CROP AREA ASSESSMENT USING REMOTE SENSING ON THE NORTH CHINA PLAIN

Dong Qinghan[1], Eerens Herman[1], Chen Zhongxin[2]

[1]Flemish Institute for Technological Research (VITO), Boeretang 200, 2400 Mol, Belgium - qinghan.dong@vito.be
[2]Chinese Academy of Agricultural Sciences, P.R. China - zxchen@mail.caas.net.cn

Commission VI, WG VI/4

KEY WORDS: Crop acreage, LANDSAT TM, IRS-P6 AWiFS, SPOT-VEGETATION, Sub-pixel classification

ABSTRACT:

Crop acreage assessment is of particular importance in China where total crop land area is shrinking following the urbanization, directly threatening the policy of grain self-sufficiency. This study demonstrated the application potential of remote sensing to estimate large crop areas on the North China Plain (NCP). The cropping pattern in this plain comprises two growth seasons per year. The harvest of winter wheat is usually followed by the plantation of maize. Two mapping approaches were applied. One consists of the hard classification of multi-temporal high resolution images. Whereas LANDSAT TM data were classically used, the appearance of high resolution sensors with large swath widths such as IRS-P6 AWiFS provides data appropriate for crop mapping on much larger areas. The second approach uses sub-pixel classification of medium or low resolution imagery with the aim to improve the cost-efficiency of assessment. In this study, the sub-pixel classification of SPOT-VEGETATION data was calibrated using the hard classification products of two TM frames of 2005. On the other hand, the classification of two AWiFS scenes of 2007 led to a crop area estimation for nearly the whole study area.

1. INTRODUCTION

Crop area assessment and yield estimation are the two major components of crop production monitoring. In the context of China's mega-growth today, accurate and early crop area estimates have their particular importance. Facing the boost of agricultural commodity prices, Chinese economic minders are worried that fast-shrinking farmland could affect grain supply in the near future. According to the official statistics, China's total cropland area was dropping in 2006 (http://www.agri.gov.cn/), looming close to the benchmark 120 million hectares considered necessary to sustain national food security. Therefore a cost-efficient area assessment methodology providing early estimations is of great importance for the region.

The North China Plain, or Huabei Plain, is the largest alluvial plain of eastern Asia (Fig. 1). This smooth, yellow soil plain is the main producing region of wheat, maize and cotton in China. The three most important wheat-producing provinces, namely Shangdong, Henan and Hebei, are located in the plain and they are responsible for about half of the entire Chinese wheat production. The North China Plain is also one of the regions with the highest population density. The region of interest of this study includes the provinces Beijing, Tianjing, Shandong, Hebei, Henan and Shanxi. In total they comprise sixty districts which are the basic administrative unit of the area assessment in this study.

The predominant cropping system applied in the region consists in the annual repetition of winter wheat and summer maize. The winter wheat is planted in the beginning of November and harvested at the end of May. The season of maize begins in June and ends in October. Other major agricultural crops include cotton, vegetables and orchards.

Classical high resolution imagery such as LANDSAT TM (30m resolution) can produce spatial cropping details. However with a low temporal resolution (16 days for Landsat), these sensors produce image data with low ability to discriminate the crops based on phenology. As the coverage per scene of these sensors is also limited, crop mapping on large areas using high number of scenes is not a cost-efficient approach. On the other hand , the coarse resolution sensors such as NOAA-AVHRR, SPOT-VEGETATION, EOS-MODIS or ENVISAT-MERIS have near-daily global coverage, but their low spatial resolution pixels (250m to 1km) mostly cover several fields with different land-use classes or crops. It becomes therefore difficult to assign each pixel univocally to one single class.

Two approaches exist to achieve a crop mapping at larger scale with relatively higher accuracy. Fuzzy or sub-pixel classification tries to avoid this "one pixel–one class" assignment by determining the area fraction covered by each land-use class or crop in coarse (and most often mixed) pixels. Although the spatial locations of the class fractions within each mixed pixel are unknown, the percentage or the distribution of each class may well be evaluated. The second approach takes advantage of the high resolution imagery provided by a wide-swatch sensor such as IRS-P6 AWiFS.

