

**A FIELD-BASED CROP AND LAND USE MAP OVER SANJIANG PLAIN IN HEILONGJIANG
DERIVED FROM MULTI-TEMPORAL IMAGERY OF LANDSAT7-ETM+**

Herman EERENS¹, Qinghan DONG¹, Pierre OZER² and Xiufen LI³

¹ Vlaamse Instelling voor Technologisch Onderzoek
Centre for Remote Sensing and Atmospheric Processes
Boeretang 200, B-2400 Mol, Belgium

Tel: ++32(0)14336844; Fax: ++32(0)14322795; Email: herman.eerens@vito.be

² Fondation Universitaire Luxembourgeoise
Avenue de Longwy 185, B-6700 Arlon, Belgium
Tel: ++32(0)63 230975, Fax: ++32(0)63 230800

³ Heilongjiang Province Institute for Meteorological Sciences
Diantan Street 71, Xiangfang District, Harbin, China, 150030
Tel.: ++86-451-5320631, Fax: ++86-451-5301013

Abstract

Multitemporal Landsat7-ETM+ imagery and supplementary field survey data were used to establish a crop and land use map for an area of $\pm 20\,000$ km² of Sanjiang Plain in China's north-eastern province Heilongjiang. The "field-based" map is based on data of the year 2002 and is created by combining the results of a supervised Maximum Likelihood classification with the estimated field boundaries extracted by an automatic segmentation procedure. No sufficient ground truth was available to perform an independent validation, and some classes are certainly confused. Nevertheless, given the minimal investments and the unexplored potentials for improvement, the approach seems very promising and cost-effective. By superimposing the classification with the map of the administrative counties, regional statistics were derived with the acreage distributions of the concerned crops and land use types.

Keywords: Maximum Likelihood classification, segmentation, post-classification, per-field filter, CGMS, China.

Résumé

L'imagerie multitemporelle Landsat7-ETM+ et des données complémentaires acquises durant des missions de terrain sont utilisées pour établir une carte d'utilisation du sol à des fins agronomiques sur une région d'approximativement 20 000 km² dans la plaine de Sanjiang, au nord-est de la province de Heilongjiang, Chine. La carte se base sur des données de terrain récoltées durant la saison agricole 2002 et est créée en combinant les résultats d'une classification supervisée basée sur un maximum de probabilité avec une estimation des limites des parcelles extraite par une procédure de segmentation automatique. Un nombre suffisant de données de vérification terrain a posteriori n'était pas disponible pour réaliser une validation indépendante de la carte d'utilisation du sol et il est donc probable que certaines classes soient confuses. Cependant, étant donnés les investissements financiers limités et les progrès potentiels de cette méthodologie, l'approche semble très prometteuse et rentable. En superposant cette classification avec la carte des comtés administratifs, des statistiques régionales ont pu être obtenues avec les distributions des superficies vouées à différents types d'utilisation du sol et de cultures.

Mots clés: classification par maximum de probabilité, segmentation, post-classification, filtre par parcelle, CGMS, Chine.

1. INTRODUCTION

Heilongjiang, China's most north-eastern province, has an area of 460 000 km² and a population of about 40 millions. The moderate monsoon climate is characterised by harsh, dry winters, and warm, wet summers. The province has a wealth of natural resources. The low mountain areas host China's largest forest reserves, while huge amounts of soybean, rice, maize and spring wheat are produced on the rich black soils of two vast, alluvial plains: Songnen Plain in the Centre to Southwest and Sanjiang Plain in the Northeast. The short summer only allows for a single harvest per year, and channel irrigation is mainly applied for rice. The province is administratively subdivided in 79 "counties" (see figure 1).

Given the importance of the agricultural sector for Heilongjiang and for entire China, it is obvious that provincial and national authorities give high priority to the correct and timely monitoring of the crop productions. As a department of the Heilongjiang Province Weather Bureau in Harbin, the Heilongjiang Province Institute of Meteorological Sciences (HPIMS) is responsible for the yield forecasting of the main crops at the levels of the counties and the entire province. To this goal, it developed a statistical model, mainly driven by weather data (Zu, 1998; Jiang *et al.*, 2003). Since 2001, HPIMS also collaborates with three Belgian institutes (VITO, FUL and CRA) to implement the European yield forecasting system CGMS in Heilongjiang (Tychon *et al.*, 2003).

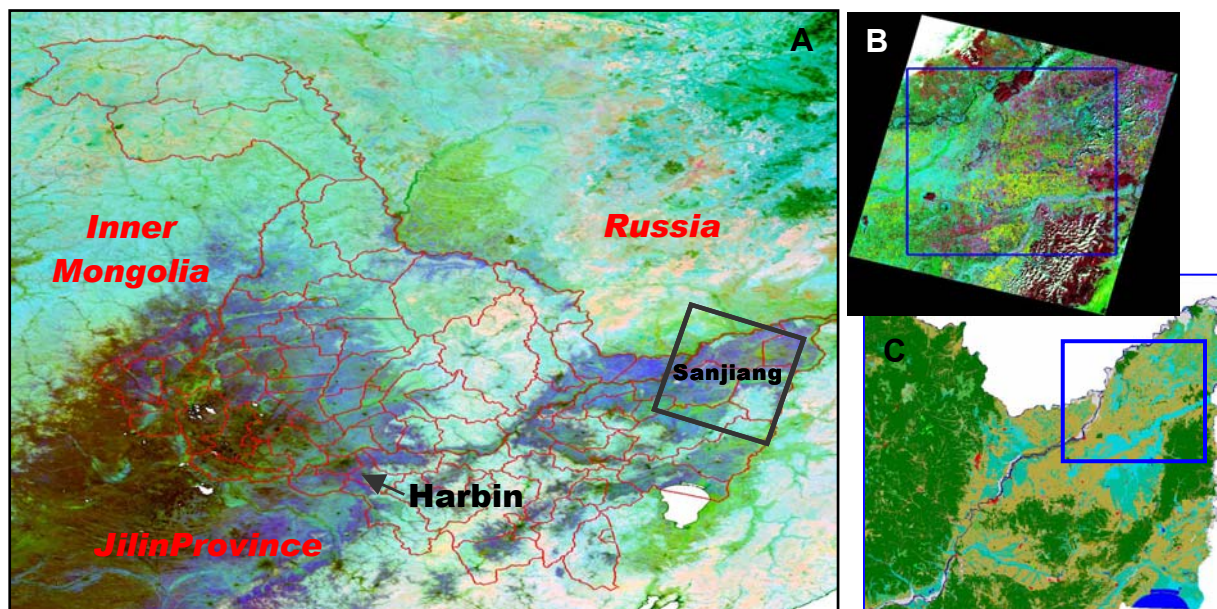


Figure 1: A. Heilongjiang province: 79 counties and position of the studied Landsat7-frame (114/027) in Sanjiang Plain. B. Original Level1G registration with the position of the classified area (blue rectangle). C. Extract of the general land use map of 1995 (Chinese Academy of Sciences).

Obviously, the production for each crop is equal to the product of the cultivated area with the mean yield level. Unfortunately, crop acreage information for the Heilongjiang Province is

scattered over different departments and hard to obtain. The Chinese Academy of Sciences (CAS) developed a general land use map of China, on the base of monotemporal imagery registered throughout the year 1995 by the high-resolution satellite sensor Landsat5-TM. The CAS map is quite similar to the CORINE land cover map of the European Union, but for this agricultural application it is too generalised and outdated. Hence the Sino-Belgian collaboration was extended with a feasibility study on the potentials of the more recent sensor Landsat7-ETM+ for the production of detailed crop and land use maps, from which the requested area statistics can be derived.

During this first phase the procedure was tried out on a single test area in the north-eastern Sanjiang Plain. As indicated by figure 1, Sanjiang Plain is wedged between the Heilongjiang river (Amur) and its southern affluent Wusuli (Ussuri), which respectively form the northern and eastern Sino-Russian borders. The area is rather scarcely populated and mainly occupied by agricultural cropland and remnants of the original wetlands, while the hills in the south-east are still covered by deciduous forest massifs.

2. GENERAL METHODOLOGY

The methodology is outlined in figure 2. Multispectral registrations of different dates are acquired and separately submitted to a number of enhancements at the levels of calibration, atmospheric correction, geometric rectification, etc. All pre-processed scenes can be superimposed to form a “multivariate image set”, which contains two types of information: spectro-temporal and spatio-contextual. The spectro-temporal information is extracted by the supervised classification, which results in a first raster: the per-pixel classification. The segmentation extracts the spatio-contextual information contents and creates a map (in raster or vector format) with the delineation of all uniform image zones (segments). In the best case, these segments should correspond with parcels on the ground. The post-classification per-field filter combines both results, and creates a segment-based (or field-based) map with enhanced accuracy and legibility. This final result can be used in various applications, including the extraction of regional crop acreage statistics.

