Classification methodology and operational implementation of the land cover database of the Netherlands

H.A.M. Thunnissen and E. Noordman

Report 124

DLO Winand Staring Centre, Wageningen (The Netherlands), 1996

ABSTRACT

Thunnissen, H.A.M. and E. Noordman, 1996. *Classification methodology and operational implementation of the land cover database of the Netherlands.* Wageningen (The Netherlands), DLO Winand Staring Centre. Report 124. 88 pp.; 1 Figs; 23 Tables; 47 Refs.

Timely and accurate information on land cover on regional and national scales is required to support environmental policy and for physical planning purposes. In 1987 the LGN1 land cover database was produced with satellite images. An improved classification method has been developed, consisting of the integrated use of satellite images, digital geographical data, reference data, and expert knowledge in geographical information systems. The LGN1 database has been updated (LGN2 database) and the production of the LGN3 database has been started. To achieve commercial implementation or to continue operational implementation, the advantages of the LGN database over other digital geographical databases should be exploited.

Keywords: geographical information system, remote sensing, satellite image

ISSN 0927-4537

This report has also been published as BCRS report

Project 7224 [Rep124.HM/01.97]

Contents

Preface

In 1987 it was decided to produce a land cover database of the Netherlands (further to be mentioned 'LGN1 database'), using satellite images. The classification accuracy of the LGN1 database showed a large variation over the country. Altough the classification results were disappointing in some areas, evaluation of the project showed that users of the LGN1 database were interested in an up-date of the database. However, the classfication results should be improved considerably. In 1992 the 'LGN Research Project' was started, aiming to develop an improved classification method and to investigate the possibilities to up-date the LGN1 database and to implement the database operationally, c.q. commercially. The results of this project are presented in this report. The project was carried out in the framework of the National Remote Sensing Programme (NRSP-2), under responsability of the Netherlands Remote Sensing Board.

W.W.L. van Rooij is acknowledged for the analysis of the ERS-1 images and for the investigation of the possibilities of satellite images to classify buildings in agricultural area.

Summary

Backgrounds
Because of the increasing concern about the impacts of man's intervention on the environment, timely and accurate information on land cover at regional and national scales is required to support environmental policy and for physical planning purposes. Therefore, in 1987 it was decided to produce a land cover data base of the Netherlands (further to be mentioned 'LGN1 database'), using satellite images. The LGN1 database was produced by automatic classification of manually stratified single-date satellite images from 1986. The classification result showed a large variation over the country due to spectral confusion between different land cover classes. A comprehensive evaluation of the LGN1 database showed that users were interested in an up-date of the database. However, the classification result should be improved considerably. In 1992 the 'LGN Research Project' was started, aiming to develop an improved classification method and to investigate the possibilities to up-date the LGN1 database and to implement the database operationally, c.q. commercially. In 1993 the up-dating of the LGN1 database was started (the up-dated version of the LGN1 database will further be described as 'LGN2 database'). The up-dating was partially performed simultaneously with the LGN Research Project and was finished at the end of 1995. Dependent on the progress of the up-dating, results of the LGN Research Project which could applied operationally were incorporated in the current classification method or will be applied in future up-dates of the LGN database. Conversely, some problems met during the up-dating of the LGN-database were included in the LGN Research Project for additional research. In the framework of the LGN Research Project Landsat TM, SPOT, and ERS-1 satellite images were used.

Significant improvements of the classification result could be expected by a reduction of the spectral confusion between the different land cover types. Similarity in spectral reflectances at the image acquisition date impedes consistent identification and mapping of a large number of important land cover types when using single-date satellite images. However, spectral signatures of a wide range of cover classes, such as agricultural crops or natural vegetation, vary throughout the year. By that, classes which appear very similar in spring, may become separable at other stages of the phenological cycle. It is therefore expected that multi-temporal approaches provide important means to improve classification accuracy. Further classification improvement of spectral overlapping land cover classes may be expected by the use of other digital geographical data. These data often contain useful additional spatial or temporal information on land cover classes. Moreover, the discrimination between different land *use* classes may be impossible because they posses similar spectral reflectance. For example, a short herbaceous cover may represent agricultural use, or recreational use, or residential use. In these cases discrimination between different land cover/use classes can only be achieved by using other digital geographical data or by visual image interpretation techniques.

The (integrated) use of additional digital geographical data and multi-temporal satellite data for obtaining land cover information was an important research item

in the LGN Research Project. Further, much attention was paid to the development of a validation procedure for the LGN database.

Combined use of satellite images and other digital geographical data

In the Netherlands several nation-wide digital geographical databases are available or will be available in near future. The most relevant databases are the 'CORINE Land Cover database', the 1 : 50 000 soil map (including water-table classes), the 'Land Use Database' of the State Department for Physical Planning (the so-called 'BARS Land Use Database'), the 'Land Use Database' of the Central Bureau of Statistics (the so-called 'CBS Land Use Database'), topographic maps at scales 1 : 50 000, and 1 : 25 000/10 000 (the topographic maps on the scales 1 : 25 000 and 1 : 10 000 contain the same information), and the 'Agricultural Statistics' of the CBS. The possibilities to use these databases in combination with satellite images in order to improve the classification result were assessed.