In this paper, the application of both approaches to the North China Plain are discussed. The sub-pixel classification of 1km resolution SPOT-VEGETATION time series for the year 2005, by means of neural networks, was carried out in two stages: the detailed crop mapping on a limited area (calibration area) was

* Contact: qinghan.dong@vito.be

first performed using a Landsat TM dataset. This detailed information was then used in the second stage as reference data to calibrate the neural network. On the other hand, crop mapping for the year 2007 using two IRS-P6 AWiFS scenes is described.

The software package GLIMPSE developed by VITO and commercial packages ENVI (ITT Visual Information Solution) and ArcGIS (ESRI) were used in this study.

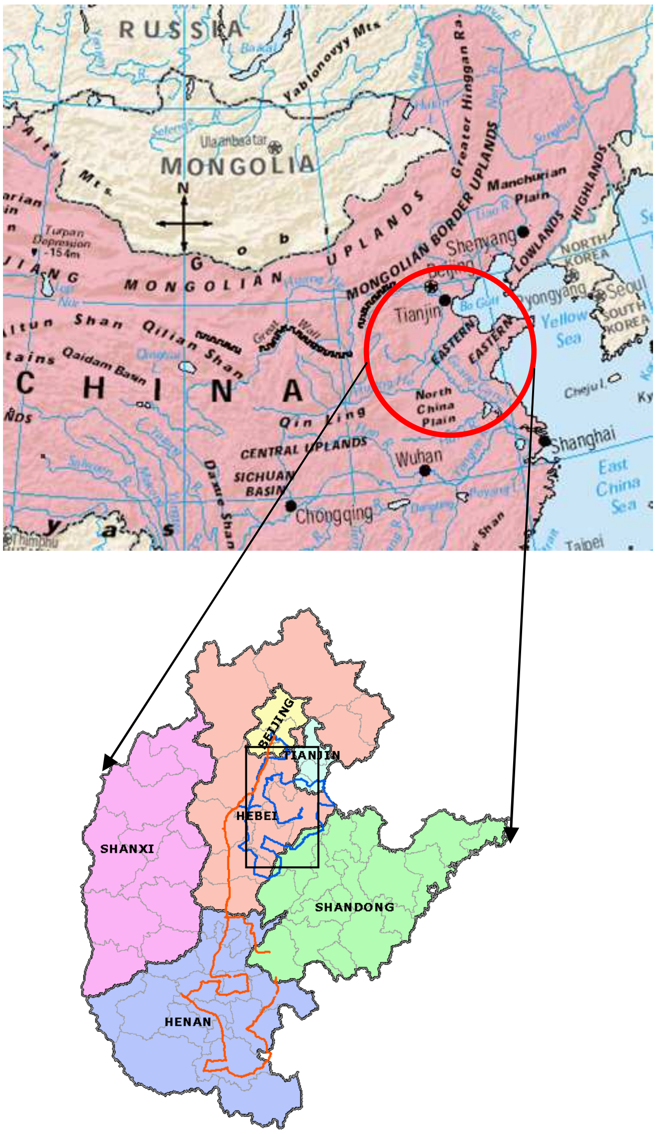


Fig. 1 Location of the North China Plain (Huabei Plain). The region includes six provinces or sixty counties. The figure also shows the tracks of the field surveys conducted in 2005 (blue) and 2007 (orange). The black box roughly indicates the area covered by the two Landsat TM frames (NE Hebei and SE Hebei, see also Fig. 4), which are rather spatially limited.

2. DATA AND METHODS

2.1 Data

2.1.1 Image data: The high resolution images used in this study are listed in the Table 1. The Landsat TM sensor produces image scenes at 30m resolution and with a coverage of 180km x 180 km. Three spectral bands (RED, NIR, SWIR1) out of seven running by the TM sensor were used. The IRS-P6 AWiFS sensors operates in four spectral bands and has a ground swath of 740 km. The similar band combination (RED, VNIR, SWIR) was selected for classification purpose.