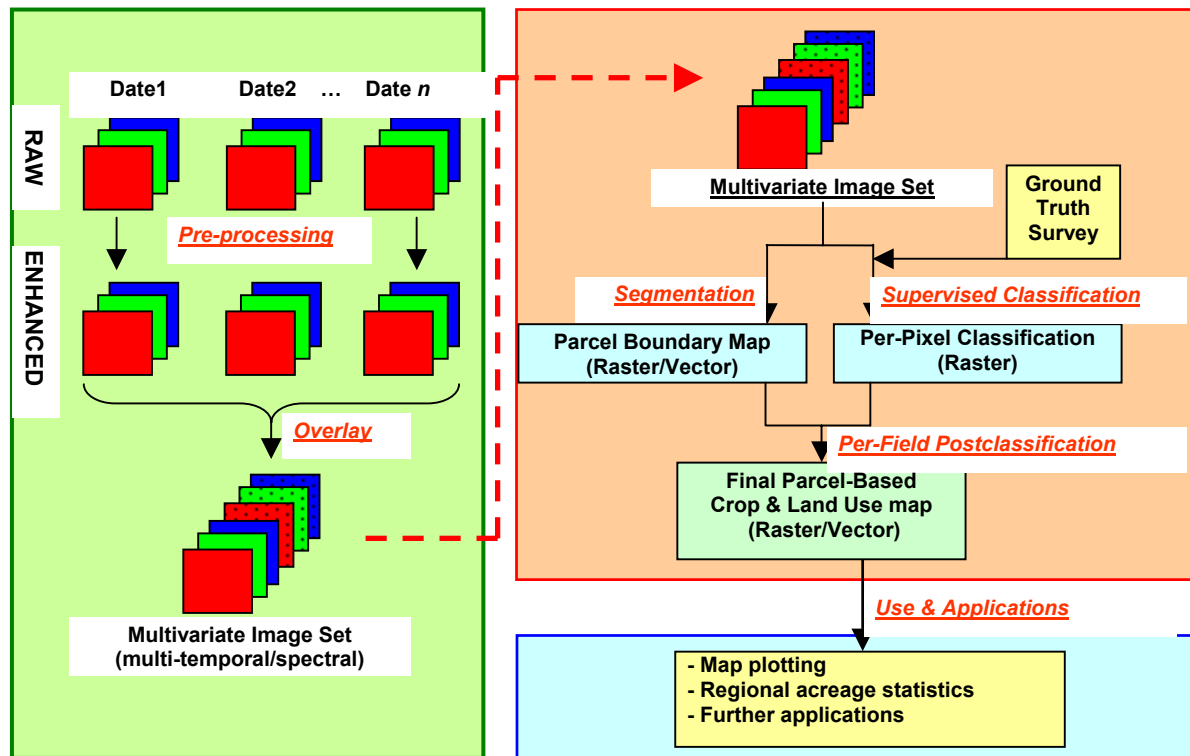


Figure 2: General flowchart of the image-based procedure for crop and land use mapping.

In practice, the per-pixel classification was created with the Maximum Likelihood algorithm. Its principles are briefly summarised below (see for instance: Duda & Hart 1973).

Suppose the pre-processed, multivariate image set comprises N_v image layers, labelled as $v=1 \dots N_v$, while the number of pixels amounts to N_p ($p=1 \dots N_p$). Each pixel p is thus associated with a measurement vector $\mathbf{x}_p = \{x_{p,1}, x_{p,2}, \dots, x_{p,N_v}\}$, which represents its set of image values, and which points to a specific location in the N_v -dimensional measurement space. The classification has to assign each pixel to one (and only one) of the N_k (mutually exclusive) classes in the used legend or key ($k=1 \dots N_k$). In Bayesian decision theory, this is achieved by computation (per pixel) of N_k aposteriori probabilities $P_k(\mathbf{x}_p)$, and assignment of the pixel to the class (k_{est}) with highest aposteriori value:

$$P_k(\mathbf{x}_p) = A_k(p) \cdot f_k(\mathbf{x}_p) \qquad P_{k_{est}}(\mathbf{x}_p) = \underset{k=1}{\overset{N_k}{MAX}} \{P_k(\mathbf{x}_p)\}$$

The decision is thus based on two types of inputs. $A_k(p)$ represents the apriori knowledge which is often available, either from earlier classifications or from agro-statistics. If not, all classes receive equal apriori weight ($A_k(p)=1$ for all classes). On the other hand, $f_k(\mathbf{x}_p)$ is the probability of obtaining the image measurement \mathbf{x}_p , if the concerned pixel p would belong to class k . In general, $f_k(\mathbf{x}_p)$ is a N_v -dimensional distribution function which describes the specific clustering of the pixels of class k over the measurement space. For instance, for the vector $\mathbf{x}_p = \{117, 33, \dots, 243\}$ one might find that $f_1(\mathbf{x}_p)=0.1$ and $f_2(\mathbf{x}_p)=0.00001$, which indicates that this measurement is very likely for pixels of class 1 but rather exceptional for those of class 2. The aposteriori probabilities $P_k(\mathbf{x}_p)$ thus represent the chance that pixel p belongs to class k , given the apriori weight of that class and the plausibility of the measurement \mathbf{x}_p . The pixel is then assigned to the class k_{est} with the highest aposteriori probability.

In practice, the Bayesian decision theory is mostly implemented via the Maximum Likelihood classifier, which assumes that each class' specific clustering can be approximated (and

mathematically described) by means of a multinormal density function (the multivariate version of the classical univariate Gauss distribution):

$$f_k(\mathbf{x}_p) = \frac{\exp(-r^2/2)}{(2\pi)^{N_v/2} \cdot |\mathbf{C}_k|^{1/2}} \quad \text{with} \quad r^2 = (\mathbf{x}_p - \boldsymbol{\mu}_k)^T \cdot \mathbf{C}_k^{-1} \cdot (\mathbf{x}_p - \boldsymbol{\mu}_k)$$

Here $\boldsymbol{\mu}_k$ is the mean vector of class k , which points to the class centroid in the N_v -dimensional measurement space, and \mathbf{C}_k is its covariance-matrix, a symmetrical ($N_v \times N_v$)-matrix with the univariate variances on the diagonal and the bivariate covariances on the off-diagonal places. \mathbf{C}_k^{-1} is the inverse and $|\mathbf{C}_k|$ the determinant of this covariance matrix. The deviation vector $(\mathbf{x}_p - \boldsymbol{\mu}_k)$ – or its transposed version $(\mathbf{x}_p - \boldsymbol{\mu}_k)^T$ – shows the difference between the concerned measurement \mathbf{x}_p and the centroid of class k . Finally, r^2 is the squared mahalanobis distance, a statistically weighted, non-euclidean distance which takes into account the specific clustering of the class.

Samples drawn from a multinormal distribution tend to fall in a single cloud or cluster. The centre of the cluster corresponds with the mean vector $\boldsymbol{\mu}_k$, while its shape is determined by the covariance-matrix \mathbf{C}_k . The iso-probability contours are hyper-ellipsoids (if $N_v=2$: ellipses) with constant mahalanobis distance to $\boldsymbol{\mu}_k$. The volume of these hyper-ellipsoids is a measure for the scatter around the mean, and it is proportional with the determinant $|\mathbf{C}_k|$. The Maximum Likelihood algorithm thus assumes that the specific clustering of each class k can be fitted fairly well with such a multinormal density function.

The multinormal density $f_k(\mathbf{x}_p)$ is completely defined by $N_v + N_v(N_v+1)/2$ parameters, i.e. the elements of the mean vector $\boldsymbol{\mu}_k$ and the independent elements of the covariance-matrix \mathbf{C}_k . In this supervised classification variant, these parameters are estimated in advance, and separately for each class, from a sample of training pixels with known land cover type (typically 0.1% to 1% of the total population of image pixels). The knowledge gained in this calibration step is then extrapolated over the entire image by application of the above Bayesian decision equations.

For general land cover classifications with a limited number of broad categories (water, forests, cropland,...), the required training pixels can easily be defined and delineated by photo-interpretation of the pre-processed imagery. However, the production of crop maps with more elaborate legend systems requires more detailed reference data, which mostly have to be acquired by means of dedicated field surveys (“ground truthing”). For this classification, the study area in Sanjiang Plain was effectively surveyed in June 2002 (Fig. 3).