Because of the deviating nomenclature and scale of the CORINE Land Cover database this database is not suitable to be used for stratification or postclassification sorting in the framework of the LGN project. The BARS and CBS Land Use Databases and the digitized topographic databases enable, in principle, discrimination between the main land use types in the LGN2 database, i.e. agricultural area, built-up area, and natural area and forest. However, because of the thematic classes, cost, accuracy and continuity of the data, use of the CBS Land Use Database is preferred. The CBS Agricultural Statistics enable a further subdivision of agricultural area. In the framework of the LGN project, use of these ancillary digital data for stratification purposes is preferred to use for postclassification sorting. Stratification is rather simple, during the classification it is easier to deal with smaller areas, it decreases spectral variation and confusion and enables to better focus the discrimination process on problem classes. At last, stratification may separate different land use classes which are spectrally similar. On the contrary, the recoding of misclassified spectral classes by postclassification sorting can be rather troublesome. When ancillary data are not available before the classification, use of these data for postclassification sorting may be efficient. Because it is applied after classification, misclassifications can be corrected by recoding, avoiding a time comsuming new classification.

Some LGN land cover classes, especially greenhouses, orchards, roads and buildings in agricultural areas, are also included in other digital geographical databases (especially the CBS Land Use Database and the digitized topographic databases). Classifying these classes again by interpretation of satellite images is, generally, waste of time. However, satellite images can sometimes be used for up-dating these classes. Further, the use of satellite images can also be preferred because of financial reasons.

When land cover statistics for strata are available (e.g. CBS Agricultural Statistics) class-based a-priori probabilities, estimated by the relative areas of the land cover classes, can simply be included in the (maximum likelihood) classification process. However, even for spectral overlapping classes the increase in classification accuracy is small. In stead of providing a-priori probabilities, the CBS Agricultural Statistics seem more suitable to be used for stratification and validation purposes. The possibilities of using *conditional* a-priori probabilities, based on soil type and water table classes, in the maximum likelihood classification process are rather poor. Generally, the increase in classification accuracy is small. Moreover, estimation of conditional a-priori probabilities requires often a considerable additional sampling effort. The cost of acquiring these data does not balance the expected classification improvement. On very wet soils (water table classes I, II and II*) grassland is the only crop grown. Water table classes could also be used to discriminate between wet and dry natural areas. In these cases water table classes should be used for additional stratification, rather than for providing a-priori probabilities.

A field-based classification could considerably improve the classification result for agricultural crops. The 1 : 10 000 topographic database may be used as a base to obtain actual field boundaries by automatic and/or visual interpretation of satellite images. However, it has to be investigated if the required field boundaries can be obtained in an operational and cost effective way.

Classification methodology and optimal acquisition dates and availability of satellite images

An improved classification methodology for the LGN database has been developed, consisting of the integrated use of satellite images (i.c. Landsat TM and SPOT), digital ancillary data, reference data, and expert knowledge. The classification method is characterized by a stratified approach, i.e. every stratum is separately classified. For an optimal classification result the following strata have to be distinguished: agricultural area, urban area, less densely built on area, dry natural area (including forest), wet natural area (including forest) and water. If the stratification is outdated with respect to the acquisition dates of the satellite images the strata can be up-dated by visual interpretation of the satellite images, supported by simultaneous consultation of topographic maps and aerial photographs. The CBS agricultural regions should be used for a further subdivision of the agricultural strata.

In general, mono-temporal classification using Landsat TM images, obtained during the period mid-May to late September, provide good classification results (> 90%) for most non-agricultural classes. For an accurate classification of most agricultural crops the use of multi-temporal satellite data is required. It is preferred to use the original spectral bands for image classification and to classify each satellite image separately. Subsequently, the classified images can be combined in a GIS to form the final classification result, using conditional 'IF-THEN' statements. The use of a vegetation index (e.g. NDVI) may be useful for the discrimination between bare and vegetated fields, especially in spring. In practice, visual interpretation often appears to be a valuable tool, complementary to automatic classification. Advanced hardware and software enable the simultaneous interpretation of different satellite images, while the interpretation result can directly be stored in digital form by on screen digitizing.

The extent to which use of multi-temporal data improves the classification result is dependent on the cover types involved, crop growth conditions and the spectral resolution, number and acquisition dates of the used satellite images. In general, Landsat TM images are preferred to SPOT images because of the presence of middle-infrared TM bands required for accurate land cover classification. Moreover, Landsat TM images are considerably cheaper than SPOT images. However, the ageing sensors on board Landsat 5 and the loss of Landsat 6 at launch threaten continuity of Landsat imagery. On the other hand SPOT 4, planned for launch in 1996, will be equiped with a middle infrared band. The possibility of SPOT of pre-programmed, off-nadir, imaging will increase the chance of getting suitable images. Phenological data for the main crops growing in a stratum (e.g. planting/sowing date, ripening) and cultivation practices (e.g. conversion of grassland into arable land and vice versa, harvesting, after growth) should be taken into consideration when selecting optimal image acquisitions. For most agricultural areas, it is advised to use (Landsat TM) images obtained in several periods of the growing season. In general, the increased classification result counter-balances the additional cost for purchase and processing of additional images. When images from suboptimal periods or images with poor spectral resolution are used, spectral confusion may result in classification accuracies and reliabilities below 70%. To ensure the required minimum classification result, mixed agricultural classes (e.g. maize/sugar beet) have to be defined. Mixed classes, however, hamper efficient classification of the satellite images and operational application of the LGN database. Therefore, it is prefered to avoid using suboptimal images as much as possible, even if one would have to wait another growing season for more suitable images. For training the classifier and validation of the classification result suitable reference data are of great importance. In order to optimize the gathering of reference data, high quality quick look data should be made available on line within 24 hours of acquisition of the image.