The coarse 1km resolution time series used for sub-pixel classification were extracted from the 10-daily composites of SPOT-VEGETATION (VGT-S10) kindly delivered by the EC-Joint Research Centre at Ispra, Italy.

Table 1. Landsat and IRS-P6 scenes used

Sensor	Coverage	Path	Row	Registration date
Landsat 5 TM (30 m)	NE Hebei	123	33	2005-05-06
				2005-05-22
				2005-07-25
	SE Hebei	123	34	2005-05-06
				2005-05-22
IRS-P6 AWiFS (56 m)	NCP	131	046	2007-03-04
		132	046	2007-02-05

2.1.2 Survey data: The ground truth data needed for the calibration of supervised classifications were collected in two field surveys in 2005 and 2007 (Fig. 1). The Trimble Geoexplorer 3 GPS handset and the associated software Pathfinder Office 2.70 were used for the surveys. During the field surveys, large parcels or plots corresponding different land-use classes or crops were selected and digitalized to form a ground truth vectorial GIS. As mentioned earlier, double cropping system is practiced on the North China Plain, The dominant plantation of wheat in winter and spring is followed by the cultivation of maize in summer. Two mappings per year are therefore more appropriate to estimate the areas of all crops.

From the field survey of 2005 that took place in the province Hebei, 231 polygons were registered representing 14 classes of land use or crop types: maize, cotton, sorghum, rice, cash crops (vegetables), herbs for medicine, wetland, grassland, water, orchards, poplar, built-up, waste-land and saltpan. Observation was made that the sizes of parcels were very small in the region and mixed cropping pattern (mixed crop types on one parcel, or even mixed agro-forestry structure) was quite generalised. The survey of 2007 concentrated on the province of Henan enabled for the registration of 282 ground parcels distributed in 11 land-use or crop classes: wheat, garlic, brassica, rice, vegetables, deciduous forest, conifer forest, waste-land, bare-soil, built-up and water. In all cases, more polygons representing stable land-use classes (built-up, water) were added by means of visual inspection of satellite images.

2.1.3 Pre-processing of high resolution images: The Landsat TM data were pre-processed at the levels of calibration, atmospheric correction and geometric rectification. three classification layers (RED, NIR, SWIR1) of different registrations were extracted to form a "multivariate images set" ready for the classification.

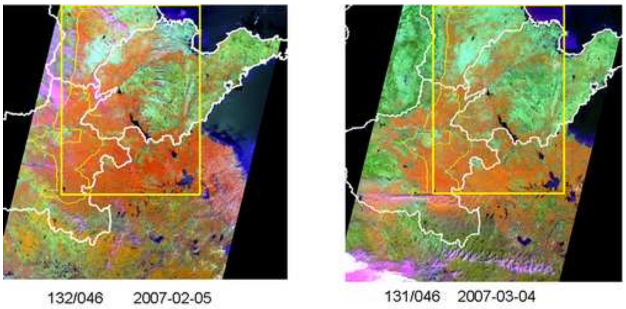


Fig. 2 The overlapping region from two neighbouring AWiFS frames covering a large part of the Plain was extracted .

The pre-processing of AWiFS images (Table 1) consisted of the extraction of the overlapping region from two neighbouring frames (Fig. 2). The band combination (RED, VNIR, SWIR) of two registrations formed a multi-spectral and multi-temporal dataset. The geometric correction was carried out using the GPS-tracks collected during the field survey as well as terrain-corrected TM images.

For the coarse resolution images of SPOT-VEGETATION, NDVI monthly composites (S30) were computed according to the maximum NDVI-criteria. Nine of these S30, from March till November 2005, were used in the sub-pixel classification analysis.

