Efficiency and accuracy of per-field classification for operational crop mapping

A.J.W. de Wit and J.G.P.W. Clevers

Short title: Efficiency and accuracy of per-field classification

Centre for Geo-information

Wageningen-UR

P.O. Box 47

6700 AA Wageningen

The Netherlands

E-mail: allard.dewit@wur.nl

Tel: +31-317-474761

Fax: +31-317-419000

Abstract

A crop map of the Netherlands was created using a methodology that integrates multi-temporal and multi-sensor satellite imagery, statistical data on crop area and parcel boundaries from a 1:10,000 digital topographic map. In the first phase a crop field database was created by extracting static parcel boundaries from the digital topographic map and by adding dynamic crop boundaries using on-screen digitising. In the next phase the crop type was determined from the spectral and phenological properties of each field. The resulting crop map has an accuracy larger than 80% for most individual crops and an overall accuracy of 90%. By comparing cost and man-hours it was demonstrated that per-field classification is more efficient than per-pixel classification and decreased the effort for classification from 1500 to 500 man-hours, but the effort for creating the crop field database was estimated at 2300 man-hours. However, the total cost could be lowered by outsourcing the digitising of the crop boundaries. The use of image segmentation techniques for deriving the crop field database was discussed. It was concluded that image segmentation cannot replace the use of a large-scale topographic map, but in future segmentation may be used to map the dynamic crop boundaries within the topographic parcels.

1. Introduction

In the Netherlands timely and accurate information on land cover/use at regional and national scales is required by national and regional governmental agencies to support environmental policy and physical planning purposes. Until the end of the eighties information on land cover/use was usually obtained from land use statistics and topographical maps. However, land use statistics were only available at the level of administrative units (e.g. municipalities or provinces) and could not be obtained for areas with deviating boundaries (e.g. river basins and groundwater protection areas). Moreover, topographical maps often excluded some land cover/use classes, they were often outdated and they were not available in digital form until recently. Therefore, in 1987 it was decided to produce a land cover database of the Netherlands in a raster format (further to be mentioned 'LGN database'), using satellite images (Thunnissen et al. 1992a and 1992b). Four versions of the database are available (LGN1, 2, 3 and 4), based on satellite images from 1986, 1992/'94, 1995/'97 and 1999/2000. Today, the LGN database discriminates 39 land use classes including crop types, forest types, water, various urban classes and semi-natural vegetation types.

An important aspect of all LGN versions is that the database provides an overview of the crops that were grown in that period in the Netherlands. The mapping of the crop types is clearly the most time-consuming and expensive task of the updating of the LGN database and much effort has been spent on deriving efficient methodologies for classifying crops from satellite images. A distinct evolution in the crop mapping methodology can be recognised from LGN1 to LGN3. The crop classification methodology used for LGN1 was based on a combination of supervised per-pixel classification using a single Landsat TM image and manual delineation of urban areas (Thunnissen et al. 1992a). Because this approach yielded poor classification results (60% overall accuracy), it was replaced by a multi-temporal approach during the production of the LGN2 (75% overall accuracy). An even higher accuracy was needed for the LGN3 database and this requirement was met using a stratified multi-temporal approach that was heavily based on visual interpretation for the crop classification. Although this approach yielded an 85% overall accuracy, it was clear that the use of visual interpretation was time-consuming and expensive.

In 2000 the production of LGN4 was initiated and a new methodology for crop classification was required in order to reduce labour cost while maintaining high classification accuracy. A distinct advantage over the preceding versions of the LGN database was that since 1998 the 1:10,000 digital topographic vector database of the Netherlands is available. This vector database nowadays is being used as the common geometric base for most geographical databases in the Netherlands. LGN4 should therefore comply with this database as much as possible. Besides the necessity of standardising the LGN4 database to a common geometric base, the use of the 1:10,000 digital topographic vector database also provides new methodological opportunities because it can be used as a starting point for mapping all crop fields. The prospect of having these crop fields in a digital vector database makes the application of per-field classification techniques feasible.

This paper focuses on the crop classification methodology that was developed in order to update the crop classification in the LGN database. The LGN database contains many other classes but the classification methodologies for these classes are largely based on visual interpretation in combination with interpretation of ancillary data and are therefore considered out of scope. Readers that are interested in the methodologies used for the non-agricultural classes are advised to look at Thunnissen and De Wit (2000) and Thunnissen and Noordman (1996).

The objective of this research was to develop an efficient per-field crop classification approach that makes optimal use of multi-temporal satellite data, the 1:10,000 digital topographic vector database and statistical data on crop acreage in the Netherlands. The final product was defined as a crop map for the entire Netherlands using the 1:10,000 digital topographic vector database as its geometric base (further to be mentioned 'LGN4 crop map'). The accuracy requirement for the LGN4 crop map was similar to that for LGN3 (85% overall accuracy, 75% accuracy for individual classes).

2. Data, pre-processing and nomenclature

2.1. Satellite imagery and optimal acquisition periods

In general, the use of multi-temporal satellite acquisitions is required for an accurate classification of most agricultural crops (Jewell 1989, Murakami et al. 2001). The optimal acquisition periods of optical satellite imagery are determined by the phenological characteristics of the main crops (figure 1) and by cultivation practices like harvesting, after-growth for green manuring, mowing and conversion of grassland to arable land and vice versa. A minimum Landsat TM image dataset should include at least one image from spring (April, May) and one image from summer (end of July, August). The optimal image dataset should include an additional image obtained between the previous two periods (June or beginning of July).

[Insert figure 1 about here]

The classification of agricultural crops in the LGN4 crop map is based on a multi-temporal approach. Optical satellite imagery obtained by Landsat TM and IRS-LISS3 has been used as well as radar observations from the ERS2-SAR (table 1). In general, Landsat TM imagery is the preferred type of imagery due to the low cost per unit area and the presence of bands in the middle-infrared part of the electromagnetic spectrum.

[Insert table 1 about here]

Three classification zones can be distinguished over the Netherlands that were classified using different combinations of imagery (figure 2). Zone 1 was classified using six Landsat TM images from three different dates in 1999. All TM images in zone 1 were based on path 198. Zone 2 was classified using six Landsat TM images from two different dates in 2000 and two IRS-LISS3 images of one date in 2000. All TM images in zone 2 were based on path 197. The IRS-LISS3 images were used to fill the temporal gap between the two Landsat TM image dates. Zone 3 was classified using the same Landsat TM images that were used for zone 2. For this zone, three ERS2-SAR radar images were used to fill the temporal gap between the two Landsat TM image dates because no cloudfree IRS-LISS3 image was available for this area.

[Insert figure 2 about here]

2.2. The digital 1:10,000 topographic map of the Netherlands (TOP10-vector)

The Netherlands Topographic Service (TDN) produces the 1:10,000 digital topographic map of the Netherlands (further to be mentioned 'TOP10-vector'). The nomenclature of the TOP10-vector consists of a few hundred entities, which are related to polygon, line and point features. Since 1998, the entire Netherlands is covered by around 1350 map sheets, which cover an area of 5 km by 6.25 km each. The 1:10,000 digital topographic map mainly contains information on land cover. The functional use of many classes can only be determined by contextual information, e.g. grassland and forest located in urban areas (parks or sport grounds).

2.3. Ancillary data

The PIPO system. LASER (Landelijke Service bij Regelingen) is an implementing body of the Dutch Ministry of Agriculture, Nature Management and Fisheries which has the task of supervising the allocation of agricultural subsidies in the framework of the Common Agricultural Policy (CAP) of the European Union. An administrative system (PIPO) has been developed which uses GIS technology to check all acreage-based applications for subsidies. In this system, all subsidised crops, their acreage’s and the concerning farmers are linked with topographical parcels obtained from TOP10-vector. One parcel from TOP10-vector can contain multiple crops and/or different farmers. PIPO has been operational since 1997 and samples of the information stored in PIPO are checked for fraud by farmers using remote sensing. Within the framework of the LGN4 crop map, small samples from the PIPO database have been used to assess the accuracy of the LGN4 crop map.

Agricultural Statistics. The General Census of Agriculture, organised by the Central Bureau for Statistics (CBS), is carried out annually in May. Among other things, the General Census of Agriculture provides information on the acreage of crops grown. The CBS Agricultural Statistics contain cultivated areas of around 100 crop types, not including roads, ditches and hedges less than 4 m wide. The CBS Agricultural Statistics are published per municipality, per province and per 'agricultural region' (further to be mentioned 'CBS regions'). An example of the crop statistics at the level of the CBS regions for 1999/2000 is given in Table 2. The CBS regions are more or less homogeneous areas as far as soil type and agricultural land use are concerned. The Netherlands is subdivided into 66 CBS regions (figure 2).

[insert table 2 about here]

2.4. Pre-processing of satellite images

Channels 3,4,5 and the panchromatic channel were selected from the Landsat TM data because these channels essentially capture the information content of Landsat TM data. Moreover, much experience in crop classification was available for this band combination. The use of the IRS-LISS3 satellite data was restricted to channels 1,2,3 because the 4th mid-infrared channel has a lower spatial resolution. No radiometric or atmospheric corrections were applied to the imagery because the basic requirement for classification is to separate one class from the others and the use of calibrated image values does not significantly help in that respect.