The classification of greenhouses, orchards and buildings in agricultural area prove to be troublesome. Special classification techniques were developed for these classes. In order to get accurate information on the location of greenhouses, they should be digitized from maps or copied out of existing databases. Satellite images may be useful for up-dating of greenhouses in existing digital databases.

The use of satellite images obtained late in the growing season provides best classification results for orchards. Visual interpretation of these satellite images, supported by topographic maps, provides in most areas sufficient classification results for the LGN database.

Buildings in agricultural areas can not be sufficiently accurately classified by automatic classification of individual optical satellite images. Visual interpretation proves more succesful. Visual interpretation is not only guided by tone but also by size and situation (e.g. with regard to roads) of the buildings. A specific classification method has been developed, using these specific characteristics of buildings. The method exists of the combined use of multi-temporal NDVI images and ancillary data and the application of specific GIS techniques. By applying this method most (large) farms and clusters of small buildings are correctly classified, while scattered small buildings (dwelling-houses, sheds and the like) are only partly correctly classified. The backscatter values of ERS-1 SAR images are not suitable for the classification of buildings in agricultural area. It is interesting to investigate the possibilities of ERS multi-pass SAR inferometry for the classification of buildings in agricultural area.

In (near) future LGN classes like greenhouses, orchards and buildings in agricultural area, can simply be copied out of other available geographical databases. However, there may be financial or copyright contraints. Moreover, in practice, data in these geographical databases will often be outdated compared with the acquisition dates of the satellite images. In the latter case, satellite images could be used for up-dating of the concerning LGN classes.

The ERS-1 SAR, which operates in the microwave part of the spectrum, is not hindered by cloudiness or haze. Especially the multi-temporal classification of agricultural crops in the LGN database requires a regular data acquisition. A *field-based* multi-temporal classification of ERS-1 images provides, in principle, good classification results. A field-based classification requires the availability of digital field boundaries. These data are, however, not available and digitization of field boundaries for large areas is too expensive. A *pixel-based* multi-temporal classification of ERS-1 SAR images leads for all crops to significant lower classification results than the field-based classification.

Following the classification, different postprocessing techniques could be applied to further improve the classification result. To remove noise and to improve the overall classification accuracy a 3 x 3 pixel majority filter has to be applied on the output from the automatic classifier. The disappearance of small objects is not considered to be a problem for applications on a regional scale. In order to prevent pixels from being converted to very dissimilar classes, each stratum has to be filtered separately. The influence of mixed classes on the classification result could possibly be decreased by application of a *selective* majority filter. In the case of mixed classes, which contain only two crop types and have a relatively high reliability (at least 70%), application of a 3 x 3 selective majority filter could effect a considerable improvement of the classification result. The filter should be applied three times in succession. Only under specific conditions other window sizes may be applied and mixed classes which contain more than two crop types may be filtered.

Locally, misclassifications could be corrected by application of postprocessing techniques like 'CLUMP' and 'SIEVE' operations, specific developped filters or postclassification sorting.

As a final step in the post-classification processing it is important to perform a visual check of the classification result. In this way obvious classification errors can be interactively corrected.

Collection of reference data and classification accuracy assessment procedure

In order to create un unbiased and statistical valid sample of pixels, simple random sampling or stratified random sampling schemes are preferred. Due to time and budget constraints the accuracy assessment of the non-agricultural classes in the LGN2 database has been based on a stratified, *systematic* sampling scheme. For the concerning classes the results of this sampling scheme are comparable with the results of a random sampling scheme. By reason of efficiency, the sampling is concentrated on the most important (groups of) non-agricultural LGN2 classes. The accuracy assessment of the non-agricultural LGN2 classes that are less important and/or show little variability or are (partly) copied out of other high quality databases, is only performed qualitatively and/or in a non-site-specific way.

The gathering of a statistical valid sample of pixels for the agricultural crops in the entire Netherlands is estimated to take 150 à 200 working days. So due to time and money constraints a random or systematic sampling and even a random cluster sampling per stratum is not feasible. Therefore, in the framework of the LGN database a 'controlled' cluster sampling is proposed. The 'controlled' cluster sampling is a sampling method in which all (agricultural) plots bordering on a number of selected sections of roads are sampled. For all agricultural strata a number of outwardly representative sections of roads have to be selected on the basis of topographic maps and satellite images. For the main agricultural crops in a stratum at least 10 plots have to be sampled. In order to maximize the information derived from cluster sampling, the sampled pixels used for the accuracy assessment consist of clusters of 3 x 3 pixels in the centre of the plots. For small plots only the centre pixel is sampled. Comparison of the accuracy assessments of the classification result of a test site, performed with reference data from both a controlled cluster sampling and a systematic sampling, shows that individual classes may show some deviations. However, the overall classification accuracies of both approaches are comparable. To better found the reliability of the controlled cluster sampling, it is advised to compare the results of the controlled cluster sampling with the results of the random or systematic sampling in some other test sites.

The CBS Agricultural Statistics are suitable for a non-site-specific accuracy assessment of the LGN2 database. One should take into consideration that the CBS Agricultural Statistics contain net cultivated areas, while the agricultural stratum in the LGN2 database contains the total agricultural area inclusive of ditches, (minor) roads, hedges, farm yards, farms and other buildings.

