Remote sensing recognition of paddy waterlogging using change vector analysis model

Xiaohe Gu[1], Jingcheng Zhang[1], Peng Xu[1], Yingying Dong[1], Yansheng Dong[1*]

[1] Beijing Research Center for Information Technology in Agriculture, Beijing 100097, China

Abstract. It is significant to monitor crop waterlogging range timely and correctly for later field management, agricultural insurance and yield prediction. The change regularity of paddy injured growth was analyzed and the sensitive parameters of growth stress were screened. The model of identifying the paddy waterlogging range based on change vector analysis (CVA) was developed by using the HJ-1/2 CCD images around waterlogging. At last the accuracy of the developed model was evaluated by in-situ sample data. Results showed that the waterlogged paddy mainly scattered around the Huaihe River system. The spatial distribution pattern was in conformity with the occurrence tendency of paddy waterlogging provided by local agricultural department as a whole. By evaluating the accuracy of the model with in-situ samples, the overall accuracy of the model developed in the study reached 87.5%, while the Kappa coefficient reached 0.737. The change vector analysis model could identify the waterlogged paddy and normal paddy correctly and efficient in Huaihe River Basin.

1   Introduction

Flood and waterlogging present such features as high frequency and wide range in China, which have greatly affected national economic development. The flood and waterlogging mainly locate around major rivers with rich soil and developed agriculture. According the statistics, every year more than 7.33 million hectares farmland suffered floods and waterlogging in China. It is significant to monitor crop waterlogging range timely and correctly for later field management, agricultural insurance and yield prediction.

Most researches on loss assessment of flood and waterlogging mostly focused on statistical data such as population and economy, but lack of special loss assessment aiming at crop using remote sensing technology. The loss assessment of crop waterlogging not only needs to monitor the characteristics information of flood, but also needs to consider the factors such as crop variety, water-fertilizer, growing stage

---

* Corresponding author: dongys@nercita.org.cn

and soil texture. Li et al (2003) achieved damage evaluation of flood by using remote sensing and spatial economy database, including affected farmland area and affected population. Zhang et al (2007) developed the climate identifying index of drought and flood, analyzed the regularity of outbreak and its influence of crop yield. Recent research on evaluating waterlogging in relation to crop loss mostly focused on experience models, but rarely on mechanistic model (Profeti, et al., 1997 & Srikantha, et al., 1998;).

Because paddy field is usually flat and long-term in water, it is different to distinguish waterlogging region and natural water covering region through water spectrum change with the influence of paddy coronal. In addition, remote sensing images with appropriate resolution are hardly acquired for the cloudy weather when flood and waterlogging happens. So it is different to identify the paddy waterlogging directly using up-to-date images within the flood period. In this study the change regularity of paddy injured growth was analyzed and the sensitive parameters of growth stress were screened. By using the HJ-1/2 CCD images in the former and later periods of waterlogging, the model of monitoring the paddy waterlogging range based on change vector analysis (CVA) was developed. At last the accuracy of the developed model was evaluated by in-situ sample data.

2   Study area and materials

2.1   Study area

The Huaihe River Basin is chosen as study area, where locates at transverse zone of semi-humid area to humid area in the east of China. The weather caused flood in the Huaihe River Basin is very changeable and the space-time location of the precipitation is not well balanced. The average annual precipitation is 883 mm and mainly concentrates on the period from June to September (Bi, et al, 2004). By analyzing the interannual variations, the precipitation of high flow years reaches five times more than that of dry years. The Huaihe River Basin is one of the highest frequency basins of flood in China. The variation law of summer precipitation with 2-3a cycles makes the basin suffering flood once every two or three years. In addition, because of low-lying land and inadequate ability of draining waterlogged fields, the Huaihe River Basin often holds out the situation of outside flood and inside waterlogging. For example, the flood happened in 2007 caused 1.25 million hectares of crop waterlogged in Anhui Province, including 360000 hectares of no harvest land and 44 hundred million yuan of agricultural loss.

The paddy waterlogging is directly related with the weather condition, especially short-term precipitation. Near Huoqiu and Shouxian county of Huaihe River Basin suffered continuous heavy precipitation from 26 to 30 August 2002 on the occasion of mid filling stage of local paddy. According to the records of local weather stations, the precipitation reached 74.3 mm in a short period of time. The observations of local department of agriculture discovered the severe waterlogging locally for hysteretic drainage management.