2.2 Methods

2.2.1 Supervised classification of multi-temporal high resolution imagery: The general approach for supervised classification is outlined in Fig. 3. Different registrations for a specific sensor, after pre-processing, form a "multivariate image set". These superimposed data contain two types of information: spectral-temporal and spatial-contextual. The spectral-temporal information can be extracted with a supervised Maximum Likelihood algorithm (Duda *et al*. 2000), resulting in a per-pixel classification. The spatio-contextual information can be derived by means of a segmentation procedure. On the ideal situation, these segments correspond with parcels on the grounds. In this study the segmentation was applied with the commercial package eCognition (Definiens Imaging GmbH). In the final stage, the outputs of the per pixel classification and the segmentation are combined using a per-segment mode filter. This procedure determines for each parcel the pre-dominant class and assign all pixels of the parcel to this modal class. By reducing speckle and errors in the vicinity of field boundaries, this application enhances the accuracy and legibility of the final map (Eerens *et al.*, 2003).

2.2.2 Sub-pixel classification of SPOT-VEGETATION time series: Sub-pixel classifications aim to evaluate the proportion of each land-use or crop class in a coarse pixel. This can be namely achieved using neural networks, which should be first calibrated or trained with the reference data. These reference data are usually land–use maps derived from a vectorial GIS or a raster classification product covering the calibration zones. The reference datasets are subsequently divided into two subsets (CAL and VAL) for calibration and validation purposes respectively. The input data are composed of time series of a coarse resolution imagery. Reference reviews in this domain have been made for example, by Paola and Showengerdt (1995) or by Atkinson and Tatnall (1997).

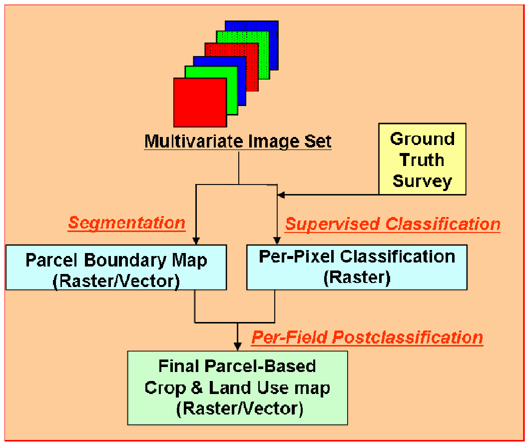


Fig. 3 The flowchart followed for the supervised classification of high resolution images (from Eerens et al. 2003)

In order to calibrate neural networks, "area fraction images" or AFI's of the reference data are created as reported in details by Verbeiren (2008). These AFI's are spatially congruent with the input images of coarse resolution sensors. They have thus same pixel size (at coarse resolution) and represent the area fraction occupied by a specific class in the pixels. The calibration step establishes a "relationship" between the land-use distribution of the reference data and the spectral proprieties of the input times series by means of neural networks weights. Only the subset of the reference data CAL is used for the calibration.

The extrapolation of the calibrated neural network produces the estimated area fractions for the entire region.

Quantitative validation can be carried out by means of class-specific scatter plots, computing the linear regression $REF=a+b \cdot EST$, where REF and EST represent respectively the values of the reference and the estimated area fractions. Only the pixels in the subset VAL are included in this step.

3. RESULTS AND DISCUSSION

3.1 Supervised classification

3.1.1 Supervised Classification of TM images (2005):

The results of the classification on the frames of TM "NE Hebei" and "SE Hebei" from the year 2005 are discussed here. The Maximum Likelihood classifier was applied on the multi-temporal datasets of the TM sensor (Table 1) to perform a classification on pixel level, after the pre-processing described above.

Maps with 151235 and 14937 segments were retrieved by the respectively for the frames "NE Hebei" and "SE Hebei". Per field mode filter was applied for the per pixel classification for both frames. The per-field classification was achieved by the application of the filter on the product of per-pixel classification (Fig. 4). 14 land-use or crop classes were represented in the final map.

Unfortunately no solid validation for this supervised classification could be carried out. Theoretically, the ground truth pixels have to be divided into two sets (as performed with the sub-pixel classification), one is used for the calibration purpose, the validation procedure has to be carried out with the second set. However insufficient ground truth data were available for an independent validation approach. The back-validation as shown in Table 2 (only for the frame "NE Hebei") was the only solution for an accuracy estimation. The overall accuracy resulted in the back-validation for the classification on the frame "NE Hebei" amounted to 88%, which was probably overestimated since the same set of the ground truth data was used for calibration, as well as for validation. However from the table, it can be concluded that the major crop or land-use classes, such as maize, cotton, water, and built-up, were correctly classified. On the other hand, confusions or errors remained with minor classes, for example with sorghum, vegetables, herbs and to some lesser extent with orchards.