Application, cost and benefit analysis and implementation of the LGN database

The main users of the LGN(2) database are national and regional governmental agencies. Because of its digital format the LGN database can be easily combined with other digital information. It has frequently been used for different purposes in the fields of environmental protection, water management, nature conservation and physical planning on regional and national scales. Mostly, the LGN data are combined with other geographical information, such as soil type, water-table, the occurrence of seepage, meteorological data, and the application of animal manure, fertilizers and pesticides.

The cost of production of the LGN2 database amounted to Dfl 1 120 000 (i.e. circa Dfl. 0.35 per hectare). The selling price of the LGN2 database has been determined on the basis of the cost, the expected number of users and negotiations with (potential) users. The selling price of the LGN2 database depends on the area required, the number of classes, the spatial resolution and the number of applications. The earnings from sale of the LGN2 database amount to circa Dfl 1 170 000, inclusive of warrants at Dfl 350 000 (situation June 1996).

The LGN database is used for many applications. According to some users, the benefits of using the LGN database exceed the cost. In practice, the LGN database, once being available, appears often to be used for all kinds of unintended applications. An accurate estimation of the benefits, however, is troublesome. Information from the LGN database is mainly used by governmental agencies which are responsible for policies in the fields of environmental protection, water management, nature conservation and physical planning on regional and national scales. The effects of using the LGN database on the quality of the pursued policy are difficult to assess.

Operational implementation of the LGN database has been achieved when the database or parts of the database are up-dated at regular intervals and the up-dating is largely paid by the users of the database. Commercial implementation implies that the up-dating is completely paid by the users. At this moment operational implementation of the LGN database has been achieved in contrast with commercial implementation. The chance of commercial implementation or continuity of operational implementation of the LGN database in near future is determined by the need of land cover data and the cost of gathering these data. In this framework the LGN database has to compete with other available nation-wide digital land cover/use databases, especially the topographic databases and the CBS Land Use Database. Evaluation of the different databases shows that (a part of) the LGN database may distinghuish itself favourably from the CBS Land Use database and/or the topographic databases, especially concerning cost, thematic classes, timeliness and dataprocessing. That means that, altough there may be some overlap, the databases largely supplement each other. In practice, combined use of different databases will give a surplusvalue to the separate databases. In order to achieve commercial implementation or to continue operational implementation of the LGN database, the advantages of the LGN database with respect to other digital geographical databases should be exploited as much as possible. Further, in order to meet customers' needs more satisfactorily, the LGN database should become a more flexible product and the classification accuracy and reliabilitiy of some classes should be improved. Finally, more users should be found. In order to reach these objectives, the following activities have to be performed: setting up a subscription system, variation of up-date intervals for different classes and/or areas, no inclusion of mixed classes by applying only optimal images, intensifying the marketing and reducing of the production cost. At last, the possibility of using the 1 : 10 000 topographic database for performing a field-based classification should be investigated.

1 Introduction

The need of land cover data
Because of the increasing concern about the impacts of man's intervention on the environment, timely and accurate information on land cover at regional and national scales is required by national and regional governmental agencies to support environmental policy and for physical planning purposes. For environmental purposes especially agricultural land cover data is required. Pollution from diffuse sources (e.g. application of pesticides and manure surplusses) threatens the soil and groundwater quality in The Netherlands. These pollution loads vary with the various agricultural crops. Information on agricultural land cover can be obtained from land use statistics and topographical maps. However, land use statistics are only available for restricted areas (e.g. municipalities or provinces) and cannot be derived for areas with deviating boundaries (e.g. river basins and groundwater protection areas). Topographical maps do often not contain all required land cover classes, are often outdated and were till recently not available in digital form.

Historical background
In 1987 it was decided to produce a land cover database of the Netherlands (further to be mentioned 'LGN database'), using satellite images (Thunnissen et al., 1992a and 1992b). The spatial resolution of images obtained with both the LANDSAT Thematic Mapper (TM) and the French satellite SPOT is in general sufficient for the recognition of the individual agricultural fields in The Netherlands. The objectives of the land cover classification project were:
- Realization of a national land cover data base, containing information on the spatial distribution of main agricultural crops, deciduous and coniferous forest, water, natural area and built-up area.
- Evaluation of the possibilities to get accurate land cover information for the different physiographic units in The Netherlands by interpretation of satellite images.
- Definition of additional activities and research required for operational implementation of the LGN data base.

The first version of the LGN database (further to be mentioned 'LGN1 database') was produced by automatic classification of manually stratified single-date satellite images from 1986. The classification accuracy showed a large variation over the country due to spectral confusion between different land cover classes. The overall classification accuracy for 18 reference areas varied between 50 and 84%, while the average overall classification accuracy amounted to 67%. The extent of spectral confusion in a stratum was dependent on the occurring land cover, size and shape of the land cover units, spectral resolution and acquisition dates of the satellite images and crop development. Regionally, crop development was strongly influenced by stress due to drougth or water logging. Further, the classification result was influenced by the applied stratification and the limited availability of field reference data. Another disadvantage of the classification, applied mainly on the basis of spectral signatures, was that various land use classes could not be differentiated because they possessed similar spectral properties. For example, a short herbaceous cover may represent agricultural use, or recreational use, or residential use.

Altough the classification results of the database were disappointing in some areas, evaluation of the results of the LGN project showed that users of the LGN1 database (i.e. national and regional governmental agencies) were interested in an up-date of the database (Thunnissen et al., 1992a). However, the classification results should be improved considerably.

