated the impact of diseases and insect pests.

On the one hand, pesticide spraying increased the production cost. However, the abuse of pesticides might cause crop drug damage and economic loss, even brought some hidden dangers to the safety of food consumption, and increased the load of pesticide residues on the farmland ecological environment. The excessive use of pesticides not only increased the production cost of growers, but also seriously damaged and pollutes the ecological environment [70][71], and at the same time seriously endangered the national food safety and human health.

Therefore, by acquiring the spatial distribution and occurrence degree of diseases

and insect pests in real time and accurately, managers could be instructed when and where to spray pesticides and how to reasonably determine the scope and dosage of spraying chemicals, so as to effectively reduce the amount of chemicals used, reduce the cost of production and management, and reduce the negative impact of pesticides on crops and the environment.

From the perspective of national and regional agricultural macro management and decision-making, especially in the period when the occurrence of diseases and insect pests was more serious, timely and accurate understanding of the scope and extent of the occurrence of diseases and insect pests could also provide the basis for the relevant parts in formulating acquisition policies and agricultural insurance rates and claims.

Muhammed et al. [72] used the canopy hyperspectral data of wheat diseases to extract the disease characteristics; Huang et al. used ASD spectrometer to measure the canopy spectrum of wheat stripe rust, and constructed the spectral index suitable for monitoring the disease; In [73], Jiang et al., used hyperspectral differential index to monitor the winter wheat stripe rust, and found that the differential vegetation index can monitor and retrieve the disease information of wheat stripe rust; In [74], Wang et al., captured the canopy hyperspectral data of different severity of wheat stripe rust, and used the support vector machine (SVM) algorithm to classify and identify the severity of wheat stripe rust with high accuracy. Chen Bing et al. [75] estimated the spectral characteristics and severity of cotton verticillium wilt by using the hyperspectral data collected from the ground. Liu et al., used similar instruments to determine the spike spectrum of rice glume blight and establish the relationship model between the spectrum and disease level [76].

For hyperspectral imaging, Luo Juhua et al. used push scan hyperspectral PHI aerial image to construct a monitoring model based on the combination of waveband sensitive to wheat stripe rust, realized mapping of the susceptible range and severity of wheat stripe rust in the field, and got a good inversion result [77];Zhang et al., recognized and distinguished crop diseases and other stresses based on imaging hyperspectral technology, a satisfied result has been achieved [78].

Jones et al., employed an imaging spectrometer (Cary 500, Varian Inc., Palo Alto, CA, USA) with a spectral range of 200-2500nm to detect tomato leaf lesions [79]; Fiore et al., successfully identified maize yellow curve disease (Aspergillus flavus) [80]; Zhang et al., constructed a spectral knowledge base through the ground hyperspectral and Airborne Hyperspectral Imaging PHI, and preliminarily studied the multispectral remote sensing monitoring method of wheat stripe rust [81].

Yang et al., have compared the sensitivity of 16 band multi spectral vegetation index and Landst 5 TM vegetation index to wheat aphid, the results showed that 16 band multi spectral vegetation index was less sensitive to wheat aphid than TM wide band vegetation index; Jonas and Menz studied wheat powdery mildew and wheat stripe rust recognition based on QuickBird image after image change, and the recognition accuracy reached 88.6% [82].

Yang et al., have investigated the effect of multispectral image and hyperspectral image in monitoring cotton root rot and found that the multispectral image can also achieve satisfactory accuracy [83]; Luo et al., firstly identified and distinguished the damage degree of wheat aphids by obtaining the aphid damage level of the ground sample points, and then extracted the two-dimensional feature space constructed by the surface temperature and the normalized water index from the multi temporal Landsat 5 TM multi spectral data, and good recognition results have been achieved [84].

Wheat aphids were the main pests in China's wheat production areas, accounting for more than 90% of the total area every year, which harmed wheat by absorbing the juice of Wheat by adults and larvae. In addition, the honeydew discharged by aphids falled on the wheat leaves, which seriously affected its photosynthesis, thus causing the wheat yield reduction. According to statistics, aphids could reduce the wheat yield by more than 8% all year round, and the yield reduction in serious years was up to 20-30%. According to the damage degree of wheat aphid, the 1000 grain weight measured in different conditions was 7.8-34.6% lower than that of uninjured wheat. Wheat aphids not only reduced the yield of wheat, but also seriously damaged the nutritional quality of wheat, and reduced the content of crude protein, hydrolyzed

protein amino acid, vitamin B and vitamin C in wheat flour.

In order to effectively control the harm of wheat aphid, timely, accurate and large-scale monitoring and prediction of the occurrence of wheat aphid was the premise of effective control work. The application of remote sensing technology could improve the method of wheat aphid monitoring and prediction and improved the timeliness of wheat aphid monitoring and prediction.

Aphids were one of the most destructive pests, causing damage to greenhouse crops [85]. They could reproduce in a few days, with nymphs hiding on the lower surface of the leaves. Quantitative measurement of harmful organisms was time-consuming and error prone [86]. Although it had been proposed to improve strategies such as counting pest subsets to estimate the total population [87][88], it was still expensive to apply manual counting in large-scale practice. Therefore, automatic methods should be developed to support rapid decision-making.

The application of pesticides was considered to be able to effectively control aphids and minimise the yield loss of winter wheat [89], but the efficacy and cost of pesticides were questionable for long-term use [90]. Because of the poor regularity of wheat aphid infestation, the automatic spraying system was usually applied to overdose pesticides, which inevitably increased the production cost and environmental impact of wheat [91]. Therefore, it was important to monitor aphid infection at the critical moment of crop growth and obtain the spatial distribution information of aphid invasion in specific fields.

The conventional method of obtaining density infection information in the field was artificial field investigation, but it had been proved that this method was expensive, time-consuming and difficult to carry out in large cultivation areas. Fortunately, because the cost of hyperspectral remote sensing technology was relatively low, and it could be installed on airborne and airborne platforms, it might be an effective alternative method in large area, which could obtain the spatial distribution information of large area aphid invasion. Many studies have found that hyperspectral remote sensing had the potential to detect crop stress, such as disease [92], water stress [93], toxic industrial chemical stress [94].

Image-based pest automatic segmentation and case counting have been widely studied. Some researchers used sticky traps to collect pests before counting them. The color of the sticky trap was fixed, so color model transformations (such as lab [95], HSV [96], YCbCr [97] and YUV [98]) and thresholding methods were applied to segmentation. However, for nymphs on leaves, the color threshold method might classify the veins or lesions as pests. The researchers applied a well-designed post-processing program to go to areas other than pest areas.

Barbedo [99] considers the area within a specific eccentricity range as aphid nymph. Maharlooei and his colleagues [100] used size filtering to get rid of the segmentation results of aphid objects. The misclassification rate of these strategies was low, but it needed to be corrected manually in some cases. A similar approach was used by Solis-Sánchez et al. [101] [102]. As suggested by Barbedo, designing more effective features to describe and detect targets could further improve segmentation.

2.3 Summary

In this chapter, application of artificial neural network and prediction of crop growth by remote sensing data have been introduced. The applications of ANN in various fields have been introduced, especially in the fields of wheat growth prediction, yield assessment and aphid validation. Afterwards, a brief summary of multiple vegetation indexes and regression algorithms for crop prediction will be presented. Based on the presented work, the designed methods will be also given afterwards.

Chapter 3.

Theoretical Background

3.1 Introduction

The fundamental theories required to understand the work done in this thesis are introduced in this chapter.

In Section 3.2 of Chapter 3, the background theories for the various regression algorithms are introduced. In Section 3.3 of Chapter 3, vegetation indexes which can reflect crop growth information through linear or nonlinear combination of remote sensing spectral data of different wavebands are introduced. In Section 3.4 of Chapter 3, Spectral characteristics of vegetation, which are important basis to distinguish vegetation from non-vegetation, vegetation types, and to monitor the growth of vegetation, are introduced. In Section 3.5 of Chapter 3, the definitions for the evaluation criteria are introduced. Adjusted MSE and un-centered R-square are presented for practical application.

In Section 3.6 of Chapter 3, cross-validation for algorithm training is introduced. Cross- validation is employed throughout this thesis for algorithm training to obtain reliable and stable models for the prediction of wheat growth, yield and aphid. Mathematical background is presented in sufficient details to understand the usage and implications. Finally, a brief summary is provided in section 3.7 of Chapter 3.

3.2 Regression analysis

Regression analysis and prediction method are to establish regression equation between variables on the basis of analysing the correlation between independent variables and dependent variables. The regression equation is taken as the prediction model to predict the dependent variables according to the quantity change of independent variables in the prediction period. In statistics, regression analysis refers

to a statistical analysis method to determine the quantitative relationship between two or more variables. Regression analysis can be divided into single regression analysis and multiple regression analysis according to the number of variables involved; straight forward regression analysis and multiple regression analysis according to the number of dependent variables; linear regression analysis and nonlinear regression analysis according to the type of relationship between independent variables and dependent variables.

The steps of regression analysis prediction method are as follows:

1. Determine the independent variable and dependent variable. In one experiment, the variable that the experimenter actively manipulates and may have an impact on the response of the subjects is independent variable, which is independent of the behavior of the subjects. Dependent variable is the variable changed by the change of independent variable, which is observed by the experimenter.

2. Establish regression prediction model. Based on the historical statistics of independent variables and dependent variables, the regression analysis equation is established.

3. Check the regression prediction model and calculate the prediction error. Whether the regression prediction model can be used for the actual prediction depends on the verification of the regression prediction model and the calculation of the prediction error. The regression equation can be used as a prediction model only if it passes various tests and the prediction error is small.

4. Calculate and determine the predicted value. The regression prediction model is used to calculate the prediction value, and the comprehensive analysis of the prediction value is carried out to determine the final prediction result.

There are various regression techniques for prediction, which are briefly discussed as follows.

1. Linear regression

It is one of the most familiar modeling techniques. Linear regression is usually one of the most preferred techniques in learning prediction model. In this technique, the

dependent variable is continuous, the independent variable can be continuous or discrete, and the property of regression line is linear [103][104][105].

Linear regression uses the best fitting line (i.e. regression line) to establish a relationship between the dependent variable (y) and one or more independent variables (x). Schematic diagram of linear regression is shown as Figure 3.1.

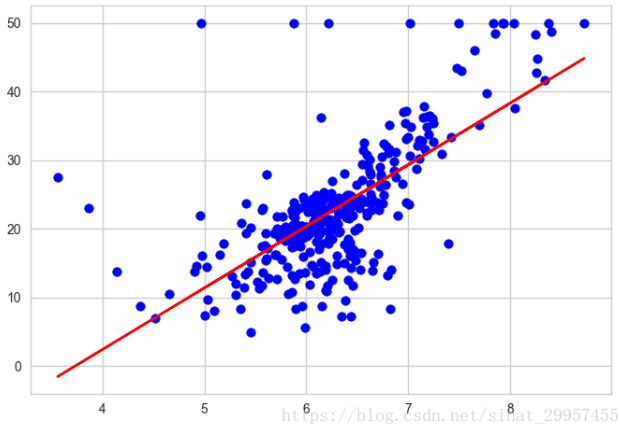


Figure 3.1: Schematic diagram of linear regression

2. Polynomial regression

For the case of not satisfying linear regression, polynomial regression [106][107] can be used by adding polynomials.

$$y = w_0 + w_1 x + w_2 x^2 + \cdots + w_n x^n \tag{3.1}$$

In this regression technique[108], the best fit line is not a straight line. It is a curve used to fit data points. Schematic diagram of polynomial regression is shown as Figure 3.2.

3. Using Neural Networks for Regression

Generally, neural network is often used for supervised learning, classification and regression. Neural network can help to group unlabeled data, classify data, or output continuous values after supervised training. The typical application of neural network in classification is to use logical regression classifier at the last layer of the network to convert continuous values to a discrete value such as 0/1. For example, given a person's height, weight, and age, with or without a heart attack can be determined. The real regression is to map one set of continuous inputs to another set of continuous

outputs [109][110][111].

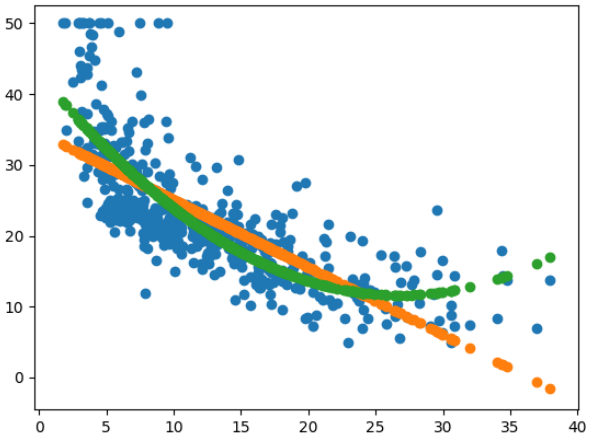


Figure 3.2: Schematic diagram of polynomial regression

For example, given a house's age, area, and distance to a good school, we will predict the price of the house. This is continuous input mapped to continuous output. There is no 0/1 in the classification task, but only independent variable x mapped to continuous output y. Regression structure of neural network is shown as Figure 3.3.

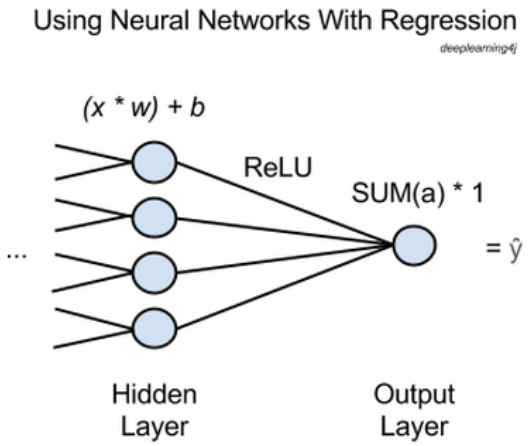


Figure 3.3: Regression structure of neural network

In Figure 3.3, x represents the input, and the feature propagates forward in the layer in front of the network. Multiple x's are connected with each neuron in the hidden layer, and each x will be multiplied by a corresponding weight w. the sum of these products plus a bias is sent to an activation function ReLU (= max (x , 0)), a

widely used activation function, which will not saturate like the sigmoid activation function.

For each hidden layer neuron, activation function ReLU inputs an activation value a. At the output node of the network, the sum of these activation values is calculated as the final output. There will be an output node using the neural network for regression, and this node only adds the activation values of the previous nodes. The obtained $\hat{y}$ is the independent variable obtained from all x mappings.

The training process of neural network is as follows. In order to carry on the network back propagation and the network training, we can simply use the network output $\hat{y}$ to compare with the real value y and adjust the weight and bias to minimise the network error. The root means squared error (RMSE) can be used as the loss function.

3.3 Vegetation indexes

Crop yield estimation by remote sensing is a technology based on the special spectral reflection characteristics of crops, which uses remote sensing to monitor and forecast crop yield. Any objective has the characteristics of absorbing and reflecting electromagnetic waves of different wavelengths, which are the basic characteristics of the objective. Crop yield estimation refers to a series of methods to identify crop types, monitor crop growth and predict crop yield before harvest, based on the collection and analysis of different spectral characteristics of various crops in different growth periods and through the surface information recorded by sensors on the platform. It includes crop identification, sowing area extraction, growth monitoring and yield prediction.

The spectral information of the image can be used to retrieve the growth information of crops (such as LAI, biomass). The yield information of crops can be obtained by establishing the correlation model between the growth information and the yield (combining some agricultural models and meteorological models). In

practical work, vegetation index (a mathematical index which can reflect crop growth information through linear or nonlinear combination of multispectral data) is often used as a standard to evaluate crop growth.

Vegetation index is a linear or non-linear combination of different remote sensing spectral bands. It is a marker of the relative abundance and activity of green vegetation (dimensionless). It is a comprehensive reflection of LAI, coverage, chlorophyll content, green biomass and absorbed photosynthetic effective radiation (APAR) of green vegetation. The vegetation index mainly reflects the differences between the visible light, near-infrared reflection and soil background. Each vegetation index can be used to quantitatively describe the growth of vegetation under certain conditions.

When learning and using vegetation index, there should be some basic knowledge:

1. The reflection of healthy green vegetation in NIR and R is quite different, because R is strongly absorbed by green plants, and NIR is highly reflected and highly transmitted.

2. The purpose of establishing vegetation index is to effectively integrate all relevant spectral signals, enhance vegetation information and reduce non-vegetation information.

3. Vegetation index has obvious regionality and timeliness, which is affected by vegetation itself, environment, atmosphere and other conditions.

Plant leaves have obvious absorption characteristics in visible light band and obvious reflection characteristics in near infrared band, which is the physical basis of vegetation remote sensing monitoring. Different vegetation indexes can be obtained by different combinations of the measured values of these two bands.

Some commonly used vegetation indexes are given below (as Figure 3.1):

a. NDVI normalized difference vegetation index

NDVI = (NIR-R) / (NIR+R).

1. Application of NDVI: detection of vegetation growth status, vegetation coverage and elimination of some radiation errors;

2. – 1<=NDVI<=1, negative value means that the ground is covered with clouds, water, snow, etc., which reflects the visible light highly; 0 means that there is rock or bare soil, etc., NIR and R are approximately equal; positive value means that there is vegetation coverage, which increases with the increase of coverage.

3. The limitation of NDVI is that the contrast of the reflectivity of NIR and R is enhanced by nonlinear stretching. For the same image, when calculating RVI and NDVI separately, it can be found that the increasing speed of RVI value is higher than that of NDVI, that is, NDVI has low sensitivity to high vegetation area.

4. NDVI can reflect the background impact of plant canopy, such as soil, wet ground, snow, dead leaves, roughness, etc., and is related to vegetation coverage.

b. RVI ratio vegetation index

RVI = NIR / R

1. The RVI of green and healthy vegetation covered area is far greater than 1, while the RVI of the ground (bare soil, artificial building, water body, dead vegetation or serious insect pest) without vegetation coverage is near 1. The RVI of vegetation is usually greater than 2.

2. RVI is a sensitive indicator parameter of green plants, which has a high correlation with LAI, leaf stem biomass and chlorophyll content. It can be used to detect and estimate plant biomass.

3. Vegetation coverage affects RVI. When the vegetation coverage is high, RVI is very sensitive to vegetation; when the vegetation coverage is less than 50%, the sensitivity is significantly reduced.

4. The RVI is affected by atmospheric effect which greatly reduces the sensitivity of vegetation detection. Therefore, atmospheric correction is required before calculation of RVI.

c. DVI difference vegetation index

DVI =NIR-R

DVI can well reflect the change of vegetation coverage, but it is sensitive to the change of soil background. When the vegetation coverage is 15% ~ 25%, DVI

increases with the increase of biomass; and when the vegetation coverage is more than 80%, the sensitivity of DVI to vegetation decreases.

d. SAVI soil adjusted vegetation index

$$SAVI = ((NIR - R) / (NIR + R + L))(1 + L)$$

Among them, L is the parameter with the change of vegetation density, the value range is 0~1. Obviously, if L=0, SAVI = NDVI.

1. SAVI is to explain the changes of the optical characteristics of the background and correct the sensitivity of NDVI to the soil background. Compared with NDVI, the soil regulation coefficient L determined according to the actual situation is increased. When L=0, it means that the influence of soil background is zero, that is to say, the vegetation coverage is very high. This kind of situation will only appear in the place covered by tall trees with dense canopy.

2. SAVI is only applicable when the soil line parameter a=1, b=0 (i.e. very ideal state). There are TSAVI, ATAVSI, MSAVI, SAVI2, SAVI3, SAVI4 and other improved models.

e. PVI Perpendicular vegetation index

In the two-dimensional coordinate system of R-NIR, the vertical distance between vegetation pixel and soil brightness line.

PVI = $((S_R\text{-}V_R)^2+(S_{NIR}\text{-}V_{NIR})^2)^{1/2}$, S is the soil reflectance, V is the vegetation reflectance.

1. The influence of soil background is eliminated, and the sensitivity to atmosphere is less than other VI.

2. PVI is the simulation of GVI in R-NIR two-dimensional data, both of which have the same physical meaning.

3. $PVI = (DN_{NIR} - b)\cos\theta - DN_R \bullet \sin\theta$, b is the intercept between soil baseline and NIR, θ is the angle between soil baseline and R.

f. GVI green degree vegetation index

After K-T transformation, it represents the component of green degree.

1. The spectral characteristics of vegetation and soil were separated by K-T transformation. The spectrum of vegetation growth process is in the shape of so-called "spike cap", while the soil spectrum constitutes a soil brightness line. The spectral changes of soil moisture content, organic matter content, particle size, mineral composition, surface roughness and other characteristics are generated along the direction of the soil brightness line.

2. After KT transformation, the first component represents soil brightness, the second component represents green degree, and the third component represents different meanings with different sensors. For example, the third component of MSS represents yellow degree, which has no definite meaning, and the third component of TM represents humidity.

3. The first two components contain more than 95% of the information, which can reflect the differences between vegetation and soil spectral characteristics.

