Optik 125 (2014) 5660–5665

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Predicting the probability of wheat aphid occurrence using satellite remote sensing and meteorological data

Juhua Luo [a,*], Wenjiang Huang [b], Jingling Zhao [c], Jingcheng Zhang [c], Ronghua Ma [a], Muyi Huang [d]

[a] The State Key laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, China
[b] Key Laboratory of Digital Earth Sciences, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, China
[c] Beijing Research Center for In-formation Technology in Agriculture, Beijing, China
[d] Department of Environmental and Energy Engineering, Anhui Jianzhu University, Hefei, China

ARTICLE INFO

Article history:
Received 28 September 2013
Accepted 26 May 2014

Keywords:
Wheat aphid
Logistic regression
Remote sensing image
Meteorological data
Probability

ABSTRACT

The occurrence and prevalence of wheat aphid has severe impact on wheat quality and bring about wheat yield loss. Infestation models commonly based on in situ meteorological data can be effective, but are usually lacking spatialized information, which could be provided using multispectral remote sensing datasets. The purpose of this study was to develop a prediction model for aphid occurrence probability by combining remote sensing images and meteorological data in a logistic regression based framework. In such framework, the predictor variables are: land surface temperature (LST), normalized difference vegetation index (NDVI) and perpendicular drought index (PDI) derived from satellite remote sensing image and temperature, precipitation and wind speed coming from meteorological stations. Logistic regression estimated predictor coefficients showed that LST was the most important factor inducing aphid occurrence, followed by NDVI, temperature, precipitation, wind speed and PDI. The prediction accuracy of the model was evaluated by calculating receiver operating characteristics (ROC) against reference data, scoring a value of area under the ROC curve of 0.993, and resulting in an overall predicted accuracy of 94.4% for the estimated subset. In addition, the results suggested that the probability of aphid occurrence obtained by the logistic regression model had a positive correlation with aphid damage level from field data. Based on the validation subset, it was observed that there were lower omission error with 23.68 and commission error with 23.81% and higher overall accuracy with 75.61% when the cut-off probability (p) value was 0.45. The model developed can be a useful tool to predict aphid occurrence and incidence over winter wheat with some advance, and it could be effective in protecting winter wheat from aphid infestation.

1. Introduction

Wheat aphid (*Sitobion avenae* F.) is one of the main aphid species infesting winter wheat, and is also one of the most destructive pests for agricultural activities in Northwest China. Aphid infestation can cause significant yield loss. In high enough densities, wheat aphids can deprive plants of nutrients, bringing about potential reduction of the number of heads, the number of grains per head, and a reduction seed weight, especially during the flowering and filling stages in wheat phenology [1]. It was reported that a density of 10 aphids per tiller resulted in 7% yield loss and 40 aphids per tiller led to 11% yield loss [2]. Consequently, predicting the probability of wheat aphid infestation in an effective way could facilitate timely implementation of preventive strategies, which are critical to enhancing the viability of wheat production industry in China, in particular for the large scale area application.

The most important driving forces in occurrence and development of crop disease and pests infestation are considered to be meteorological data, including temperature, rainfall, and wind speed, usually acquired by weather stations [3,4]. As a consequence, the models for predicting the incidence and probability of crop disease and pest infestation are generally established based on series of historical meteorological data, and traditional statistical treatment [5,6]. Such prediction models are effective in forecasting the overall occurrence condition of disease and pest in a given area and temporal range according to meteorological data, but they are of little help in predicting spatial occurrence of disease and pest due to limited density of weather sites.

* Corresponding author.
E-mail address: luojuhua@126.com (J. Luo).

http://dx.doi.org/10.1016/j.ijleo.2014.06.010

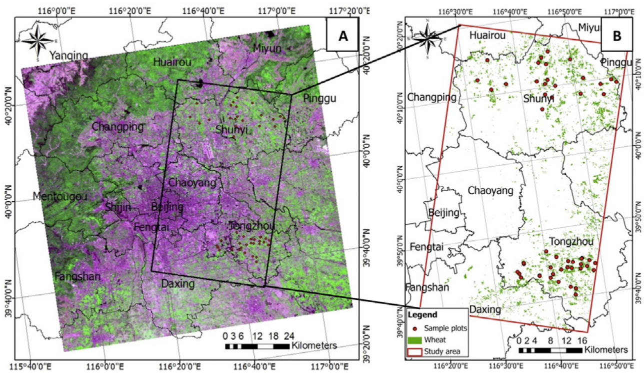

Fig. 1. The location of study areas (A, HJ-CCD image covering study area; B, wheat planting area of this study).