All optical satellite images were geometrically corrected to the Dutch Reference System using ground control points obtained from TOP10-vector. In general, a geometric accuracy of one pixel was obtained. During the geometrical correction, the images were resampled to either 25 meter resolution (TM multispectral, IRS-LISS3) or 12.5 meter resolution (TM panchromatic) using cubic convolution resampling. The choice for this resolution was made because it allows the satellite pixel grid to be aligned exactly with the LGN grid. Moreover, choosing a grid size of 25 meter simplifies aligning the grid origin during vector to raster operations because the pixel centres of the target grid are always located at known values (xxxx12.5, xxxx37.5, xxxx62.5, xxxx87.5) and the rasterising window can thus be easily specified. This careful procedure for aligning the grids is necessary to avoid shifting of the grid origin during further processing.

The choice for cubic convolution resampling was made because this resampling technique avoids the disjointed appearance of nearest neighbour resampling and makes the satellite data easier to interpret for the human eye (Lillesand and Kiefer 1994). The latter was an important aspect because the discrimination of many non-agricultural LGN classes depends on visual interpretation. Other authors (Fuller et al. 2002) have also documented the preference for cubic convolution resampling over nearest neighbour resampling. Moreover, the discrimination of some crop types can be improved using the appearance (texture) of the field. For example, flower bulb fields can often be recognised by a striped pattern in the field which is caused by alternating patches of different varieties of flower bulb species with slightly different spectral properties. It is often difficult to recognise this striped pattern when the satellite data have been resampled using nearest neighbour resampling. Although cubic convolution resampling changes the original pixel values, it affects the per-pixel classification of the crop types only marginally and, according to the author's experience, has no practical consequences for the crop classification.

For the Landsat 7 ETM images only, the panchromatic channel was merged with the multi-spectral channels (3,4,5) using a Brovey transform, in order to obtain an image product that was optimal for visual interpretation.

The radar images were geometrically corrected using ground control points from TOP10-vector and resampled to 12.5-meter spatial resolution using nearest neighbour resampling. Nearest neighbour resampling was chosen because it preserves the original pixel values so that a speckle filter could be applied optionally. However, a speckle filter was not applied because it did not improve the interpretation of the radar images in our case. Finally, the radar images were rescaled from 16-bit to 8-bit radiometric resolution in order to reduce the datasize and combined into a multi-date radar composite.

2.5. Nomenclatures and classification scheme

The classification scheme of the LGN4 crop map distinguishes 6 different crops (grassland, maize, potatoes, sugar beet, cereals and flower bulbs) and a category 'other crops'. In theory, the LGN classification scheme could be extended to include some additional crop types that are now classified as 'other crops'. In practice, many of these crops have a very small acreage that does not justify extending the classification scheme or the timing of the satellite images that are needed to discriminate these crops is too critical to rely on.

The use of many different sources of data in the classification methodology can easily lead to problems with regard to the compatibility of the nomenclatures of the different data sources. It often occurs that nomenclatures are incompatible when (geographic) databases are compared that have differences in scale or that were created with a different purpose. With regard to the different sources of crop data that were used for deriving the LGN4 classification methodology (CBS statistics, TOP10-vector, PIPO database), it can be stated that the nomenclatures are compatible with the classification scheme of the LGN4 crop map. This is mainly because all databases refer to the same geographic entity, a cropped field, which is well-defined in terms of thematic class and spatial delineation.

3. Classification methodology

3.1. Per-field classification for crop mapping

Research on crop classification in the past 20 years has demonstrated that the spectral and spatial information delivered by high-resolution optical sensors like Landsat TM, SPOT-HRV and IRS-LISS3 is generally sufficient to recognise different agricultural crops. Highly accurate crop classifications can be obtained using data from these satellite sensors, particularly when the satellite data are combined in a multi-temporal dataset with an appropriate timing of the satellite data acquisitions over the growth cycle. Despite these good results, there are still two frequently occurring classification problems that can strongly deteriorate the classification result of a per-pixel crop classification (Smith and Fuller 2001).

The first effect is the spectral variability of the canopy reflectance within an agricultural field due to, for example, variations in soil moisture conditions, nutrient limitations or pests and diseases. This effect causes part of the cropped field to have different spectral properties compared with the rest of the field. The classifier then may assign these pixels to another class depending on the actual spectral variation and the spectral distance of the other classes. The second effect is the mixed-pixel effect that occurs when a pixel is located at the boundary of two fields. In many cases the spectral signature of the mixed-pixel will not resemble the spectral signature of one of the two crops of which it consists, but it will resemble the spectral signature of another crop causing the classifier to label these pixels erroneously.

The principle of per-field classification provides a simple and elegant solution for eliminating the effect of the spectral variability within the field as well as the mixed pixel effect. The basic idea behind per-field classification or, more generally, object-based classification, is that the satellite image is divided into segments (objects) using knowledge of the 'real-world' objects on the earth's surface (Kettig and Landgrebe 1976, Mason et al. 1988). With regard to crop classification this means that the location and extent of each field is known. The final class of the entire object is then assigned on the bases of one or more statistical properties of the collection of pixels that the object consists of, instead of determining the class for each pixel separately.

From this point onward two object-based classification strategies emerge. One group of authors determines the statistical properties of the object on the bases of the satellite spectral information (Baker and Drummond 1984, Pedley and Curran 1991). Usually the average reflectance and standard deviation per channel are calculated for each object, and are used as input for the classifier. Another group of authors applies a per-pixel classification to the satellite images and determines the statistical properties of the object on the bases of the thematic per-pixel classification (Janssen et al. 1990, Mattikali 1995, Aplin et al. 1999). In this case the final class of the object is usually assigned from the majority class of the pixels within the object. Hybrid algorithms that use both approaches for classifying objects have also been developed (Cross et al. 1988, Smith and Fuller 2001).

During the development of object-based classification techniques it has been demonstrated that the map accuracy obtained with object-based classification is usually superior to that of per-pixel classification (Pedley and Curran 1991, Lobo et al. 1996, Shandley et al. 1996). However, only for some time these techniques are being used in operational mapping projects (Fuller et al. 2002). The main difficulty with applying object-based classification techniques is that the land surface objects are often difficult to acquire (Smith and Fuller 2001). This major obstacle is gradually being removed due to two developments. First, many countries now have a digital topographic vector database available on scales 1:10,000 to 1:25,000 that can be used to delineate the land surface objects. Secondly, operational software packages are becoming available that can apply a segmentation algorithm to a satellite image in order to generate the land surface objects from the satellite data itself. With regard to the LGN database it was the availability of the 1:10,000 digital topographic map of the Netherlands that triggered the development and implementation of an object-based (per-field) classification methodology.

3.2. Overview of the LGN4 per-field classification methodology

Three phases can be distinguished in the LGN4 classification methodology: object creation, object classification and object post-processing (figure 3). The objective in the first phase is to obtain a vector database containing all crop fields using TOP10-vector as a base map. In the second phase the crop type is determined for each field (object) in a two-stage approach. First, the phenological behaviour of each field is classified using multi-temporal NDVI images. Next, the images are classified using a per-pixel maximum-likelihood classification. The final class for each field is determined from the combination of the phenological class and the spectral class and stored as an attribute in the crop field database. In the third phase post-processing is carried out on the crop field database by correcting the label of small or elongated fields using criteria based on field size and shape. A more elaborate description of the three phases will now be given.

[Insert figure 3 about here]

3.2.1. Phase 1: Creating the crop field database

We used TOP10-vector as a starting point for creating a crop field database covering the entire Netherlands. The first step consisted of selecting all polygons from TOP10-vector that were labelled ‘grassland’ or ‘arable land’. This selection yields all relevant parcels but also yields many parcels that are not relevant for deriving the crop field database (urban parks, sports fields, recreational areas, etc.). Therefore, we used the LGN3 database to reselect only those parcels that were located in agricultural areas. The underlying assumption is that no land is converted from other use into grassland or arable land.

The result of this process was a database containing all topographic agricultural parcels. The boundaries of these parcels are defined by static topographic elements, which are not likely to change from year to year. Examples of these types of elements are roads, hedges, ditches, streams and fences. Within a topographic agricultural parcel various crops can be grown depending on the particular cropping scheme of the farmer.

The second processing step consisted of adding crop boundaries to the topographic agricultural parcels. The crop boundaries are dynamic because the location depends on how the farmer subdivides his parcel into fields, which can differ from year to year. We decided that visual interpretation was the most appropriate method to add the crop boundaries to the topographic agricultural parcels. During this process the topographic agricultural parcels were displayed on a computer screen with the 12.5-meter merged Landsat TM image on the background (figure 4a). An operator added the crop boundaries manually (figure 4b). Note that 'dangling arcs' occur in figures 4a which correspond to topographic elements like hedges and ditches that do not span the entire field. In some cases these dangling arcs have been extended (figure 4b) because they coincide with crop boundaries and in other cases they are left untouched. Nevertheless, these dangling arcs do not cause problems for feature extraction as long as the polygon topology is valid.

After digitising, the vector sheets were stored in a library structure that allowed easy handling of large GIS datasets.

[Insert figure 4 about here]

3.2.2. Phase 2: crop field classification

The usual method for training a classifier to recognise different crops is by sampling fields in the image whose crop type is known from reference data. In our case, reference data for training was not available for the years 1999/2000. However, large reference datasets were available which were gathered for the classification of previous LGN versions. Therefore the 'new' satellite images could be interpreted using knowledge about crop temporal and spectral characteristics from the historic reference data. Training samples were thus assigned on the bases of experience because the spectral and temporal properties of crops do not change a lot over the years.