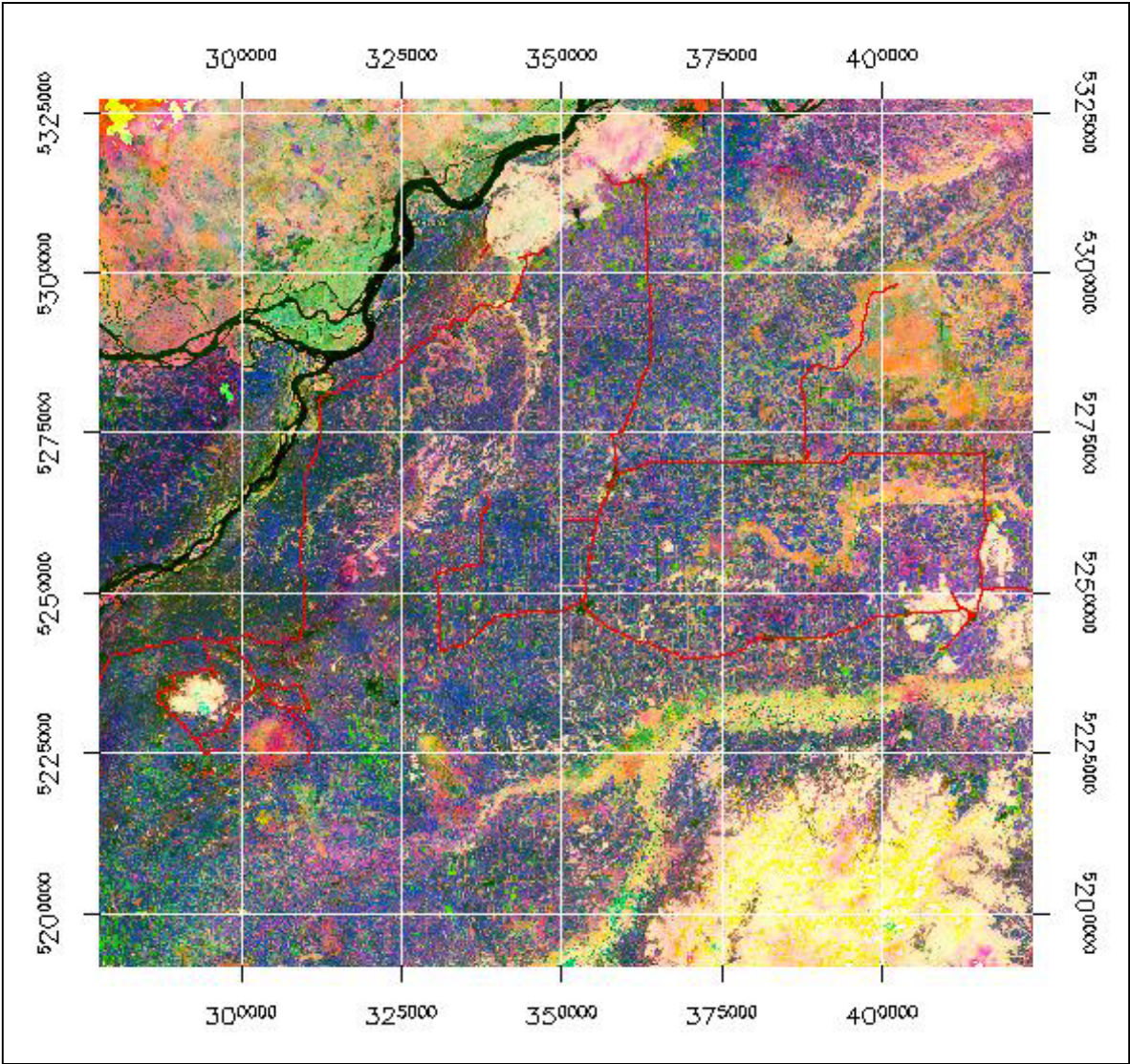


Figure 3: Output frame for the cartography in Sanjiang Plain (WGS84-UTM53N projection, 30 m resolution). In overlay the GPS-registered trajectory followed during the field survey (21-25 June, 2002).

3. PRE-PROCESSING OF LANDSAT7-ETM+ IMAGERY

ETM+ (Enhanced Thematic Mapper Plus) is a high resolution, multispectral, imaging sensor on board of the orbital satellite Landsat7, which was launched in 1999. As indicated in table 1, ETM+ provides multispectral imagery in 8 spectral bands, with resolutions (pixel sizes) ranging between 15m and 60m. Each registration covers an area of about 180km x 180km (32 200 km²).

Table 1: Spectral bands of the sensor Landsat7-ETM+ (NIR, SWIR, TIR = Near, ShortWave and Thermal InfraRed).

Spectral band	Bandwidth (nm)	Resolution (m)
PAN	520-900	15
BLUE	450-520	30
GREEN	530-610	30
RED	630-690	30
NIR	780-900	30
SWIR1	1550-1750	30
SWIR2	2090-2350	30
TIR	10400-12500	60

Although Landsat7 has a repetition cycle of 26 days, most images are spoiled by clouds. In practice the analysis was based on 5 relatively cloudfree ETM+ images, all over the same area in Sanjiang Plain (path 114/row 027, see Fig. 1): 3 were registered in 2001 (20 May, 5 June, 9 September) and 2 in 2002 (7 May, 11 August).

In practice, Level1G data were acquired from USGS-NLAPS in the FAST-L7A format. Each spectral layer is stored in a separate, binary, byte image file, while all the ancillary information is listed in a single text file. The Level1G scenes are already geometrically corrected to the WGS84-UTM map system, with the appropriate UTM zone (for Sanjiang: 53N). This correction is performed with an automatic procedure which removes all systematic distortions. However, small translations in the X and Y-directions remain between the registrations of different dates. As can be seen in figure 1B, the geometric correction also gives rise to triangular, empty zones in the corners of the Level1G-images. For this feasibility study, a more restricted output area was defined in advance. This central zone is totally covered by the 5 subsequent images and it covers an area of 145.8km x 135.6km (19 770.5 km² - see the blue frame in figures 1B/C, and Fig. 3).

In a spectral sense, the classification exercise is only based on three of the 30m-bands of each registration: RED, NIR and SWIR1. The spatial resolution of the selected output area (and of the final classification) will thus be 30m. The two other visual channels (BLUE, GREEN) and the SWIR2 band were skipped because they are strongly correlated with the RED and SWIR1. As described below, the panchromatic and thermal infrared (TIR) layers were only used sideways.

Each individual registration was separately pre-processed with the following steps:

- *Geometrical rectification:* To eliminate the mentioned translations, all registrations were warped towards the pre-defined (30m resolution) output area using 15 manually defined control points, a first order polynome and nearest neighbour resampling. Actually, the first registration was corrected in an absolute way (image to map) by means of the GPS-trajectory registered during the field survey (Fig. 3). The later registrations were then

warped onto this reference scene (image to image). The control points were always defined on the most detailed panchromatic image (15m pixels). The warps were performed for the three classification layers (RED, NIR, SWIR1) and also for the TIR-band in view of the cloud masking. In the last case, the resolution was thus promoted from 60m to 30m.

- *Calibration and atmospheric correction:* By means of the band-specific calibration factors, mentioned in the FAST-L7A text-file, the shortwave scenes (RED, NIR, SWIR1) were converted to top-of-atmosphere radiances. New images with (estimated) surface reflectances were then computed by means of the “Simple Method for Atmospheric Correction” (SMAC), developed by Rahman & Dedieu (1994). In the absence of better alternatives, the SMAC model-parameters of the similar sensor Landsat5-TM were used, while standard clear sky conditions were always assumed for the atmospheric parameters (ozone, aerosols and water vapour).
- *Addition of NDVI-layers:* For each registration, an additional NDVI-layer was computed: $NDVI = (NIR - RED) / (NIR + RED)$. As a combination of the reflectances in the RED and NIR, this Normalised Difference Vegetation Index gives a fair estimation of the amount of green vegetation. The NDVI of land surfaces gradually increases with soil cover, from 0.15 for bare soils to 0.85 for the densest canopies. (When the NDVI is directly computed from the raw digital values it has completely different and unstable values.) As these NDVI-layers do not add new information, they were not used in the classification process. However, they provide an interesting data compression and hence they were used for the segmentation and for other tasks (for instance: delineation of training polygons), which require the visualisation of multitemporal information (see also Fig. 4).