Remote sensing technologies have a great potential in providing spatially continuous observations of environmental variables on large scales [7,8]. Satellite imagery, in particular, provides better spatial coverage than in situ surface meteorological station data and is available in most parts of the world on a regular basis. For this reason, the satellite-derived variables were more and more applied in agrometeorological crop simulation models for monitoring the effect of weather conditions on crop growth and for predicting crop yields form regional to continental scales [9]. Moreover, recent studies have reported that the occurrence and development of crop disease and insect pest are related to some variables that can be derived from satellite remote sensing, such as LST, NDVI, modified normalized difference water index (MNDWI) [10]. Luo et al. (2012) found that the incidence of wheat yellow rust had relationship with LST, while another study showed that aphid occurrence and the damage degree were related to LST, NDVI and MNDWI [11]. This suggests that the occurrence and development of crop disease and pest may be related with habitat and crop growth factors, derived from both in situ and remote sensing data.

Logistic regression has become a widely used and accepted method for the analysis of presence/absence dependent variables, and it can predict the probability for the state of a dichotomous dependent variable based on the predictor variables. It has been applied in many fields, ranging from tree and stand survival estimation in forest competitive environment to ground subsidence hazard mapping in geological field [12,13]. In the field of remote sensing, it has been used for various purposes ranging from land-cover change detection [14] to mapping insect tree defoliation due to insect infestation [15].

By using logistic regression, the objective of this work is to implement a multivariable model for predicting the probability of aphid occurrence based on both meteorological data and satellite remote sensing data.

2. Materials and methods

2.1. Study area

The study areas are located in Shunyi district (116°28′–116°58′ E, 40°00′–40°18′ N) and Tongzhou district (116°32′–116°56′ E, 39°36′–40°02′ N) in Beijing, China (Fig. 1a). The study areas are characterized by flat topography, with elevation ranging from 20 m to 40 m, and semi-humid warm temperate climate. The annual precipitation and average temperature are 625 mm and 11.5 °C in Shunyi district, 620 mm and 11.3 °C in Tongzhou district, respectively. The two districts are considered to be the major winter wheat planting areas in Beijing, where aphid infestations occur almost every year.

2.2. Field inventory and data pre-processing

Field campaigns were carried out during the growing seasons of winter wheat in 2010. The winter wheat in the study areas were planted approximately from September 25 to October 7, 2009, and harvested from June 19 to 25, 2010. The field data were collected during different stages of winter crop growth: joining (May 10, 12), heading (May 20, 21), filling (June 3, 4). Based on the combination scheme of representative sampling and random sampling scheme, 85 sample plots sized of 0.36 ha (60 m × 60 m) each one were collected. At each plot, five quadrats (1 m × 1 m) were selected to survey aphid density through diagonal sampling. The tillers in each quadrat were randomly selected and the aphids were counted. The aphid densities were then estimated by using following formula: aphid density = total aphids/10 tillers. The average aphid density of five quadrats was used for characterizing the aphid density of the whole sample plot. Aphid damage was assessed, based on six levels of severity related to aphid density according to the plan protection research specification of China [16].

At the same time, the geographical coordinates of each plot were recorded by a global positioning system (GPS) (GeoExplorer 3000 GPS, with a sub-metre positioning accuracy) at the centre point of each plot.

The sample plots were divided into two groups. A total of 55 plots surveyed on May 10, May 12, May 20 and May 21 were assigned to use as input subset for implementing the prediction model of aphid, while the other 30 plots surveyed on June 3 and June 4 were used for validating the model.

2.3. Meteorological data and pre-processing

Meteorological data including the daily average temperature, precipitation and maximum wind speed, were collected from eight meteorological stations of Beijing suburb from May 1 to June 1, 2010.

For the purpose of this study, the original metrological variables were also obtained, namely: (1) Temperature – T1: the average temperature from May 1 to May 13, 2010; (2) Temperature – T2: the average temperature from May 13 to May 20, 2010; (3) Precipitation – T1: the average precipitation from May 1 to May 13, 2010; (4) Precipitation – T2: the average precipitation from May 13 to May 20, 2010; (5) Wind – T1: the average wind from May 1 to May 13, 2010; (6) Wind – T2: the average wind from May 13 to May 20, 2010.

Inverse distance weighted (IDW) interpolation algorithm was used to spatialize meteorological data across the study area starting from eight in situ stations points using ARCGIS software.

2.4. Satellite imagery acquisition and pre-processing

Remote sensing dataset used was from the environment and disaster reduction small satellites (HJ-1A/B). HJ-1A and HJ-1B satellites were launched by China on September 6th, 2008. The CCD sensor onboard of the two satellites has similar spectral and spatial characteristics of Landsat 5 TM. However, it has a more frequent revisit interval (2 days) by constellation of the two satellites, and wide swath width (360 m), which is considered of great importance for agricultural monitoring and observation. The infrared scanner (IRS) sensor, onboard of HJ-1B satellite has one infrared band (10.5–12.5 μm), with radiometric resolution higher than 1.0 K. The nadir spatial resolution of the thermal band is 300 m and the swath width is 720 km (4 days repeat cycle). In this study, HJ-1A/B data were used to derive normalized difference vegetation index (NDVI), land surface temperature (LST) and perpendicular drought index (PDI) corresponding to each sample plots.