Gross classification errors were avoided by using statistics on crop acreage to ‘guide’ the crop classification. On the basis of previous LGN classifications it was known that the CBS statistics were quite a good indication of the true crop acreage, particularly if the acreage for a certain crop in the area is large. Therefore, it was assumed that when the crop acreage in the classified image deviated more then 15% from the acreage reported by the statistics, then the operator probably assigned incorrect training samples. A prerequisite for this technique is that it is the area for which the crop statistics were gathered is the same as the one being classified. This was accomplished by clipping the crop field database with the boundaries of the CBS regions, thus treating each CBS region as a separate classification unit.

We used the optical satellite data to create a multi-temporal set of NDVI images for each CBS region. The field boundaries in the crop field database were used to calculate the average NDVI per field so that the seasonal behaviour of the NDVI could be analysed for each field. For each image date and CBS region, an operator determined a suitable NDVI threshold to separate bare soil from vegetated fields. Adapting the NDVI threshold values for each image date and CBS region effectively eliminates differences in NDVI between image dates due to spectral or radiometric variations and also compensates phenological variations over the country.

A decision tree then clustered the fields into four phenological groups: ‘evergreen crops’ with high NDVI all year (mainly grasslands), ‘early crops’ with high NDVI in spring (winter cereals and spring flowerbulbs), ‘late crops’ with high NDVI in summer (maize, sugarbeet, consumption potatoes, etc.) and a category ‘bare soil’ for fields that did not show a seasonal NDVI pattern. This last group of fields mainly concerns crops with a short phenological cycle so that emergence and harvesting were in between satellite overpasses (for example seed potatoes and many horticultural crops). The acreage of the four phenological groups was compared with the crop statistics in order to determine if the NDVI thresholds were set correctly.

The field boundaries in the crop field database could now be used to clip the satellite image that was most suitable for classifying each phenological group. So, the boundaries of the fields that were labelled as ‘early crops’ were used to clip the Landsat TM images of spring, while the boundaries of the fields that were labelled as ‘late crops’ were used to clip the Landsat TM images from summer. Fields with the phenological class ‘evergreen crops’ were best classified using a Landsat TM image from summer because occasionally some fields with early crops like winter cereals ended up in this phenological group. This confusion was present mainly in the phenological classifications that were based on Landsat TM data from 1999 (zone 1) because the acquisition dates of all Landsat TM images were within the phenological cycle of winter-wheat (figure 1).

The actual classification of the satellite images was carried out for each phenological group separately using a combination of visual interpretation and maximum likelihood classification. The phenological groups ‘evergreen crops’ and ‘early crops’ were generally so homogeneous in crop type that only a few fields needed to be assigned to another class and this could be done more efficiently using visual interpretation. The phenological group ‘late crops’ was generally classified using a per-pixel maximum likelihood classification because this phenological group consists of a large number of different crops with sufficiently distinct spectral signatures.

The phenological group ‘bare soil’ needed a special approach because this group consisted of fields that did not show a seasonal NDVI pattern at all and thus no information on crop type could be gathered from the satellite data. In zones 1 and 2 these fields were generally classified as ‘other crops’ because knowledge of the crop calendar excludes all other classes in the LGN4 nomenclature. In zone 3, however, confusion was present between seed potatoes and spring cereals because the phenological cycle of these crops (figure 1) is in between the Landsat TM overpasses. We used the multi-date radar composite to separate these two crops. Due to the different canopy structure, the two crop types show a large difference in radar backscatter and can be easily discriminated by visual interpretation of the multi-date radar composite.

In the last classification stage, the four classification results corresponding to the four phenological groups were combined into a single classified image. The class for each field in the crop field database was determined by taking the majority class of all pixels within that field. Each field's class was stored as an attribute in the crop field database for further processing.

3.2.3. Phase 3: post-processing of the crop field database

The scale difference between the crop field database and the satellite data is one of the implications of using a 1:10,000 base map as a starting point for the crop field database. As a result, the crop field database contains a large number of very small or elongated polygon features (often smaller then 0.5 ha) that consist of a few satellite pixels only. These polygon features cannot be properly classified because the class that is assigned to these polygon features depends on the class of only one or two pixels. Most of these small polygon features consist of grass-covered areas like riverbanks, yards and borders. It was therefore decided to assign these features to the LGN class 'grassland' and select them using three selection criteria: shapefactor, TOP10-vector label and polygon area. The shapefactor was calculated for all polygon features using:

$$shapefactor = \frac{\sqrt{4 \cdot \pi \cdot area}}{perimeter}$$

This equation will return a maximum value of '1' for a perfect circle and values smaller than 1 for polygons with other shapes. Narrow, elongated polygons in particular will obtain small values for this shapefactor.

The following queries were carried out independently in order to select small and elongated polygons that should be reclassified to the LGN class 'grassland':

1. (Shapefactor < T_{shp}) AND (Top10-vector label = 'grassland')

2. (Polygon-area < T_{area}) AND (Top10-vector label = 'grassland')

T_{shp} and T_{area} are visually defined thresholds for the shapefactor and polygon-area. These thresholds were adjusted for each CBS region. Generally, values for T_{shp} were between 0.35 and 0.45 depending on the parcel shape and size in the particular CBS region. Furthermore, it was found that 0.5 hectare was a suitable threshold on polygon-area (T_{area}).

3.3. Accuracy assessment

Accuracy assessment of the LGN4 crop map was carried out using validation data from the PIPO database. The crop information in this database is gathered from the farmers themselves for the allocation of agricultural subsidies. The PIPO database therefore cannot be regarded as 100% correct because fraudulent practices can occur. However, the advantage of using the PIPO database consists of the abundance of validation data which is expected to outweigh the disadvantage of errors in the PIPO database due to fraudulent farmers.

The validation dataset included 53 validation sites across the Netherlands with a total of 15865 fields available for validation. The nomenclature of the PIPO database was aggregated and converted to the LGN4 nomenclature. Both the validation dataset and the LGN4 crop map were converted to raster format with a cell size of 25 meter. Error matrices were generated for each validation site separately and aggregated to provincial and national level.

4. Results

4.1. Pixel-based validation

The error-matrix (table 3) demonstrates that at the national level the classification has an overall accuracy of 90.4%. The classes 'grassland', 'maize', 'sugar beet' and 'cereals' were classified with a user’s and producer’s accuracy of over 85%. The class ‘potatoes’ has been classified with high user’s accuracy (85.9%) but somewhat lower producer’s accuracy (79.8%). Potatoes are mainly confused with ‘other crops’ which can be explained by the deviating phenological behaviour of seed potatoes.

The class ‘flower bulbs’ has been classified with moderate user’s accuracy (55.2%) and somewhat better producer’s accuracy (69.9%). This accuracy assessment is based on few validation pixels because flower bulb cultivation is concentrated in certain areas in the Netherlands (Table 2) and these areas are not represented by the validation sites. Moreover, flower bulb cultivation is not a subsidised crop and consequently there is little information about flower bulb fields in the PIPO database anyway. Our current validation sample for ‘flower bulbs’ thus consists of isolated fields in areas where flower bulbs are scarcely cultivated and thus easily overlooked. On the basis of experience it is known that flower bulb fields can be classified relatively easily from satellite data in areas where flower bulb fields are abundant. Therefore, it is expected that the user’s and producer’s accuracy for flower bulbs is in the order of 85% in areas where flower bulb cultivation is abundant.

Classification results are relatively poor for the class ‘other crops’ with 49.9% user’s accuracy and 52.6% producer’s accuracy. Similar classification results for the class ‘other crops’ were obtained with the LGN2 and LGN3 databases. There are several reasons for this poor classification result. Firstly, the class consists of a large number of crops with different spectral and phenological properties, which makes them difficult to recognise. Secondly, many of the crops in this group are horticultural crops that are often grown in fields that are small compared to the pixel size of Landsat TM. Thirdly, the assignment of a field to the class ‘other crops’ is in many cases not based on a positive recognition of the crop but merely based on the fact that all other crop types are ruled out. This causes classification errors in one of the other classes to accumulate in the class ‘other crops', while classification errors in the class 'other crops' are distributed over the 6 other crop types.

[Insert table 3 about here]

The overall classification accuracy at the provincial level (figure 5) demonstrates that the classification accuracy is homogeneous over the entire country and for most provinces between 80% and 95%. The classification accuracy is relatively low in the province of Flevoland (71.1%). This low accuracy is in contrast to what we expected because Flevoland is characterised by large-scale agriculture with large, rectangular fields and a homogeneous soil type. Therefore, we suspect that the poor classification accuracy is caused by non-representative sampling. This hypothesis is supported by the small number of reference pixels in this province compared to most of the other provinces.

[Insert figure 5 about here]

4.2. Comparison of LGN4 statistics to the CBS agricultural statistics

We compared the crop acreage derived from the LGN4 crop map with the acreage reported by the CBS agricultural statistics. This must be regarded as a qualitative validation because the two data sources are not strictly independent. Nevertheless, it gives insight in how well crop statistics can be predicted from the LGN4 crop map.

We selected two scatterplots that demonstrate the relationship between the LGN4 and CBS statistics for grassland (figure 6) and cereals (figure 7). For the class grassland there is a very good relationship, but the LGN4 crop map systematically overestimates the acreage of grassland. The overestimation is due to a difference in measurement between the statistics and the LGN4 crop map. In the CBS statistics only grassland is counted that is in use for agricultural production, while the LGN4 crop map also includes yards, banks and borders that are not counted in the CBS statistics. Furthermore, the census that the CBS carries out only counts farms larger than a certain threshold. The majority of the small farms excluded are dairy farms and this will cause an underestimation of the true area of grassland.