2.2 Data acquisition and preprocessing

Three images of HJ-1/2 CCD with mid resolution were used in the study, of which the temporal included July 13, August 26, September 3 and September 5, 2009. The resolution and response feature of HJ-1/2 CCD image were very close to these of LandSat5 TM image, which could meet accuracy requirement of paddy identification, growth retrieval and yield loss evaluation. The covering cycle is two days with a width of 360 kilometers, of which the advantages of multi-temporal and large scale is help to monitor the crop growth.
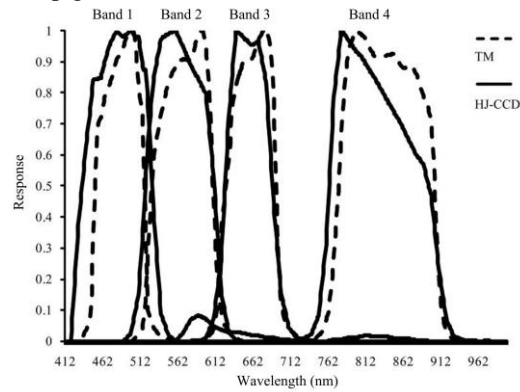

Fig. 1. Response function of former four band between HJ CCD and TM image

The in-situ data used in the study included 42 samples surrounding Fuyang, Huainan and Bengbu city in Anhui Province, which collected from September 17 to 20 in 2009. The collecting indexes of sample parcels included GPS position, planting density and waterlogged level. The representative plant samples of three square meters were gathered, of which the yield were determined in the lab.

In addition, phenomena data in Anhui Province was used to analyze the spectrum feature of fall crops and provide prior knowledge for decision tree classification stat.

The preprocess steps of remote sensing image included geometric correction, atmospheric calibration and projection transformation. The method of atmospheric calibration used in the paper was FLAASH model. The geometric correction used the polynomial model of bilinear interpolation, which could insure the error less than 0.5 pixels. The projection type of all spatial data was Universal Transverse Mercator (UTM), while the geography coordinate system was WGS 84.

3 Method and technical process

3.1 Identifying spatial distribution of paddy

The decision tree classification was used to identify paddy by analyzing spectrum

feature of multi-temporal HJ CCD images with in-situ sample point data. The confused crops in the study area mainly includes peanut.

By consulting the former studies, the spectrum information and vegetation indexes of two temporal HJ CCD images were used to identify paddy, including the NDVI of HJ-CCD image of August 26 (named T_1), the reflectivity of the fourth band of August 26 image (named T_2), the reflectivity of the fifth band of August 26 image (named T_3) and the NDVI of July 13 image. The classification was carried in the Ecognition7.0.

3.2   Physiological mechanism of growth loss of paddy waterlogging

Water management of paddy field is of the utmost importance for growth and maturity of paddy. Paddy growing can not live without water during most growing periods, especially flowering stage and filling stage.

Maintain reasonable water level from heading to mid filling stage is helpful for paddy field. The water caused by heavy rains is apt to flow into paddy field for the restraints of topography and terrain. On account of the limited ability of flood discharge, the paddy field is apt to suffer waterlogging, even the flag leaf and ear is submerged. When the inversion second and third leaf are submerged, paddy waterlogging is considered to be suffered. Because of the depressed photosynthesis and damaged rice embryos, the waterlogging suffered at the filling stage will reduce grain plumpness and affect the yield evidently. If the ears of paddy are submerged more than 48 hours, the grain will stop filling and make for aborted kernels probably.

The spatial distribution of paddy waterlogging in the study was identified by the change vector analysis model with the two HJ CCD images around flood. According the information provided by local agricultural department, the leaves of some submerged rice plants turned white and were apt to lodge after a week of waterlogging. These variations provided the basis for identifying spatial distribution of waterlogging paddy.

3.3   CVA model for identifying waterlogged paddy

The method of change vector analysis is widely applied for processing multi-bands images. The change vector could be used to describe the direction and amount of spectrum variety from one temporal to another temporal. Assuming the vector of pixel spectrum of temporal t_1 and t_2 represent G = (g_1 , g_2 , ..., g_k)^T and H= (h_1 , h_2 , ..., h_k)^T respectively, the change vector between t_1 and t_2 will be described as follows.

$$\Delta G = G - H = \begin{pmatrix} g_1 - h_1 \\ g_2 - h_2 \\ \\ g_k - h_k \end{pmatrix}.$$

(1)

The ΔG includes the all change information of two temporal image, of which the change intensity depends on the module of ΔG.

$$\|\Delta G\| = \sqrt{(g_1 - h_2)^2 + (g_2 - h_2)^2 + \cdots + (g_k - h_k)^2}.$$

(2)

According to the definition of change intensity, the value of $\|\Delta G\|$ discloses the difference of two images and the change possibility. By setting proper threshold values of $\|\Delta G\|$, the changed and unchanged pixels could be distinguished. The method of change vector analysis is usually applied for identifying the pixels of land use and coverage change. The translated types of land use are depended on the direction of ΔG.