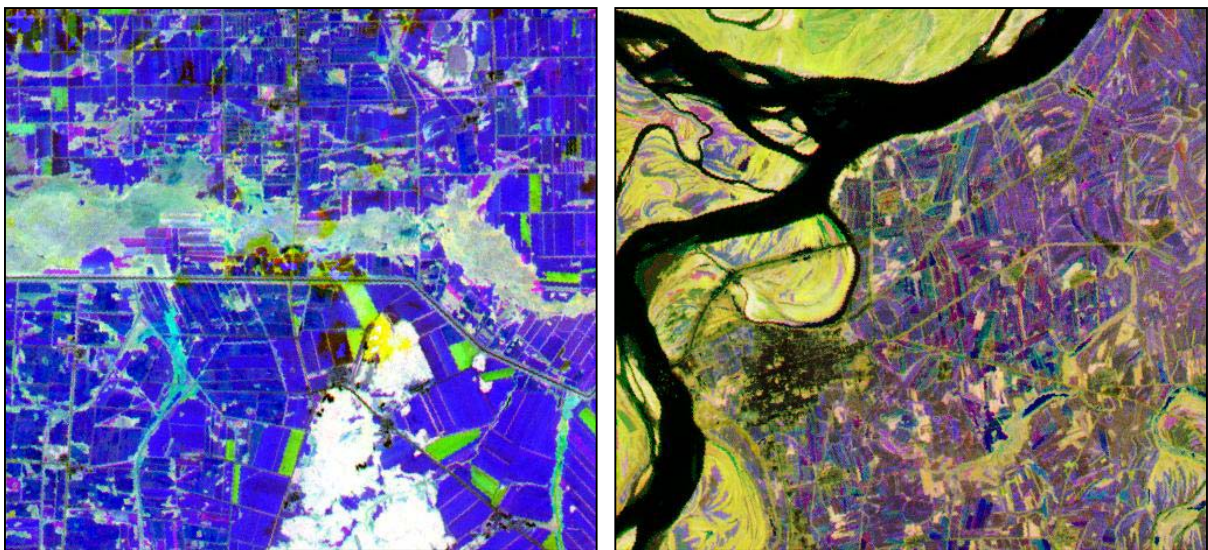


Figure 4: Colour-composites with the 3 NDVI-images of 2001 (Red=May, Green=June, Blue=September). At the left, a typical rural area in Sanjiang Plain, at the right Tonjiang city at the confluence of the Sonhua and Heilongjiang rivers. Such NDVI-composites are easy to interpret: white zones are always vegetated (forests, grasslands, some wetlands,...), blue spots are only “green” in May (not in June-September), black zones must be water, etc.

- *Cloud masking:* For each registration a cloud (and cloud shadow) mask image was created, in a manual-interactive way, by thresholding the available spectral layers. An example is given in figure 5. So far these masks were not really used in the classification process, but in a fully operational application they should be created automatically and effectively used.

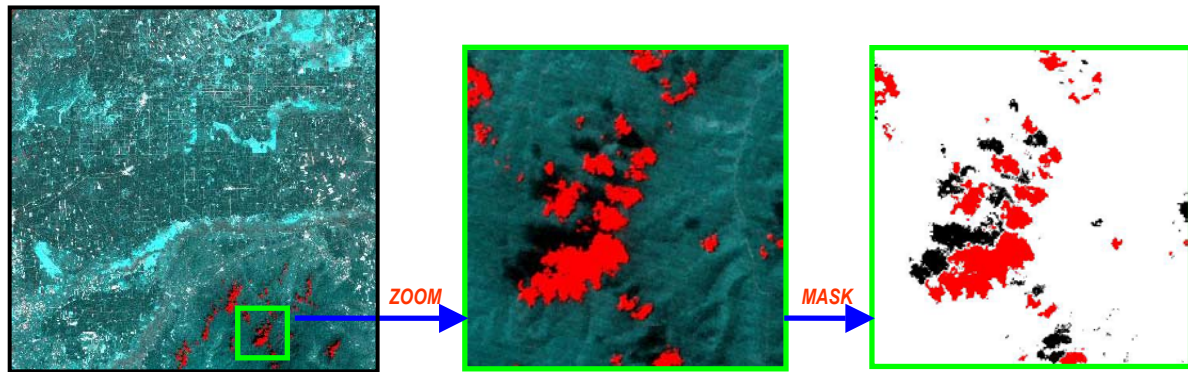


Figure 5: Cloud and cloud shadow masking by means of the RED and TIR bands of the ETM+ registration of 11 August, 2002 (south-eastern part). In the RGB-composite on the left, the R-plane is defined by the RED-band, while the TIR-band is projected on the GB-planes (Green+Blue=Cyan). Cloudfree pixels have low Red-reflectance and normal temperatures, and hence appear as Blue-Cyan. Cloudy pixels are Red because they have very low temperature and high reflectance. Cloud shadows have an intermediate position.

The pre-processing thus generated a multivariate set of congruent images, which all cover the same output area of 19 770.5 km², projected in the WGS84-UTM53N system with a resolution of 30m. The multivariate set comprises all relevant layers (RED, NIR, SWIR1, NDVI, TIR) of the 5 registrations spread over the years 2001 and 2002.

4. PER-PIXEL CLASSIFICATION

Figure 3 shows the GPS-registered trajectory followed during the survey, which took place from 21 to 25 June, 2002. Along the road, 261 relatively large parcels or plots were selected, covered by 16 different crops or other land cover types (Tab. 2). Afterwards, the boundaries of these ground truth (GT) polygons were digitised over the pre-processed imagery to form a vectorial GIS, and the corresponding class codes were stored in the associated database. This GIS-information was then converted to raster format to yield the so-called “GT parcel raster” and “GT class raster”, in which each pixel is respectively labelled with the ID-number (1-261) of the parcel and the code k (1-16) of the class to which it belongs. The large majority of the pixels in both GT-rasters was not visited during the survey, and hence bears the special code 0. Some examples are shown in figure 6.

Table 2 gives an overview of the 16 classes which could be discerned on the field, together with the corresponding number of parcels and pixels available in the ground truth rasters. The tabulated acreage distributions of the survey (sample) are certainly not representative for the entire area (population). Some easy classes such as rice and soybeans might be overestimated, while for others the amount of ground truth is rather critical. In the survey, no fields were encountered for some crops such as sugar-beets, flax, and hemp, which according to literature should be present in Sanjiang Plain. These crop are thus implicitly withdrawn from the classification.

As mentioned in §3, the per-pixel classification will only be based on the three most relevant spectral layers (RED, NIR, SWIR1) of the 5 available registrations. Hence, the number of image variables per pixel amounts to $N_v = 5 \times 3 = 15$. By superimposing the GT class raster with the multivariate set of N_v pre-processed image layers, all the image statistics required by the Maximum Likelihood classifier could be extracted: the mean vector μ_k and covariance matrix C_k for each class k (with $k=1 \dots N_k$ and $N_k=16$). Some partial (bivariate) examples of true pixel

distributions are shown in figure 7, together with the corresponding multinormal density approximations (iso-probability contours).

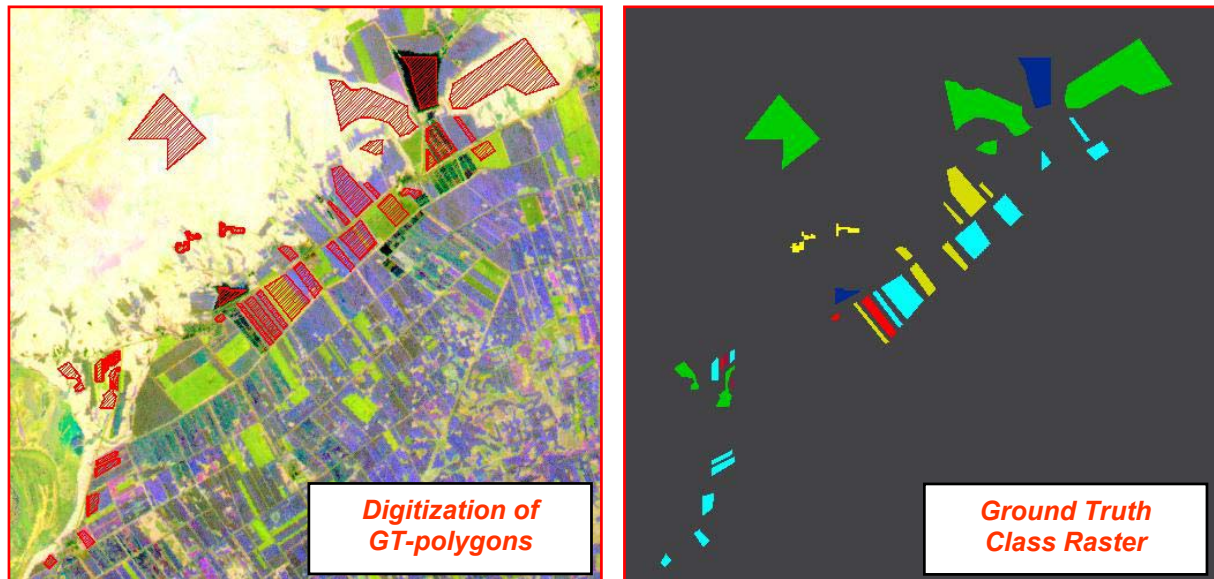


Figure 6: Zoom-in on an area near the forests in the north of the study area. Left: Ground truth polygons in vectorial form. Right: Corresponding class raster (each colour corresponds with a different class – the grey background pixels were not surveyed).