The difference in measurement and class definition does not play a role for the class cereals. In this case the relationship between the LGN4 statistics and the CBS statistics is almost perfect and without a bias towards overestimation. The scatterplots for most other crops are similar to the scatterplot for cereals.

[Insert figure 6 and figure 7 about here]

4.3. Efficiency of per-field classification

The crop classifications in the LGN3 and LGN4 databases have been obtained using different classification techniques. For LGN3 semi-automated per-pixel classification was applied which relied heavily on visual interpretation, while LGN4 was classified using a per-field classification approach. Both LGN versions have a similar accuracy and we can therefore compare the effort that was spent in order to evaluate if a per-field classification approach really is more efficient.

We made an estimate of the total man-hours that was spent for both LGN versions, in order to derive a crop map from the satellite imagery. In the case of the LGN3 database, the man hours consisted only of hours for classification, while for the LGN4 database the total man-hours were divided into hours for digitising and hours for classification (table 4).

The results demonstrate that with regard to classification, the total man-hours could be reduced from about 1500 to 500 hours. These results confirm that per-field classification is more efficient and speeds up the classification process with a factor three. However, the total man-hours that was spent in deriving the crop field database (digitising) is in the order of 2300 hours, which almost doubles the total man-hours spent for mapping the crops in the LGN4 database compared to the LGN3 database.

The total cost of mapping the crops in the LGN4 database (table 4) was less than the cost of mapping in the LGN3 database, despite the large effort spent on digitising field boundaries. The explanation is that digitising field boundaries is a relatively simple job and doesn't need to be carried out by a skilled remote sensing operator. Digitising can thus easily be contracted out, thereby reducing cost.

[insert table 4 about here]

5. Conclusions and discussion

We have implemented a methodology for the per-field mapping of crops in the Netherlands. The methodology integrates multi-temporal and multi-sensor satellite imagery, statistical data on crop acreage and parcel boundaries from the 1:10,000 digital topographic vector database. In the first stage of the methodology we created a crop field database covering the entire Netherlands by selecting the static boundaries of agricultural parcels from the 1:10,000 digital vector database and by adding the dynamic crop boundaries from satellite images using on-screen digitising. In the next stage, the crop discrimination procedure made use of the phenological and spectral properties of each field in order to determine the crop type. The use of the crop field boundaries effectively eliminated the classification errors due to the within-field spectral variability and mixed-pixels along the boundaries of fields.

Field information on crop type is needed for obtaining a thorough knowledge of the spectral and temporal properties of different crops. This methodology demonstrates that, once this knowledge has been collected, it can be transferred to new satellite datasets. Crop classifications were carried out using 'interpreted' training samples and statistical data on crop acreage were used to guide the classification process. The accuracy assessment of the LGN4 crop map demonstrates that this approach leads to a highly accurate crop map.

The error matrix at the national level demonstrates that we achieved the aim of 85% overall accuracy and 75% accuracy at class level for most crops. A logical explanation of the relatively poor classification results for those crops that did not meet the objective could be given ('flower bulbs' and 'other crops'). At the provincial level, the validation results illustrate that the use of different combinations of satellite imagery for different parts of the country has not lead to large spatial variation of the overall map accuracy.

There is a good relationship between the crop acreage derived from the LGN4 crop map and the acreage reported by the CBS statistics for all classes. These results indicate that the LGN4 crop map can predict crop statistics well. The LGN4 crop map is depicting the individual fields which is a distinct advantage over the CBS agricultural statistics because the CBS statistics are only available for administrative units.

The evaluation of the efficiency of the per-field classification versus a per-pixel classification showed that the use of field boundaries decreased the effort for classification from 1500 to 500 man-hours. However, the effort spent on digitising was estimated at 2300 man-hours, which almost doubled the total time effort for the LGN4 crop classification compared to the LGN3 crop classification. Given the large effort that had to be spent on digitising the crop field boundaries, the question can be raised whether image segmentation techniques are a more efficient alternative for deriving field segments.

From a practical point of view, image segmentation was of no use for the LGN4 crop map because, at the start of the project, segmentation algorithms were not implemented in operational software packages and were difficult to implement in a map production environment. However, from a methodological point of view there are several disadvantages to the use of segmentation techniques.

First of all, the results of segmentation algorithms are often unpredictable and in many cases they fail to collect all boundaries that can be discriminated by the human eye. Also, the regions that a segmentation algorithm creates do not necessarily relate to meaningful entities (Hill 1999) because the algorithm cannot take contextual information (parcel shape and occupation pattern) into account. In cases where segmentation was successfully applied for obtaining crop fields (Janssen and Molenaar 1995) segmentation was only proven to work for areas with simple field geometry (large rectangular fields).

A further disadvantage of applying image segmentation to Landsat TM images without using a large-scale topographic map, is that sub-pixel elements like small roads, ditches, hedges and streams will be included in the field segments. This causes an overestimation of the true cropped area. Particularly if crop fields are small, like in many areas in Europe, the crop area overestimation can be significant. Estimation of this effect using the LGN4 crop map demonstrates that the overestimation varies between 3% to 9% of the total crop area. A crop field database derived from a large-scale topographic map does not have this effect because the static parcel boundaries have been derived from aerial photography and have a much higher spatial accuracy.

The most promising way to apply segmentation techniques for crop mapping is by combining image segmentation techniques with a large-scale topographic map. In such a methodology the segmentation is carried out within the parcel boundaries that are defined by the large-scale topographic map. An additional problem that needs to be solved in such a methodology is that image segmentation provides raster segments. Combining vector boundaries obtained from a large-scale topographic map with blocky vector boundaries obtained from vectorising the raster segments is not a trivial task, and can easily lead to under- and overshoots and polygon errors.

Still, even if this methodology proves to be successful, it has to be tested whether it is more cost-efficient than on-screen digitising. Preliminary analyses of the current crop field database with older satellite imagery indicate that the 'dynamic' crop boundaries are less dynamic than thought, and many boundaries can be re-used. We therefore expect that updating the crop field database for classifying a new series of satellite images involves considerably less effort compared to creating the initial crop field database. In such a scenario, the use of image segmentation techniques really needs to be highly cost-efficient compared to on-screen digitising before it can be successfully implemented.

References

Aplin, P., Atkinson, P.M. and Curran, P.J., 1999, Per-field classification of land use using the forthcoming very fine spatial resolution satellite sensors: problems and potential solution. In: P.M. Atkinson and N.J. Tate (Editors), Advances in remote sensing and GIS analyses. Wiley, Chichester, pp. 273.

Baker, J.R. and Drummond, J.E., 1984, Environmental monitoring and map revision using integrated Landsat and digital cartographic data. ITC-Journal (1), pp-10-19.

Cross, A.M., Mason, D.C. and Dury, S.J., 1988, Segmentation of remotely-sensed images by a split-and-merge process. International Journal of Remote Sensing, 9(8): 1329-1345.

Fuller, R.M., Smith, G.M., Sanderson, J.M., Hill, R.A. and Thomson, A.G., 2002, The UK Land Cover Map 2000. Construction of a parcel-based vector map from satellite images. The Cartographic Journal, 39(1), 15-25.

Hill, R.A., 1999, Image segmentation for humid tropical forest classification in Landsat TM data. International Journal of Remote Sensing, 20(5), 1039-1044.

Janssen, L.L.F., Jaarsma, M.N. and Van der Linden, T.M., 1990, Integrating topographic data with remote sensing for land-cover classification. Photogrammetric Engineering & Remote Sensing, 56: 1503-1506.

Janssen, L.L.F. and Molenaar, M., 1995, Terrain objects, their dynamics and their monitoring by the integration of GIS and remote sensing. IEEE Transactions on Geoscience and Remote Sensing, 33(3), 749-758.

Jewell, N., 1989, An evaluation of multi-date SPOT data for agriculture and land use mapping in the United Kingdom. International Journal of Remote Sensing, 10, 939-951.

Kettig, R.L. and Landgrebe, D.A., 1976, Classification of multispectral data by extraction and classification of homogeneous objects. I.E.E.E. Transactions on Geoscience Electronics, 14, 19-26.

Lobo, A., Chic, O. and Casterad, A., 1996, Classification of Mediterranean crops with multisensor data: per-pixel versus per-object statistics and image segmentation. International Journal of Remote Sensing, 17(12), 2385-2499.

Lillesand, T.M. and Kiefer, R.W., 1994, Remote sensing and image interpretation. John Wiley & Sons, New York.

Mason, D.C., Corr, D.G., Cross, A., Hogg, D.C., Lawrence, D.H., Petrou, M. and Tailor, A.M., 1988, The use of digital map data in the segmentation and classification of remotely-sensed images. International Journal of Geographical Information Systems, 2(3), 195-215.

Mattikali, N.M., Devereux, B.J. and Richards, K.S., 1995. Integration of remotely sensed satellite images with geographical information systems. Computers and Geosciences, 21: 947-956.

Murakami, T., Ogawa, S., Ishitsuka, K. and Saito, G., 2001, Crop discrimitation with multitemporal SPOT/HRV data in the Saga Plains, Japan. International Journal of Remote Sensing, 22(7), 1335-1348.