4. GVI is the weighted sum of the radiance values of each band, and the radiance is the comprehensive result of atmospheric radiation, solar radiation and environmental radiation, so GVI is greatly affected by the external conditions.

Table 3.1: VI classification [112]

Type	Typical representative	Characteristics
Linear type	DVI	When LAI is low, the effect is better; when LAI is increased, it is sensitive to soil background.
Ratio type	NDVI, RVI	The reflection contrast between soil and vegetation is enhanced.
Vertical type	PVI	When LAI is low, the effect is better; when LAI is increased, it is sensitive to soil background.

3.4 Spectral characteristics of vegetation

The spectral characteristics of vegetation are the important for distinguishing vegetation from non-vegetation, vegetation types, and monitoring the growth of vegetation. Pigment absorption determines the spectral reflectance of visible light, cell

structure determines the spectral reflectance of near-infrared, and water vapor absorption determines the spectral reflectance of shortwave infrared. Spectral response characteristic curve of green plants is shown as Figure 3.4.

In general, vegetation has the following typical spectral characteristics in the range of 350-2500nm.

1). 350-490 nm spectrum: As 400-450 nm spectrum is obvious absorption band of chlorophyll, 425-490 nm spectrum is obvious absorption band of carotenoid, and there is weak absorption band of atmosphere near 380 nm wavelength, the average reflectivity of 350-490 nm spectrum is very low, generally no more than 10%, and the shape of reflection spectrum curve is very gentle;

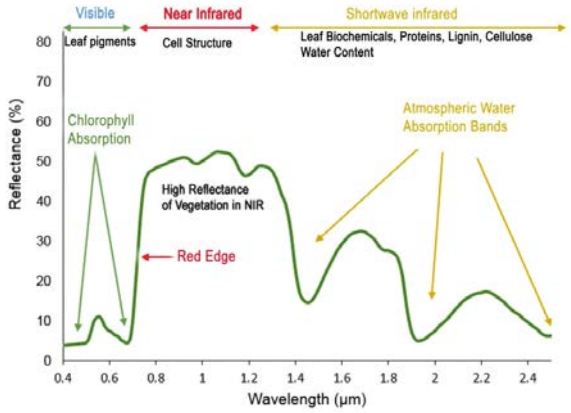


Figure 3.4: Spectral response characteristic curve of green plants [113]

2). 490-600mn spectrum: because there is a obvious reflection peak area of chlorophyll near the wavelength of 550nm, the reflection spectrum curve of vegetation in this band has a wave peak shape and a medium reflectivity value (about 8-28%);

3). 600-700nm spectrum: 650-700nm is the obvious absorption band of chlorophyll, 610 and 660nm are the main absorption bands of phycocyanin, so the reflection spectrum curve of vegetation at 600-700nm has the shape of trough and very low reflectivity value (usually no more than 10% except for the plant community at the defoliation stage);

4). 700-750nm spectrum: the reflection spectrum curve of vegetation rises sharply in this spectrum segment, which has the shape of steep and nearly straight line. Its slope is related to the content of chlorophyll per unit area, but it tends to be stable when the content exceeds 4-5mg/cm^2.

5). 750-1300nm spectrum: vegetation in this band has the characteristics of obvious reflection (it can be understood as the self-defense instinct of plants against burns), so it has a high reflectivity value. In this band, the average reflectivity measured in the laboratory is mostly between 35% and 78%, while that measured in the field is mostly between 25% and 65%. Because of the narrow absorption band of water or oxygen near the wavelength points of 760nm, 850nm, 910nm, 960nm and 1120nm, the reflection spectrum curve of vegetation in the 750-1300nm range also has the characteristics of undulation.

6). 1300-1600nm spectrum: it is related to 1360-1470nm spectrum which is the obvious absorption band of water and carbon dioxide. The reflection spectrum curve of vegetation in this spectrum has the shape of wave trough and low reflectivity value (mostly between 12-18%):

7). 1600-1830nm spectrum: it is related to the spectral characteristics of plants and their water content. The reflection spectrum curve of vegetation in this band has the shape of wave crest and higher reflectivity value (mostly between 20-39%);

8). 1830-2080nm spectral segment: this spectral segment is a obvious absorption zone of water and carbon dioxide in plants, so the reflection spectral curve of vegetation in this spectral segment has the form of trough and very low reflectivity value (mostly between 6-10%);

9). 2080-2350nm spectral band: it is related to the spectral characteristics of plants and their water content, and the vegetation is in this wave.

The reflection spectral curve of the segment has peak shape and medium reflectivity value (mostly between 10% and 23%):

10). 2350-2500mn spectral section: this spectrum section is a obvious absorption zone of water and carbon dioxide in plants, so the reflection spectrum curve of vegetation in this spectrum section has the shape of trough and low reflectivity value

(mostly between 8-12%).

Vegetation spectral reflection is the basis of vegetation remote sensing, and the spectral characteristics of green vegetation mainly depend on its leaves. The main spectral response characteristics of green plant leaves are shown in the Figure 3.4. For the reflection spectrum curve of green plant leaves, in the blue band with a central wavelength of 450nm and the red band with a central wavelength of 670nm, chlorophyll obviously absorbs radiation energy (> 90%) and forms absorption valley, the latter is the energy basis of photosynthesis, while the absorption between the two absorption valleys (near the 540nm wavelength) is relatively reduced, forming a green reflection peak area (10-20%), which is also the optical principle of green plants [114].

With the change of wavelength, the reflectance of green plants increases sharply from 740nm, forming a sharp and prominent peak. This is mainly because there are many cavities in the structure of spongy tissue in the plant mesophyll, which has a large reflective surface, and the chlorophyll in the cell is water-soluble colloidal state, which has a obvious infrared reflection. In addition, the plant spectrum often shows two weak water absorption bands at 960nm and 1100nm. In the shortwave infrared spectrum (beyond 1300nm), the reflection spectrum characteristics of plants are controlled by the water absorption band with the centers of 1400nm, 1900nm and 2700nm, and the attenuation curve is in a falling state.

As the background of green plants, the "peak valley" change of the reflection spectrum curve of soil is relatively weak, and its effect on electromagnetic wave is mainly reflected and absorbed, with less transmission. Therefore, by using the difference of near-infrared high reflection and red band obvious absorption of vegetation, the information of vegetation is highlighted, the difference between green plants and soil water body is distinguished, and the difference between vegetation and environment is strengthened.

In short, the reflected spectrum of vegetation is based on the following characteristics:

1). There is a small reflection peak at 0.55um in the visible band, and there are

two absorption bands at 0.45um (blue) and 0.67um (red) on both sides. This characteristic is the effect of chlorophyll.

2). In the near infrared band (0.7-0.8um), there is a reflected "steep slope" (known as "red edge"), and there is a "peak" near 1.1um, forming the unique characteristics of vegetation. This feature is caused by vegetation structure.

3). In the mid infrared band (1.3-2.5um), the reflectance is greatly reduced, especially in the center of 1.45um and 1.95um is the absorption band of water, forming a trough.

In addition to the special spectral response characteristics of plants controlled by the pigment, water content and its structure, the physiological, shape and structure of plants will change from germination, growth, flowering and fruiting to senescence and death in the growth cycle, such as the change of leaf structure, i.e. the increase or decrease of mesophyll cell gap, the accumulation, decline and disappearance of chlorophyll and other pigments, and the change of water content in leaves,, the germination and growth of branches and leaves, and the change of plant coverage to the ground. The periodic change of this plant takes season as its cycle, and its chemical, physical and biological properties also show seasonal change seasonal rhythm. The seasonal rhythm of the plant is reflected from the micro structure of the plant cell to the macro structure of the plant population, which will inevitably lead to the periodic change of the physical and optical characteristics of a single plant or plant population, and its spectral characteristics change accordingly.

Because the spectral reflectance of vegetation is related to various factors such as vegetation type, species composition, vegetation coverage, chlorophyll content, water content, soil physical characteristics, atmospheric conditions, etc., the image characteristics of different plant types are different, and even the same plant, in different growth and development stages, its spectral reflectance is slightly different, so it can be changed according to the vegetation reflectance. To monitor crop growth, phenology and crop type identification. It can be seen that the study of vegetation parameters is actually a good foundation for the further study of vegetation remote sensing and agricultural remote sensing.

3.5 Evaluation Criteria

3.5.1 Error Analysis

Regression is different from classification. The regression method firstly predicts a series of values. After the prediction is completed, it is necessary to evaluate the predicted results. The following are some common methods for assessment.

1. SSE：The sum of squares due to error

This statistical parameter is to calculate the sum of the squares of the errors of the fitting data and the corresponding points of the original data. The calculation formula is as follow,

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{3.2}$$

The closer the SSE is to 0, the better the model selection and fit, and the more successful the data prediction.

2. MSE: Mean squared error

This statistical parameter is to calculate the mean of the sum of the squares of the corresponding point errors of the predicted data and the original data. The calculation formula is as follow,

$$MSE = SSE/n = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{3.3}$$

3. RMSE: Root mean squared error

The statistical parameter, also called the standard deviation of the regression system, is the square root of the MSE. The calculation formula is as follow,

$$RMSE = \sqrt{MSE} = \sqrt{SSE/n} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{3.4}$$

4. R-square: Coefficient of determination

Due to the different dimensions of different data sets, it is difficult to compare different models through the above three methods. A third party can be used as a reference to calculate the R-square value based on the reference, so that the models

can be compared.

This reference is the mean model. For example, a data set has a mean value, the house price data set has an average price, and the student transcript has an average score. This mean is now treated as a benchmark reference model, also called the baseline model. This mean model has the same predictive value for any data, and it is conceivable that the model is naturally inferior. Based on this we will want to find patterns from the data set to build a better model. The calculation formula for R-square is as follows.

Firstly, the other two parameters SSR and SST is introduced, because the coefficient is determined by two of them:

SSR：Sum of squares of the regression，that is, the sum of the squares of the difference between the predicted data and the mean of the original data, the formula is as follow,

$$SSR = \sum_{i=1}^{n} (\hat{y}_i - \overline{y}_i)^2 \tag{3.5}$$

SST：Total sum of squares，that is, the sum of the squares of the difference between the original data and the mean, the formula is as follow,

$$SST = \sum_{i=1}^{n} (y_i - \overline{y}_i)^2 \tag{3.6}$$

Obviously, SST = SSE + SSR. R-square is defined as the ratio of SSR to SST below:

$$R^2 = \frac{SSR}{SST} = \frac{SST - SSE}{SST} = 1 - \frac{SSE}{SST} \tag{3.7}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \overline{y}_i)^2} \tag{3.8}$$

According to its value, how it evaluates the quality of the model can be better understood. There are several kinds of typical values, which are discussed as follows.

When R-square = 1, it reaches the maximum. That is, the numerator is 0, meaning that the predicted value and the true value are completely equal without any error. That is to say, the model we build perfectly fits all the real data, and it is the best

model, and the R-square value is also the largest. But usually the model is not so perfect, there will always be errors. When the error is small, the numerator is less than the denominator, the model will approach 1. It is still a good model. As the error gets larger and larger, R-square will be further and further away from the maximum value of 1.

R-square = 0. At this point the numerator is equal to the denominator and each predicted value of the sample is equal to the mean. That is to say, the model we have worked hard to train is exactly the same as the model of the mean value mentioned above. It's better not to train, let the prediction value of the model take the mean value directly.

R-square < 0: The numerator is larger than the denominator, and the error generated by the training model is larger than that generated by using the mean value, that is, the training model is not as good as the direct averaging effect. When this happens, usually the model itself is not linear, and we misuse the linear model, resulting in large errors.

0<R-square<1: When R-square increases, the model is more accurate and the regression effect is more significant. The closer R-square is to 1, the better regression fitting effect. It is generally considered that the goodness of fit of the model is higher when R-square exceeds 0.8.

3.5.2 Adjusted MSE and Un-centered R-Square

3.5.2.1 Adjusted MSE

The mean square error is the mean of the sum of the squares of the point errors of the predicted data and the original data. The root mean square error is the standard deviation of the regression system and is the square root of the mean square error (MSE), which is the square root of the mean of the sum of the squared errors of the predicted data and the original data. The closer the calculation results of these two errors are to 0, the better the model selection and fitting, and the more successful the data prediction.

The mean square error and the root mean square error value are closely related to the number of samples n and the dimension. Their size is directly related to the value of the raw data. The larger the average of the original data, the larger the mean square error and the root mean square error, and vice versa. When the dimension is different, it is difficult to measure.

For example, the model predicts an error RMSE of $50,000 on a house price dataset and a score of 10 on another student grade dataset. With these two values, it is difficult to judge which data set the model is applied to is better.

So how to compare the predictions of models under different dimensions? Here is a modified MSE model:

$$Adjusted\ \ MSE = MSE / \sum (y_i)^2 \tag{3.9}$$

Obviously, the Adjusted MSE is not affected by the dimension of the observed data. Its normal range is [0-1]. The closer to 1, the better the model fits the data.

3.5.2.2 Un-centered R-Square

R-square is defined as the regression squared sum of SSR divided by the total deviation squared sum SST, which is based on the equation SST = SSR + SSE, and this is only true if the MODEL has a constant term. Linear regression and multiple linear regression models can meet these requirements.

When there is no constant term, the original defined SST=SSE+SSR does not hold, and the definition of R-square changes, and R-square is negative.

At this time, Un-centered R-Square is used to evaluate the fitting effect of the output data. Un-centered R-square is the proportion of dependent variable variation that can be explained by the variation of explanatory variables. When the regression process does not contain a constant term, it is used instead of the decision coefficient R-square. It is through the change of data to characterize the quality of a fit. Its normal value range is [0-1], and the closer to 1, the better the model fits the data.

This is the new version of SST,

$$SST = \sum_{i=1}^{n} \hat{y}_i^2 + \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{3.10}$$

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} \hat{y}_i^2 + \sum_{i=1}^{n} (y_i - y_i)^2} \tag{3.11}$$

3.6 Cross-validation

In the related research of pattern recognition and machine learning, the dataset is often divided into two parts: a training set and a testing set. The former is used to build the model, while the latter is used to evaluate the accuracy of the model in predicting unknown samples. The formal term is the generalization ability. How to divide the complete data set into a training set and a testing set, the following points must be observed:

1. Only the training set can be used in the training process of the model. The testing set must be used to evaluate the merits of the model after the model is completed.

2. The number of samples in the training set must be sufficient, generally at least 50% of the total number of samples.

3. The two sets of subsets must be evenly sampled from the complete set.

The last point is particularly important. The purpose of uniform sampling is to reduce the bias between the training set/testing set and the complete set, but it is not easy to do. The general practice is random sampling. When the number of samples is sufficient, the effect of uniform sampling can be achieved. However, randomness is also the blind spot of this method, and it is often possible to modify the data. For example, when the recognition rate is not ideal, a set of training sets/testing sets are resampled until the recognition rate of the testing set is satisfactory.

The purpose of cross-validation is to obtain a reliable and stable model. Cross-validation is a common method used in machine learning to build models and validate model parameters. Cross-validation, as the name implies, is the repeated use of data.

The obtained sample data is segmented, combined into different training sets and testing sets, the training set is used to train the model, and the testing set is used to evaluate the prediction of the model. On this basis, multiple sets of different training sets and testing sets can be obtained. A sample in a training set may become a sample in the testing set next time, which is called “cross”.

Cross-validation is used when the data is not sufficient. For example, in a daily project, for a moderately moderate problem, if the data sample size is less than 10,000, cross-validation will be used to optimise the training model. If the sample size is larger than 10,000, the data is generally randomly divided into three, one is the training set, the other is the validation set, and the last is the testing set. The training set is used to train the model, and the validation set is used to evaluate the quality of the model prediction and the model selection and its corresponding parameters. The resulting model is used in the testing set and ultimately determines which model to use and the corresponding parameters.

Back to cross-validation, according to the different methods of division, cross-validation is divided into the following three types: The first type is simple cross-validation, so-called simplicity, as opposed to other cross- validation methods. Firstly, the sample data is randomly divided into two parts (for example: 70% for the training set and 30% for the testing set), then the training set is used to train the model, and the model and parameters are verified on the testing set. The sample is then scrambled, the training set and testing set are reselected, and data training and model checking are continued. Finally, the loss function is selected to evaluate the optimal model and parameters.

The second type is S-Folder cross-validation. Different from the first method, the sample data in the S-fold cross-validation is randomly divided into S parts, and each time S-1 parts are randomly selected as the training set, and the remaining 1 part is used as the testing set. When this round is completed, S-1 parts are randomly selected again for training data. After several rounds (less than S), the loss function is selected to evaluate the optimal model and parameters.

The third type is Leave-one-out cross-validation, which is a special case of the

second case, where S is equal to the number of samples N, so that for N samples, N-1 samples are selected each time. To train the data, a sample is left to verify the prediction of the model. This method is mainly used in cases where the sample size is very small. For example, for a moderate problem, when N is less than 50, a cross-check is generally used.

Through repeated cross-validation, the loss function is used to measure the quality of the obtained model, and finally a better model can be obtained. As for the above three cases, which method should we choose? In a word, if you just make a preliminary model for the data, instead of doing an in-depth analysis, simple cross-validation is fine. Otherwise, use S-fold cross-validation. When the sample size is small, a special case of S-fold cross-validation is used to leave a cross-validation.

There is also a special cross-validation method, which is also used when the sample size is small. It is called bootstrapping. For example, there are m samples (m is small), and each time one sample is randomly collected in the m samples, put into the training set, and the sample is put back after sampling. This acquisition is repeated m times to obtain a training set consisting of m samples. Of course, there is a high probability that there will be duplicate sample data in these m samples. At the same time, the testing set is taken with samples that are not sampled. This is followed by cross-validation. Since our training set has duplicate data, this will change the distribution of the data, so the training results will have an estimated bias. Therefore, this method is not very common unless the amount of data is really small, such as less than 20.

3.7 Summary

This chapter introduced the fundamental theories to be used in this thesis. The background theories for various regression algorithms were introduced. Regression analysis and prediction models could be used to establish the regression analysis between variables on the basis of analysing the correlation between both the independent variables and dependent variables. The steps of regression analysis

prediction were described. Vegetation indexes could reflect crop growth status through linear or non-linear combination of remote sensing spectral data at different wavebands. The fusion of vegetation parameters was useful for the further study of vegetation remote sensing and agricultural remote sensing. Various vegetation indexes were presented and compared. Each vegetation index could be used to quantitatively describe the growth of vegetation under certain conditions. Spectral characteristics of the vegetation, which are important basis to distinguish vegetation from non-vegetation, vegetation types, and to monitor the growth of vegetation were introduced. To compare the predictions of models under different dimensions, adjusted MSE and un-centered R-square were established for practical applications. The adjusted MSE was unaffected by the dimension of the observed data. Un-centered R-square could be explained by the variation of the explanatory variables. Finally, cross-validation for model training was introduced. The purpose of cross-validation was to obtain a reliable and stable model. Through repeated cross-validation, the loss function was used to measure the quality of the obtained model, and finally a better model could be determined.

Chapter 4.

Prediction of Wheat Growth by Satellite Image Data

Through the combination of different spectral bands, the satellite remote sensing technology can retrieve or extract the characteristic factors of crop growth process, which can comprehensively reflect the crop growth and its change dynamics. In order to further improve the stability, efficiency, and accuracy of BP network training for prediction of the wheat growth, the genetic neural network algorithm and particle swarm neural network algorithm are introduced into the prediction of surface vegetation parameters. Firstly, the basic framework of BP network is established. The input of the network is the vegetation parameters based on TM data, and the output are the measured LAI and SPAD values. The learning and training process of neural network is optimised by genetic algorithm and particle swarm optimisation algorithm, and the optimal solution of network connection weights and thresholds is calculated. Finally, the optimal connection weights and thresholds are inversely tested on the unknown model. Compared with the single neural network algorithm, the genetic neural network algorithm and particle swarm optimisation algorithm perform better in the network. The prediction results of various regression algorithms are compared.

4.1 Introduction

In the last few decades, the crop growth monitoring has become an essential issue in the macro-management of agricultural production. With the indispensable information provided by the crop growth monitoring, the crop growth could be estimated far before the crop harvest, which gave the possible prospect of improving the yield. The crop growth monitoring based on satellite and UAV image data could qualitatively and quantitatively analyse crop growth status and seedling growth trend in a wide range.

There have been some outstanding achievements in the related prediction of agricultural crops. UAV-based RGB imaging data sources were widely used for precision agriculture applications[115][116]. Schirrmann et al. [117] studied the prospects for monitoring the biophysical parameters and nitrogen status in wheat crops with low-cost imagery acquired from unmanned aerial vehicles (UAV) over an 11 hectare field. Wang et al., [118] used the ecological parameters and remote sensing data to construct the winter wheat remote sensing quality model based on the research on the relationship between different vegetation indexes (VIs) at different growth stages of winter wheat. Lelong et al. [119] carried the filter and digital camera on the UAV, monitored the wheat experimental fields in southwestern France, and analysed the relationship between the spectral index and the biophysical parameters measured in the field based on the spectral images in the visible-near infrared band.