Objectives of the current study

In 1992 the 'LGN Research Project' was started, aiming to develop an improved land cover/use classification method, using satellite images and other digital geographical databases, and to investigate the possibilities to up-date the LGN1 database and to implement the database operationally, c.q. commercially.

Approach

Significant improvements of the classification result could be expected by a reduction of the spectral confusion between the different land cover types. Similarity in spectral reflectances at the image acquisition date impedes consistent identification and mapping of a large number of important land cover types when using single-date satellite images. However, spectral signatures of a wide range of cover classes, such as agricultural crops or natural vegetation, vary throughout the year. By that, classes which appear very similar in spring, may become separable at other stages of the phenological cycle. It is therefore expected that multi-temporal approaches provide important means to improve classification accuracy. Further classification improvement of spectral overlapping land cover classes may be expected by the use of other digital geographical data. Discrimination between different land use classes which possess *similar* spectral properties can only be achieved by using other digital geographical data or by visual image interpretation techniques. During the last decade several nation-wide digital geographical databases have become available. The (integrated) use of digital geographical data and (multi-temporal) satellite data was an important research item in the LGN Research Project. In this framework also attention was paid to optimal acquisition periods and the availability of satellite images which are both of outmost importance for the classification results, using multi-temporal satellite images.

For an effective use of land cover data derived from remote sensing images, it is of importance to have knowledge of the accuracy of these data. Therefore, a study was made of available classification accuracy assessment procedures, taking into account time, cost and practical restrictions associated with the LGN project.

In contrast to 10 years ago (LGN1 database) in 1996 several other digital land cover/use databases are available. The chance of operational, c.q. commercial implementation of the LGN database in near future is determined by the need of land cover data and the cost of gathering these data. In this framework the LGN database has to compete with other available nation-wide digital land cover/use databases, especially the topographic databases and the CBS Land Use Database. The LGN database can be competitive when the database distinguishes itself favourably from the other databases. The different databases were mutually compared and a cost and benefit analysis of the LGN database was performed.

In 1993 the up-dating of the LGN1 database was started (the up-dated version of the LGN1 database will further be described as 'LGN2 database'). The up-dating was partially performed simultaneously with the LGN Research Project and was finished at the end of 1995 (Noordman et al., 1996). Dependent on the progress of the up-dating, results of the LGN Research Project which could applied operationally were incorporated in the current classification method or will be applied in future up-dates of the LGN database. Conversely, some problems met during the up-dating of the LGN1-database were included in the LGN Research Project for additional research. The project was carried out in the framework of the National Remote Sensing Programme (NRSP-2), under responsability of the Netherlands Remote Sensing Board.

Contents of the report

In the framework of the LGN Research Project and the up-dating of the LGN1 database a large number of satellite images were obtained and reference data were gathered. Chapter 2 contains a review of the satellite images and reference data used in the LGN Research Project. Moreover, the legend of the LGN2 database is presented. Chapter 3 draws up an inventory of available nation-wide digital geographical databases. The possibilities to use these databases in combination with satellite images in order to improve classification accuracy are assessed. Chapter 4 discusses the improved classification methodology. Attention is paid to stratification, automatic and visual classification, postprocessing techniques and optimal acquisition dates and the availability of satellite images. Chapter 5 discusses different accuracy assessment procedures, considering time, cost and practical restrictions associated with the LGN project. An adapted sampling scheme for the accuracy assessment of the LGN database has been proposed. Finally, in Chapter 6 applications, cost-benefit analysis and operational and commercial implementation of the LGN database are discussed.

2 Data sources and nomenclature

2.1 Data acquisition and test sites

For the execution of the LGN Research Project satellite images and reference data gathered for the the production of the LGN2 database were used. Moreover, additional satellite images and reference data were gathered for test sites in the provinces of Drenthe and Flevoland (Fig. 1). The classifications of the test sites in the provinces of Drenthe and Flevoland were performed independent of the LGN2 database. The LGN2 database was produced by classification of images obtained with both the Landsat-5 (Thematic Mapper) and the SPOT (multispectral mode) satellite (Tables 1 and 2). For the production of the LGN2 database mainly satellite images obtained in 1992 and 1994 were used (Noordman et al., 1996). Reference data, required for training the classifier and assessing the accuracy of the classification, were obtained by interpretation of topographic maps and aerial photographs and by field surveys. The sampling schemes applied in the LGN2 project are discussed in 5.4.