Pedley, M.I. and Curran, P.J., 1991, Per-field classification: an example using SPOT-HRV imagery. International Journal of Remote Sensing, 12(11), 2181-2192.

Shandley, J., Franklin, J. and White, T., 1996, Testing de Woodcock-Harward image segmentation algorithm in an area of southern California chaparral and woodland vegetation. International Journal of Remote Sensing, 17(5), 983-1004.

Smith, G.M. and Fuller, R.M., 2001, An integrated approach to land cover classification: an example in the Island of Jersey. International Journal of Remote Sensing, 22, 3123-3142.

Thunnissen, H.A.M., Jaarsma, M.N. and Schouwmans, O.F., 1992a, Land cover inventory in the Netherlands using remote sensing; application in a soil and groundwater vulnerability assessment system. International Journal of Remote Sensing, 13, 1693-1708.

Thunnissen, H.A.M., Olthof, R., Getz, P. and Vels, L., 1992b, Grondgebruiksdatabank van Nederland vervaardigd met behulp van Landsat Thematic Mapper opnamen. Report 168, DLO Winand Staring Centre, Wageningen.

Thunnissen, H.A.M. and Noordman, E., 1996, Classification methodology and operational implementation of the land cover database of the Netherlands. Report 124, DLO Winand Staring Centre, Wageningen.

Thunnissen, H.A.M. and Wit, A.J.W., de, 2000, The national land cover database of the Netherlands. Geoinformation for all; XIX congress of the International Society for Photogrammetry and Remote Sensing (ISPRS), 17-21 July 2000 (Amsterdam: GITC), pp. 223-230.

List of figures

Table 1. Overview of the satellite images used to produce the LGN4 crop map.

Sensor	Date	Path	Row	Classification Zone
Landsat TM5	01-04-1999	198	23	1
Landsat TM5	01-04-1999	198	24	1
Landsat TM5	03-05-1999	198	23	1
Landsat TM5	03-05-1999	198	24	1
Landsat TM5	29-07-1999	199	24	1
Landsat TM7	30-07-1999	198	23	1
Landsat TM7	30-07-1999	199	24	1
Landsat TM7	06-05-2000	197	23	2/3
Landsat TM7	06-05-2000	197	24	2/3
Landsat TM7	06-05-2000	197	25	2/3
IRS-LISS3	09-06-2000	18	31	2
IRS-LISS3	09-06-2000	18	32	2
ERS2-SAR	04-06-2000	26787	2533	3
ERS2-SAR	09-07-2000	27288	2533	3
ERS2-SAR	13-08-2000	27789	2533	3
Landsat TM7	26-08-2000	197	23	2/3
Landsat TM7	26-08-2000	197	24	2/3
Landsat TM7	26-08-2000	197	25	2/3

Table 2. Area of the main agricultural crops in the Netherlands for the 66 CBS regions (hectares).

CBS Region number	CBS Region name	Grassland	Maize	Seed potatoes	Consumption potatoes	Sugar beet	Winter cereals	Spring cereals	Other crops	Spring flowerbulbs	Autumn flowerbulbs	Other flowerbulbs
2001	De Marne	2078	213	2487	214	1359	2866	729	801	50	0	0
2002	Centraal weidegebied in Groningen	22487	970	932	432	860	2792	680	584	6	0	0
2003	Oostelijke bouwstreek in Groningen	12817	1933	1254	6125	4557	17178	3365	3539	31	0	2
2004	Westerwolde en Groninger Veenkol.	4464	1574	185	11104	5033	2241	5623	1309	0	31	2
2005	Groninger zuidelijk Westerkwartier	14034	1151	17	92	8	54	67	16	0	6	0
2006	Oostelijk Hogeland	5969	418	3667	694	2522	5724	1005	1612	29	0	1
2101	Noordelijk Friesland	32884	1842	6065	739	4084	1163	3617	2872	141	38	0
2102	Weidestreek in Friesland	82010	3382	465	182	318	106	365	236	71	60	0
2103	De Wouden	75708	7095	565	694	212	47	326	264	0	294	0
2104	Eilanden	3573	25	0	1	0	0	0	1	0	0	0
2201	Weidegebied van het Noorderveld	6488	1184	117	1062	209	209	436	42	0	45	0
2202	Smilde & Centraal zandgeb. Drenthe	15756	3921	704	6236	2777	1270	4011	809	6	255	0
2203	Zuidwestelijk weidegeb. Drenthe	22409	4040	219	972	237	388	731	159	15	101	4
2204	Zuidelijk zandgebied in Drenthe	12328	3661	276	4201	2049	797	2213	1047	0	11	2
2205	Drentse Veenkoloniën en Hondsrug	10045	2811	863	16463	9240	3019	8282	1486	0	54	0
2301	Weidegebied in Overijssel	54629	6967	88	1122	215	488	490	242	20	191	0
2302	Noordoost Overijssel	17841	6857	280	4515	1459	671	1957	778	22	210	5
2303	Twente	54754	21585	208	1119	345	1538	1187	670	5	45	0
2304	Salland	22505	6247	69	1074	668	733	720	234	4	47	0
2401	Noordoostelijke Polder	3793	795	6597	4165	6286	1672	2588	8457	2029	371	84
2402	Zuidelijke IJsselmeerpolders	8703	3089	1801	10365	9003	5989	2699	10808	174	0	0
2501	Oostelijke Veluwe	17511	2942	6	349	141	288	257	108	2	6	0
2502	IJsselstreek	17554	4398	72	587	545	808	329	124	6	0	0
2503	Zuidelijk Gelderland	14649	4099	138	398	937	1824	234	612	0	0	0
2504	Oostelijke Betuwe en Nijmegen	3354	558	35	206	375	978	239	102	0	0	0
2505	Veluwezoom en Betuwe	25997	3728	24	751	992	2058	675	674	1	0	0
2506	Bommelerwaard	6244	1051	3	127	64	188	68	204	0	0	0
2507	Westelijke Veluwe	22174	6221	4	271	211	496	391	214	14	23	19
2508	Achterhoek	57194	22110	1152	2747	1453	2995	2671	667	217	71	0
2601	Kromme Rijn-streek en Heuvelrug	9599	1439	0	47	79	120	66	100	0	0	1
2602	Westelijk weidegebied in Utrecht	31718	1608	0	19	39	37	43	47	0	0	0
2603	Centraal tuinbouwgebied in Utrecht	4402	255	15	34	108	0	21	80	8	0	1
2604	Eemland	6800	552	0	26	28	18	34	45	0	0	0
2605	Zandgebied in Utrecht	7966	1861	1	35	30	54	79	26	2	0	2
2701	Wieringen en Wieringermeer	1996	497	3136	1314	3011	1599	1085	3105	446	166	22
2702	Haarlemmermeer	977	177	98	1804	1910	1980	922	1020	141	6	82
2703	Amstelland en Aalsmeer (- A'dam)	1515	93	29	223	216	208	47	39	2	0	6
2704	Texel en Land van Zijpe	9587	691	1838	267	1295	458	1245	1488	3932	1515	617
2705	West-Friesland en omgeving	22750	1054	1759	570	344	362	546	5478	3725	865	168
2706	Waterland + NH Droogmak. (+ A'dam)	27934	992	244	748	787	477	615	699	272	2	45
2707	t Gein en Gooiland (excl. A'dam)	5675	155	0	1	26	71	36	59	0	0	0
2708	Kennemerland	4062	185	10	20	69	43	11	178	616	143	183
2801	Voorne-Putten en Hoeksche Waard	6317	678	146	5835	4150	3953	4140	5308	4	0	1
2802	Rotterdam en omgeving	1896	59	0	425	238	123	721	1111	0	0	1
2803	Goeree-Overflakkee	1922	561	303	3390	2425	901	2136	3380	380	2	1
2804	Westelijk Rijnland	10415	745	251	446	433	216	367	581	32	57	10
2805	Boskoop en Rijneveld	2522	152	128	246	175	179	204	394	0	0	0
2806	Krimpenerwaard en Oostel. Rijnland	22433	566	18	112	67	25	27	140	0	2	0
2807	Alblasserwaard en Vijfherenlanden	25295	758	0	67	53	36	47	42	0	0	0
2808	Bollenstreek	2300	28	2	11	9	19	9	197	2048	52	285
2809	Westland en ZH Droogmak. (- R'dam)	7445	445	365	929	799	665	794	1289	0	21	5
2901	Noordelijk Zeeland	3889	1038	270	5950	4320	2239	4622	7798	56	5	50
2902	Walcheren en Zuid-Beveland	6045	1649	244	5095	5105	4454	5301	6861	126	0	1
2903	Zeeuwsch-Vlaanderen	5072	2057	1930	6798	7792	6358	6934	14262	116	7	2
3001	Noordwesthoek	2246	1003	1139	3439	2859	2902	1848	5515	50	23	0
3002	Westelijke Langstraat	6222	2930	39	870	793	416	505	1477	0	0	0
3003	Biesbosch	1377	341	7	1109	886	952	452	1867	1	0	0
3004	Oostelijke Langstraat	8234	3528	19	1098	762	530	494	1356	0	0	0
3005	Westelijke Zandgronden	8189	4128	407	1862	1522	539	885	4835	52	105	2
3006	Land van Breda	13011	3049	0	133	119	38	148	1124	0	3	0
3007	De Kempen	14752	13467	73	3785	2581	235	373	3680	18	8	7
3008	Midden Noord-Brabant	18253	12601	87	1410	1077	243	686	1799	6	8	4
3009	Maaskant en Land van Cuijk	13908	11956	81	2104	2264	487	965	1791	106	80	6
3010	Westelijk Peelgebied	25104	23571	90	2187	2099	568	874	3371	68	20	23
3101	Noord-Limburg	23369	16224	159	6446	7657	3548	2779	9975	381	565	152
3102	Zuid-Limburg	11667	4794	41	1951	4218	5815	866	878	0	0	0
	Total	1013803	240753	41223	137811	116505	99449	86877	130064	15430	5515	179

Table 3. Error matrix for the LGN4 crop map at the national level.