Satellite remote sensing image could record crop growth in different stages in real time, obtained time series images of the same location. Due to the straightforward accessibility, remote sensing data from satellites have also been applied into the prediction of crop growth. P. Yang et al. [120] presented the research result on estimating winter wheat yield in North China Plain, by assimilating multitemporal Landsat TM images into the GIS-based EPIC model. Zhao, Yu [121] applied NDVI curve retrieved from satellite images of MODIS 8-days 250m surface reflectance to monitor soybean's yield.

At present, various models have been implemented to predict crop growth. Yuan et al.[122] evaluated the performance of RF, ANN, and SVM models in LAI inversion using spectral data and corresponding ground data for heterogeneous soybean crops. Han et al. [123] used random forest (RF) and support vector machine (SVM) methods to invert the canopy LAI of apple trees. Omer et al. [124] used artificial neural network (ANN) and SVM methods to predict the LAI of six endangered tree species. Verrelst et al. [125] used ANN, SVM, nuclear ridge regression (KRR), and Gaussian process regression (GPR) methods to predict LAI. Mustafa et al. [126] combined the Bayesian network (BN) method with a forest growth model to improve LAI prediction accuracy.

Because each regression method had its own characteristics in principle, one method was not necessarily superior to the other, and the hybrid optimisation method could often synthesize their respective advantages to obtain better prediction results and better adaptability.

Taking BP network algorithm as an example, there were some drawbacks in the back propagation learning algorithm based on gradient method in network training. For example, the performance depended on the initial weight, and the possibility of solution achieving global optimisation was not assured. In the last decades, the Genetic algorithm (GA) has been successfully employed in overcoming the limitations of the BP learning algorithm via optimising the BP network in various fields [127][128].

In order to overcome the problems of slow convergence of network learning in traditional BP network and easy to fall into local minimum, the particle swarm optimisation algorithm PSO will be introduced. By using the particle swarm optimisation (PSO) algorithm instead of the gradient descent method in BP network, the connection weight between each layer of BP network can be better optimised, and the generalization ability and learning ability of BP network are improved, and its convergence speed is greatly improved.

4.2 Methodology

4.2.1 Multiple-Linear Regression (MR) Algorithm

The regression equation is used to quantitatively describe the linear dependence of a dependent variable and multiple independent variables, called multiple linear regression (referred to as multiple regression). The general purpose of multiple regression (the term was first used by Pearson, 1908) is to learn more about the relationship between several independent or predictor variables and a dependent or criterion variable.

Multiple linear regression (MR), also known simply as multiple regression, is a

statistical technique that uses several explanatory variables to predict the outcome of a response variable. The goal of multiple linear regression (MR) is to model the linear relationship between the explanatory (independent) variables and response (dependent) variable.

In essence, multiple regression is the extension of ordinary least-squares (OLS) regression that involves more than one explanatory variable. The Formula for Multiple Linear Regression is given by

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots \beta_p x_{ip} + \varepsilon \tag{4.1}$$

Where,

y_i = dependent variable
x_i = expanatory variables
β_0 = y-intercept (constant term)
β_p = slope coefficients for each explanatory variable
ϵ = the model's error term (also known as the residuals)

The coefficient of determination (R-squared or R^2) is a statistical metric that is used to measure how much of the variation in outcome can be explained by the variation in the independent variables. R-square always increases as more predictors are added to the MR model even though the predictors may not be related to the outcome variable.

R-square by itself can't thus be used to identify which predictors should be included in a model and which should be excluded. R-square can only be between 0 and 1, where 0 indicates that the outcome cannot be predicted by any of the independent variables and 1 indicates that the outcome can be predicted without error from the independent variables.

When interpreting the results of a multiple regression, beta coefficients are valid while holding all other variables constant ("all else equal"). The output from a multiple regression can be displayed horizontally as an equation, or vertically in table form.

4.2.2 Back Propagation (BP) Neural Network Algorithm

BP network is also known as back-propagation neural network. Through training of sample data, it constantly revises the weights and thresholds of the network so that the error function decreases along the direction of negative gradient, and pushes forward the expected output. BP network has strong ability of non-linear simulation.

Figure 4.1 is a BP network schematic diagram. The first layer is the input layer, and the number of nodes N is determined by the dimension of the input vector; the middle layer is the hidden layer, and the number of nodes is optional; the last layer is the output layer, and the number of nodes M is determined by the dimension of the output vector.

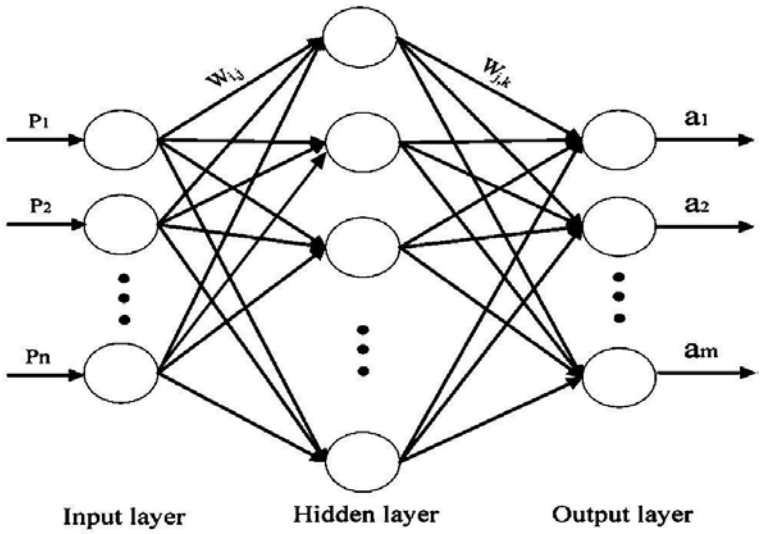


Figure 4.1: Schematic diagram of BP network

When forward propagation occurs, input samples enter the network from the input layer, and are transferred to the output layer through the hidden layer. If the actual output of the output layer is different from the expected output, the error back propagation occurs. If the actual output of the output layer is the same as the expected output, the learning algorithm is terminated.

When back propagation occurs, the output error (the difference between the expected output and the actual output) is calculated according to the original path, and the error is distributed to each unit of each layer through the hidden layer to the input

layer. The error signal of each unit of each layer is obtained and used as the basis for correcting the weight of each unit. This calculation process is completed by gradient descent method. After continuously adjusting the weights and thresholds of each layer of neurons, the error signal is reduced to the minimum.

The process of constant adjustment of weights and thresholds is the learning and training process of the network. After the forward propagation of signals and the back propagation of errors, the adjustment of weights and thresholds is repeated until the pre-set training times or the output errors are reduced to the allowable extent.

4.2.3 Genetic Algorithm (GA)

When dealing with complex nonlinear system, BP is faced with a series of problems, such as slow convergence speed, unstable network and falling into local optimum, because the initial weights set depend on the designer's experience and repeated experiments in sample space.

Genetic algorithm is an iterative adaptive search algorithm based on natural selection and genetic mechanism. It has strong global search ability and strong stability. Combining BP network algorithm with genetic algorithm, any non-linear system could be theoretically mapped and global optimal result could be obtained, thus forming a more effective non-linear inversion method. Firstly, genetic algorithm could be used to optimise the "BP network" to find a better search space, and then the optimal solution in a smaller search space could be searched for BP network.

The genetic algorithm optimises the BP network as follows:

1). According to the problems to be solved, the network structure is initialized to determine the number of nodes in input layer, output layer and hidden layer, which are recorded as N, M and H respectively.

2). Initial population P(t) is generated, genetic population size is set, and gene coding is performed for each individual in the population. The coding length of each chromosome is $(N \bullet H + H \bullet M + H + M) \bullet L$, and L is the digits number of variables.

3). Calculate the fitness of each individual in the population, and select the

individuals with good fitness according to certain rules to form a new population P(t). The fitness function is E(k) which is the prediction error of the network. The smaller the value of the function, the greater the fitness and the better the performance of the given neural network.

4). The individuals in the new population are randomly paired, and each pair of individuals is crossed according to a certain crossover probability to produce two new individuals; and each individual in the new population is mutated and evolved according to the set mutation probability to produce new individuals.

5). Add the new individuals into the population to form a new population P(t+1). The population size remains unchanged. The individual fitness of the new population is calculated. If the iteration termination condition is satisfied, the next step will be taken, otherwise repeat (3)-(5) steps.

6). When the maximum genetic generation or the error requirement is reached, the best individual is selected as the genetic result, and the connection weights and thresholds of the neural network are initially assigned. Then the network training is carried out to adjust the weights and thresholds until the accuracy requirement or the maximum iteration times are met, and the algorithm is finished.

4.2.4 Particle Swarm Optimisation (PSO) Algorithm

Particle swarm optimisation was developed in 1995 by Kennedy and Eberhart, inspired by the behaviour of social organisms in groups, such as bird and fish schooling or ant colonies.

Particle swarm optimisation is similar to the genetic algorithm technique for optimisation in that rather than concentrating on a single individual implementation, a population of individuals (a “swarm”) is considered instead. Each individual in the swarm has a position and velocity defined, the algorithm looks at each case to establish the best outcome using the current swarm, and then the whole swarm moves to the new relative location.

To solve the problem in BP network with PSO algorithm, the key points are as

follows:

1). The dimensional values of the particles in the particle swarm algorithm correspond to the number of connection weights in the BP network. The dimension components of each particle in the PSO are corresponding to each of the connection weights in the BP network. The key in the PSO algorithm is the particle. Many particles form a multi-dimensional vector, and the number of connection weights is consistent with the dimension of the particle.

2). Replace the adaptive function of the particle swarm algorithm with the mean square error function in the BP network. After that, the powerful iterative search capability of the particle swarm algorithm is used to minimise the mean square error of the BP network.

The specific learning process of the BPNN-PSO network is as follows: Firstly, the number of weights in the BP network is calculated, which is recorded as the dimension of the individual vector of the particle swarm algorithm; then the particles are randomly generated under the dimension condition, and the learning iteration is performed. According to the particle swarm algorithm, the vector of the individual optimised by the algorithm is transformed into the connection weights in the BP network, and the weights are assigned to the BP network; after the network is determined, all the sample data are trained through the BP network. The mean square error is calculated; the program does not stop until the error is less than the accuracy value set by the system (not exceeding the set maximum number of iterations).

If the structure of the BP network is certain, then we only need to encode the connection weights between the neurons, and then determine the dimension of the particles, and then generate new particles and determine the network. The mean square error of the BP network output is used as the fitness function to evaluate the quality of the generated individual, and the survival of the fittest is achieved. The goal is to achieve the mean square error value or the fitness function value to a minimum.

Here are the details of the steps.

1). The particle swarm algorithm begins by creating the initial particles, and assigning them initial velocities.

2). Evaluates the objective function at each particle location, and determines the best (lowest) function value and the best location.

3). Chooses new velocities, based on the current velocity, the particles' individual best locations, and the best locations of their neighbors.

4). Then iteratively updates the particle locations (the new location is the old one plus the velocity, modified to keep particles within bounds), velocities, and neighbors. Iterations proceed until the algorithm reaches a stopping criterion.

4.3 Model construction

4.3.1 Samples

The experimental field is located at north latitude $32^0$25', and East longitude $119^0$31'. The total area of the test area is 50 mu (hm2). The experimental data are from the LANDSAT satellite data on May 7, 2015 and the ground manual measurement data.

The experimental field is divided into 1078 plots. 147 representative breeding plots are measured by SPAD502 chlorophyll content analyser and LAI-2200C plant canopy analyser. The information of Winter Wheat Canopy Chlorophyll and leaf area index are collected. The cell distribution is shown in the Figure 4.2.

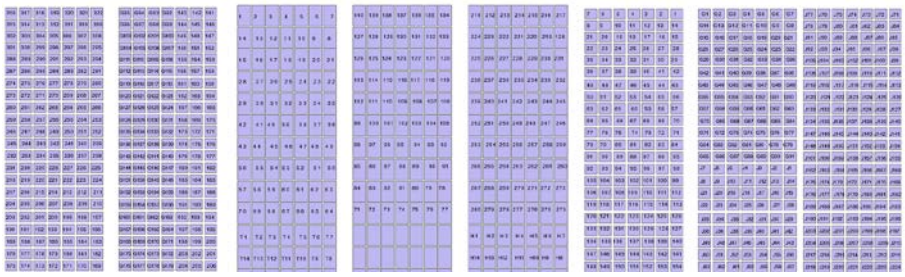


Figure 4.2: The experimental field and cell distribution

The TM image refers to the multi-band scan image acquired by Thematic Mapper on the US Landsat satellite 4 to 5. The spatial resolution of the image is 30 meters except for the thermal infrared band (120 meters), and the image width is 185×185km2. Because of its high spatial resolution, spectral resolution, extremely

rich information volume and high positioning accuracy, TM images have become an important source of remote sensing data for earth resources and the environment, which has been widely used in countries in the mid-to-late 1980s. TM image can meet the relevant analysis of agriculture, forestry, water, soil, geology, geography, surveying and mapping, regional planning, environmental monitoring.

The four bands (TM1, TM2, TM3, TM4) of Landsat satellite TM data correspond to visible blue, green, red and near infrared bands.

The image characteristics of each band are as follows:

TM1 is a blue band of 0.45-0.52 micrometers. This band is located at the location where the water body attenuation coefficient is the smallest. With the largest penetrating power to the water body, it can be used to distinguish the water depth, study the topography of shallow sea water, turbidity of water body, etc., and map the water system and shallow waters;

TM2 is a green band from 0.52 to 0.60 microns. This band is located near the reflection peak of green plants and is sensitive to reflections from healthy lush plants. It can be used to identify plant types and evaluate plant productivity. With a certain penetrating power to the water bodies, it can reflect underwater terrain, sandbars, coastal sand dams and other characteristics;

Table 4.1: The parameters and formulas used in this thesis

Parameters	Formula
Redness intensity	R
Greenness intensity	G
Blueness intensity	B
Near infrared intensity	NIR
Normalized difference vegetation index, NDVI	(NIR-R)/(NIR+R)
Ratio vegetation index, RVI	NIR/R
Green ratio vegetation index, GRVI	NIR/G-1
Normalized redness intensity, NRI	R/(R+G+B)
Normalized greenness intensity, NGI	G/(R+G+B)

Note: R, G, B are pixel values in the red, green and blue channels, respectively.

TM3 is a red band from 0.63 to 0.69 micrometers. This band is located in the main absorption zone of chlorophyll, which can be used to distinguish plant type, coverage, and judge plant growth. In addition, this band provides abundant plant information for bare ground, vegetation, lithology, stratigraphy, structure, geomorphology, hydrology and other characteristics.;

TM4 is 0.76~0.90 micron, which is in the near-infrared band. This band is located in the high reflection area of plants and reflects a lot of plant information. It is mostly used for plant identification and classification. It is also located in the obvious absorption area of water body and can be used for mapping water boundary, identifying water-related geological structures, landforms, etc. More detailed vegetation index parameters (such as NDVI, RVI, GRVI, NRI, NGI) can be obtained based on above band values. The parameters and formulas are shown in Table 4.1.

4.3.2 Network structure design

The values of the vegetation parameters (R, G, B, NIR, and NDVI) are used as input with LAI & SPAD as output, so the number of nodes in the input layer is 5 and the number of nodes in the output layer is 2.

Relevant studies show that a neural network with a hidden layer can approximate a non-linear function with arbitrary accuracy as long as there are enough hidden nodes. Therefore, a three-layer multi-input and two-output with a hidden layer is adopted to establish a prediction BP network. There is a common empirical formula for determining the number of nodes in the hidden layer. $h = \sqrt{m+n} + a$, where h is the number of hidden layer nodes, m and n are the number of input layer and output layer nodes respectively, and a is the adjustment constant between 1 and 10. In this experiment, the number of nodes in the hidden layer is set to 6.

The schematic diagram of the network structure is shown as Figure 4.3. The neural network toolbox in MATLAB is used to train the network. The training sample

data are normalized and input into the network. The transfer functions of input layer, output layer and hidden layer are set as tansig, tansig and purelin function (default) respectively. The training function is set as trainglm function (default). The number of network iterations epochs is 5000, the expected error goal is 0.00000001, and the learning rate LR is 0.01.

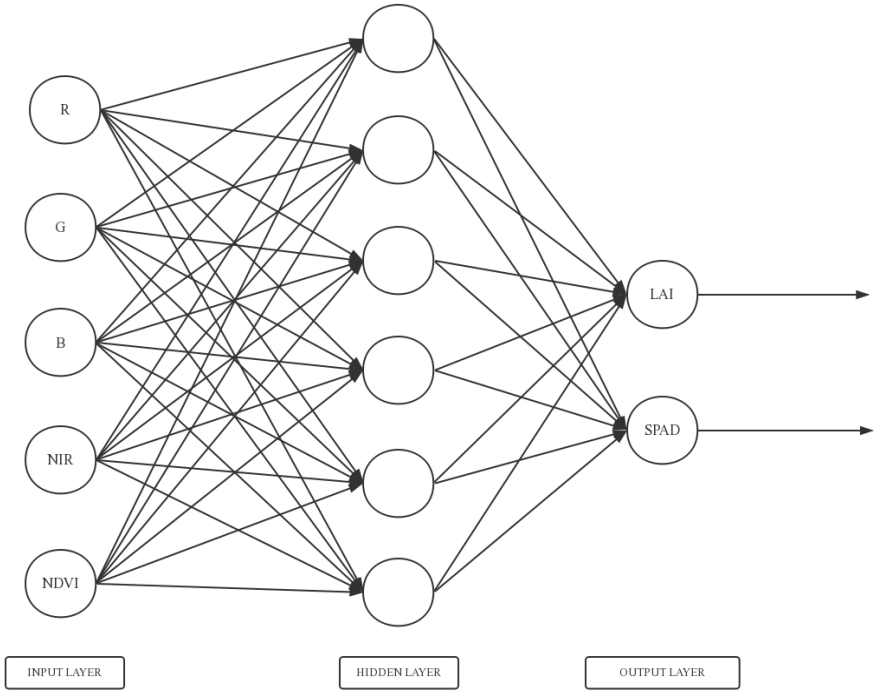


Figure 4.3: Schematic diagram of the network structure

4.3.3 Optimising process of Genetic Algorithms

Population size refers to the total number of individuals in any generation, which is set manually. The larger the population size is, the more likely it is to find a global solution, but the operation of the program is also relatively time-consuming. Generally, the value is between 40 and 100. The choice of the crossover probability and mutation probability is generally between 0 and 1. The evolutionary generation is just the number of iterations.

Without loss of generality, the initial parameters of the neural network optimised by the genetic algorithm are as follows: population size 100, crossover probability 0.3, mutation probability 0.1 and evolutionary generations 10. All these settings are the

initial parameters. Generally, within the appropriate value range, these settings have limited influences on the results.

The basic flow of genetic algorithm to optimise the initial weights and thresholds of neural networks is shown in Figure 4.4.

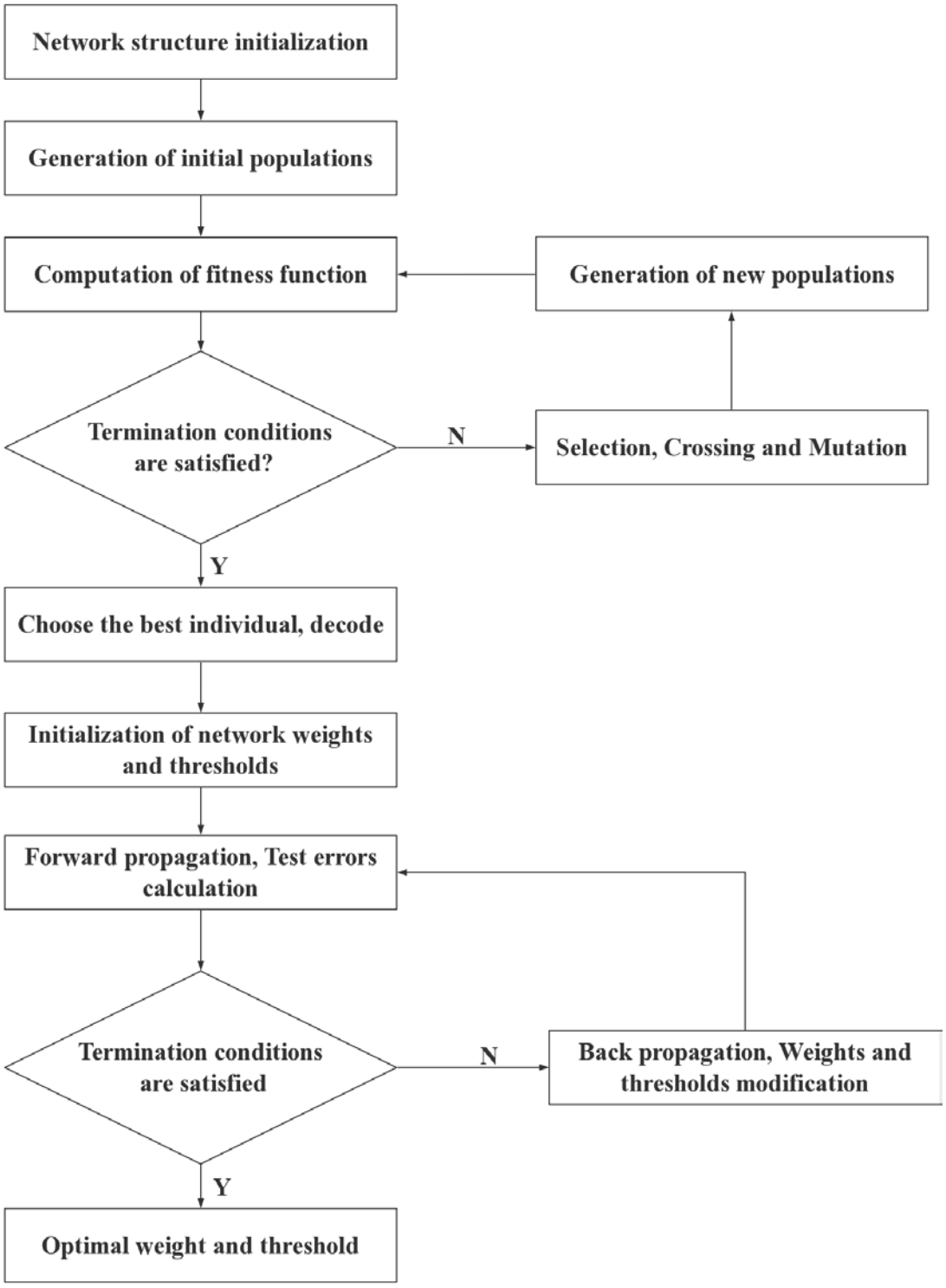


Figure 4.4: Flow chart of neural network optimised by genetic algorithm

4.3.4 Optimising process of PSO Algorithms

The basic flow of PSO algorithm to optimise the initial weights and thresholds of neural networks is shown in Figure 4.5.