Table 2: Detailed classification legend or key with $N_k=16$ classes. N_{par} and N_{pix} indicate the number of parcels and pixels available in the ground truth rasters and used for the training of the Maximum Likelihood classification (each pixel is $30 \times 30 m^2 = 0.09$ ha).

CLASS		N_{pix}	N_{par}
AGRO	- Soybean	6 759	62
	- Maize	482	5
	- Wheat	2 315	21
	- Melons	252	12
	- Potatoes	10	1
	- Rice	11 394	78
WETLANDS		3 779	13
WATER		2 279	4
GRASSLAND		198	3
FOREST	- Poplar	735	11
	- Deciduous	21 106	20
	- Pine	79	5
	- Coniferous	811	9
URBAN	- Open	605	8
	- Bare soil	231	5
	- Dense	2 302	4
TOTAL		53 337	261

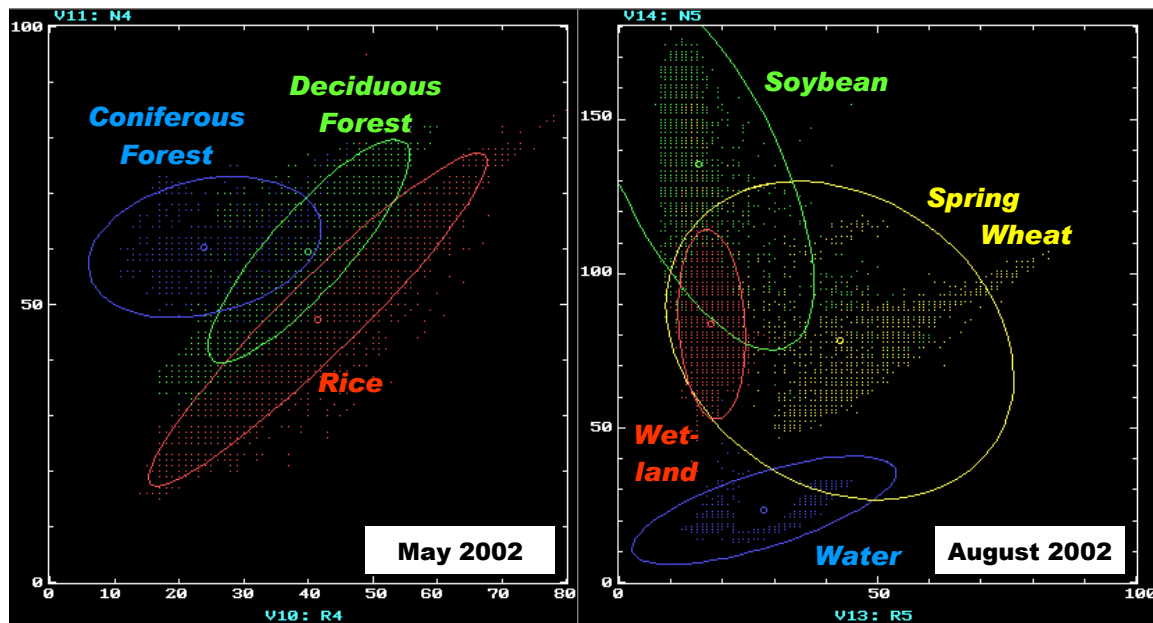


Figure 7: NIR versus RED scatterograms with the bivariate projections of some class-conditional distributions $f_k(x_p)$ for the two registrations of 2002: true clusterings (points) and multinormal density approximations (95% confidence ellipses). All values derived from the ground truth data set (table 2).

All this information suffices to apply the Maximum Likelihood classifier on the entire image set. However at this stage it was noticed that the direct processing of the described data set ($N_v=15$ layers, $N_k=16$ classes) as such did not yield the best results. Given its experimental nature, every classification has to be optimized by repeated application of the classical cycle “calibration → extrapolation → validation” for a wide range of possible variants (image subsets, simplified legends,...). This sensitivity analysis pointed out that the best classification results were obtained by the three steps procedure outlined in figure 8. It is based on the following principles:

- The discriminative power of any classifier increases with N_v , the number of image layers - and thus with the number of registration dates. In our case, this pleads for the systematic use of all 5 registrations together.
- The general land use classes, which by definition remain rather stable in time (at least for a few years), can indeed best be identified with the imagery of both available years (5 registrations). On the other hand, within the general AGRO-class, the acreage proportion of the individual crops may vary drastically from one year to another. For this reason, crop classifications are year-specific and may only be derived from the imagery of the concerned year (in our case 2002).
- Like all other classifiers, the ML-algorithm runs into problems for pixels with incomplete measurement vectors x_p . In this case, the missing values are mainly due to cloudy measurements in one or more of the involved registrations. Although literature is overwhelmed with papers on different classification algorithms, few attention has been paid to this practical problem of missing values. Although cloud masks were prepared for each registration (see Fig. 5), they were not practically used because this would involve very complex, stratified processing schemes. For reasons of simplicity, it was preferred to classify each pixel as such, as if the cloud problem did not exist. Amazingly enough, this

did not raise too much problems, as long as the cloudy measurements (very high reflectances in the 3 retained spectral bands) was limited to a single registration. But the probability (for any pixel) of obtaining multiple cloudy measurements obviously increases with the number of involved registration dates. Such pixels have atypical measurement vectors \mathbf{x}_p and low aposteriori probabilities $P_k(\mathbf{x}_p)$ for all the considered classes. By putting an upper threshold, these deviate/cloudy pixels can be separated and assigned to a “garbage class” ($k_{est}=0$).

STEP	CLASSIFY WHAT?	LEGEND	IMAGE DATES	RESULT
1	All pixels	General (11 classes)	5 (2001-2002)	Much “holes” (clouds)
2	Only “holes” of step 1	General (11 classes)	2 (only 2002)	General map ($N_k=11$)
3	Only AGRO-pixels	Detailed (6 classes)	2 (only 2002)	Detailed map ($N_k=16$)

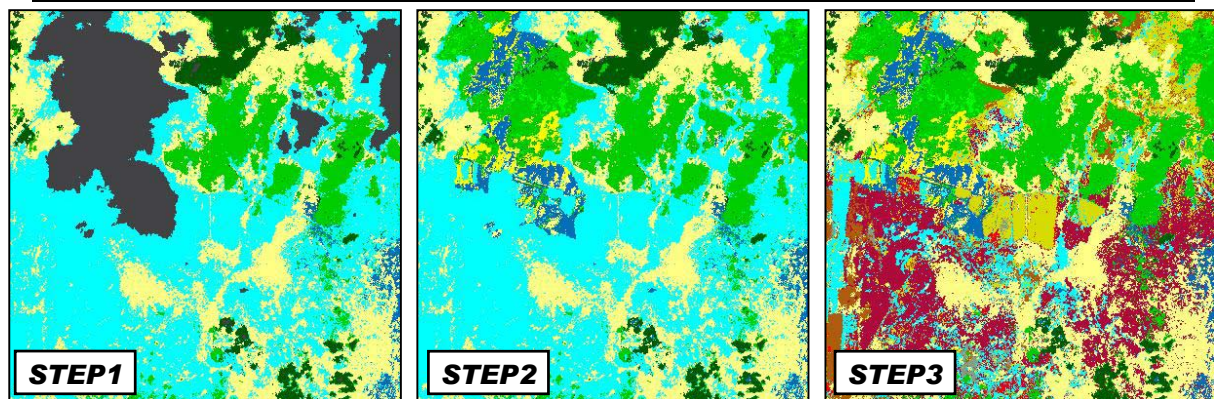


Figure 8: Three-step procedure followed for the per-pixel classification. The grey zones resulting from step 1 are non-classified “holes” which are due to clouds (missing values) in one or more of the 5 registrations. In step 2, most of them have disappeared. The cyan zones resulting from steps 1/2 represent the compound cropland class. In step 3, this mixed class has been split up into its pure crop components, on the base of the registrations of 2002.

The final scheme, presented in figure 8, involved the three steps, described below. In advance we however mention that, given the complete lack of external, apriori information, all classifications were performed with equal class weights ($A_k(p)=1$ for all classes and pixels).