LGN4	Reference data: grassland	maize	potatoes	sugar beet	cereals	other crops	flower bulbs	user's accuracy
grassland	462520	10216	1562	635	5107	9498	196	94.4%
maize	2884	135913	4478	2046	476	944	141	92.5%
potatoes	5183	2084	80866	598	1467	3853	89	85.9%
sugar beet	227	1387	1117	48124	80	950	0	92.8%
cereals	2064	249	1368	249	111202	6199	124	91.6%
other crops	3920	1826	11796	676	5612	24234	474	49.9%
flower bulbs	1080	330	103	69	0	355	2382	55.2%
producer's accuracy	96.8%	89.4%	79.8%	91.8%	89.7%	52.6%	69.9%	90.4%

Table 4. Comparison of man-hours and cost involved in creating the LGN3 and LGN4 crop maps.

	LGN3	***LGN4***
Man-hours		
Classification	1500	520
Digitising	-	2230
Cost (€)		
Classification	144000	50000
Digitising		40000

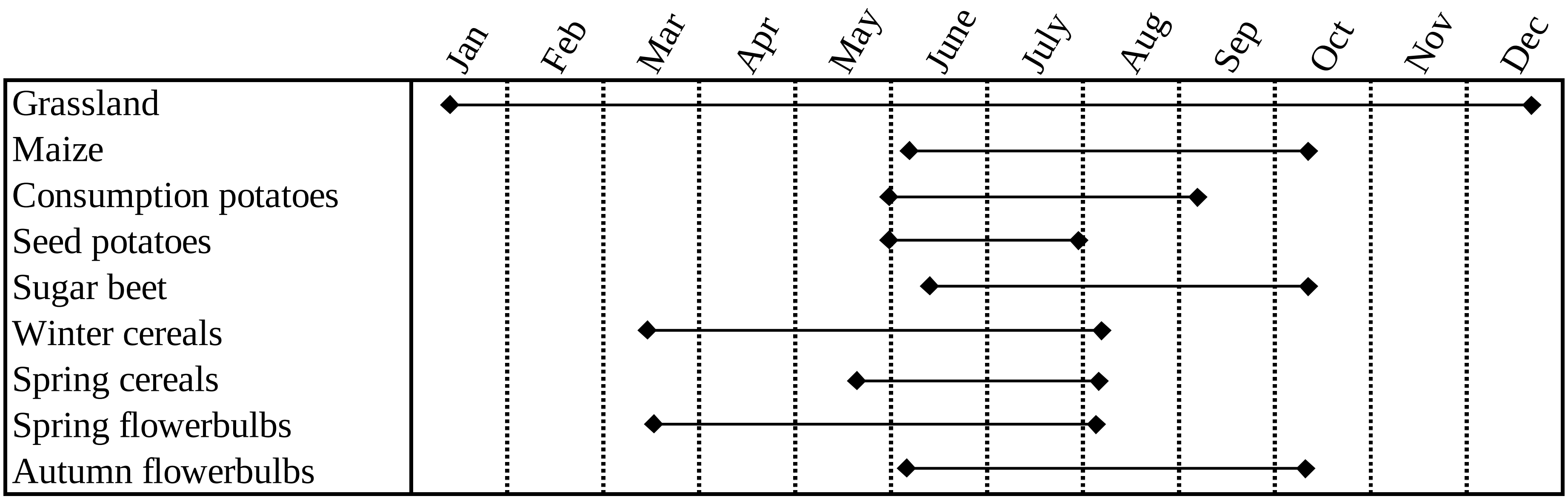

Figure 1. Crop calendar for satellite mapping of the most important crops in the Netherlands

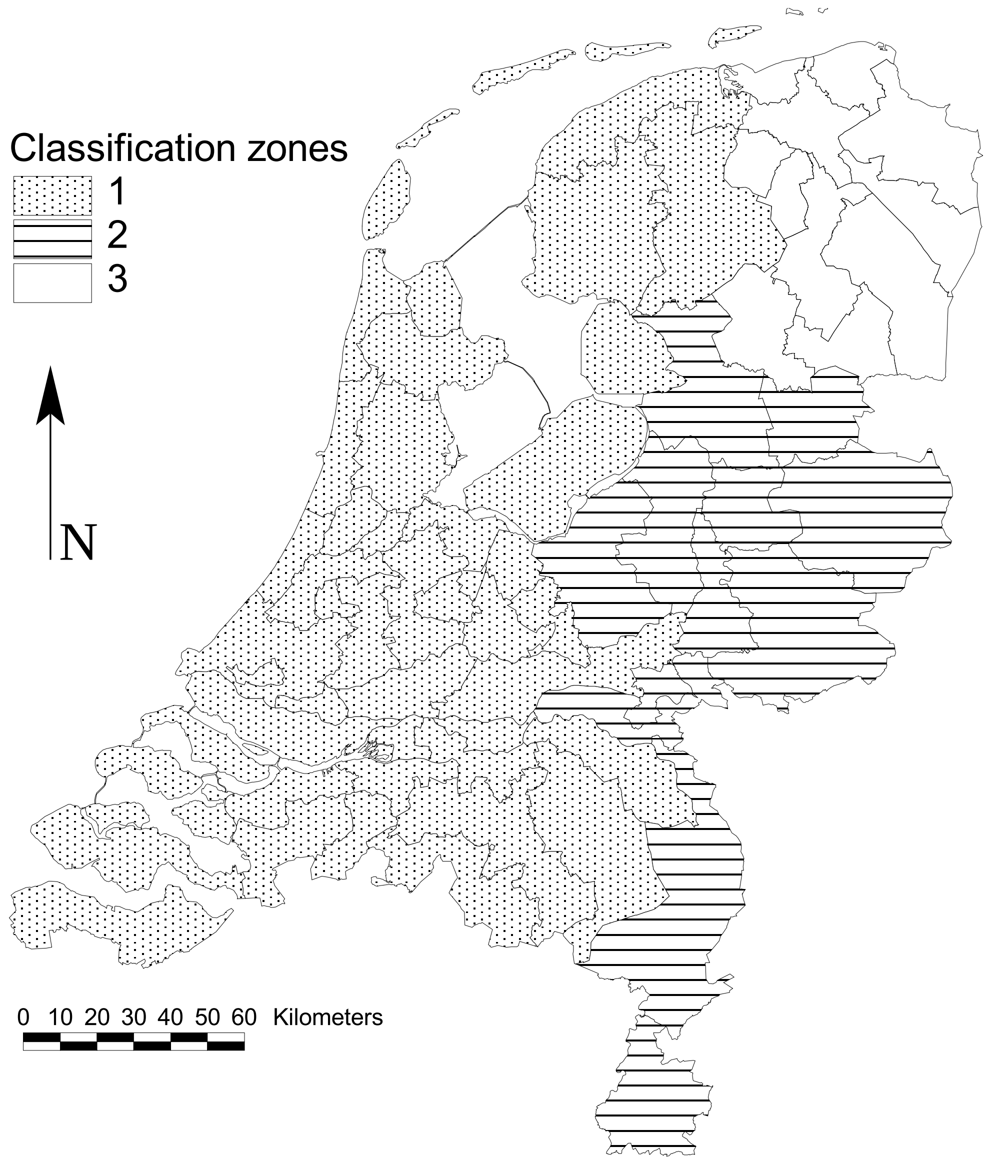

Figure 2. Boundaries of the 66 CBS regions in the Netherlands and the division in classification zones.