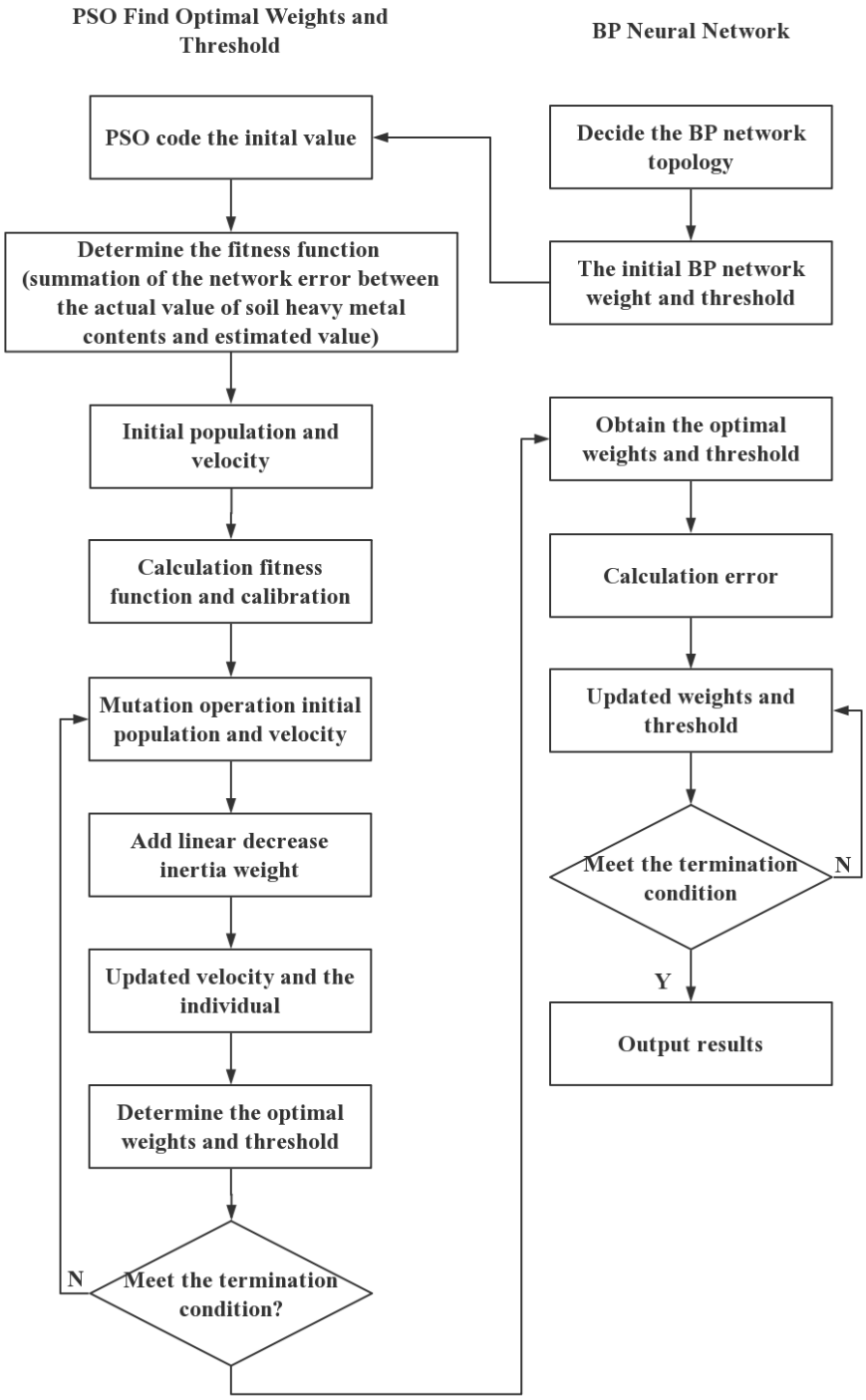


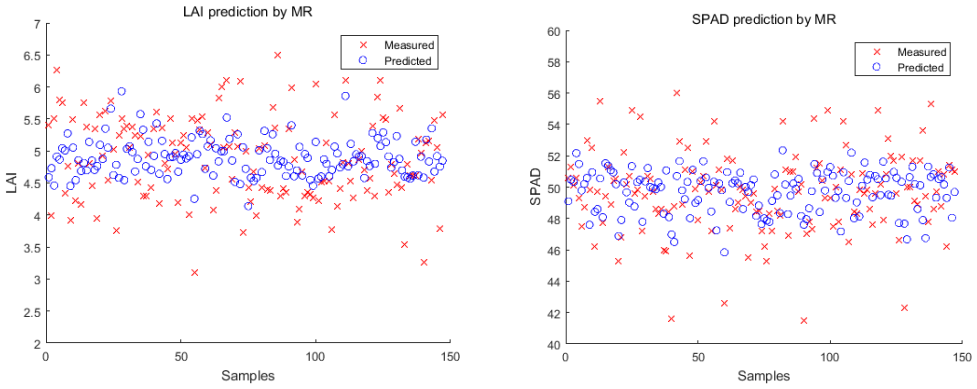
Figure 4.5: Flow chart of neural network optimised by PSO algorithm

4.4 Results and analysis

Five parameters of R, G, B, NIR of TM data and NDVI obtained by operation are used as input of the network, and LAI and SPAD parameters measured on the ground are used as output of the network. There are total 147 sets of data, of which 100 sets are chosen randomly to train the algorithm of BP network (BPNN), genetic neural network (BPNN-GA) and Particle Swarm Optimisation neural network (BPNN-PSO). The remaining 47 sets of data are validated by the model. According to the cross-validation, the above operations were carried out 20 times, and the optimal RMSE is finally selected. Besides, multiple regression (MR) algorithm is used to predict the total 147 sets of data directly.

4.4.1 Comparison of Results

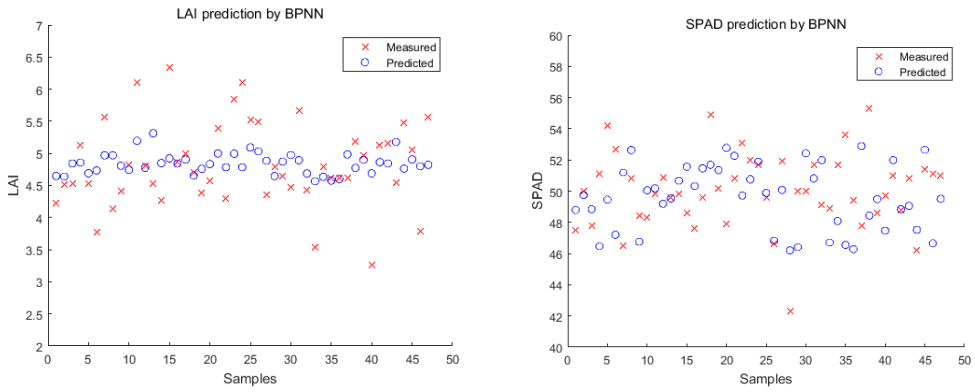
The results of predicting LAI and SPAD by multiple regression (MR) algorithm are shown in Figure 4.6.



(a) (b)

Figure 4.6: The results of predicting by MR. (a) LAI and (b)SPAD

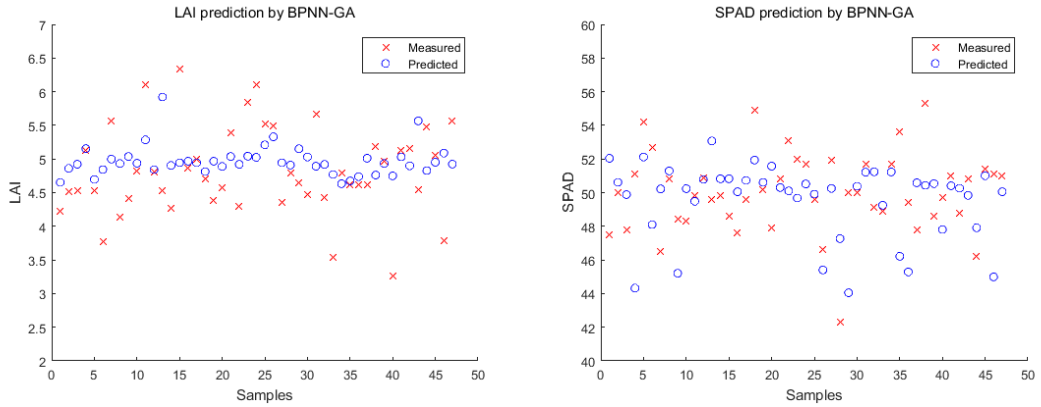
The results of predicted LAI and SPAD by BPNN algorithm are shown in Fig. 4.7.



(a) (b)

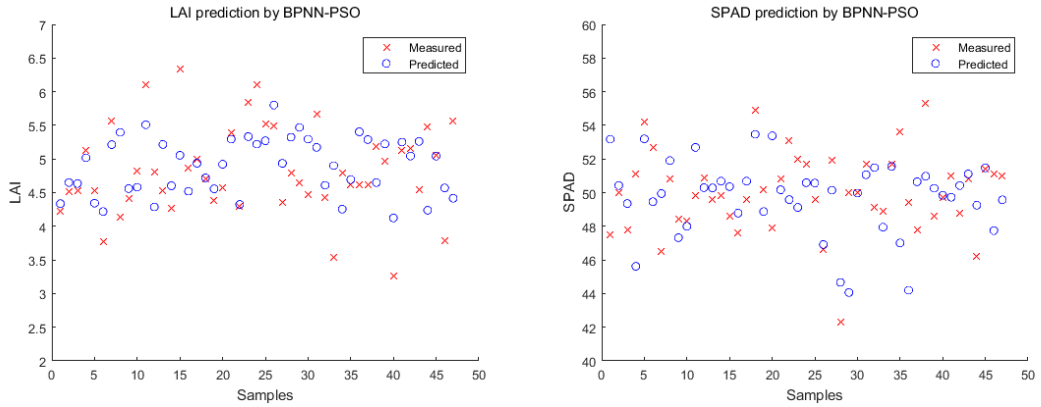
Figure 4.7: The results of predicting by BPNN. (a) LAI and (b)SPAD

The results of predicting LAI and SPAD by BPNN-GA algorithm are shown in Figure 4.8.



(a) (b)

Figure 4.8: The results of predicting by BPNN-GA. (a) LAI and (b)SPAD



(a) (b)

Figure 4.9: The results of predicting by BPNN-PSO. (a) LAI and (b)SPAD

The results of predicting LAI and SPAD by BPNN-PSO algorithm are shown in Figure 4.9.

4.4.2 Results analysis

From the comparison among Figure 4.6, Figure 4.7, Figure 4.8 and Figure 4.9, the comparisons between predicted data and measured data for LAI prediction and SPAD prediction by Multiple Regression, BPNN, BPNN-GA and BPNN-PSO are shown, which can lead to the conclusion that the prediction of Multiple Regression and BPNN and BPNN-GA have more errors than prediction of BPNN-PSO, especially in the prediction of LAI.

On the other hand, the predicted data of BPNN-PSO are more convergent and stable. When the measured data has a significant change, the predicted data of BPNN-PSO will have more logical change rather than a jump in prediction like Multiple Regression and BPNN. It means that the measuring error has less effect on BPNN-PSO. Therefore, all this data shows that in practical application and in big data application the BPNN-PSO will be more accurate and stable.

RMSE between model output and measured data for different algorithms are given in Table 4.2.

Table 4.2: RMSE between model output and measured data

	LAI	SPAD
MR	0.0153	0.0029
BPNN	0.0155	0.0035
BPNN-GA	0.0184	0.0035
BPNN-PSO	0.0144	0.0020

From Table 4.2, Whether LAI or SPAD prediction, BPNN-PSO has Minimum RMSE. This result also confirms that BPNN-PSO has Minimum prediction error. The predicted data of BPNN-PSO is better than the predicted data of Multiple Regression

or BPNN or BPNN-GA for both LAI and SPAD. The PSO algorithm is used to optimise the initial weights and thresholds of BP network, which improves the convergence rate to a certain extent.

R-square between model output and measured data for different algorithm are as Table 4.3.

In Table 4.3, either LAI or SPAD prediction, BPNN-PSO has the Maximum R-square. This result also confirms that BPNN-PSO has Minimum prediction error. Also, the PSO algorithm improves the convergence rate to a certain extent.

Table 4.3: R-square between model output and measured data

	LAI	SPAD
MR	0.9852	0.9971
BPNN	0.9847	0.9964
BPNN-GA	0.9827	0.9964
BPNN-PSO	0.9856	0.9980

4.5 Further validation

Four additional parameters such as RVI, GRVI, NGI and NRI are also used as input of the network, so the number of nodes in the input layer is 9. The number of nodes in the hidden and output layer remains unchanged.

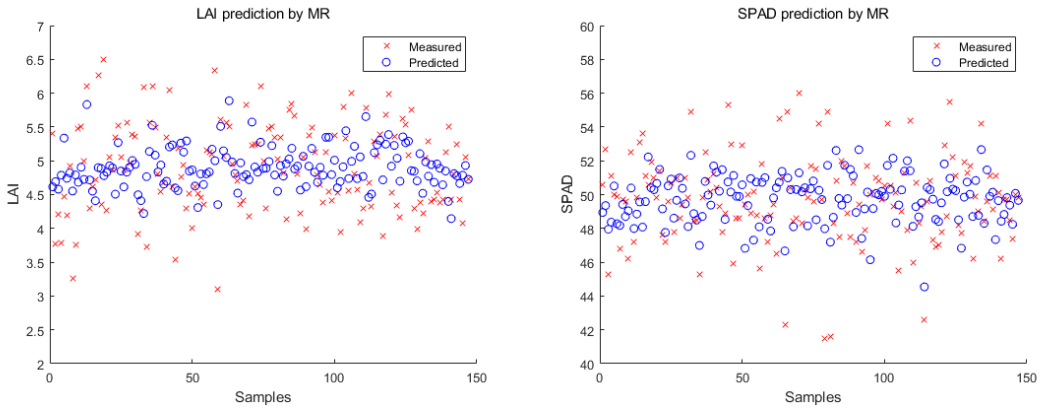
As above-mentioned, there are total 147 sets of data, of which 100 sets are chosen randomly to train the algorithm of BP network (BPNN) and genetic neural network (BPNN-GA) and Particle Swarm Optimisation neural network (BPNN-PSO). The remaining 47 sets of data are validated by the model.

According to the cross-validation method, the above operations were carried out 20 times, and the optimal RMSE is finally selected. Besides, multiple regression (MR) algorithm is used to predict the remaining 47 sets of data directly.

4.5.1 Comparison of Results

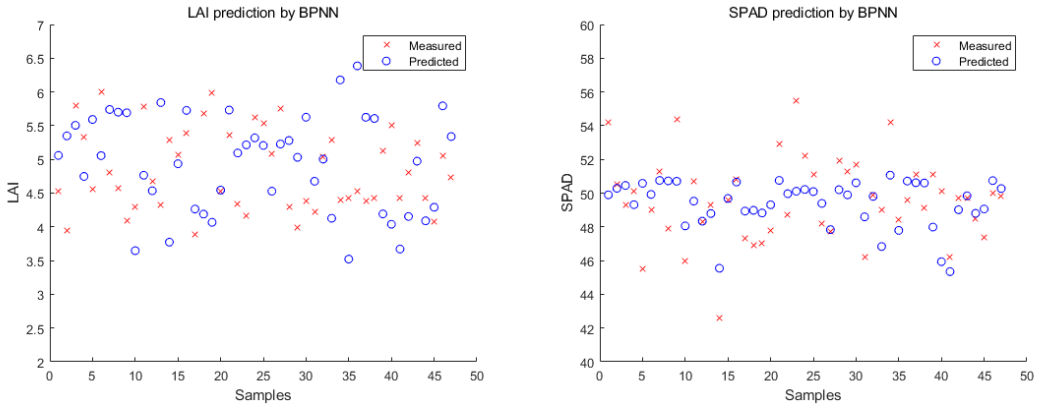
The results of predicting LAI and SPAD by multiple regression (MR) algorithm are shown in Figure 4.10.

The results of predicted LAI and SPAD by BPNN are shown in Fig. 4.11, and the results of predicted LAI and SPAD by BPNN-GA algorithm are shown in Figure 4.12.



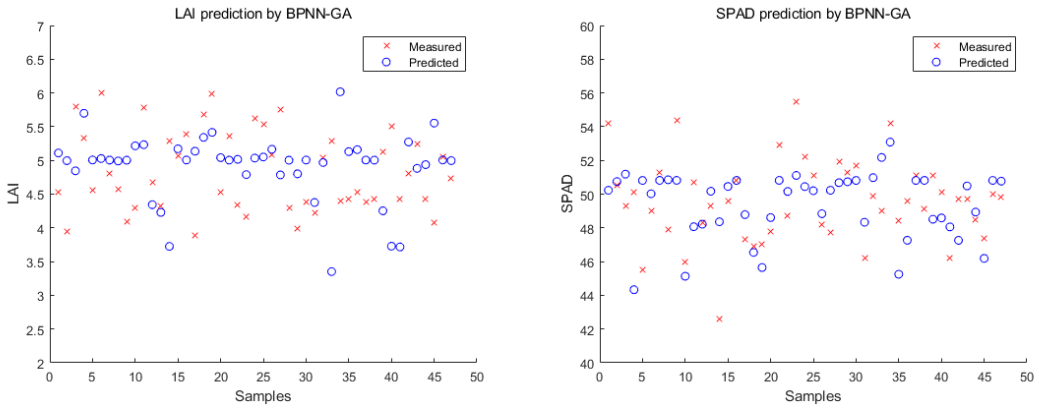
(a) (b)

Figure 4.10: The results of predicting by MR. (a) LAI and (b)SPAD



(a) (b)

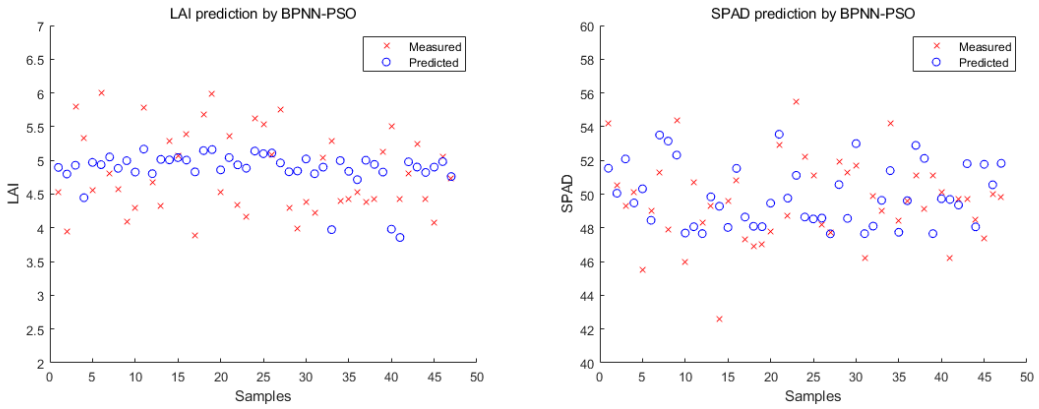
Figure 4.11: The results of predicted by BPNN. (a) LAI and (b)SPAD



(a) (b)

Figure 4.12: The results of predicting by BPNN-GA. (a) LAI and (b)SPAD

The results of predicted LAI and SPAD by BPNN-PSO algorithm are shown in Figure 4.13.