- First, a general land use classification was made according to a simplified legend which was obtained by combining six units of the original key (see table 2) into a single category “non-irrigated cropland”. These crop classes are: soybean, maize, wheat, melons, potatoes and also bare soil/urban (but not rice!). The Maximum Likelihood statistics (μ_k, C_k) for this compiled class were derived from the merged training sets of the individual classes. In the first step, the entire image was classified according to this generalised legend ($N_k=11$). Because the retained units are rather stable land cover classes, the spectral information of the five registration dates was used ($N_v=5 \times 3=15$). The results were excellent, except for the presence of a significant amount of garbage pixels, which remained unclassified due to cloudy measurements.
- The second step only treated these “holes” in the map ($k_{est}=0$) and classified them with only the two registrations of the year 2002 ($N_v=2 \times 3=6$). The use of only two instead of five registration dates drastically reduced the fraction of pixels with multiple cloudy measurements. In this way, the general land use classification could be completed.

- The third step only classified the pixels of the merged category “non-irrigated cropland” and assigned them to the pure units. Because the crop surfaces rapidly change over the years, this step was only based on the imagery of 2002. The final result consists in a per-pixel classification according to the detailed legend of table 2 ($N_k=16$).

5. SEGMENTATION AND PER-FIELD CLASSIFICATION

Per-pixel classifications often have rather low accuracy and contain a lot of speckle. Much errors occur in the vicinity of field boundaries, where the mixed pixels have an unreliable signal. But mistakes can also take place in the field centres due to local effects (isolated trees, variations in plant density, soil moisture) which affect the optical signals but which are irrelevant for the objective of field-based crop mapping. If one disposes of a map with the parcel boundaries, these errors can be largely eliminated, by application of a post-classification per-field mode filter (for instance: Pedley & Curran 1991). This simple technique, which is outlined in figure 9, searches for each parcel the predominant class (according to the per-pixel classification) and then assigns the entire parcel (with all its pixels) to this modal class. By reducing the spatial variability and speckle, the method clearly enhances the legibility of the final map. But accuracy is only improved if the per-pixel classification is at least correct for the major part, and if the parcel boundary map is reliable.

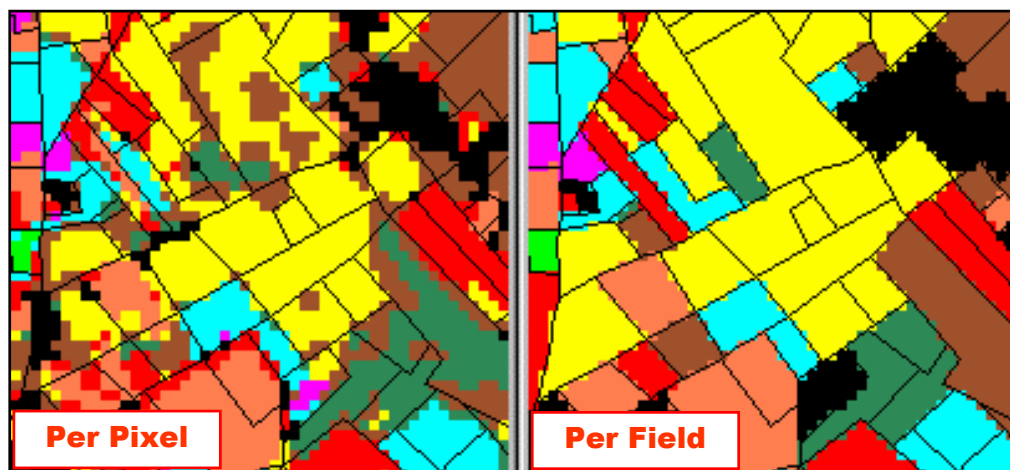


Figure 9: Left: Per-pixel classification with parcel boundaries in overlay. Right: Per-field classification (Example from another study- the colours represent different classes).

Because the parcel boundaries are seldom available, it is often attempted to extract them directly from the data set by means of image segmentation. This procedure detects and delineates groups of neighbouring pixels which are homogeneous in colour and texture, and which together form logical shapes. In the optimal case, the obtained image segments should correspond with parcels on the ground. The segmentation can be realized in a manual way, by digitizing all uniform structures visible on high-contrast colour composites. Although this approach still gives the best results, it is very time-consuming and unrealistic for large areas. In practice, only a 15x15km area in the centre of the study area (Qianjin farm) was manually segmented (see Fig. 10A).

For the treatment of the entire study area, use was made of an automatic segmentation procedure provided by the commercial software *eCognition*. This routine is very fast and stable, it simultaneously inspects different spectral bands, and it retrieves acceptable results

(both in raster and vector format). In practice, the automatic segmentation was performed with the two NDVI-images of 2002 (May, August). As shown in figure 10, very different results can be obtained depending on some user-specified parameters which define the relative weight of the different factors (colour, texture, shape) in the recognition process. The “over-segmentation” in figure 10B divided the entire image area ($\pm 20\,000\text{ km}^2$) into 450 980 segments with an average size of 4.4 ha. On the other hand figure 10C shows an example of “under-segmentation” which only discerned 19 348 very large segments (102.2 ha per segment).

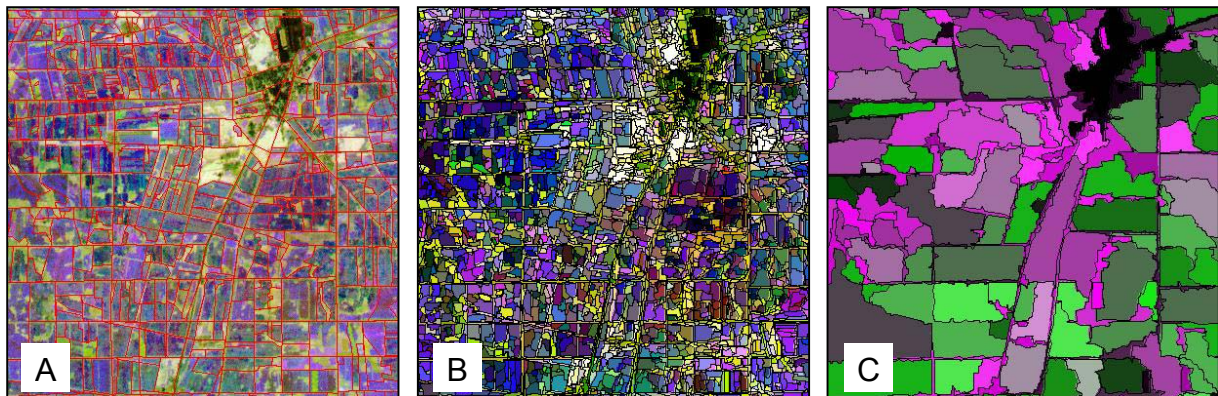


Figure 10: Segmentation results for a 15x15km²-block in the central Qianjin area. A: manual segmentation. B/C: Automatic segmentation with the eCognition software.

If per-field filtering is the goal, over-segmentation obviously is the right choice. For instance, it does not matter if a wheat field is subdivided in two or more segments, because they all will be post-classified as wheat (at least if the per-pixel classification was not too bad). In the final map the segments are thus regrouped as a single wheat parcel (though possibly joint with adjacent wheat parcels). On the other hand, under-segmentation is detrimental. For instance, if two parcels with different crops are combined in a single segment, one of both unavoidably will be post-classified in the wrong way.

6. RESULTS, VALIDATION AND ACREAGE STATISTICS

Hence, in practice the per-field mode filter was applied over the entire study area by means of the per-pixel classification (see §4, $N_k=16$) and the map with the 450 980 segments retrieved by the *eCognition* software (example of figure 10B). The post-classification result is a field-based crop and land use map valid for the year 2002. This final product is available both in raster format and in the form of a vectorial GIS (450 980 polygons, with the corresponding area and estimated land cover class in the associated database). As an example, figure 11 shows some map extracts.

The derived version shown in figure 12, covers the entire study area, but here the 16 original classes were compiled into only six broad categories. This generalised map can be considered as an updated version of the land cover map produced by the Chinese Academy of Sciences (see Fig. 1C).

Unfortunately, no sound validation of the map could be performed so far. Normally, the collected ground truth is split in two fractions: the procedures are then calibrated with the first part, and independently validated with the second. However, insufficient ground truth was available for this approach.