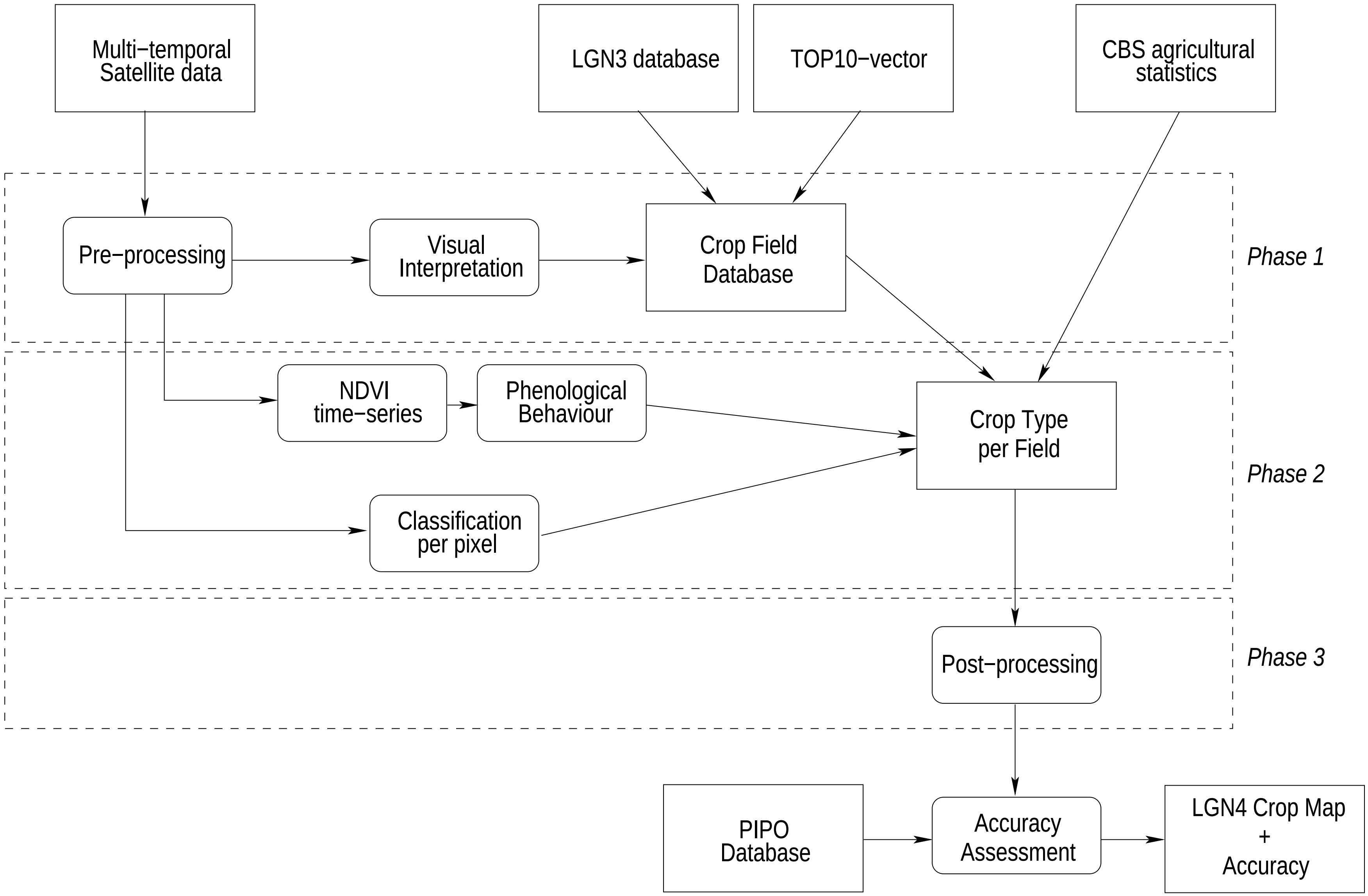

Figure 3. Schematic overview of the processing steps that were carried out to derive the LGN4 crop map.

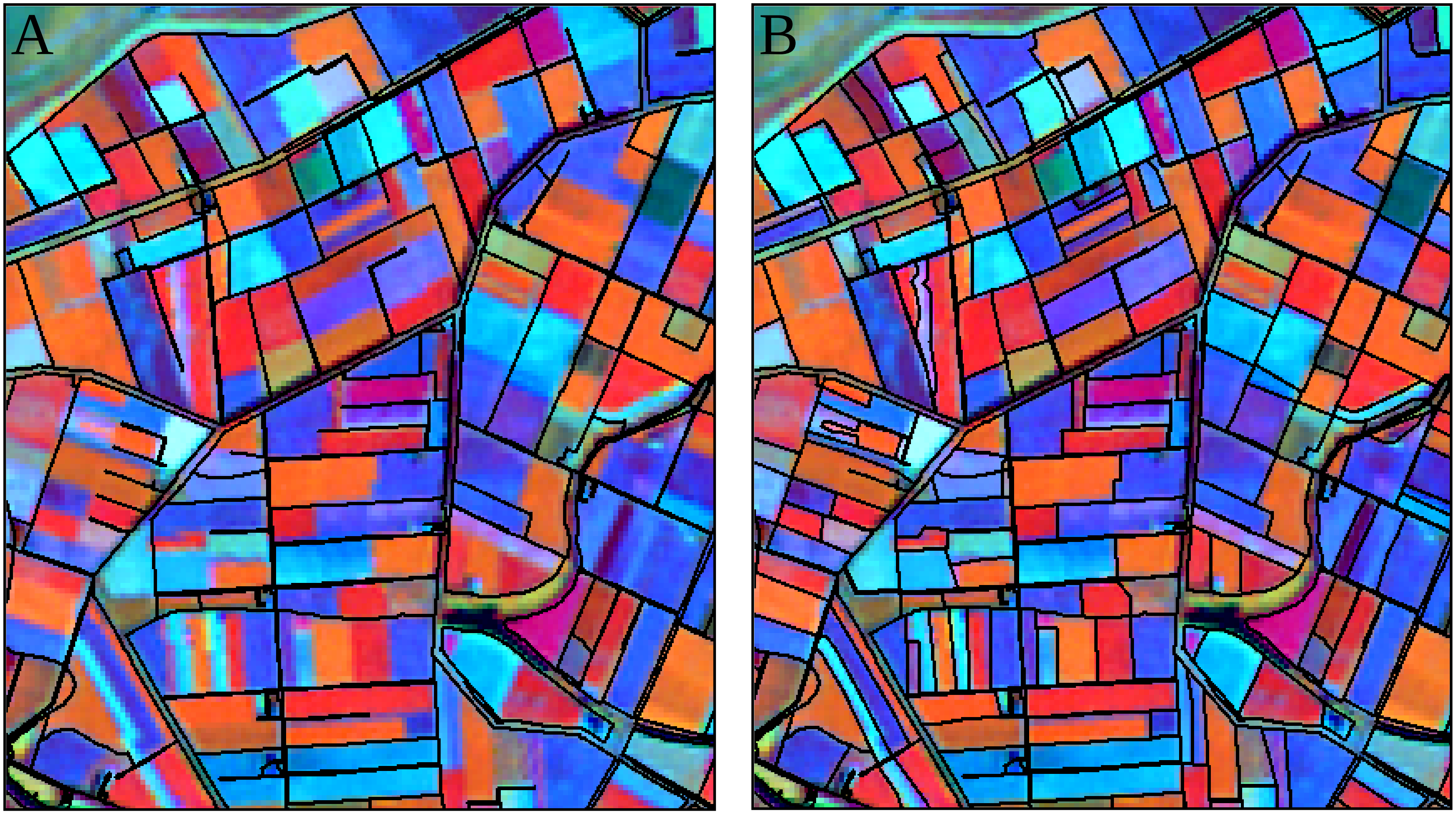

Figure 4. Landsat 7 ETM images overlayed with the topographic agricultural parcels (A) and the crop field database (B).

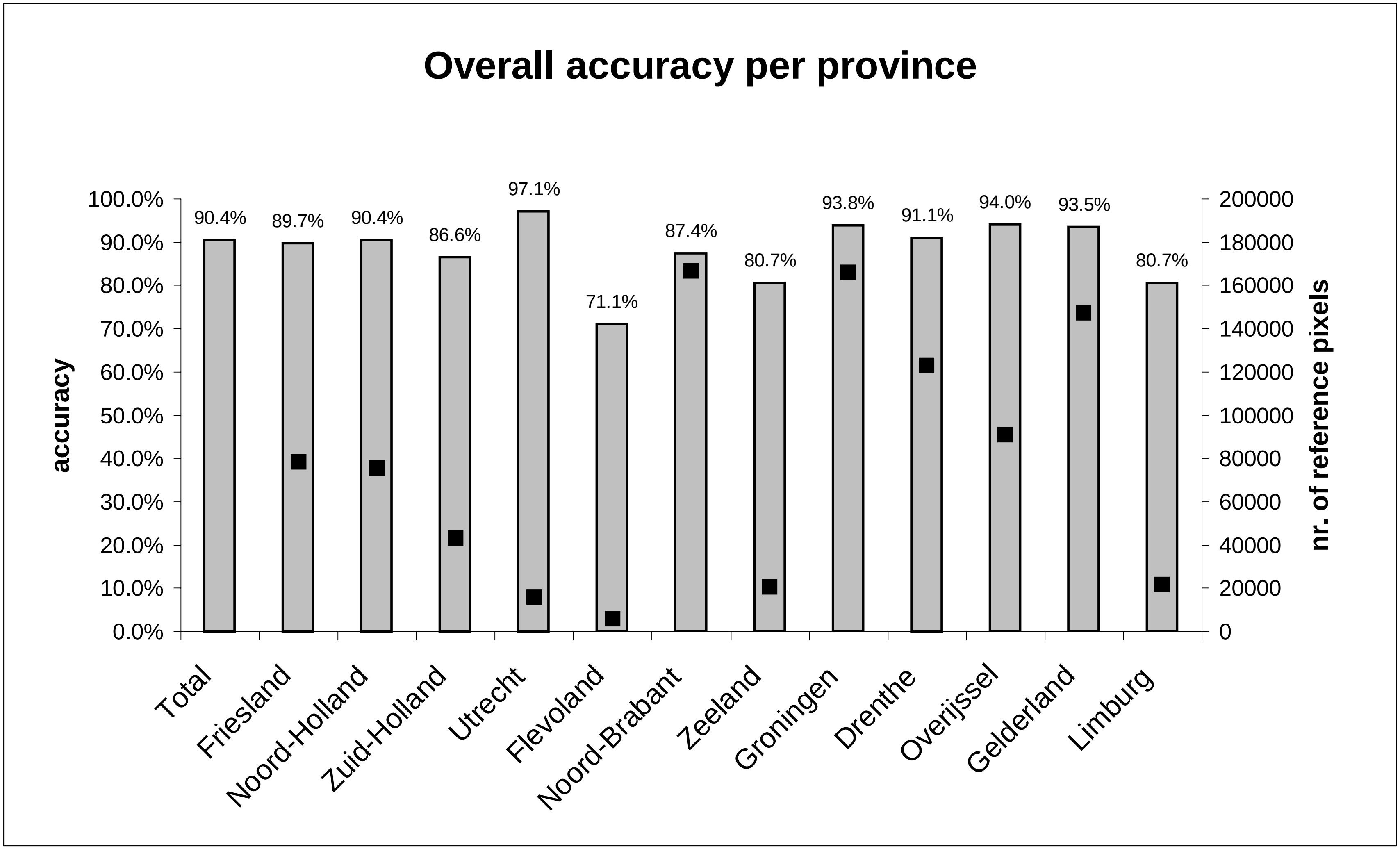

Figure 5. Overall accuracy of the LGN4 crop map per Dutch province. Bars indicate the overall accuracy on the primary y-axis, black squares indicate the number of reference pixels on the secondary y-axis.