(a) (b)

Figure 4.13: The results of predicting by BPNN-PSO. (a) LAI and (b)SPAD

4.5.2 Results analysis

From the comparison among Figure 8, Figure 9 and Figure 10, it is also shown that the prediction of Multiple Regression and BPNN and BPNN-GA are more discrete than the results from BPNN-PSO.

On the other hand, the predicted result of BPNN-PSO are more convergent and stable, which means that the BPNN-PSO is not affected by the measurement error. The compared results have confirmed that the prediction of BPNN-PSO is more

accurate and stable than the prediction of Multiple Regression, BPNN and BPNN-GA.

RMSE between model output and measured data for different algorithm are as Table 4.4. From Table 4.4, Whether LAI or SPAD prediction, BPNN-PSO has also Minimum RMSE as in Table 4.2. This result also confirms that BPNN-PSO has Minimum prediction error. Also, the PSO algorithm improves the convergence rate to a certain extent.

Compared with Table 4.2, with the increase of nodes of input layer, corresponding RMSE is even less. It means that the increase of network input parameters is helpful to achieve higher prediction accuracy.

Table 4.4: RMSE between model output and measured data

	LAI	SPAD
MR	0.0165	0.0019
BPNN	0.0401	0.0023
BPNN-GA	0.0269	0.0021
BPNN-PSO	0.0141	0.0017

R-square between model output and measured data for different algorithm are as Table 4.5.

Table 4.5: R-square between model output and measured data

	LAI	SPAD
MR	0.9838	0.9981
BPNN	0.9634	0.9977
BPNN-GA	0.9744	0.9979
BPNN-PSO	0.9859	0.9983

In Table 4.5, for both LAI and SPAD prediction, BPNN-PSO has also Maximum R-square as in Table 4.3. This result has confirmed that BPNN-PSO has Minimum

prediction error. Also, the PSO algorithm improves the convergence rate to a certain extent.

4.6 Summary

By reasonably designing the parameters of input layer, hidden layer and output layer of the network, a neural network structure suitable for vegetation prediction has been constructed. After the training process, the trained network can accurately reflect the characteristics of surface vegetation and achieve better prediction results.

The genetic algorithm(GA) and particle swarm optimisation (PSO) algorithm are utilized to optimise the weights and thresholds of the neural network. With the aid of the PSO, the trained network can be better converged and the prediction accuracy has been improved. The prediction results have further verified the feasibility and validity of particle swarm optimised neural network algorithm in the prediction of surface vegetation parameters.

In the case of small sample data-set and uncomplicated distribution, MR algorithm is indeed simple and effective. As shown in Table 4.5, even MR is found to be good enough. In the case of a large number of data and complex distribution, the advantages of the neural network and optimisation algorithm will be reflected. In the case of limited data at present, the prediction accuracy of several algorithms has little difference. From Table 4.2 to 4.5, all the four algorithms are very close. But this does not mean that BPNN, BPNN-GA, and BPNN-PSO algorithms should not be studied anymore. This thesis aims at wheat growth prediction based on agricultural big data. At present, the comparison and validation of several algorithms are only preliminary.

Chapter 5.
Predicting Wheat Yield and Aphid using Ground Spectral Data

5.1 Introduction

When the absolute temperature is above 0K, all objects will emit electromagnetic radiation, absorb and reflect the radiation emitted by other objects. Spectral remote sensing accurately records the change of the interaction between electromagnetic wave and matter with wavelength and provides rich information of surface features by reflecting the difference of interaction. This information is determined by the macro and micro characteristics of surface features.

In recent years, hyperspectral remote sensing has been widely used in many fields. It is a remote sensing technology that uses remote sensing instruments to obtain continuous spectral images of ground objects in a specific spectral region with high spectral resolution (resolution at 3-6nm).

Hyperspectral remote sensing makes remote sensing applications focus on spatial information expansion in the spectral dimension to obtain more fine spectral information and quantitative analysis of the biophysical and chemical processes and parameters on the earth's surface. It is characterized by high spectral resolution and strong band continuity (hundreds of bands in the range of 0.4-2.5 μm). It can not only be used to improve the recognition ability of crops and vegetation types, but also can be used to monitor the growth of crops and retrieve the physical and chemical characteristics of crops, and then combine the physical and chemical characteristics of crops with the yield, and finally reverse the performance of crop yield, and further establish spectral remote sensing yield estimation model.

Hyperspectral remote sensing uses a number of narrow electromagnetic waves, i.e. spectral bands, to produce continuous spectrum. Hyperspectral remote sensing

technology has become the leading technology in the field of remote sensing, which has new characteristics different from traditional remote sensing.

1) Multiple bands: dozens/hundreds or even thousands of bands can be provided;
2) Narrow spectral range: the band range is generally less than 10nm;
3) Continuous band: some sensors can provide almost continuous ground spectrum in the solar spectrum range of 350 ~ 2500nm;
4) Large amount of data: with the increase of waveband number, the amount of data increases exponentially;
5) Increase of information redundancy: due to the high correlation between adjacent bands, the redundant information is also relatively increased.

The remote sensing estimation of crops is an emerging technology developed in recent decades which has the advantages of macroscopic, objective, rapid, economical and large amount of information. Spectral remote sensing has become a powerful tool for crop yield estimation and an effective tool for observing surface vegetation. Spectral remote sensing can provide more detailed spectral information which have a good correlation with vegetation index, chlorophyll content, vegetation coverage and vegetation biomass, while crop yield and crop physical and chemical features also have a good correlation. Therefore, based on the physical and chemical properties of crops, a correlation can be established between crop yield and spectral remote sensing. Finally, the data collected by high spectral resolution remote sensing can be used to reverse the yield of the crop being tested.

Spectral remote sensing technology is one of the most advanced means to monitor the degree of crop diseases and insect pests in the world. It is based on the physiological changes of cell activity, water content and chlorophyll content in green leaves caused by crop diseases and insect pests, which shows the differences in the spectral characteristics of crop reflected spectrum, especially in the visible light area and the shortwave infrared area. Therefore, the occurrence degree of crop diseases and insect pests can be monitored by comparing the spectral characteristics of some characteristic wavelengths of damaged plants with those of healthy plants.

Vegetation index is a remote sensing method to monitor the growth and distribution of plants on the ground. Qiao et al., [129] have analysed the reflective spectral characteristics of winter wheat canopy at the early stage of different degrees of wheat aphid damage in the field. The results showed that the red edge slope of wheat canopy changed dramatically in the near-infrared band after wheat aphid damage, and its value decreased gradually with the aggravation of the damage degree. The normalized vegetation index (NDVI) was used for regression analysis of 100 wheat aphids. Luo et al. [130] determined the spectral characteristics of aphid damage in the field during the filling stage of winter wheat and established a remote sensing inversion model of spectral index and aphid damage level.

5.2 Prediction of wheat yield by ground remote sensing

5.2.1 Samples

The location of experimental site is the same as one described in Figure 4.3. All field spectral measurements are made using Analytical Spectral Devices Fieldspec instruments Pro2500 Spectroradiometer (ASD, Boulder, CO, USA), which gather data between 350–2500 nm on May 7, 2015. For wheat crops, there are a total of 119 data points for which spectral data are available.

The red edge is the most important spectral feature of vegetation features. The position of the red edge is affected by the growth state and physical and chemical parameters of vegetation. However, for normal growing wheat crops, the red edge parameters of the reflectance spectrum of the canopy change with the development of the growth period, and have the following rules:

1). In the early stage of wheat growth, i.e. the vegetative stage of wheat, the red edge will move towards the long wave direction, i.e. the so-called "red shift" phenomenon, with a red shift range of about 10 nm, i.e. from 710 nm to 720 nm; when the wheat enters the reproductive stage, especially the milk ripening stage, the

red edge will move rapidly towards the short wave direction, i.e. the so-called "blue shift", with a moving range of 720nm to 710nm.

2). The change of the position of red edge with the growth period is similar to that of red edge, except that the change of its spectral position is relatively large, especially from rejuvenated to tillering and milk ripening. From turning green to the end of tillering, the position of red valley increased rapidly from about 678 nm to 688 nm. In the milk ripening stage, the position of red valley will decrease rapidly from about 690nm to about 675nm.

3). Red edge width, that is, the difference between the red edge position and the red valley position. In the early stage of wheat vegetative growth, the red edge width is reduced from 34nm to 29nm, which means that the red edge spectrum is more and more steep. From heading stage to milk stage, the width of red edge gradually increased, and the increase was up to 10 nm, that is, from 29 nm to 39 nm.

Obviously, the yield of wheat is closely related to the red edge characteristics of wheat. According to the regularity of wheat growth and development and spectral characteristics, taking the influence of red edge into full consideration, a few new spectral vegetation indexes are analysed based on spectral data, and a few spectral prediction models of wheat yield are established to promote the application of spectral remote sensing technology in wheat growth monitoring and yield estimation.

5.2.2 Model construction

5.2.2.1 Yield prediction model by visible spectral index (VHI model）

Visible waveband spectral data between 400–690nm are total used as visible hyperspectral index to predict the yield of wheat crop, as shown in the Figure 5.1.

5.2.2.2 Yield prediction model by hyperspectral vegetation index (HVI model)

Two-band normalized difference optimal spectral vegetation indexes are computed for biophysical characterization. The HVI is computed using the standard equation:

$$HVI_{ij} = \frac{R_j - R_i}{R_j + R_i} \tag{5.1}$$

where, i and j are the two waveband centers for reflectance.

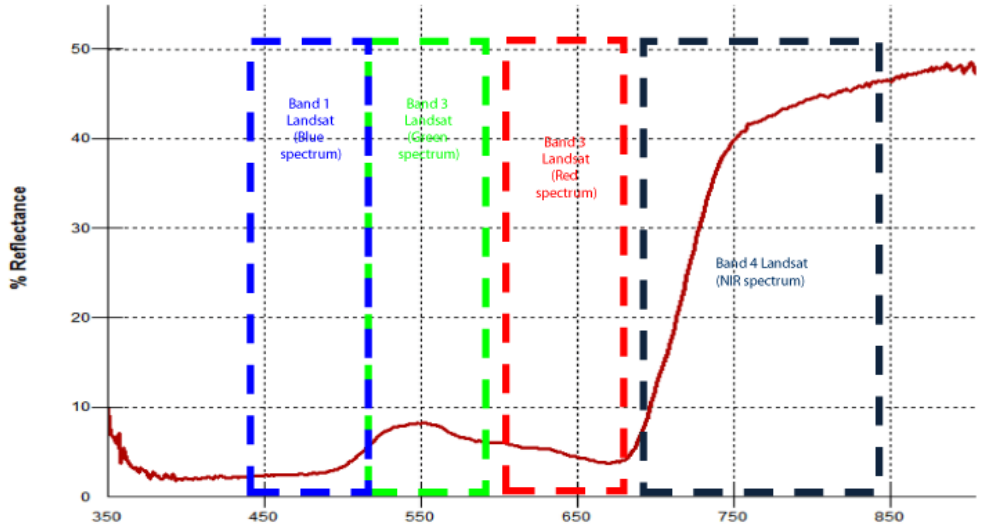


Figure 5.1: Visible waveband spectral data between 400–690nm [131]

Fourteen best two-band HVIs are composed by corresponding spectral narrow bands (i.e. normalized differences) for biophysical characterization of biomass, LAI, plant height, plant density, and grain yield [132].

HVI REDND1=(R855-R687)/(R855+R687)

HVI REDND2=(R855-R650)/(R855+R650)

HVI REDND3=(R760-R687)/(R760+R687)

HVI REDND4=(R760-R650)/(R760+R650)

HVI GREENND1=(R550-R687)/(R550+R687)

HVI GREEDND2=(R550-R650)/(R550+R650)

HVI FNIRND1=(R1045-R687)/(R1045+R687)

HVI FNIRND2=(R1045-R650)/(R1045+R650)

HVI FNIRND3=(R1245-R687)/(R1245+R687)

HVI FNIRND4=(R1245-R650)/(R1245+R650)

HVI SWIRND1=(R1650-R687)/(R1650+R687)

HVI SWIRND2=(R1650-R650)/(R1650+R650)

HVI SWIRND3=(R2205-R687)/(R2205+R687)

HVI SWIRND4=(R2205-R650)/(R2205+R650)

5.2.2.3 Yield prediction model by difference hyperspectral index (DHI model)

In order to analyse the relationship between canopy spectral and wheat yield, the following difference vegetation index is used to predict wheat yield according to the characteristics of canopy spectrum at maturity [133].

$$DHI_i = R_2 - R_1 \tag{5.2}$$

DHI1=R1200-R680

DHI2=R990-R680

DHI3=R800-R680

DHI4=R1200-R550

DHI5=R990-R550

DHI6=R800-R550

DHI7=R1200-R440

DHI8=R990-R440

DHI9=R800-R440

5.2.2.4 Yield prediction model by normalized hyperspectral index (NHI model)

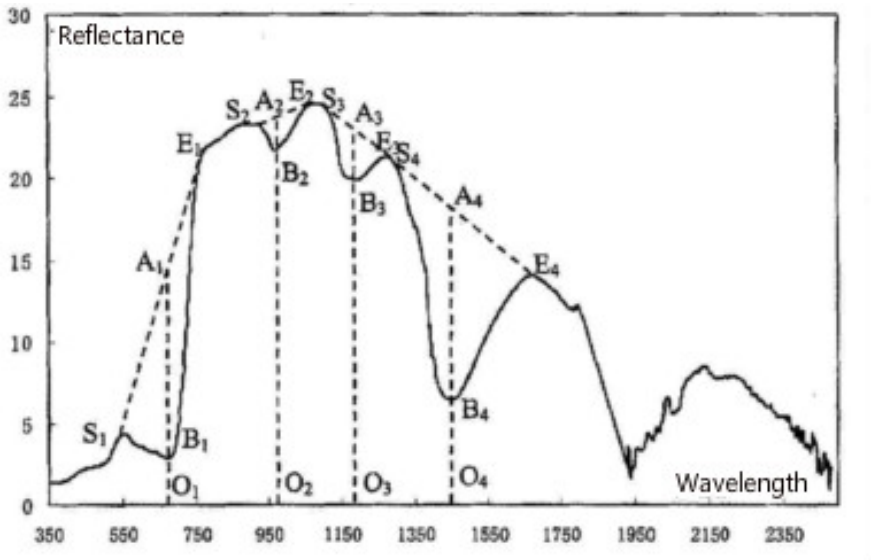


Figure 5.2: Four spectral absorption characteristics of winter wheat canopy in near infrared (NIR) [132].

According to the reflectance spectrum of wheat canopy, four absorption characteristics are selected, as shown in the Figure 5.2.

The position of characteristic spectrum is shown in Table 5.1. According to the reflectance spectrum of wheat canopy, six emission characteristics are selected, as shown in the Figure 5.3.

Table 5.1: The position of characteristic absorption spectrum

No.	central wavelength (nm）	spectral range (nm)
1	670	560-760
2	980	920-1080
3	1190	1120-1280
4	1450	1280-1675

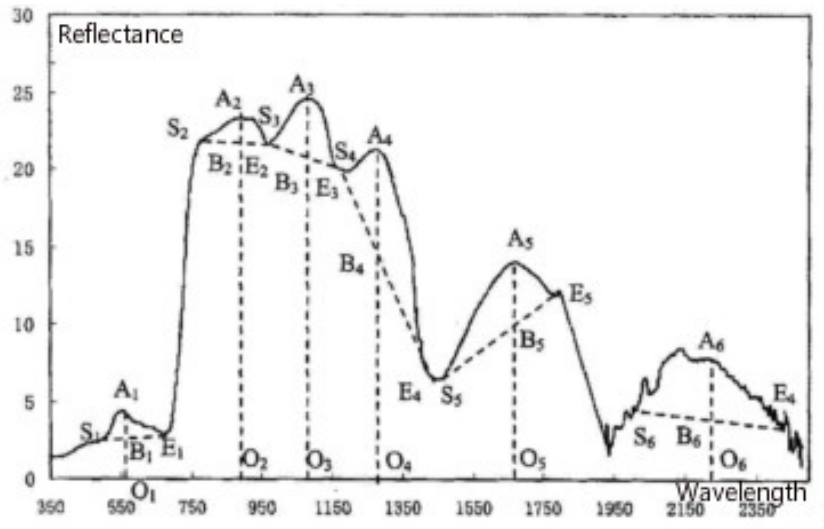


Figure 5.3: Six spectral emission characteristics of winter wheat canopy in NIR [132].

The position of characteristic spectrum is shown in the Table 5.2. According to the above spectral characteristics of wheat crops, we have defined 7 kinds of normalized spectral indexes, as shown in the Table 5.3 [134]. The normalized spectral index is defined as follows:

$$NHI_i = \frac{\left|R_i(B1) - R_i(B2)\right|}{R_i(B1) + R_i(B2)} \tag{5.3}$$

where *i* is the spectral index; B1 and B2 are the corresponding two wavelengths.

Table 5.2: The position of characteristic emission spectrum

No.	central wavelength (nm）	spectral range (nm)
1	560	500-670
2	920	780-980
3	1100	980-1200
4	1280	1200-1480
5	1690	1480-1780
6	2230	2000-2400

Table 5.3: Normalized spectral indexes

No.	1	2	3	4	5	6	7
B1(nm)	560	670	890	920	857	820	820
B2(nm)	670	890	980	980	1210	1650	2200
Function	Vegetation green peak	Vegetation red edge	Water index		Water index		

5.2.3 Network structure design

A three-layer multi-input and two-output with a hidden layer is adopted to establish a prediction BP network to train is above 4 models. As discussed in 4.3.2, a common empirical formula is adopted for determining the numbers of nodes in the hidden layer. $h = \sqrt{m+n} + a$, where h is the number of hidden layer nodes, m and n are the number of input layer and output layer nodes respectively, and a is the adjustment constant between 1 and 10.

For the VHI model, visible waveband spectral data between 400–690nm are used as the input and the yield as output, so the number of nodes in the input layer is 291,

the number of nodes in the output layer is 1, and the number of nodes in the hidden layer is 17.

For the HVI model, 14 two-band normalized difference optimal spectral vegetation indexes are used as input and the yield as the output, so the number of nodes in the input layer is 14, the number of nodes in the output layer is 1, and the number of nodes in the hidden layer is 7.

For the DHI model, 9 different hyperspectral indexes are used as the input and the yield as the output, so the number of nodes in the input layer is 9, the number of nodes in the output layer is 1, and the number of nodes in the hidden layer is 6.

For the NHI model, 7 normalized hyperspectral indexes are used as the input and the yield as the output, so the number of nodes in the input layer is 7, the number of nodes in the output layer is 1, and the number of nodes in the hidden layer is 6.

The neural network toolbox in MATLAB is used to train the network. The training sample data are normalized and input into the network. The transfer functions of input layer, output layer and hidden layer are set as tansig, tansig and purelin function (default) respectively. The training function is set as trainglm function (default). The number of network iterations epochs is 5000, the expected error goal is 0.00000001, and the learning rate LR is 0.01.

5.2.4 Results and analysis

The above four models are used to predict the yield of wheat using the BP network. There are total 119 sets of data, of which 80 sets are used to train the algorithm of BP network. The remaining 39 sets of data are validated by the model.

5.2.4.1 Results comparison

1. VHI model

For VHI model, 291 measured spectral data are used as input of the network, and wheat yield is used as output of the network. The result of predicting wheat yield by VHI model with BPNN algorithm is shown in Figure 5.4.

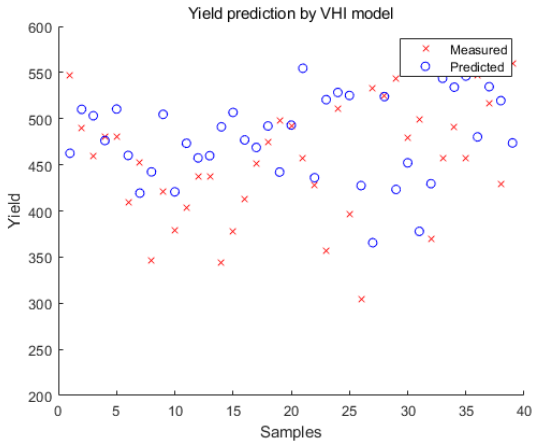


Figure 5.4: The result of yield predicting by VHI model

2. HVI model

For HVI model, fourteen indexes are used as input of the network, and wheat yield is used as output of the network. The result of predicting wheat yield by HVI model with BPNN algorithm is shown in Figure 5.5.