In the study, the difference of growth between waterlogged paddy and normal paddy coheres with the spectrum variation caused by waterlogging. For this reason the method of change vector analysis is introduced to identify the spatial distribution of waterlogged paddy.

3.4   Accuracy evaluation

Through the spatial overlay and analysis between in-situ samples and waterlogged paddy monitoring by remote sensing, the error matrix was established. The accuracy of change vector analysis model was evaluated with overall accuracy and Kappa coefficient.

4   Results analysis

On account of the degraded chlorophyll level and enlarged water information derived from submerged paddy, normalized difference vegetation index (NDVI), reflectivity of near infrared band (NIR) and greenness normal difference vegetation index (GNDVI) were input the model. The NDVI represented crop growth, which was declined with the waterlogging. The NIR and GNDVI were used to indicate water content of plant, which ascended for submerging plants. So the CVA model was transformed as followed function.

$$\|\Delta G\| = \sqrt{(ST\Delta NDVI - ST\Delta Nir - ST\Delta GNDVI)^2} \quad . \tag{3}$$

The STΔNDVI, STΔNir and STΔGNDV indicated the differences of three normalized eigenvectors. The normalization was intended to eliminate the difference from dimension of eigenvectors and improved the comparability. The normalization function was as follows.

$$ST\Delta F^2 = \frac{x_1 - \overline{x}}{\sigma} \quad . \tag{4}$$

Here, the x_1 means the difference values of eigenvectors of every pixel. The $\overline{x}$ means the average difference values of eigenvectors of every pixel. The σ means standard deviation. The differences of three eigenvectors from two temporal images were calculated by followed function.

$$\Delta NDVI = NDVI_{0826} - NDVI_{0903\text{-}05} \quad . \tag{5}$$
$$\Delta Nir = Nir_{0826} - Nir_{0903\text{-}05} \quad . \tag{6}$$
$$\Delta GNDVI = GNDVI_{0826} - GNDVI_{0903\text{-}05} \quad . \tag{7}$$

The spatial distribution of $\|\Delta G\|$ was mapped as figure 2.

Because two temporal images were used in the study, the change vector had only

two directions, including magnifying and reducing. According to the magnifying modulus, the waterlogged paddy and normal paddy were identified by the threshold. In the study, the threshold was determined with statistical analysis of pixel values of two temporal images, which attributed to improve the generality of the method with different region and images. Because paddy submerged in small scale in study area, the mean and standard deviation value (SD) were used to set threshold through statistical analysis of paddy pixels.

$$\text{Threshold} = \text{Mean}_{\|\Delta G\|} + 2 * SD_{\|\Delta G\|} \qquad (8)$$

If the modulus of change vector was above the threshold, the change extent of waterlogged paddy pixels exceeded that of normal paddy pixels, that is $\text{Mean}_{\|\Delta G\|} \pm SD_{\|\Delta G\|}$. Because the direction of change vector in the study was fixed and conservative, the threshold was determined as $\text{Mean} \pm \text{SD}$. The threshold setting was 5.76, which was used to map the waterlogged paddy as figure 3.
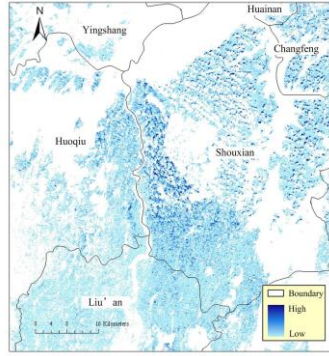

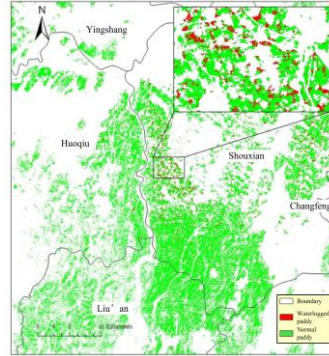**Fig. 2.** Spatial distribution of $\|\Delta G\|$ **Fig. 3.** Spatial distribution of waterlogged paddy

The figure 3 showed that the waterlogged paddy (colored red) scattered everywhere in the study, mainly around the Huaihe River system. The waterlogging in Shouxian County was comparatively serious. The spatial distribution pattern was in conformity with the occurrence tendency of paddy waterlogging provided by local agricultural department as a whole.