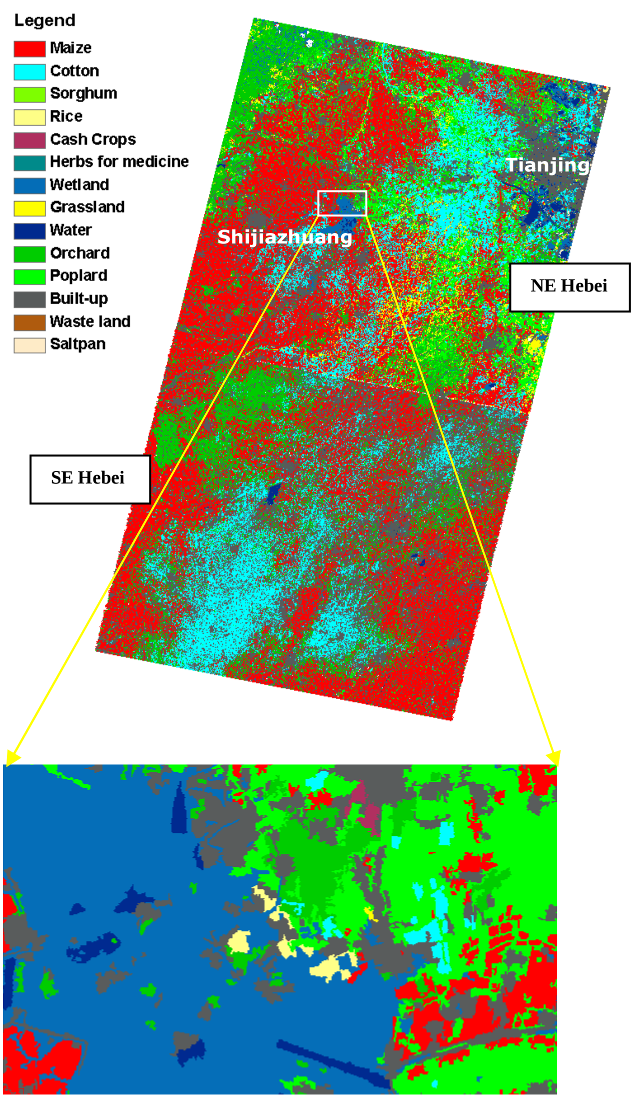


Fig. 4 Per-field classification of two Landsat TM frames "NE Hebei" and "SE Hebei" resulted in two crop maps, which include 14 land-use or crop classes.

3.2 Sub-pixel classification of SPOT-VEGETATION time series

An elementary back-propagation neural network was applied for this sub-pixel classification analysis. The reference data necessary for calibrating the neural network were derived from the products of the supervised classification of two LANDSAT TM frames described above. The classes derived from the TM classification were combined into seven groups: maize, cotton, trees, grassland, built-up (urban), wetland and water. Seven AFI's corresponding to these groups were accordingly created (Fig. 5). The reference dataset was subsequently divided in two subsets: CAL including 10% or 4796 randomly selected pixels and VAL containing 90% remaining pixels.

As the input data was composed of NDVI-S30 time series, the calibration step, using the CAL subset, tried to make connection between the temporal reflectance pattern of the NDVI time series and the land-use composition (area fractions) in the calibration area. This was achieved by defining neural network weights. In this application the input layer comprised 9 nodes corresponding 9 monthly NDVI's. The output layer contained 7 nodes corresponding 7 estimated area fractions.