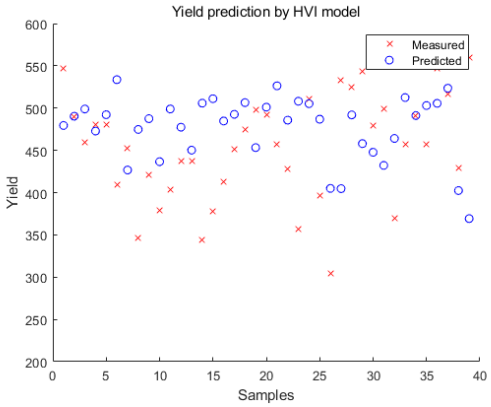


Figure 5.5: The result of yield predicting by HVI model

3. DHI model

For DHI model, nine indexes are used as input of the network, and wheat yield is used as output of the network. The result of predicting wheat yield by DHI model with BPNN algorithm are shown in Figure 5.6.

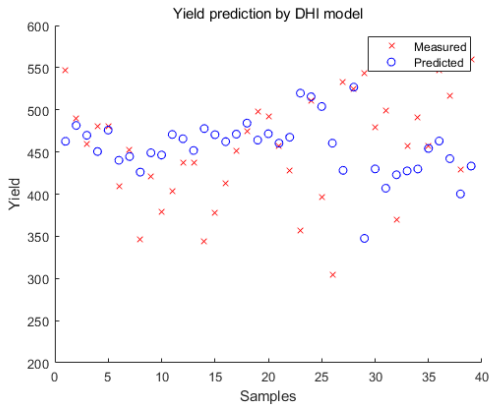


Figure 5.6: The result of yield predicting by DHI model

4. NHI model

For NHI model, seven indexes are used as input of the network, and wheat yield is used as output of the network. The result of predicting wheat yield by NHI model with BPNN algorithm is shown in Figure 5.7.

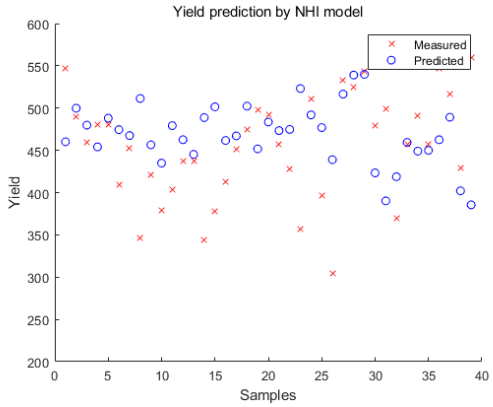


Figure 5.7: The result of yield predicting by NHI model

5.2.4.2 Results analysis

From the comparison among Figure 5.4, Figure 5.5, Figure 5.6, and Figure 5.7, which show the contrast between predicted data and measured data for yield prediction by four models, it leads to the conclusion that the predicted data of NHI model are more convergent and stable, which means that obviously this model is not affected by the measurement error has the most accurate predicting results. RMSE between model output and measured data for different model are as Table 5.4.

Table 5.4: RMSE and R-square between model output and measured data

	RMSE	R-square
VHI model	0.0352	0.9674
HVI model	0.0349	0.9704
DHI model	0.0386	0.9670
NHI model	0.0282	0.9737

From Table 5.4, it is shown that NHI model has Minimum RMSE and Maximum R-square. This result also confirms that NHI model has Minimum prediction error.

5.3 Comparison of yield predicting by different algorithm

NHI model is used to predict the yield of wheat by the algorithms of BPNN, BPNN-GA and BPNN-PSO respectively. The network structure is the same as one for NHI model in 5.2.3. There are total 119 sets of data, of which 80 sets are chosen randomly to train the algorithms. The remaining 39 sets of data are validated by the model. According to the cross-validation, the above experiments are repeated 20 times, and the optimal RMSE is finally selected for each algorithm.

5.3.1 Comparison of Results

1. BPNN:

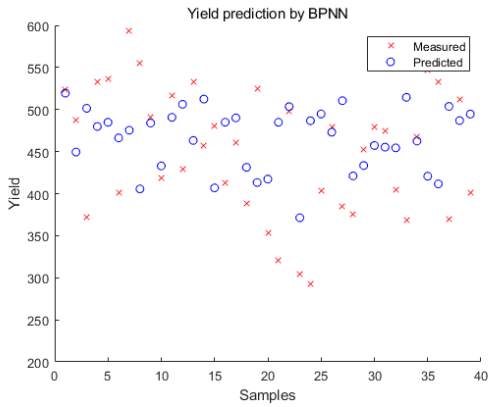


Figure 5.8: The result of predicting by BPNN

The results of predicted wheat yield by BPNN algorithm are shown in Figure 5.8.

2. BPNN-GA

The results of predicted wheat yield by BPNN-GA algorithm are shown in Figure 5.9.

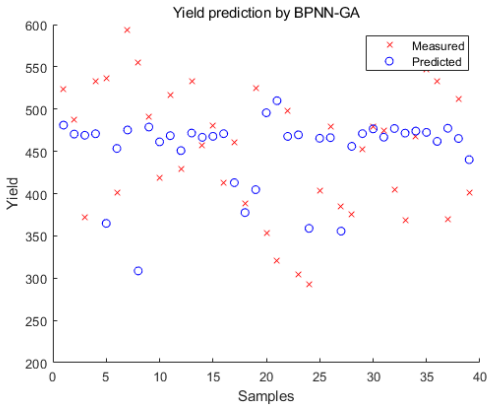


Figure 5.9: The result of predicting by BPNN-GA

3. BPNN-PSO

The results of predicted wheat yield by BPNN-PSO are shown in Figure 5.10.

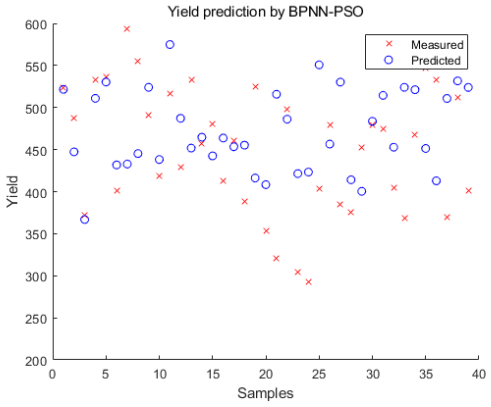


Figure 5.10: The result of predicting by BPNN-PSO

5.3.2 Results analysis

From the comparison among Figure 5.8, Figure 5.9 and Figure 5.10, the comparison between predicted data and measured data for yield prediction are shown by BPNN, BPNN-GA and BPNN-PSO respectively. It leads to the conclusion that the prediction of BPNN and BPNN-GA are more discrete than prediction of BPNN-PSO. The results of BPNN and BPNN-GA are more erroneous.

On the other hand, the predicted data of BPNN-PSO are more convergent and stable, which means that the BPNN-PSO is not affected by the measurement error. When the measured data has a significant change, the predicted data of BPNN-PSO will have more logical change rather than a jump in prediction like BPNN and BPNN-GA. The large amount of data will contain more measuring error, which will make the BPNN-PSO better than BPNN and BPNN-GA. RMSE between model output and measured data for different algorithm are as given in Table 5.5.

Table 5.5: RMSE and R-square between model output and measured data

	RMSE	R-square
BPNN	0.0175	0.9822
BPNN-GA	0.0150	0.9849
BPNN-PSO	0.0147	0.9858

As shown in Table 5.5, BPNN-PSO has Minimum RMSE and Maximum R-square. This result also confirms that BPNN-PSO algorithm has Minimum prediction error.

5.4 Prediction of wheat aphid infection by ground remote sensing

5.4.1 Methodology

In order to further optimise the parameters of neural network, two new algorithms such as ant colony optimisation algorithm and cuckoo search algorithm are introduced.

5.4.1.1 Ant colony optimisation

The Ant colony optimisation algorithm (ACO) is inspired by the mechanism of biological evolution in the early 1990s by Italian scholars Dorigo M and Maniezzo V [135][136].

A new simulated evolutionary algorithm is realized by simulating the behavior of ants searching path in nature. This algorithm can solve some difficult problems in system optimisation by using the optimisation ability of ant colony in the process of searching food source.

The basic idea of ant colony algorithm is to imitate ant's dependence on pheromone and guide each ant's action by positive feedback between ants. The feasible solution of the problem to be optimised is represented by the ant's walking path, and all the paths of the whole ant colony constitute the solution space of the problem to be optimised. In the process of movement, ants will leave a kind of substance called pheromone on the path they pass. The ants behind can choose the path according to the pheromone left by the ants in front. The more ants they pass on the path, the stronger the pheromone they leave behind, the greater the probability that the ants behind will choose it, and the shortest path to food through this information exchange between ants.

The ant colony optimisation algorithm just uses this optimisation mechanism for reference. It represents the solution of the problem as an ordered sequence of points, and the ants move away from the adjacent nodes to release pheromones, which guides the ants to select the next adjacent node, and gradually build a feasible solution, and finally find the optimal solution through the information exchange and cooperation between individuals.

Ant colony optimisation algorithm has the following characteristics:

1). The positive feedback mechanism is used to make the search process continually converge and finally approach the optimal solution;

2). Each individual may change the surrounding environment by releasing pheromones, perceive the real-time changes of the surrounding environment, and communicates indirectly through the environment;

3). The search process adopts the distributed computing mode, and multiple individuals carry out parallel computing at the same time, which greatly improves the computing power and operation efficiency of the algorithm;

4). The heuristic probability search method is not easy to fall into the local optimum, and easy to find the optimal solution.

The basic idea of training neural network weights and thresholds with the ant colony optimisation algorithm is as follows:

1). Assume that there are m parameters in the neural network, including all weights and thresholds. These are sorted as P1, P2, PM, for parameter Pi($1 \leq i \leq m$), set it to N random non-zero values within the possible value range to form a set.

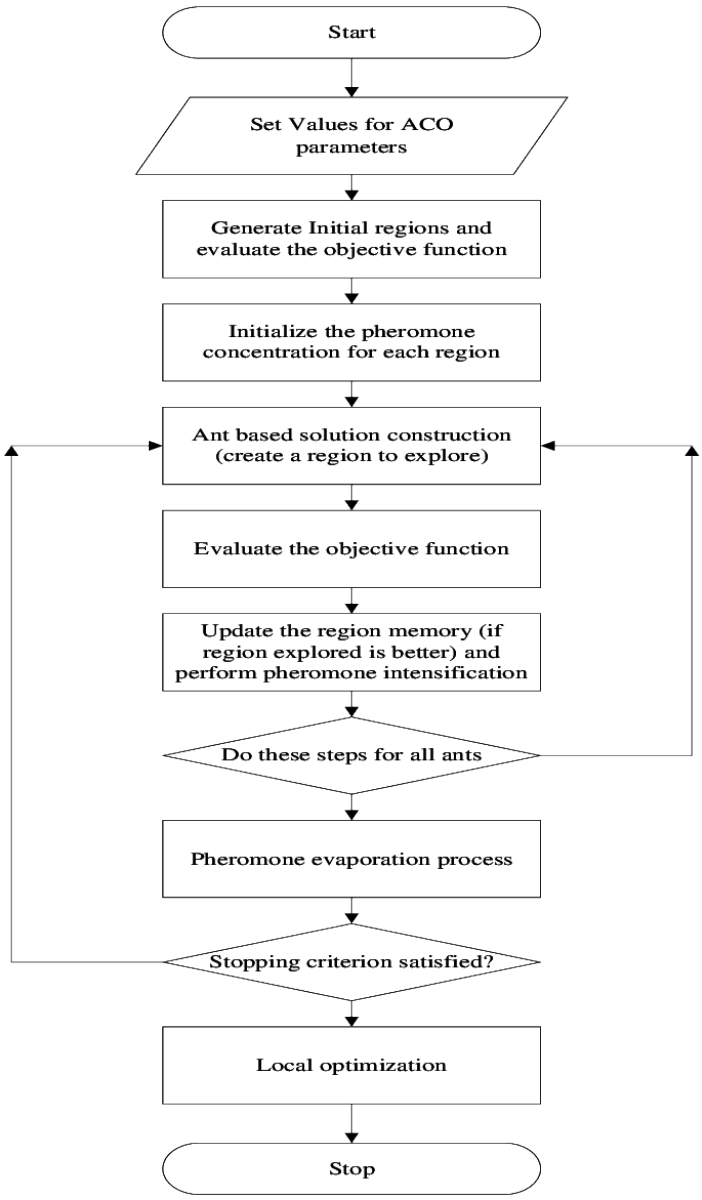


Figure 5.11: Basic flow chart of BP network trained by ACO algorithm

2). Define the number of ants as K. these ants start from the nest to find food. Each ant starts from the first set, randomly selects one element from each set according to the pheromone state of each element in the set, and adjusts the pheromone of the selected element accordingly.

3). When an ant completes the selection of elements in all sets, it reaches the food source, returns to the ant nest according to the original path, and adjusts the pheromone of the selected elements in the set.

4). This process is repeated. When all ants converge to the same path, it means that the optimal solution of network parameters is found.

5.4.1.2 Cuckoo search algorithm

Cuckoo search (CS) algorithm is a new heuristic optimisation algorithm proposed in 2009 by Yang and Deb [137], famous scholars of Cambridge University, based on the propagation characteristics and flight mechanism of cuckoo. The advantages of the algorithm are: strong global search ability, fast convergence speed, less parameters, better universality and robustness.

Cuckoo search algorithm is evolved through cuckoo's special breeding mode and flight mechanism. The cuckoo lays its eggs in the nest of other host birds for hatching. If the cuckoo's behavior is found by the host bird on the spot, there will be a fierce conflict. If the host finds that the eggs are not its own after the cuckoo lays its eggs, it will give up the eggs or nest directly. Therefore, cuckoos usually choose the nest where other birds have just laid their eggs to lay their eggs. Once the eggs of cuckoos are preserved, the eggs of cuckoos usually hatch before the rest. The new pups have the instinct to push the rest of eggs out of nests, so that the host bird will raise the cuckoo.

Cuckoo search algorithm is based on two strategies: nest parasitism and Levy flight mechanism. Cuckoos can find an optimal nest to hatch their eggs by random walk, which can achieve an efficient optimisation mode. This algorithm can be enhanced by so-called Levy flight, rather than simple isotropic random walk.

Yang's cuckoo algorithm is based on the following three assumptions:

1). Each cuckoo lays only one egg at a time, and randomly selects the nest to lay eggs at the same time;

2). The high-quality eggs in the best nest will be preserved and hatched to the next generation;

3). The probability of the host finding the alien egg. Once found, the host will give up the egg or nest directly.

Assuming that the number of bird's nests available n is fixed, the probability that the owner of a bird's nest can find an exotic egg $Pa \in [0,1]$, the path and position updating formula of cuckoo nest searching is as follows:

$$X_i(t+1) = X_i(t) + \alpha * L(\lambda) \qquad i = 1, 2, \ldots, N \tag{5.4}$$

where $X_i(t)$ is the nest position of the ith nest in the tth generation, * is the point-to-point multiplication, α is the step control quantity, $L(\lambda)$ is the levy random search path, and $L \sim u = t^{-\lambda}$, $(1 < \lambda \leqslant 3)$. After the position is updated, the random number $r \in [0,1]$ is compared with Pa. If r > Pa, $X_i(t+1)$ will be changed randomly, otherwise it will not change. At last, the nest position $X_i(t+1)$ with better test value was reserved.

The following is the specific steps of Cuckoo algorithm to optimise BP network:

1). Randomly generate N nests in a given space. Each nest represents a group of weights and thresholds of neural networks to be optimised. The parameters in the algorithm are set, optimised and trained, and calculated according to the fitness function (fitness function is the error evaluation index) to find the current optimal nest location.

2). Keep the optimal nest position of the previous generation and update the N nest position according to Levy flight mode. At the same time, the fitness value of the updated nest is calculated and compared with that of the previous generation. If it is better, the location will be updated. If it is not better, the location of the previous generation nest will still be retained.

3). The random number k is generated and compared with Pa to update all nest positions. If k > Pa, the original nest position will be kept; if k < Pa, the nest position will be updated with a random step. Comparing the new location with the original nest location, if it is better, the new nest will be kept, if it is not better, the original nest will still be used, and finally the updated n nest locations will be obtained.

4). The best bird's nest with many iterations is regarded as the optimal weight and threshold of neural network.

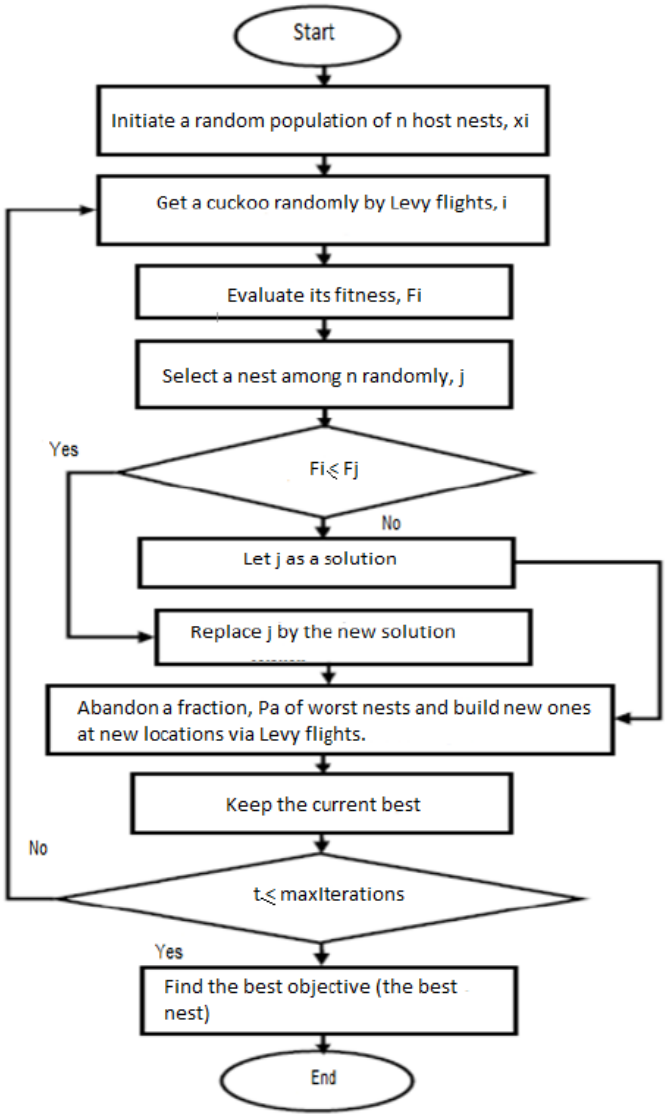


Figure 5.12: Basic flow chart of BP network trained by CS algorithm

5.4.2 Samples

In the peak period of wheat aphid (wheat flowering and filling stage), the relationship between the amount of aphid and the content of nitrogen and chlorophyll has been studied by estimating the spectrum of leaves with different amount of aphid. The spectral characteristics and mechanism of the leaves harmed by aphid were studied. The sensitive spectrum feature set of wheat aphid was extracted, and the estimation model based on the amount of aphid in leaves was constructed.

The experimental field was located at north latitude 39056', and East longitude 116016'. All field spectral measurements were made using ASD FieldSpec Pro FR spectroradiometer, which gather data between 350–2500 nm on May 13, 2011. For wheat crops, there were a total of 74 data points for which spectral data were available.

The process of aphid counting and spectral data collection is as follows:

According to the occurrence of aphids in the experimental field, the number of aphids collected (the number of aphids on each leaf sample) ranges from 0 to 120. In order to ensure the typicality and representativeness of the experimental samples, the interval of aphids selected is about 5, and the number of samples in each interval is equal.

According to the amount of aphid in the visual inspection, select the wheat leaves suitable for the experimental purpose, use scissors to cut off the leaves gently, count the aphid amount on the leaves, and record the sample number and the corresponding aphid amount. After counting, use a fine brush to gently sweep away the aphid in the leaves, and immediately measure the spectrum of the leaves.

5.4.3 Model construction

In order to analyse the relationship between canopy spectral and wheat aphid, the following nine spectral indexes which have potential in detecting plant stresses such as crop diseases, drought and insufficient of nitrogen are used to predict wheat aphid according to the characteristics of canopy spectrum at maturity[129].

1) Aphid Index(API)

   AI = (R740-R887)/(R691-R698)

2) Damage Sensitive Spectral Index1(DSSI1)

   DSSI1 = (R719-R873-R509-R537)/(R719-R873+R509-R537)

3) Photochemical Refledtance Index (PRI)

   PRI = (R731-R570)/(R731+R570)

4) Normalized Difference Water Index (NDWI)

   NDWI = (R860-R1240)/(R860+R1240)

5) Green Normalized Difference Vegetation Index (GNDVI)

   GNDVI = (R747-R537)/(R747+R537)

6) Narrow-Band Normalized Difference Vegetation Index (NBNDVI)

   NBNDVI = (R850-R680)/(R850+R680)

7) Structure Insesitive Pigment Index (SIPI)

   SIPI = (R800-R450)/(R800-R680)

8) Modified ChlorophyII Absorption Reflectance Index (MCARI)

   MCARI = [(R700-R670)-0.2(R700-R550)]*(R700/R670)

9) Red-edge Vegetation Stress Index (RVSI)

   RVSI = [(R712+R752)/2]-R732

5.4.4 Network structure design

A three-layer multi-input and two-output with a hidden layer is adopted to establish a prediction BP network to train the above model. As in 4.3.2, a common empirical formula is adopted for determining the number of nodes in the hidden layer. $h = \sqrt{m+n} + a$, where h is the number of hidden layer nodes, m and n are the numbers of the input layer and output layer nodes respectively, and the parameter a is the adjustment constant between 1 and 10.