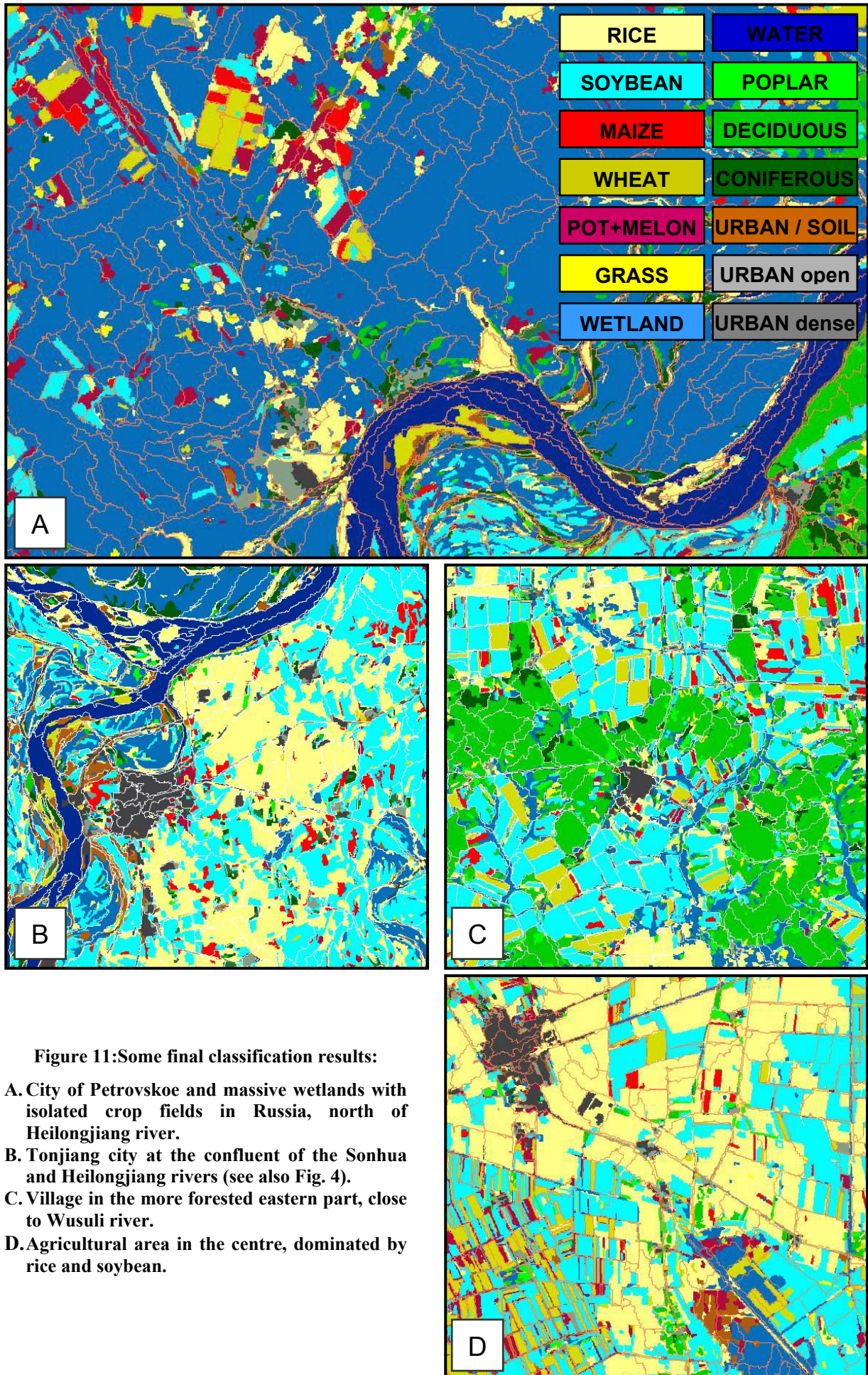


Figure 11: Some final classification results:

- A. City of Petrovskoe and massive wetlands with isolated crop fields in Russia, north of Heilongjiang river.
- B. Tongjiang city at the confluence of the Sonhua and Heilongjiang rivers (see also Fig. 4).
- C. Village in the more forested eastern part, close to Wusuli river.
- D. Agricultural area in the centre, dominated by rice and soybean.

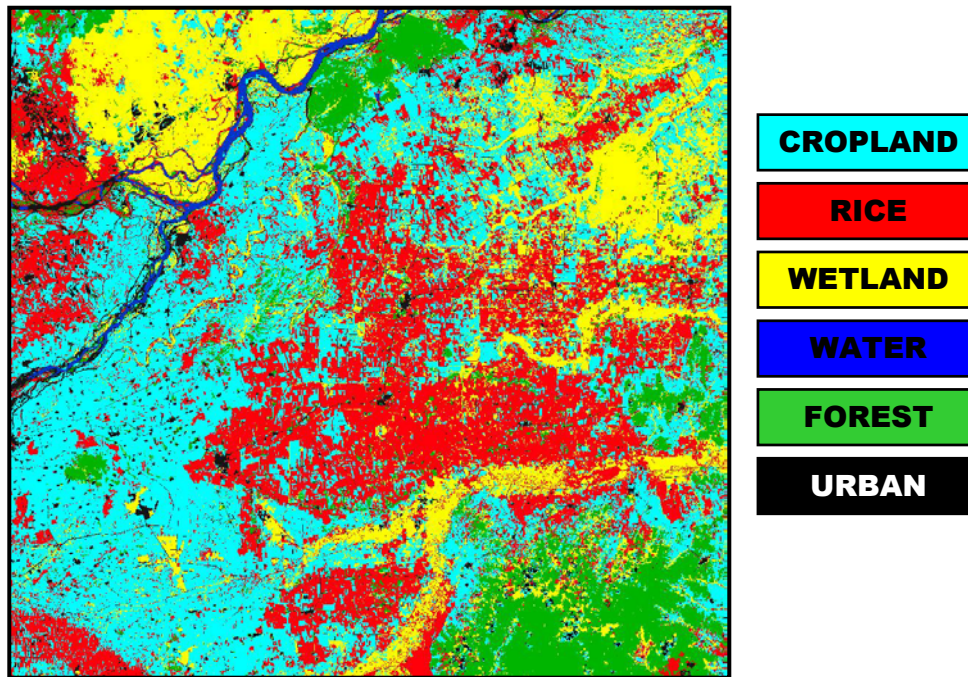


Figure 12: Generalised map with 6 compiled classes for the entire area in Sanjiang Plain.

Hence one can only rely on the back-validation in table 3 and on the general appreciation of the map. As evidenced by the examples in figures 11 and 12, the visual aspect is quite convincing. Thanks to the per-field filtering, the map looks attractive, speckle-free and “legible”. Moreover, the distribution of the land cover types seems logical and in line with the expectations. Table 3 presents the error matrix for the comparison of the true class (survey) and estimated class (final map) for the 53 337 pixels in the ground truth. The overall accuracy (sum of diagonal values) amounts to 93.7%. But as the same pixels were also used for the training of the Maximum Likelihood classifier, this is probably an overestimation. Nevertheless, the table shows that many of the remaining errors are due to the confusion of similar units (for instance different forest types, or dense and open urban areas).

Table 3: Error matrix of the final map (EST), derived from the 53 337 ground truth pixels (TRU). All values in % of total (53 337=100%). See also table 2.

TRU \ EST	SOY	MAI	WHE	MEL	POT	RIC	WET	WAT	GRA	POP	DEC	PIN	CON	Uop	Uso	Ude	Σ(TRU)
SOY	10.9	0.4	0.8			0.4	0.1				0.0			0.0		0.0	12.7
MAIZE	0.0	0.2	0.0	0.0		0.7					0.0					0.0	0.9
WHEAT	0.3	0.0	3.9	0.0		0.1	0.0							0.0	0.0	0.0	4.3
MELON	0.1			0.4		0.0											0.5
POTAT	0.0				0.0												0.0
RICE	0.4	0.0	0.0	0.0		20.1	0.7			0.0	0.0		0.0	0.0	0.0	0.0	21.4
WETLAND	0.0	0.0				0.0	6.9			0.0	0.0			0.1			7.1
WATER								4.3									4.3
GRASS									0.4		0.0						0.4
POPLAR	0.0					0.0	0.0			1.3						0.0	1.4
DECIDU.	0.0					0.1	0.2		0.0	0.4	38.3		0.5				39.6
PINE											0.0	0.1			0.0		0.1
CONIFER.						0.0					0.0		1.5				1.5
URB open	0.0		0.0			0.0								1.0	0.0	0.2	1.1
URB soil			0.0								0.0		0.0	0.1	0.1	0.3	0.4
URB dens						0.0										4.3	4.3
Σ(EST)	11.8	0.6	4.7	0.4	0.0	21.4	8.0	4.3	0.4	1.8	38.4	0.1	2.0	1.2	0.1	4.8	100.0

By superposition of the final classification with the map of the agro-statistical regions (see Fig. 13), the acreage distributions of the 16 crop and land use classes could be derived – which was the ultimate goal of this classification study. These regional statistics are listed in table 4.