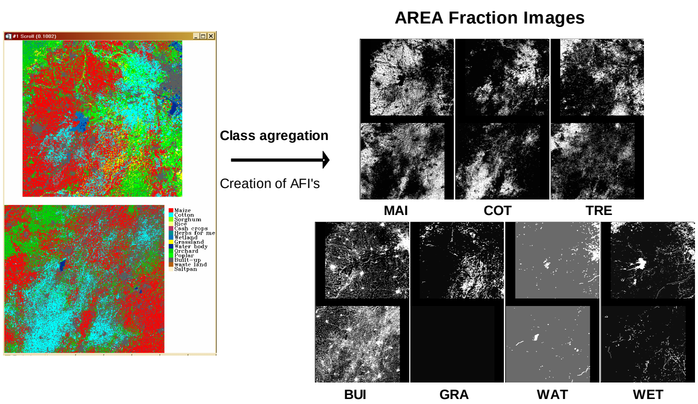


Fig. 5 Generation of reference Area Fraction Images (AFI's) starting from the hard classification (MAI=maize, COT=cotton, TRE=trees, BUI=urban areas, GRA=grassland, WAT=water bodies, WET=wetland). .

The results of the neural network application on the entire region as shown in Fig. 6A suggested that the sub-pixel classification approach could reproduce the general land-use or cropping patterns on the North China Plain. Comparison has been made between the estimated and the reference (or "true") AFI's for the two main summer crops, maize and cotton (Fig. 6B).

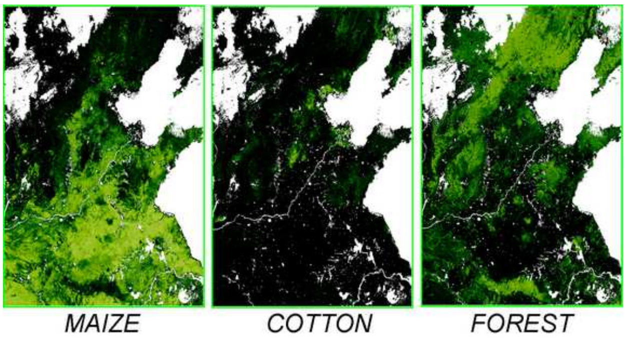
The process of quantitative validation, by means of class-specific scatter plots (REF=a+b·EST) (Fig. 7), was conducted on two different spatial levels:

- 43159 1km^2-resolution pixels
- 60 districts with mean surface of 11522km^2

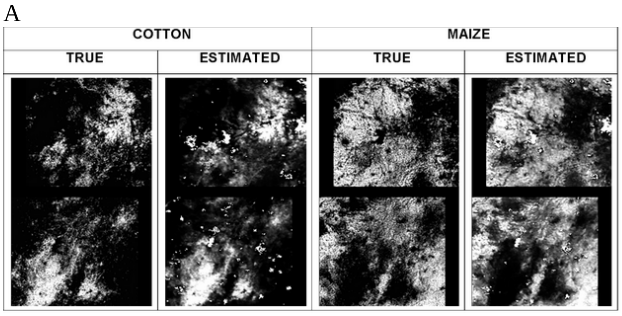
In the later case, the regressions were performed on the regional mean area fractions which were obtained by averaging all pixel values in one district. The sum of the class fractions amounted to one.

Table 2: Error matrix of the per-field classification for the frame "NE Hebei" in percentage, derived from the 10900 ground truth pixels

ESTIM. \ TRUE	MAI	COT	SOG	RIC	VEG	HER	WET	GRA	WAT	ORC	POP	BUI	WAS	SAL	Σ(TRUE)
MAIZE	**13.8**					1.8	0.5			1.0	0.3	0.1			**17.5**
COTTON	0.4	**15.7**	0.4				0.1	0.4			0.2	0.2			**17.4**
SORHGUM															**-**
RICE				**1.4**											**1.4**
VEGETABLES															**-**
HERBS															**-**
WETLAND							**4.4**		2.3						**6.7**
GRASSLAND								**1.1**		0.1					**1.2**
WATER									**24.5**						**24.5**
ORCHARD										**4.7**	0.2				**4.9**
POPLAR	0.4								0.4	1.2	**4.3**				**6.3**
BUILT-UP					0.3		0.9	0.1	0.1	0.4		**17.5**			**19.4**
WASTE LAND													**0.3**		**0.3**
SALPAN									2.2					**0.5**	**0.5**
Σ(ESTIMATED)	**14.6**	**15.7**	**0.4**	**1.4**	**0.3**	**1.8**	**5.9**	**1.6**	**27.4**	**7.4**	**5.0**	**17.9**	**0.3**	**0.5**	**100.0**



A



B

Fig. 6 (A) Estimated area fraction images covering the whole study area for the classes maize, cotton and forest. (B) Comparison between the reference and estimated area fractions for cotton and maize on the calibration zone covered by the reference data (Landsat TM frames).