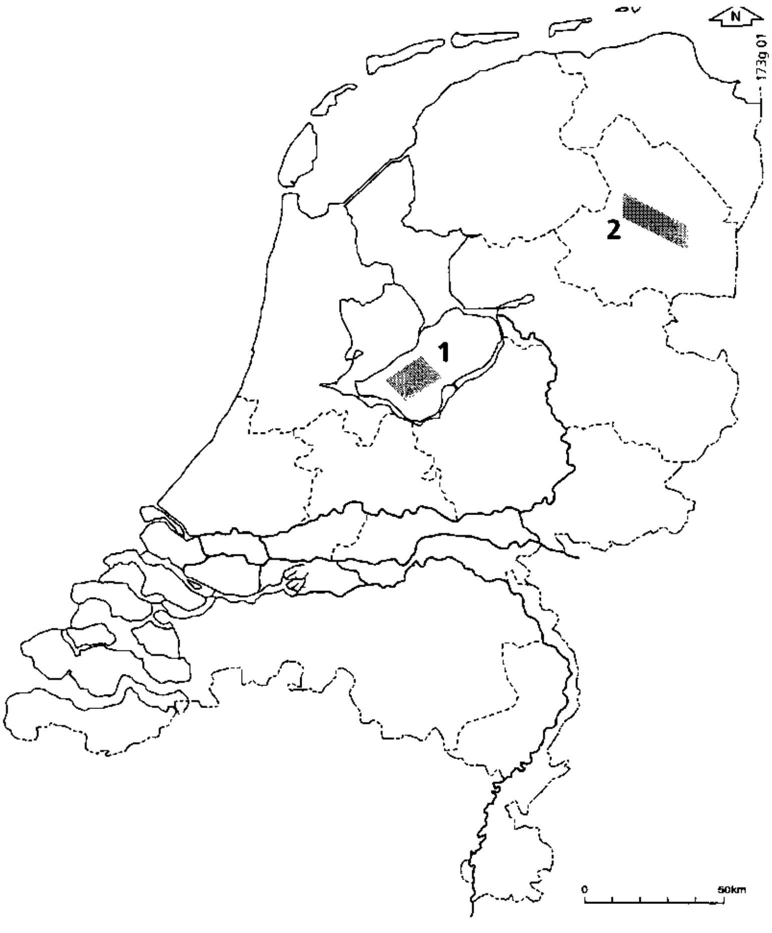

Fig. 1 Situation of the Zuid Flevoland (1) and Drenthe (2) test sites

Table 1 Wave length bands and spatial resolution of the Thematic Mapper sensor aboard the Landsat satellite

Band	Wave length (µm)	Resolution (m)	Description
1	0.45- 0.52	30	visible blue
2	0.52- 0.60	30	visible green
3	0.63- 0.69	30	visible red
4	0.75- 0.90	30	near infra red
5	1.55- 1.75	30	middle infra red
6	10.40-12.50	120	thermal infra red
7	2.08- 2.35	30	middle infra red

Table 2 Wave length bands and the spatial resolution of the SPOT satellite

Band	Wave length (µm)	Resolution (m)	Description
1	0.50- 0.59	20	visible green
2	0.61- 0.68	20	visible red
3	0.79- 0.89	20	near-infrared
panchromatic	0.51- 0.73	10	visible

The test site in the province of Flevoland, further to be mentioned 'Zuid Flevoland test site', was only used to investigate the possibilities of the first European Remote sensing Satellite (ERS-1) for crop classification. The used ERS-1 images and the reference data of the Zuid Flevoland test site are described in 4.10.

Table 3 Satellite images covering the Drenthe test site in 1991

Satellite	Scene	Acquisition date
SPOT	45 -243	3 February
SPOT	45 -242	29 July
Lansat-5 (TM)	198 /23	2 September

The test site in the province of Drenthe, further to be mentioned 'Drenthe test site', is covered by three satellite images from 1991 (Table 3). The SPOT image from 29 July 1991 contains some scattered clouds. The Drenthe test site is largely situated on the Drenths plateau and the outmost western and eastern parts are partly situated in the 'Veenkoloniën'. The Drenths plateau consists of sligthly undulating sandy soils and slopes gradually from east to west. The sandy plateau is intersected by brook valleys with peat and hystic soils. The Veenkoloniën largely consists of peat and hystic soils. The Drenthe test site is characterized by a mixed agricultural land use. Further some large forests and natural areas occur in the test site. Reference data of the Drenthe test site were collected in the field in september 1991 based on a systematic sampling scheme (Table 4). The sampling points coincide with the points of intersection of the 1 km grid lines on the Dutch topographic maps. The use of these points facilitates the location of the sampling points. To avoid problems due to positional inaccuracies of the satellite images, sampling points near agricultural field boundaries were shifted to a point at least several pixels away from the field boundary. In the reference class 'mixed forest' both deciduous and coniferous forest occupy more than 25% of the forested area. No samples were taken within built-up area. In addition to the systematic sampling, agricultural fields situated in some outwardly representative reference areas were sampled. The reference areas were visually selected on the basis of topographic maps.

Table 4 Result of the systematic sampling in the Drenthe test site

Land cover	Number of samples
Grassland	126
Maize	33
Potatoes	133
Sugar beets	64
Cereal	50
Other agricultural crops	11
Decidious forest	39
Coniferous forest	101
Mixed forest	23
Nature area with low vegetation	32

2.2 Nomenclature

It is important to distinguish between 'land cover' and 'land use'. Whereas *land use* refers to human activity of a certain kind for a given land surface, *land cover* refers to the vegetational and artificial construction occupying the land surface. To enable discrimination between the main land use types (i.e. agriculture, built-up area and natural area) the legend of the LGN2 database (Table 5) was adapted with regard to the LGN1 database. The LGN2 legend distinguishes 25 classes, grouped in a two-level hierarchy Discrimination between the main land use types enabled the definition of new classes (e.g. 'grass in urban area'). Moreover, the class 'greenhouses' was added to the LGN2 database. In this report the LGN2 classes will be mentioned 'land cover' classes. However, when the functional land use of a LGN2 class is explicitely meant, the LGN2 class will be mentioned 'land use' class.

To enable applications of the LGN2 database on a regional scale mininimum classification accuracies and reliabilities of 70% are required. However, poor spectral separability and cultivation practices such as harvesting may result in classification accuracies and reliabilities below 70%. In these cases mixed agricultural classes, containing more than one agricultural class, are defined (4.5).