A nine-parameters model is used to predict the aphid of wheat using the algorithm of BPNN，BPNN-GA，BPNN-PSO, BPNN-ACO and BPNN-CS. The number of nodes in the input layer is 9, the number of nodes in the output layer is 1, and the

number of nodes in the hidden layer is 6.

The neural network toolbox in MATLAB is used to train the network. The training sample data are normalized and input into the network. The transfer functions of input layer, output layer and hidden layer are set as tansig, tansig and purelin function (default) respectively. The training function is set as trainglm function (default). The number of network iterations epochs is 5000, the expected error goal is 0.00000001, and the learning rate LR is 0.01.

5.4.5 Results and analysis

There are in total 74 sets of data, of which 50 sets are chosen randomly to train five algorithms respectively. The remaining 24 sets of data are validated by the model. According to the cross- validation method, the above operations were carried out 20 times, and the optimal RMSE was finally selected.

5.4.2.1 Comparison of Results

The results of predicted wheat aphid by BPNN, BPNN-GA, BPNN-PSO, BPNN-ACO and BPNN-CS (e) are shown in Figure 5.13.

From the comparison among Figure 5.13, Figure 5.14, Figure 5.15, Figure 5.16 and Figure 5.17, which show the contrast between predicted data and measured data for yield prediction by BPNN, BPNN-GA, BPNN-PSO, BPNN-ACO and BPNN-CS respectively. It leads to the conclusion that the prediction of BPNN and BPNN-GA are more discrete than prediction of BPNN-PSO. Obviously, the prediction data of BPNN and BPNN-GA have more error.

On the other hand, the predicted data of BPNN-CS are more convergent and stable, which means that obviously the BPNN-CS is not affected by the measurement error. When the measured data has a significant change, the predicted data of BPNN-CS will have more logical change rather than a jump in prediction like other algorithms.

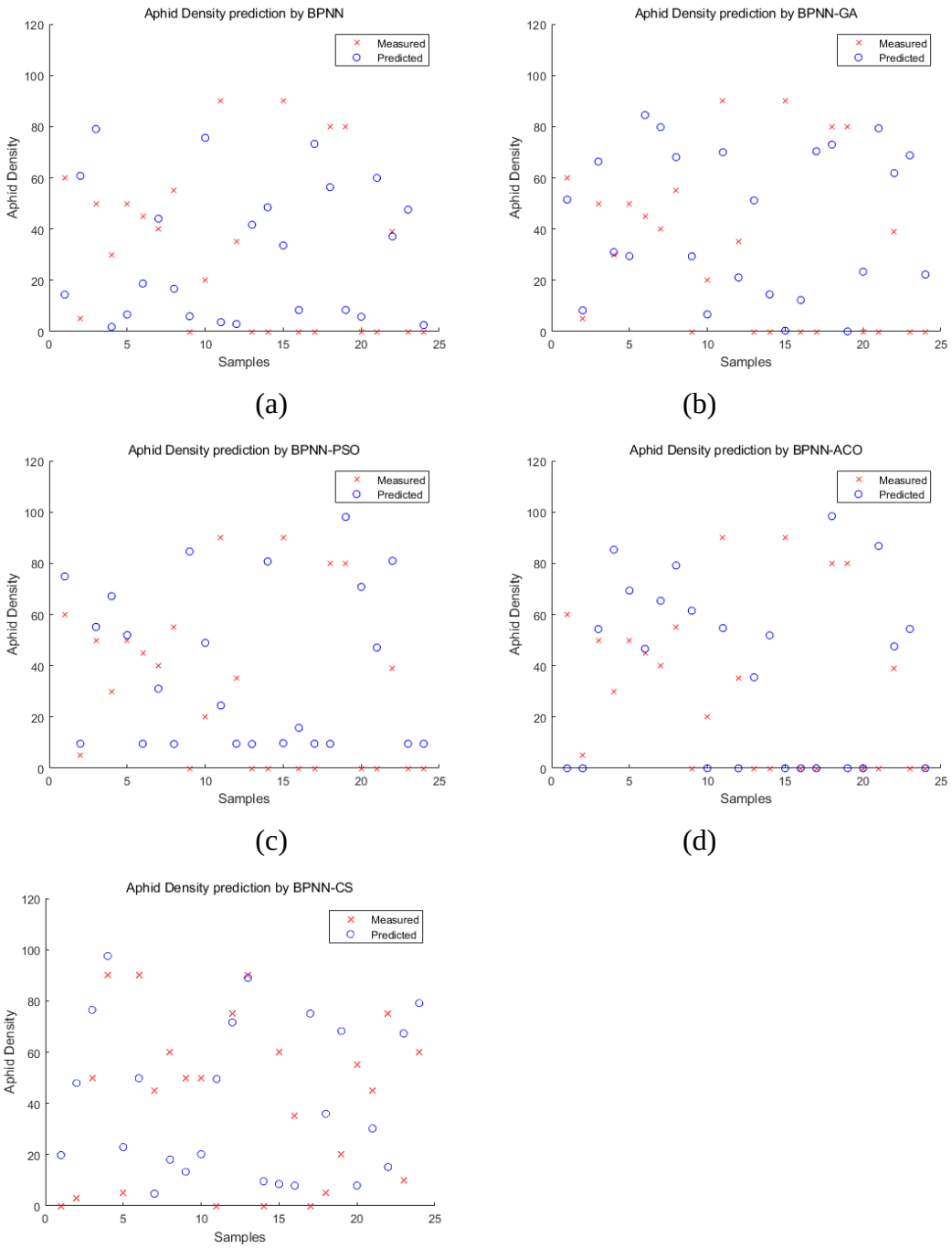


Figure 5.13: The result of predicted Aphid density from BPNN (a), BPNN-GA (b), BPNN-PSO (c), BPNN-ACO (d), and BPNN-CS (e).

The large amount of data will contain more measuring error, which will make the BPNN-CS better than other algorithms. RMSE between model output and measured data for different algorithm are as Table 5.6.

Table 5.6: RMSE and R-square between model output and measured data

	RMSE	R-square
BPNN	0.2203	0.7568
BPNN-GA	0.2352	0.7716
BPNN-PSO	0.2643	0.7172
BPNN-ACO	0.2562	0.7515
BPNN-CS	0.2159	0.8339

From table 5.6, it is shown that BPNN-CS has Minimum RMSE and Maximum R-square. This result also confirms that BPNN-CS algorithm has Minimum prediction error.

5.5 Conclusion

Four spectral index models were compared to predict wheat yield by the algorithm of BP network. For the optimal NHI model, three regression algorithms such as back propagation network (BPNN) algorithm, genetic algorithm (GA) optimised BPNN algorithm and particle swarm optimisation (PSO) optimised BPNN algorithm, were compared to predict wheat yield.

By reasonably choosing the parameters of input layer, hidden layer and output layer of the network, a BP network structure suitable for aphid prediction was constructed. BP network algorithm and other optimisation algorithm were used to optimise the weights and thresholds of the neural network. Through learning and training of the network, the trained network can accurately reflect the characteristics of aphid and achieve better prediction results. As a comparison, BPNN-CS algorithm has the best prediction result.

Chapter 6.

Conclusions and Future Work

6.1 Conclusions

The general objective of this thesis is to carry out the application of spectral data and machine learning in big data precision agriculture. These include various algorithms and models for wheat growth prediction, yield assessment and aphid validation as presented in Chapters 4 and 5 respectively.

The main contributions of the thesis are summarized as follows.

1) In Chapter 4, multiple vegetation indexes related to wheat growth based on the TM 1-4 image band data of Landsat satellite were used, where four regression algorithms including the multiple linear regression (MR) algorithm, back propagation neural network (BPNN) algorithm, genetic algorithm (GA) optimised BPNN algorithm and particle swarm optimisation (PSO) optimised BPNN algorithm were used to train the models for predicting the LAI and SPAD.

   Firstly, five vegetation index parameters of R, G, B, NIR of TM data and NDVI obtained by operation were used as network input, LAI and SPAD parameters measured on the ground were used as the output of the algorithm. Secondly, four additional vegetation index parameters including RVI, GRVI, NGI and NRI were also used as algorithm input for prediction and comparison.

   To improve algorithm training, cross-validation method was used for each algorithm to obtain reliable and stable models for the prediction. Through repeated cross-validation, the loss function was used to measure the quality of the obtained model, and finally a better model could be obtained. The predicted results were assessed for each algorithm with data measured through hand-held instruments on the ground by adjusted MSE and un-centered R-square as accurate evaluation criteria for practical application. The RMSE and R-square results of all algorithms were compared and analysed.

It had shown that BPNN-PSO algorithm had minimum RMSE and maximum R-square for LAI or SPAD prediction no matter five or nine vegetation index parameters were used. In other words, BPNN-PSO algorithm has minimum prediction error.

2) In Chapter 5, by means of spectral data, four kinds of vegetation indexes models related to wheat yield such as visible hyperspectral index (VHI) model, hyperspectral vegetation index (HVI) model, difference hyperspectral index (DHI) model and normalized hyperspectral index (NHI) model were used along with the BP networks to train the models and predict wheat yield.

For the optimal NHI model, three regression algorithms including the back propagation network (BPNN) algorithm, genetic algorithm (GA) optimised BPNN algorithm and particle swarm optimisation (PSO) optimised BPNN algorithm, were used to predict wheat yield. Similarly, cross-validation method was repeated 20 times for each algorithm to improve the prediction accuracy.

The relationship between spectral data and wheat aphid was investigated. Firstly, the basic framework of the BP network was introduced. The input of the network was the vegetation parameters based on spectral data, and the output was the measured amount of aphid. Nine vegetation indexes related to wheat aphid were analysed. Five regression algorithms including the back propagation network (BPNN) algorithm, genetic algorithm (GA) optimised BPNN algorithm, particle swarm optimisation (PSO) optimised BPNN algorithm, ant colony (ACO) optimisation algorithm optimised BPNN algorithm and cuckoo search (CS) optimised BPNN algorithm, which were used to train the model and predict wheat aphid.

Cross-validation was also applied and repeated 20 times for each algorithm to improve prediction accuracy. Through learning and training of the network, better prediction results can be achieved to accurately reflect the characteristics of aphid. The RMSE and R-square results of each algorithm were also compared and analysed. As a comparison, BPNN-CS algorithm has the best prediction result.

6.2 Future work

Followed by the presented results and conclusions, the directions for future work are summarised as follows.

1) The difficulty of applying satellite remote sensing to agriculture is that it is not easy to ensure that the data can be acquired at the requested time period. Taking TM image as an example, its orbit repetition period is 16 days, which is difficult to track short-term changes of crop growth. Obviously, the satellite image with shorter playback time is more conducive for the study of crop growth. In addition, the spatial resolution of TM image is 30 meters, which is suitable for the study of regional scale remote sensing monitoring. For crop growth prediction at canopy scale, remote sensing images with a higher spatial resolution should be used. Finally, the amount of remote sensing data and the amount of ground measured data in this thesis is relatively small, and the evaluation of the model and algorithm is preliminary and can be further improved in the future.

   The next step should be to carry out more in-depth research. Through multi-source remote sensing platforms such as satellite image, UAV spectral image and field portable monitoring, all-round crop growth data can be collected. On this basis, more crop growth models and advanced regression algorithms can be developed for more accurate crop growth and yield prediction.

2) Diseases and insect pests are the natural enemies of agricultural production, which not only reduce the yield of crops, but also greatly reduce the quality of crops. Remote sensing can effectively investigate and monitor the occurrence of diseases and insect pests with various control measures. In this thesis, the single leaf scale remote sensing monitoring of wheat aphid was carried out, using collected data of wheat aphid amount and spectral data for regression analysis and machine learning based prediction. Due to the relatively small sample size, although the prediction model of aphid damage grade established by five regression algorithms have achieved relatively good accuracy, the applicability of the model needs to be further assessed and validated.

In the future, we will use different varieties of winter wheat to further improve the prediction models of aphid amount, so as to enhance its applicability and stability. On the other hand, we only obtained the spectrum of winter wheat in the peak period of aphid occurrence, and did not obtain the spectrum data of winter wheat in the initial period and rising period of aphid. As a result, we will further analyse the characteristics of aphid damage changes in the whole growth period of winter wheat. In addition, wheat powdery mildew and stripe rust are also common diseases of wheat, which are also worthy of future study.

References

[1] A. C. Tyagi, Towards a Second Green Revolution, Irrigation and Drainage, 65(4): 388-389, 2016.

[2] R. Gebbers, and V. I. Adamchuk, Precision agriculture and food security, Science,327(5967):828-831, 2010.

[3] FAO. 2009. How to Feed the World in 2050. Rome: Food and Agriculture Organization of the United Nations.

[4] A. Kamilaris, A. Kartakoullis and F. X. Prenafeta-Boldú, A review on the practice of big data analysis in agriculture, Computers and Electronics in Agriculture, 143(1):23-37, 2017.

[5] A. Kamilaris, F. Gao, F. X. Prenafeta-Boldu and M. I. Ali, Agri-IoT: A semantic framework for Internet of Things-enabled smart farming applications, 3rd World Forum on Internet of Things (WF-IoT), 442-447, 2016.

[6] W. G. M. Bastiaanssen, D. J. Molden and I. W. Makin., Remote Sensing for Irrigated Agriculture: Examples from Research and Possible Applications, Agricultural Water Management, 46(2): 137-155, 2000.

[7] The STARS Project, a research project which is looking for ways to use remote sensing technology to improve agricultural practices in Sub-Saharan Africa and South Asia, Supported by the Bill & Melinda Gates Foundation, Netherlands, 2014.

[8] F. Dodds, and J. Bartram, The Water, Food, Energy and Climate Nexus: Challenges and an Agenda for Action, Routledge, 2016.

[9] FAO. 2015. The state of food insecurity in the world. meeting the 2015 international hunger targets: Taking stock of uneven progress.

[10] UN General Assembly, Transforming our world: the 2030 Agenda for Sustainable Development, A/RES/70/1, 21 October 2015.

[11] K. Liakos, P. Busato, D. Moshou, S. Pearson and D. Bochtis, Machine Learning in Agriculture, A Review. Sensors, 18(8):2674, 2018.

[12] Arthur L. Samuel, Some Studies in Machine Learning Using the Game of

Checkers, IBM Journal of Research and Development, 3: 210-229,1959.

[13] M. J. C. Hu, Application of the adaline system to weather forecasting, Master Thesis, Technical Report 6775-1, Stanford Electronic Laboratories, Stanford, CA, June, 1964.

[14] D. E. Rumelhart, G. E. Hinton and R. J. Williams, Learning representations by back-propagating errors, Nature, 323:533–536, 1986.

[15] P. J. Werbos, Generalization of back propagation with application to a recurrent gas market model, Neural networks, 1(4):339–356, 1988.

[16] B. George, J. Gwilym, Time Series Analysis: Forecasting and Control, San Francisco: Holden-Day, 1970.

[17] Z. Tang, C. Almeida, P. A. Fishwick, Time series forecasting using neural networks vs Box-Jenkins methodology, Simulation, 57 (5):303-310, 1991.

[18] Z. Tang, P. A. Fishwick, Feedforward neural nets as models for time series forecasting, ORSA Journal on Computing, 5 (4) :374-385, 1993.

[19] G. Q. Zhang, B. E. Patuwo, M. Y. Hu, Forecasting with artificial neural networks, The state of the art. International Journal of Forecasting 14:35–62, 1998.

[20] J. Koscak, R. Jaksa, R. Sepesi, P. Sincak, Weather forecast using Neural Networks, 9th Scientific Conference of Young Researchers, 2009.

[21] M. P. Naeini, H. R. Taremian, H. B. Hashemi, Stock market value prediction using neural networks, IEEE 2010 International Conference on Computer Information Systems and Industrial Management Applications (CISIM) - Krackow, Poland (2010.10.8-2010.10.10), 2010.

[22] B. Ji, Y. Sun, S. Yang and J. Wan, Artificial neural networks for rice yield prediction in mountainous regions, Journal of Agricultural Science, 145: 249-261, 2007.

[23] A. Irmak, J. W. Jones, W. D. Batchelor, S. Irmak, K. J. Boote, J. O. Paz, Artifical neural network model as a data analysis tool in precision farming, American Society of Agricultural and Biological Engineers, 49(6): 2027−2037, 2006.

[24] J. Leng, A. Neuhaus, and L. Armstrong, A network that really works - the

application of artificial neural networks to improve yield predictions and nitrogen management in Western Australia, Proceedings of Asian Federation for Information Technology in Agriculture, 298-306, 2014.

[25] W. W. Guo, H. R. Xue, An incorporative statistic and neural approach for crop yield modelling and forecasting, Neural Comput & Applic. 21:109–117, 2012.

[26] A. Krizhevsky, I. Sutskever, G. E. Hinton, ImageNet Classification with Deep Convolutional Neural Networks, NIPS'12: Proceedings of the 25th International Conference on Neural Information Processing Systems, 1: 1097–1105, 2012.

[27] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 3431–3440, 2015.

[28] S. Ren, K. He, R. Girshick, J. Sun, Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, arXiv:1506.01497, 2015.

[29] P. Fischer, A. Dosovitskiy, T. Brox, Descriptor matching with convolutional neural networks: A comparison to sift, arXiv:1405.5769, 2014.

[30] R. Darvishzadeh, A. Skidmore, M. Schlerf, C. Atzberger, F. Corsi, and M. Cho, LAI and chlorophyll estimation for a heterogeneous grassland using hyperspectral measurements, Isprs Journal of Photogrammetry and Remote Sensing, 63:409–426, 2008.

[31] N. Aparicio, D. Villegas, J. L. Araus, J. Casadesus, and C. Royo, Relationship between growth traits and spectral vegetation indices in durum wheat, Crop Science, 42:1547–1555, 2002.

[32] N. A. Noureldin, M. A. Aboelghar, H. S. Saudy and A.M. Ali, Rice yield forecasting models using satellite imagery in Egypt, The Egyptian Journal of Remote Sensing and Space Sciences, 16: 125-131, 2013.

[33] J. M. Chen, J. Cihlar, Plant Canopy Gap-Size Analysis Theory for Improving Optical Measurements of Leaf-Area Index, Applied Optics, 34:6211-6222, 1995.

[34] Y. Knyazikhin, et al. Hyperspectral remote sensing of foliar nitrogen content, www.pnas.org/cgi/doi/10.1073/pnas.1210196109.

[35] X. D. Wu , Q. Xiao, J. G. Wen, et al., Advances in uncertainty analysis for the

validation of remote sensing products：Take leaf area index for example, Journal of Remote Sensing, 18(5):1011-1023, 2014.

[36] D. Stroppiana, M. Baschetti, R. Confalonieri, S. Bocchi, P. A. Brinio, Evaluation of LAI-2000 for leaf area index monitoring in paddy rice, Field Crops Research, 99:167-170, 2006.

[37] S. Peng, F. V. Garcia, R. C. Laza, A. L. Sanico, R. M. Visperas, and K. G. Cassman, Increased N-use efficiency using a chlorophyll meter on high-yielding irrigated rice, Field Crops Research, 47, 243–252, 1996.

[38] S. Peng, F. V. Garcia, R. C. Laza, et al., Adjustment for specific leaf weight improves chlorophyll meter's estimate of rice leaf nitrogen concentration, Agronomy Journal, 85(5):987-990, 1993.

[39] N. H. Broge, and E. Leblanc, Comparing Prediction Power and Stability of Broadband and Hyperspectral Vegetation Indices for Estimation of Green Leaf Area Index and Canopy Chlorophyll Density, Remote Sensing of Environment, 76:156-172, 2001.

[40] A. A. Gitelson, A. Vina , V. Ciganda, D. C. Rundquist, T. J. Arkebauer, Remote estimation of canopy chlorophyll content in crops, Geophysical Research Letters 32: L08403, 2005.

[41] V. Ciganda, A. Gitelson, J. Schepers, Non-Destructive Determination of Maize Leaf and Canopy Chlorophyll Content, Journal of Plant Physiology, 166:157-167, 2009.

[42] F. Li, Y. Miao, S. D. Hennig, M. L. Gnyp, et al., Evaluating hyperspectral vegetation indices for estimating nitrogen concentration of winter wheat at different growth stages, Precis. Agric., 11:335-357, 2010.

[43] P. J. Zarco-Tejada， J. R. Miller， A. Morales, Hyperspectral indices and model simulation for chlorophyll estimation in open-canopy tree crops, Remote Sens. Environ., 90(4):463-476, 2004.

[44] M. A. White, G. P. Asner, R. R. Nemani, J. L. Privette, and S. W. Running, Measuring fractional cover and leaf area index in arid ecosystems: digital camera, radiation transmittance, and laser altimetry methods, Remote Sens. Environ., 74:

45-57, 2000.