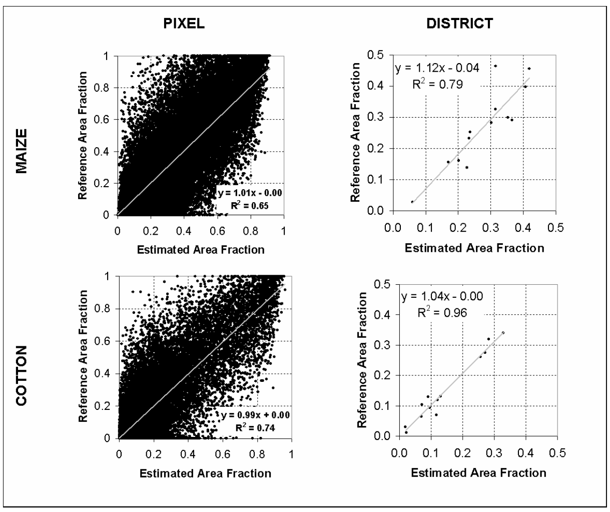


Fig. 7: Scatter plots and linear regressions (REF=a+b·EST) for the validation of maize and cotton at pixel (left) and district (right) level over the calibration area.

At the pixel level, two major summer crop classes, maize and cotton, had the two highest R^2 scores for their area estimates, around 0.7. However the accuracy increased when the pixel estimates were aggregated to district level. The R^2 score for cotton area estimate for example rose from 0.74 to 0.96 (Table 3).

Table 3 Coefficient of determination R^2, intercept a and slope b of the regressions, between the reference area fractions and the estimates delivered by the neural network

	PIXEL			DISTRICT		
	R^2	a	b	R^2	a	b
MAIZE	0.67	0.10	0.66	0.79	-0.04	1.12
COTTON	0.74	0.04	0.76	0.96	-0.00	1.04
FOREST	0.47	0.11	0.50	0.70	0.03	0.89
GRASSLAND	0.22	0.02	0.20	0.83	-0.01	1.21
BUILT-UP	0.45	0.16	0.48	0.95	-0.01	0.99
WATER	0.01	0.00	0.01	0.71	0.00	3.99
WETLAND	0.35	0.00	0.42	0.70	-0.01	1.76

3.2.1 Supervised Classification of AWiFS images (2007)

The supervised multi-temporal classification for AWiFS data of 2007 is presented. Two registrations of 2007-02-05 and 2007-03-04 were used in this analysis as described previously.

The ground truth raster was composed of the ground truth parcels collected during the field survey of 2007 complemented with plots discerned from visual inspection over the AWiFS images. In total, 63 polygons representing 8 classes (wheat, cotton, garlic, vegetables, orchard, conifer, urban and water) were included. The application of the Maximum Likelihood algorithm led to the classification as shown in Fig. 8, covering an area about 450km x 600km.

No quantitative validation has been performed. However the general cropping pattern as displayed in Fig. 8 was confirmed by the multi-annual field surveys and by an alternative crop map derived from two Landsat TM frames of year 2007 covering the eastern part of Henan province (data not shown):

wheat is the predominant crop especially in Henan province. In the south of Hebei province three major agricultural vegetations co-exist: wheat, cotton and orchards.