Table 5 Legend to the LGN2 data base of the Netherlands

Level 1	Level 2	
1 Agricultural area	1.1	grass
	1.2	maize
	1.3	potatoes
	1.4	beets
	1.5	cereals
	1.6	other agicultural crops
	1.7	bare soil in agricultural area
	1.8	greenhouses
	1.9	orchard
	1.10	bulbs
2 Forest	2.1	deciduous forest
	2.2	coniferous forest
3 Nature area	3.1	heath land
	3.2	other nature area with low vegetation
	3.3	bare soil in nature area
4 Water	4.1	inland waters
	4.2	marine waters
5 Built-up area	5.1	continuous urban area
	5.2	built-up in rural areas (exclusive of farms)
	5.3	deciduous forest in urban area
	5.4	coniferous forest in urban area
	5.5	densely forested residential area
	5.6	grass in urban area
	5.7	bare soil in rural built-up areas
	5.8	main roads and railways

3 Satellite images and other digital geographical data

3.1 Introduction

An automatic pixel-based classification of (multi-temporal) satellite images is only based on differences in spectral reflectances. However, in practice the spectral reflectances of different land cover classes are often not completely unique. Moreover, different land *use* classes cannot be differentiated because they posses similar spectral reflectance (Thunnissen et al., 1992b). For example, a short herbaceous cover may represent agricultural use, or recreational use, or residential use. In these cases (spectral confusion) land cover/use classes cannot be discriminated from satellite images alone and ancillary data are required to make an accurate distinction. In this framework especially the use of digital geographical databases is promising. In the Netherlands several nation-wide digital geographical databases are available or will be available in near future. The most relevant databases are the 'CORINE Land Cover database', the soil map of the Netherlands (scale 1 : 50 000), the 'Land Use Database' of the State Department for Physical Planning (the so-called BARS database), the 'Land Use Database' of the Central Bureau of Statistics (CBS), topographic databases at scales 1 : 50 000, 1 : 25 000 and 1 : 10 000, and the 'Agricultural Statistics' of the CBS. In this chapter the possibilities to use these databases in combination with satellite images in order to improve classification accuracy are assessed. At first different approaches to use ancillary data will be discussed.

3.2 Ancillary data

3.2.1 Stratification

Different approaches can be applied for combination of ancillary data with satellite images. Hutchinson (1982) distinguishes between incorporation of these data either before, during or after classification, through stratification, classifier operation, or postclassification sorting. Use of ancillary data prior to classification involves a division of the area to be classified into smaller, more hogeneous areas or strata based on some criterion or rule, so that each stratum may be processed independently (Hutchinson, 1982). Stratification has different advantages. During the classification it is easier to deal with smaller areas. Moreover, seperate land cover classes show probably less spectral variation within strata. The spectral characteristics of any set of objects, such as soil or vegetation types, are likely to vary over distance. As variance increases, the likelihood of confusion between spectrally similar objects also increases (Hutchinson, 1982). At last, stratification may seperate different land use classes which are spectrally similar. So, stratification increases the quality of the classification process when compared to the uniform treatment of a whole area or image. Moreover, separate classification of the different strata with a limited number of land cover classes enables to better focus the discrimination process on problem classes and to reduce misclassifications due to spectral confusion.

Stratification can be performed by interactive interpretion of satellite images, supported by ancillary data, or by using digital land use databases. The former method was applied during the production of the LGN1 database (Thunnissen et al., 1992b). The interpretation of the data and the digitization of the strata required a considerable effort. It is obvious that the latter method is preferred, especially when large areas have to be classified. When relevant geographical databases are available, stratification can simply be applied. Moreover, generally, the accuracy of the stratification is higher in comparison with manual stratification. However, at the time of the production of the LGN1 database no nation-wide digital land use database was available.

3.2.2 A-priori probabilities

An approach of using ancillary data during classification involves the specification of the a-priori probabilities in the maximum likelihood decision rule. In most classifications a-priori probabilities are assumed to be equal for all classes. However, a-priori probabilities can be specified based on knowledge either on the areas occupied by the land cover classes in a stratum or on known relations between land cover classes and ancillary data. The determination of a-priori probabilities based on the relative areas of the land cover classes in a stratum leads to equal probabilities for all pixels in a certain class. Ancillary data (e.g. a soil map) can provide spatial and/or temporal information on the occurrence of certain land cover classes. This information can be included in the maximum likelihood decision rule by means of *conditional* a-priori probabilities (Strahler, 1980 and Janssen, 1993). This means that the a-priori probability depends on the value of a conditioning variable (e.g. the crop in the preceding growing season or soil type). Application of conditional a-priori probabilities is expected to result in a higher classification accuracy than application of class-based a-priori probabilities, estimated by the relative areas of these classes (Janssen, 1993 and Van der Wel, 1993).

Use of a-priori probabilities prove to be most effective when classes have spectral overlap (Janssen, 1993). If the considered classes have a high spectral separability, the application of a-priori probabilities at its best results in a marginal increase of classification accuracy. The more different are the values of the a-priori probabilities of spectral overlapping classes, the higher is the information content and the more effect has the application of a-priori probabilities. However, when the a-priori probability becomes large and approaches 1 (since the sum of the probabilities of all classes amounts to 1 the probabilities of the remaining classes will be low) the classification will be forced into the class with high probability. For simililar reasons the a-priori probability approaching 0 may remove a class from the classification result. Therefore, extreme values of a-priori probabilities may overestimate or underestimate the occurrence of the concerning classes (Strahler, 1980 and Van der Wel, 1993).