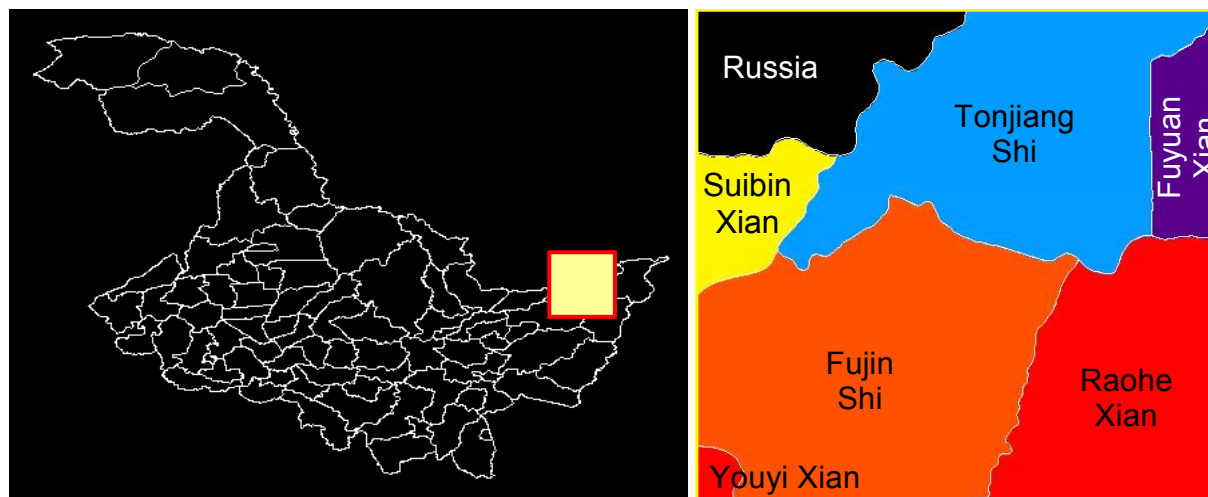


Figure 13: Map of the 79 administrative counties in Heilongjiang Province, and a detailed extract over the classified area.

Table 4: Relative acreage distribution of the 16 crops and land uses types, for Russia and the 6 counties in Heilongjiang (see Fig. 13). Values in % of the total area covered by the classification (sum=100%). Three bottom lines indicate the total areas in km² covered by the classification and by the entire county, and the ratio of both.

COUNTY	RUSSIA	FUYUAN XIAN	TON-JIANG	SUIBIN XIAN	FUJIN SHI	RAOHE XIAN	YOUYI XIAN	TOTAL
SOYBEAN	3.99	23.90	33.17	36.53	39.96	17.79	17.77	28.85
MAIZE	0.82	0.75	1.44	5.07	0.90	0.83	0.00	1.22
WHEAT	1.79	1.83	3.36	2.66	4.45	3.26	15.41	3.50
MELON	4.15	1.03	1.69	1.22	5.77	0.53	3.42	3.00
POTATO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RICE	18.78	17.77	28.24	29.04	32.39	27.77	58.08	28.18
WETLAND	51.69	48.45	19.60	2.83	9.24	13.82	1.14	19.21
WATER	4.69	0.00	0.52	5.31	0.01	0.00	0.00	0.93
GRASS	0.17	0.62	0.52	0.03	0.35	1.16	0.00	0.53
POPLAR	0.17	0.30	0.67	0.13	1.03	1.74	0.24	0.88
DECL. FOR.	4.42	1.19	6.18	1.70	1.47	20.98	0.01	6.71
PINE	0.25	0.01	0.13	0.28	0.14	0.12	0.00	0.15
CONI. FOR.	2.22	2.48	0.54	1.51	0.81	9.99	0.08	2.73
URB open	1.43	0.56	1.18	2.39	1.32	0.53	0.22	1.15
URB+SOIL	3.49	0.38	0.79	6.09	1.03	0.84	1.72	1.43
URB dense	1.95	0.71	1.96	5.20	1.14	0.62	1.89	1.54
TOTAL %	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
km ² - CLAS		1040	5357	1010	6337	3703	155	
km ² - CTY		6238	6074	3332	8068	6338	1663	
km ² %		17	88	30	79	58	9	

7. CONCLUSIONS

The 30m-resolution images registered by the Landsat7-ETM+ sensor form an invaluable source of information for the production of thematic maps on land use and crop cultivation. The followed methodology is not new and has often proved its robustness in the past. From the same, pre-processed and multivariate image set, two products are derived. The Maximum Likelihood classifier, calibrated with field-collected ground truth, creates a per-pixel classification, while the parcel structure is extracted by an automatic image segmentation procedure. Both products are then combined by application of a per-field mode filter, which results in a final map with improved legibility and accuracy. Some aspects of the procedure can certainly be improved, especially the use of the more detailed panchromatic bands (15m resolution) and the treatment of pixels with deficient measurement vectors (missing values in one or more registrations).

This study also focused on the extraction of regional acreage statistics, which can be used for the assessment of crop productions, in combination with the yield forecasts provided by Crop Growth Monitoring System (CGMS). This objective (acreage statistics) is certainly realised and generally feasible. The image-derived crop and land use maps (and the underlying procedure) can however serve for many other purposes, at the level of land use planning, control of governmental regulations, and as input for a wide range of geophysical models (erosion, changes, carbon budgeting, etc...).

Financially, the costs are divided over three main items: images, survey and analysis. For a detailed crop map, at least two images should be available, well spread over the growing season (best: May/August). Landsat7-scenes are reasonably priced, considering their spectral richness and spatial extension (32 000 km²). To minimise the costs of the analysis, major parts of the interpretation chain can be automated, especially the pre-processing, training, segmentation and per-field filtering. If the crop mapping is annually repeated, the costs of the surveys decrease over the years. The most-effective approach is when the ground truth is delivered by local, agro-statistical institutes which are actively engaged in the project, also as end-users.

Given the relative success of this feasibility study, it might be considered to apply the procedure to other areas in Heilongjiang. However, whereas Sanjiang Plain (and especially the part explored in this study) is characterised by relatively large and square-shaped parcels, in most other areas the fields are very narrow and long as shown by the aerial photograph in figure 14. Very often, the crops are even completely intermixed in very long strips. In these areas, the use of the Landsat7-ETM+ imagery might be limited, because the details of the parcel structure generally fall below the 30m-resolution.

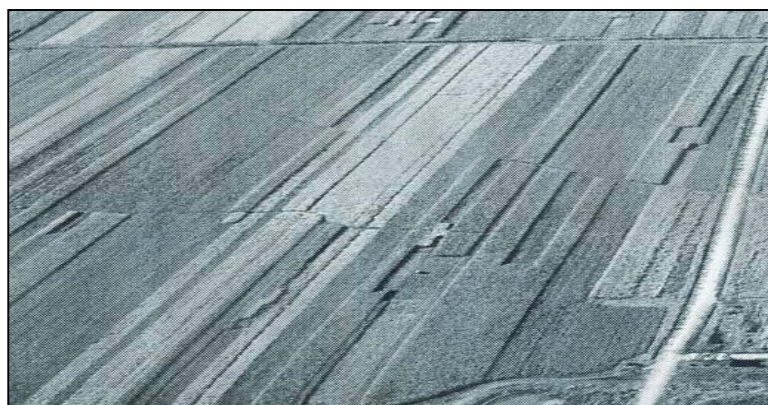


Figure 14: Typical narrow fields in Heilongjiang

REFERENCES

Duda R. & Hart P., 1973. *Pattern classification and scene analysis*, Wiley & Sons, 482p.

Jiang L., Wang Y., Li X. & Ji S., 2003. Method and application of the Crop Growth Dynamic Model to forecast yields in the Heilongjiang Province. *In: Belgian and Chinese Crop Growth Monitoring Systems: comparison, adaptation and improvement*, Tychon B. (ed.), FUL, Arlon, Belgium: 27-33.

Pedley M. & Curran P., 1991. Per-field classification: an example using SPOT-HRV imagery. *International Journal of Remote Sensing*, 12: 2182-2192.

Rahman H. & Dedieu G., 1994. SMAC: A simplified method for the atmospheric correction of satellite measurements in the solar spectrum. *International Journal of Remote Sensing*, 15: 123-143.

Tychon B., Ozer P., Pecheur C., Oger R., Decrem M., Leteinturier B., Eerens H. & Dong Q., 2003. Application of the European Crop Growth Monitoring System (CGMS) in Belgium and in China's North-eastern province Heilongjiang. *Proc. CGMS Expert Meeting, 30 Oct-1 Nov. 2002, Ispra-Italy*. (in press).

Zu S., 1998. The Heilongjiang province food production forecasting decision making service system and the yield forecasting in 1997. *Heilongjiang Meteorology*, 68: 29-34 (in Chinese).

© CYBERGEO 2003

Herman EERENS, Qinghan DONG, Pierre OZER and Xiufen LI, *Cybergeog*, No. 247, 19 September 2003