By analyzing the attribute of 42 in-situ sample points, except for 2 point of non-paddy, 25 points located in normal paddy range, while 15 points in waterlogged paddy range. Through spatial overlay and analysis of the waterlogged paddy distribution with sample data, the conversion matrix was established according to waterlogging and normal data as follows.

Table 1. Conversion matrix derived from in-situ samples and RS identification

In-situ samples	Remote sensing identification		
	Waterlogged paddy	Normal paddy	total
Waterlogged paddy	13	3	16
Normal paddy	2	22	24
total	15	25	40

From the table 1, the overall accuracy of the model developed in the study reached 87.5%, while the Kappa coefficient reached 0.737.

Table 2.   Yield analysis of in-situ samples of waterlogged paddy and normal paddy

Sample type	Sample number	Average yield (kg/hm^2)	standard deviation
Waterlogged paddy	15	5812	384
Normal paddy	25	10662	562

By analyzing the actual output of 40 samples, the average value of normal paddy was 10662 kg/hm^2, 45 percent over that of waterlogged paddy. According the t-test of independent samples, the *p*-value was below 0.0001, which reached extremely significant level. Results showed that paddy waterlogging reduced the local yield considerably. It also indicated that the change vector analysis model could identify the waterlogged paddy and normal paddy correctly.

5   Conclusion and discussion

Identifying crop waterlogging range timely and correctly is essential for field management afterwards, yield prediction and agricultural insurance. Paddy submerged at filling stage will reduce grain plumpness and affect yield evidently on account of depressed photosynthesis and damaged rice embryos. In this study the change regularity of paddy injured growth was analyzed and the sensitive parameters of growth stress were screened. By using the HJ-1/2 CCD images in the former and later periods of waterlogging, the model of identifying the paddy waterlogging range based on change vector analysis (CVA) was developed. At last the accuracy of the developed model was evaluated by in-situ sample data. Results showed that the waterlogged paddy scattered everywhere in the study, mainly around the Huaihe River system. The spatial distribution pattern was in conformity with the occurrence tendency of paddy waterlogging provided by local agricultural department as a whole. By evaluating the accuracy of the model with in-situ samples, the overall accuracy of the model developed in the study reached 87.5%, while the Kappa coefficient reached 0.737. According the t-test of independent samples, the p-value reached the extremely significant level. Results showed that paddy waterlogging reduced the local yield considerably. These results indicated that the change vector analysis model could identify the waterlogged paddy and normal paddy correctly and efficient in Huaihe River Basin.

The current study requires some further improvements. First, the model of change vector analysis need to test the efficiency of identifying waterlogged paddy in the other region. Second, because the waterlogging scale was relatively small in the study. The accuracy of the method developed in the study needs to be test in great scale of waterlogging.

References

1. Profeti, G., H. Macintosh.: Flood Management through Landsat TM and ERS SAR data: a case study. Hudrological Process. 11, 1397-1408 (1997)
2. Srikantha Herath, Dushmanta Dutta: Flood inundation modeling and loss estimation using distributed hydrologic model. GIS and RS, Proceedings of International Workshop on the Utilization of remote sensing technology to natural disaster reduction, 26-28 October, Tsukuba, Japan: 239-250 (1998)
3. Ingrid Western, Ingmar Messing, Harry Liner, et al.: Controlled drainage-effects on drain outflow and water quality. Agricultural Water Management. 47, 85-100 (2001)
4. Hideo Nakasone, Muhammad Kantar Abbes, Hsiao Kuroda.: Nitrogen transport and transformation in packed soil columns from paddy fields. Paddy and Water Environment. 23, 115-124 (2004)
5. Reddy, B.B., Ghosh, B.C., Panda, M.M.: Flood tolerance of rice at different crop growth stages as affected by fertilizer application. Plant and Soil. 83, 255-263 (1985)
6. Bi Baogui, Jiao Meiyan, Liao Yaoming.: Analysis of precipitation and river f low in the Huaihe River basins during the summer of 2003. Journal of Applied Meteorological Science (in Chinese). 15(6), 681- 687 (2004)
7. ZHAO Lin-Na, YANG Xiao-Dan, QI Dan, et al.: The analysis of precipitation and flooding features in the Huaihe River Basin during the summer of 2007. Climatic and Environmental Research. 12(6), 728-737 (2007)
8. Zhang Aimin, Ma Xiaoqun, Yang Taiming, et al.: The influence of drought and waterlogging disasters on crop yields in Anhui province. Journal of Applied Meteorological Science. 18(5), 69-626 (2007)