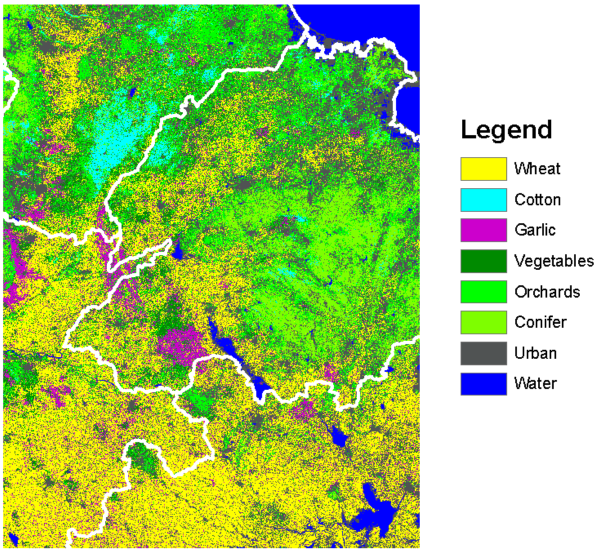


Fig. 8 Supervised classification of two superposed AWiFS registrations of 2007

In conclusion, the 30m-resolution data generated by the Landsat 5 TM sensor, even after 24 years of service, still provide precious Earth's surface feature information for crop mapping. The combination of the Maximum Likelihood classifier extracting the spectral information and the image segmentation procedure extracting textual information of the satellite images, yielded a crop map with improved legibility. However this classical approach is rather inefficient for crop mapping at large scale, because of low temporal resolution (thus insufficient data availability) and the limited swath width of the sensor. Two alternative approaches were adopted in this study to avoid these inconveniences.

The imagery provided by the new generation sensors such as IRS-P6 AWiFS could help to establish crop maps for large areas. As the cloud interference is almost inevitable for image covering large regions, improvement has still to be made in the aspect of treatment of pixels with deficient measurement vectors (missing values in one or more registrations).

The second alternative solution was the sub-pixel classification of low spatial resolution imagery. Its potential for large crop acreage estimation was confirmed in this study. In this two-stage analysis, the reference data were first derived from the classification products of Landsat TM data. These reference data were subsequently used for calibration of the neural network which was established for the sub-pixel classification of 1km SPOT-VEGETATION data. The results of the sub-pixel classification revealed that the accuracy of the area estimation at pixel level was relatively low, but that can be significantly ameliorated when a spatial aggregation to more relevant geographical unit (such as district) was applied.

One limitation of these two alternative procedures concerns the representativity of the training or calibration pixels. The representativity can only be insured when the application area is not too extended in comparison with the calibration area. Otherwise a spatial stratification must be in advance carried out and each stratum has to be trained with calibration pixels representative for its own.

References:

Atkinson, P.M., Tatnall, A.R.L., 1997. Neural networks in remote sensing. *Int. J. Remote Sens*. 18, pp. 699-709.

Duda, R., Hart, P., Stork, D., 2000. *Pattern classification (Second Edition)*. Wiley-Interscienec, 654p.

Eerens, H., Dong, Q., Ozer, P., Li, X., 2003 *A field-based crop and land use map over Sanjiang plain in Heilongjiang derived from multi-temporal imagery of Landsat7-ETM+*. Cybergeo: European Journal of Geography, N.247. http://www.cybergeo.eu/index4020.html (accessed 19 Sep. 2003).

Paola, J.D., Showengerdt, R.A., 1995. A review and analysis of back-propagation neural networks for classification. *Int. J. Remote Sens*. 16, pp. 3033-3058.

Verbeiren, S., Eerens, H., Piccard, I., Bauwens, I., Orshoven, J., 2008. Sub-pixel classification of SPOT-VEGETATION time series for assessment of regional crop areas in Belgium. *Int. J. Applied Earth Observation and Geoinformation* (in press).

Acknowledgements

The authors would like to thank the Institute for the Protection and Security of the Citizen, Joint Research Centre of European Commission for delivery of SPOT-VEGETATION imagery and the EUROPEAID Office of the European Commission for the financial support towards HUBEI-CGMS Action (2005-2007). This research was also partially granted by the National High Technology Research and Development Program of China (863 Program Nos.2006AA12010101 and 2006AA12Z103) and National Key Technologies R&D Program of China (No. 2006BAD10A06).