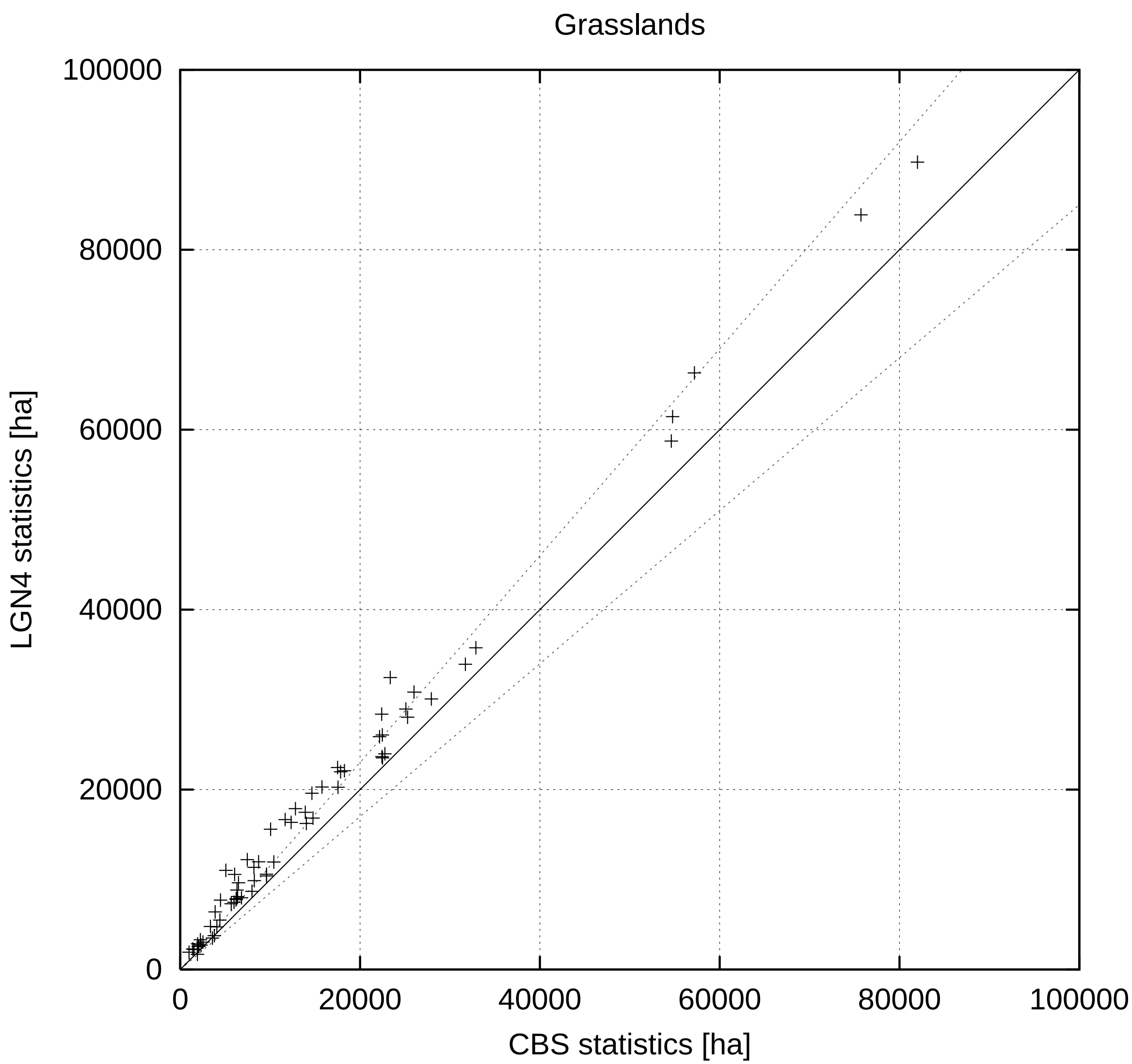

Figure 6. Comparison of area statistics obtained from the LGN4 crop map and area statistics reported by the CBS for grasslands in 66 CBS regions. The dotted lines are the lines of 15% deviation.

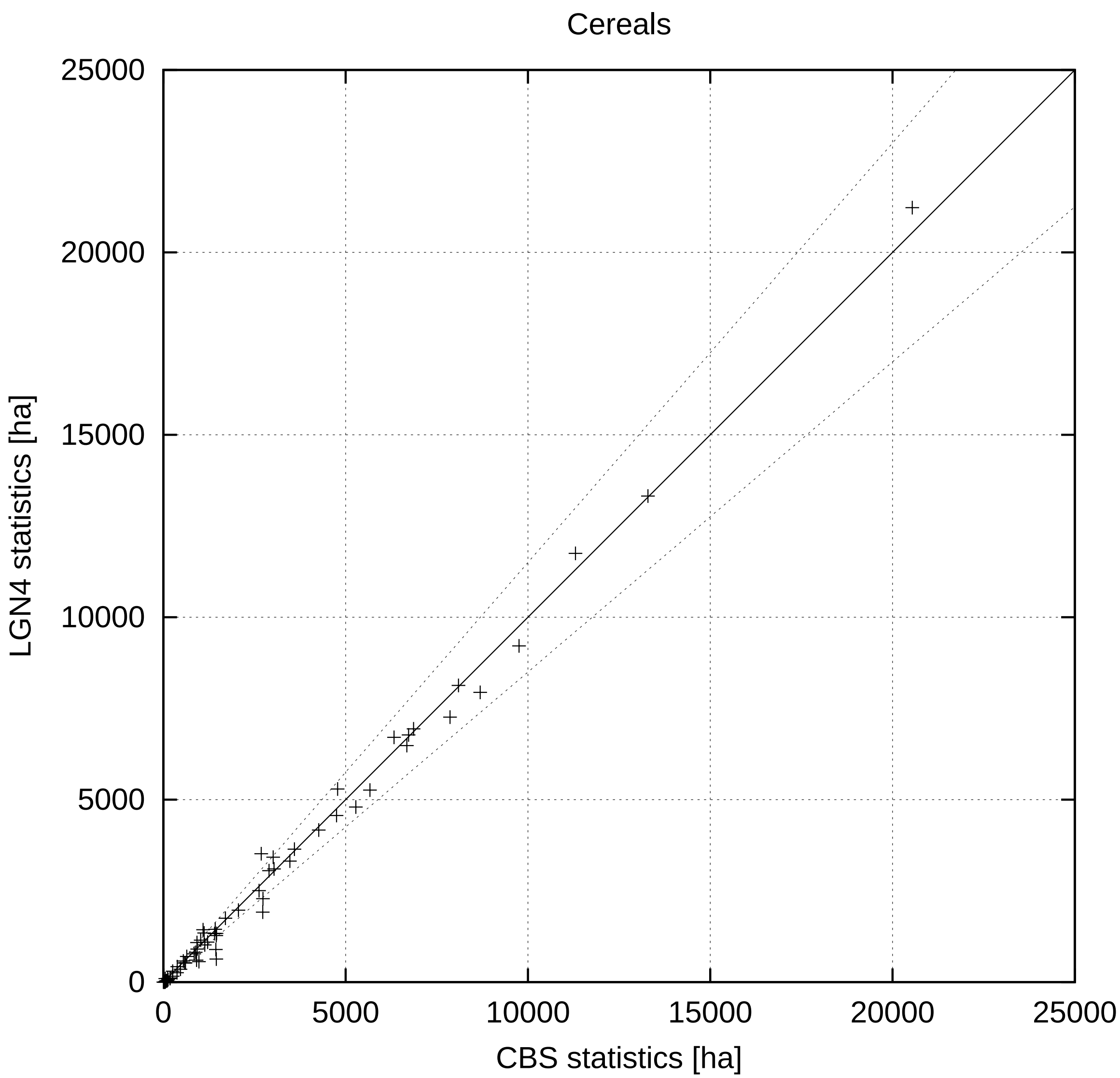

Figure 7. Comparison of area statistics obtained from the LGN4 crop map and area statistics reported by the CBS for cereals in 66 CBS regions. The

dotted lines are the lines of 15% deviation.