[45] C. S. T. Daughtry, C. L. Walthall, M. S. Kim, E. Brown de Colstoun, and J. E. McMurtrey III, Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance, Remote Sens.Environ., 74: 229-239, 2000.

[46] M. Shibayama, and T. Akiyama, A spectroradiometer for field use. VI. Radiometric estimation for chlorophyll index of rice canopy, Jpn, J. Crop Sci., 55: 433-438,1986.

[47] D. Haboudane, J. R. Miller, N. Tremblay, P. J. Zarco-Tejada, and L. Dextraze, Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture, Remote Sens. Environ., 81: 416-426, 2002.

[48] E. Takada, A. Inoue, M. Shibayama, T. Sakamoto, K. Morita, A. Kimura, and W. Takahashi, Growth condition estimation of rice by fixed point camera images, J. Jpn. Agric. Systems Soc., 25:27-34, 2009.

[49] M. Shibayama, T. Sakamoto, E. Takada, et al., Estimating Rice Leaf Greenness (SPAD) Using Fixed-Point Continuous Observations of Visible Red and Near Infrared Narrow-Band Digital Images, Plant Production Science, 15(4): 293-309, 2012.

[50] C. A. Reynolds, M. Yitayew, D. C. Slack, C. F. Hutchinson, A. Huete, M.S. Petersen, Estimating crop yields and production by integrating the FAO Crop specific Water Balance model with real-time satellite data and ground-based ancillary data, Int. J. Remote Sens., 21:3487–3508, 2000.

[51] S. S. Panda, D. P. Ames, S. Panigrahi, Application of vegetation indices for agricultural crop yield prediction using neural network techniques, Remote Sens., 2:673–696, 2010.

[52] Y. Fu, G. Yang, J. Wang, X. Song, H. Feng, Winter wheat biomass estimation based on spectral indices, band depth analysis and partial least squares regression using hyperspectral measurements, Comput. Electron. Agric., 100:51–59,2014.

[53] D. Haboudane, J. R. Miller, E. Pattey, P. J. Zarco-Tejada, I. B. Strachan, 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. 90:337–352, 2004.

[54] D. Haboudane, J. R. Miller, N. Tremblay, P. J. Zarco-Tejada, L. Dextraze, Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture, Remote Sens. Environ. , 81:416–426, 2002.

[55] M. S. Moran, Y. Inoue, E. M. Barnes, Opportunities and limitations for imagebased remote sensing in precision crop management, Remote Sens. Environ., 61:319–346, 1997.

[56] Y. P. Wang, K. W. Chang, R. K. Chen, L. Jengchung, S. Yuan, Large-area rice yield forecasting using satellite imageries, Int. J. Appl. Earth Obs. Geoinf., 12 (1):27–35, 2010.

[57] I. Becker-Reshef, E. Vermote, M. Lindeman, C. Justice, A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data, Remote Sens. Environ.,114:1312–1323, 2010.

[58] L. Xue, W. Cao, W. Luo, Rice yield forecasting model with canopy reflectance spectra, J. Remote Sens. 01:100–105, 2005.

[59] S. Pradhan, K. K. Bandyopadhyay, R. N. Sahoo, V. K. Sehgal, R. Singh, V. K. Gupta, D. K. Joshi, Predicting wheat grain and biomass yield using canopy reflectance of booting stage, J. Indian Soc. Remote Sens., 42:711–718, 2014.

[60] A. M. Sibley, P. Grassini, N. E. Thomas, K. G. Cassman, D. B. Lobell, Testing remote sensing approaches for assessing yield variability among maize fields, Agron. J., 106:24–32, 2014.

[61] C. Zhang, and J. M. Kovacs, The application of small unmanned aerial systems for precision agriculture: A review, Precision Agriculture, 13: 693-712, 2012.

[62] A. Verger, et al., Green area index from an unmanned aerial system over wheat and rapeseed crops, Remete Sens. Envieon., 152:654-664, 2014.

[63] C. Rosenzweig et al., The agricultural model intercomparison and improvement project (AgMIP): protocols and pilot studies Agric., Forest Meteorol., 170:

166–182, 2013.

[64] J. C. Ciscar, K. Fisher-Vanden and D. B. Lobell, Synthesis and review: an inter-method comparison of climate change impacts on agriculture, Environ. Res. Lett.,13(7):070401, 2018.

[65] D. B. Lobell and M. B. Burke, On the use of statistical models to predict crop yield responses to climate change, Agric. Forest Meteorol., 150(11): 1443–1452, 2010.

[66] D. B. Lobell and S. Asseng, Comparing estimates of climate change impacts from process-based and statistical crop models, Environ. Res. Lett. 12:015001,2017.

[67] W. Schlenker, and M. Roberts, Nonlinear Temperature Effects Indicate Severe Damages to US Crop Yields under Climate Change, Proceedings of the National Academy of Sciences, 106:15594-15598, 2009.

[68] M. J. Roberts, N. O. Braun, T. R. Sinclair, D. B. Lobell and W. Schlenker, Comparing and combining process-based crop models and statistical models with some implications for climate change, Environ. Res. Lett. 12:095010, 2017.

[69] D. W. Urban, J. Sheffield and D. B. Lobell, The impacts of future climate and carbon dioxide changes on the average and variability of us maize yields under two emission scenarios, Environ. Res. Lett., 10:045003, 2015.

[70] G. H. Brenchley, Aerial photography for the study of plant diseases, Annual Review of Phytopathology, 6:1-22, 1968.

[71] J. S. West, C. Bravo, R. Oberit, et al., The potential of optical canopy measurement for targeted control of field crop diseases, Annual Review of Phytopathology, 41: 593-614, 2003.

[72] H. H. Muhammed, Hyperspectral Crop Reflectance Data for characterising and estimating Fungal Disease Severity in Wheat, Biosystems Engineering, 91(1): 9-20, 2005.

[73] J. B. Jiang, Y. H. Chen, W. J. Huang, Using hyperspectral derivative index to monitor winter wheat disease, Spectroscopy and Spectral Analysis, 27(12): 2475-2479, 2007.

[74] C. J. Cai, Z. H. Ma, H. G. Wang, et al., Comparison research of hyperspectral

properties between near-ground and high altitude of wheat stripe rust, Acta Phytopa Thologica Sinica, 37(1): 77-82, 2007.

[75] B. Chen, S. K. Li, K. R. Wang, et al., Spectrum characteristics of cotton single leaf infected by verticillium wilt and estimation on severity level of disease, Scientia Agriculture Sinica, 40(12): 2709-2715, 2007.

[76] Z. Y. Liu, H. F. Wu, J. F. Huang, Application of neural networks to discriminate fungal infection levels in rice panicles using hyperspectral reflectance and principal components analysis, Computers and Electronics in Agriculture, 72: 99-106, 2010.

[77] J. H. Luo, W. J. Huang, X. H. Gu, et al., Monitoring stripe rust of winter wheat using PHI based on sensitive bands, Spectroscopy and Spectral Analysis, 30(1): 184-187, 2010.

[78] Y. D. Zhang, J. C. Zhang, D. Z. Zhu, et al., Investigation of the hyperspectral image characteristics of wheat leaves under different stress, Spectroscopy and Spectral Analysis, 31(04): 1101-1105, 2011.

[79] C. D. Jones, J. B. Jones, W. S. Lee, Diagnosis of bacterial spot of tomato using spectral signatures, Computers and Electronics in Agriculture, 74(2): 329-335, 2010.

[80] A. D. Fiore, M. Reverberi, A. Ricelli, et al., Early detection of toxigenic fungi on maize by hyperspectral imaging analysis, International Journal of Food Microbiology, 144(1): 64-71, 2010.

[81] J. C. Zhang, W. J. Huang, J. Y. Li, et al., Development, evaluation and application of a spectral knowledge base to detect yellow rust in winter wheat, Precision Agriculture, 12(5): 716-731, 2011.

[82] F, Jonas, G. Menz, Multi-temporal wheat disease detection by multi-spectral remote sensing, Precision Agriculture, 8(3):161-172, 2007.

[83] C. H. Yang, J. H. Everitt, C. J. Fernandez, Comparison of airborne multispectral and hyperspectral imagery for mapping cotton root rot, Biosystems engineering, 107(2): 131-139, 2010.

[84] J. H. Luo, C. J. Zhao, W. J. Huang, et al., Discriminating wheat aphid damage

degree using 2-dimensional feature space derived from Landsat 5 TM, Sensor Letters, 10 (1):608-614, 2012.

[85] C. D. Andrew, Machine learning methods for crop yield prediction and climate change impact assessment in agriculture, Environ. Res. Lett. 13:114003, 2018.

[86] M. H. Malais, W. J. Ravensberg, Knowing and Recognizing: The Biology of Glasshouse Pests and Their Natural Enemies, Koppert BV: Berkel en Rodenrijs, The Netherlands, 2004.

[87] Y. Sun, H. Cheng, Q. Cheng, et al., A smart-vision algorithm for counting whiteflies and thrips on sticky traps using two-dimensional Fourier transform spectrum, Biosyst. Eng., 153:82–88, 2017.

[88] K. M. Heinz, M. P. Parrella, J. P. Newman, Time-efficient use of yellow sticky traps in monitoring insect populations, J. Econ. Entomol. , 85:2263–2269, 1992.

[89] K. S. Pike, G. L. Reed, G. T. Graf, and D. Allison, Compatibility of imidacloprid with fungicides as a seed-treatment control of Russian wheat aphid (Homoptera: Aphididae) and effect on germination, growth, and yield of wheat and barley, J. Econ. Entomol., 86(2):586–593, 1993.

[90] C. M. Rush and S. D. Lyda, The effects of anhydrous ammonia on membrane stability of Phymatotrichum omnivorum, Mycopathologia, 79(3): 147−152, 1982.

[91] N. E. Ahmed, H. O. Kanan, S. Inanaga, Y. Q. Ma, and Y. Sugimoto, Impact of pesticide seed treatments on aphid control and yield of wheat in the Sudan, Crop Prot., 20(10): 929–934, 2001.

[92] J. Zhang, R. Pu, W. Huang, L. Yuan, J. Luo, and J. Wang, Using in-situ hyperspectral data for detecting and discriminating yellow rust disease from nutrient stresses, Field Crop Res., 134:165–174, 2012.

[93] S. Graeff and W. Claupein, Identification and discrimination of water stress in wheat leaves (Triticum aestivum L.) by means of reflectance measurements, Irrigation Sci., 26(1): 61–70, 2007.

[94] D. Rogge, B. Rivard, M. K. Deyholos, J. Lévesque, J. Ardouin, and A. A. Faust, Potential discrimination of toxic industrial chemical effects on poplar, canola and wheat, detectable in optical wavelengths 400–2450 nm, IEEE J. Sel. Topics Appl.

Earth Observ. Remote Sens., 5(2):563–573, 2012.

[95] J. J. Park, J. K. Kim, H. Park, K. Cho, Development of time-efficient method for estimating aphids density using yellow sticky traps in cucumber greenhouses, J. Asia-Pac. Entomol., 4:143–148, 2001.

[96] K. Espinoza, D. L. Valera, J. A. Torres, A. Lopez, F. D. Molina-Aiz, Combination of image processing and artificial neural networks as a novel approach for the identification of Bemisia tabaci and Frankliniella occidentalis on sticky traps in greenhouse agriculture, Comput. Electron. Agric., 127:495–505, 2016.

[97] C. Xia, T. S. Chon, Z. Ren, J. M. Lee, Automatic identification and counting of small size pests in greenhouse conditions with low computational cost, Ecol. Inform. , 29:139–146, 2015.

[98] J. Cho, J. Choi, M. Qiao, C. Ji, H. Kim, K. Uhm, T. Chon, Automatic identification of whiteflies, aphids and thrips in greenhouse based on image analysis, Red, 346:244, 2007.

[99] J. G. A. Barbedo, Using digital image processing for counting whiteflies on soybean leaves, J. Asia-Pac. Entomol., 17: 685–694, 2014.

[100] M. Maharlooei, S. Sivarajan, S. G. Bajwa, J. P. Harmon, J. Nowatzki, Detection of soybean aphids in a greenhouse using an image processing technique, Comput. Electron. Agric., 132:63–70,2017.

[101] L. O. Solis-Sánchez, J. J. García-Escalante, R. Castañeda-Miranda et al., Machine vision algorithm for whiteflies (Bemisia tabaci Genn.) scouting under greenhouse environment, Appl. Entomol., 133:546–55, 2009.

[102] O. S. Luis, C. Rodrigo, J. G. Juan, et al., Scale invariant feature approach for insect monitoring, Comput. and Electron. in Agric., 75(1):92–99, 2011.

[103] J. Alexander, T. Nguyen, Localized Linear Regression in Networked Data. IEEE Signal Processing Letters, 99:1-1, 2019.

[104] P. Milton, H. Coupland, E. Giorgi, et al., Spatial Analysis Made Easy with Linear Regression and Kernels, Epidemics, 2019.

[105] W. Khaled, J. G. Lin, Z. C. Han, et al., Test for Heteroscedasticity in Partially Linear Regression Models, Journal of Systems Science and Complexity,

32(4):1194-1210, 2019.

[106] C. D. Nye, J. Prasad, J. Bradburn, et al., Improving the operationalization of interest congruence using polynomial regression, Journal of Vocational Behavior, 104:154-169, 2018.

[107] M. Gentilucci, C. Bisci, P. Burt, et al., Interpolation of Rainfall Through Polynomial Regression in the Marche Region (Central Italy), 2018.

[108] A. Kumar, R. B. Chinnam, F. Tseng, An HMM and Polynomial Regression Based Approach for Remaining Useful Life and Health State Estimation of Cutting Tools, Computers & Industrial Engineering, 128, 2018.

[109] S. R. Jondhale, R. S. Deshpande, Efficient localization of target in large scale farmland using generalized regression neural network, International Journal of Communication Systems, 3: e4120, 2019.

[110] D. K. Ghose, S. Samantaray, Modelling sediment concentration using back propagation neural network and regression coupled with genetic algorithm, Procedia Computer Science, 125:85-92, 2018.

[111] S. Zhuang, X. Gong, C. Lin, et al., Estimate of daily irradiation exposure of global radiation using generalized regression neural network, Taiyangneng Xuebao/Acta Energiae Solaris Sinica, 40(1):11-16, 2019.

[112] M. Ustuner, F. B. Sanli, S. Abdikan, M. T. Esetlili, Y. Kurucu, Crop type classification using vegetation indices of rapideye imagery, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-7, 2014.

[113] A. Roman, T. Ursu, Multispectral satellite imagery and airborne laser scanning techniques for the detection of archaeological vegetation marks, In book: Landscape archaeology on the northern frontier of the roman empire at porolissum-an interdisciplinary research project, 141-152, 2016.

[114] E. Ben-Dor, Y. Inbar, Y. Chen, The reflectance spectra of organic matter in the visible near-infrared and short wave infrared region (400–2500 nm) during a controlled decomposition process, Remote Sens. Environ., 61(1):1-15, 1997.

[115] R. Ballesteros, J. F. Ortega, D. Hernandez, and M. A. Moreno, Applications

of georeferenced high-resolution images obtained with unmanned aerial vehicles. Part I: Description of image acquisition and processing, Precision Agriculture, 15(6):579-592, 2014.

[116] F. Lopez-Granados, J. Torres-Sanchez, A. Serrano-Perez, et al., Early season weed mapping in sunflower using UAV technology: variability of herbicide treatment maps against weed thresholds, Precision Agriculture, 17(2):183-199, 2016.

[117] M. Schirrmann, A. Giebel, F. Gleiniger, et al., Monitoring Agronomic Parameters of Winter Wheat Crops with Low-Cost UAV Imagery, Remote Sensing, 8(9):706, 2016.

[118] D. C. Wang, Y. F. Li, W. J. Fan, Q. M. Qin，Monitoring wheat quality protein content in critical period based division by remote sensing， IEEE International Geoscience and Remote Sensing Symposium, 2012.

[119] C. C. D. Lelong, P. Burger and G. Jubelin et al., Assessment of unmanned aerial vehicles imagery for quantitative monitoring of wheat crop in small plots. Sensors, 8: 3557-3585, 2008.

[120] P. Yang，Q. Zhou，Z. Chen, Estimation of regional crop yield by assimilating multi-temporal TM images into crop growth model, IEEE International Conference on Geoscience & Remote Sensing Symposium, 2007.

[121] Y. Zhao, Crop growth dynamics modeling using time-series satellite imagery, Proceedings of the SPIE, Volume 9260, 2014.

[122] H. H. Yuan, G. J. Yang, C. C. Li, et al., Retrieving Soybean Leaf Area Index from Unmanned Aerial Vehicle Hyperspectral Remote Sensing, Analysis of RF, ANN, and SVM Regression Models, Remote Sensing, 9:309, 2017.

[123] Z. Y. Han, X. C. Zhu, X. Y. Fang, et al., Hyperspectral estimation of apple tree canopy LAI based on SVM and RF regression, Spectrosc. Spectr. Anal., 36:800–805, 2016.

[124] G. Omer, O. Mutanga, E. M. Abdel-Rahman, E. Adam, Empirical prediction of Leaf Area Index (LAI) of endangered tree species in intact and fragmented indigenous forests Ecosystems using WorldView-2 data and two robust machine

learning algorithms, Remote Sens., ,8:324, 2016.

[125] J. Verrelst, J. Munoz, L. Alonso, J. Delegido, J. Pablo Rivera, G. Camps-Valls, J. Moreno, Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and -3, Remote Sens. Environ., 118:127–139, 2012.

[126] Y. T. Mustafa, P. E. Van Laake, A. Stein, Bayesian network modeling for improving forest growth estimates, IEEE Trans. Geosci. Remote, 49:639–649, 2011.

[127] A. Sedki, D. Ouazar, and E. El Mazoudi, Evolving neural network using real coded genetic algorithm for daily rain fall runoff forecasting, Expert Systems with Applications, 36(3):4523–4527, 2009.

[128] C. Gowda and S. G. Mayya. Comparison of Back Propagation Neural Network and Genetic Algorithm Neural Network for Stream Flow Prediction, Journal of Computational Environmental Sciences, Article ID 290127, 2014.

[129] H. B. Qiao , D. F. Cheng , J. R. Sun, et al., Effects of wheat aphid on spectrum reflectance of the wheat canopy, Plant Protection, 31, 2005.

[130] J. H. Luo, W. J. Huang, J. L. Zhao, et al., Detecting aphid density of winter wheat leaf using hyperspectral measurements, IEEE JSTARS, 6(2): 690–698, 2013.

[131] D. G. Hadjimitsis, G. Papadavid, Remote Sensing for Determining Evapotranspiration and Irrigation Demand for Annual Crops, In book: Remote Sensing of Environment - Integrated Approaches, DOI: 10.5772/39305, 2013.

[132] P. S. Thenkabail, I. Mariotto, M. K. Gumma, E. M. Middleton, D. R. Landis, and K. F. Huemmrich, Selection of Hyperspectral Narrowbands (HNBs) and Composition of Hyperspectral Twoband Vegetation Indices (HVIs) for Biophysical Characterization and Discrimination of Crop Types Using Field Reflectance and Hyperion/EO-1 Data, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 6 (2):1-13, 2013.

[133] Y. L. TANG，J. H. WANG，J. F. HUANG，R. C. WANG, Yield Estimation by Hyperserctral Data of Rice Canopies in M ature Stages, ACTA

AGRONOMICA SINICA, 30(8):780-785, 2004.

[134] L. Zhuang, J. Wang, L. Bai, et al., Cotton yield estimation based on hyperspectral remote sensing in arid region of China, Transactions of the CSAE, 27(6):176-181, 2011.

[135] A. Colorni, M. Dorigo and V. Maniezzo, Distributed optimisation by ant colonies. In: Proc. of 1st European Conf. Artificial Life, Pans, France: Elsevier, 134-142, 1991.

[136] A. Colorni, M. Dorigo and V. Maniezzo, An investigation of some properties of an ant algorithm, In: Proc. of Parallel Problem Solving from Nature (PPSN). France: Elsiver, 509-520, 1992.

[137] X. S. Yang, and S. Deb, Cuckoo Search via Levy Flights, Proceedings of the World Congress on Nature & Biologically Inspired Computing (NaBIC'09), Coimbatore, 210-214, 9-11 December 2009.

Appendix A

List of author's Publications

1.Yuxi Fang, He Sun, Yijun Yan, Jinchang Ren, Hong Yue, Tariq Durrani. Wheat Growth Assessment for Satellite Remote Sensing Enabled Precision Agriculture. 2019 International Conference on Communications, Signal Processing, and Systems (CSPS 2019).

2.Yijun Yan, Sophia Zhao, Yuxi Fang, Yuren Liu, Zhongxin Chen, Jinchang Ren. VIP-STB Farm: Scale-up Village to County/Province Level to Support Science and Technology at Backyard (STB) Program. July 13-14, 2019. The 10th International Conference on Brain-Inspired Cognitive Systems (BICS 2019).