Class-based a-priori probabilities, estimated by the relative areas of these classes, can be derived either from field sampling or from available statistics. Estimation of conditional a-priori probabilities requires often additional sampling that is considerable beyond conventional spectral classification (Strahler, 1980 and Janssen, 1993). It is up to the user to balance the cost of acquiring these data with the expected classification improvement.

3.2.3 Postclassification sorting

The use of ancillary data after classification is based on the observation that a single spectral class may often represent subsets of more than one land cover class. In postclassification sorting, individual pixels of these problem spectral classes are assigned to the appropriate land cover class using ancillary data. Hutchinson (1982) mentions several advantages of postclassification sorting. It is efficient because it deals only with problem classes. Furthermore, it is relatively simple to include several types of ancillary data in developping decision rules. Finally, because it is performed after classification, errors made in rule selection can be corrected easily as opposed to those made prior to classification using stratification.

However, in practice the recoding of spectral classes, representing more land cover classes, by postclassification sorting can be rather troublesome. Van der Laan (1988) combined a classified Landsat-TM image (10 land cover classes were distinguished) of a test site situated in the southern part of the Netherlands with a digitized topographic map (9 relevant classes were selected) in order to improve classification accuracy. The pixel by pixel comparison may result in 90 (potential) combined classes. The combined classes were recoded to the correct classes or to newly, defined classes. The recoding of the 'misclassified' pixels was based on expert judgement and additional visual interpretation. In order to improve the classification of the main roads in the LGN2 database, postclassification sorting was applied (4.11 and Noordman et al., 1996). The polygons of the main roads, including verges and vegetated areas within roundabouts, were derived from the CBS Land Use Database (3.4). The recoding of the classified pixels was not unambiguous and resulted, locally, in a poor discrimination between the road surfaces and the contiguous vegetated areas.

3.3 CORINE Land Cover database

In order to determine the European Community's environment policy, evaluate the effects of this policy correctly and incorporate the environmental dimension into other Community policies, we must have a proper understanding of the different features of the environment. It was against this background that the CORINE (CoORdination of INformation on the Environment) Programme was started to gather, coordinate and ensure the consistency of information on the state of the environment and natural resources. One of the primary thematic items of this programme is land cover.

The main objective of the CORINE Land Cover project is the gathering of coherent information on land cover for the European Community and the integration of this information into a Geographical Information System (GIS). The CORINE Land cover database of the Netherlands was produced by the DLO Winand Staring Centre (Thunnissen and Van Middelaar, 1995). The methodology consists of computer-assisted visual interpretation of earth observation satellite images, with the simultaneous consultation of additional data, into the categories of the CORINE Land Cover Nomenclature. In order to consider the complete spectrum of land cover an European nomenclature has been developed, the legend of which distinguishes 44 classes, grouped in an open three level nomenclature system. The CORINE nomenclature is primary directed on characterization of the landscape and many classes are heterogeneous and consist of different land cover types. Therefore, the nomenclature is not consistent with the LGN nomenclature. The scale of the land cover database is 1 : 100 000. The surface area of the smallest unit mapped is 25 ha. For line elements the minimum width is 100 m.

Because of the deviating nomenclature and scale of the CORINE Land Cover database this database is not suitable to be used for stratification or postclassification sorting in the framework of the LGN project.

3.4 BARS and CBS Land Use Databases

The BARS and the CBS Land Use Databases have much in common and are considered in the same section. The CBS Land Use Database provides information on land use of the total area of the Netherlands. The nomenclature consists of 33 land use classes and, especially for non-agricultural and non-natural areas, land use is decribed in detail (Table 6). The classification is largely based on functional land use. For instance 'Parks and public gardens' consist of grass, forest and public gardens. Gardens and public greens, situated in residential areas, are assigned to 'Residential areas'. The CBS Land Use database discriminates between dry and wet natural areas. Scattered houses and farms in rural areas are assigned to the surrounding area (mostly agricultural area or forest). In general, areas smaller than 1 ha are not included in the CBS Land Use Database.

Up to and including 1985 the main sources of information were municipal administrations, and the data were stored on analogue maps and published as land use statistics for grid cells of 500 m x 500 m. From 1989 onwards information on land use change is obtained by interpretation of aerial photographs (scale 1 : 10 000). In urban areas city plans are used as ancillary data. The actual land use data are stored as digitized maps in a GIS. The digital land use database based on aerial photographs from 1989 is available for the entire country. The production of the digital land use database based on aerial photographs from 1993 will be concluded in 1996. In the digital CBS Land Use Databases linear elements, like railways and roads, are included with their real width.

The BARS Land Use Database consists of a digitized land use map and was produced by the State Department for Physical Planning. The nomenclature consists of 33 land use classes. For the description of the nomenclature one is referred to Noordman et al. (1996). The main sources of information were the analogue CBS land use maps. Additional land use data were obtained from different governmental agencies and public utilities and by interpretation of topographic maps. The CBS land use classes 'Holiday recreation', Social-cultural facilities' and 'Other public facilities' were further subdivided in the BARS database. In addition to greenhouses, orchards are also included in the BARS database. Other agricultural area, including farms, is included in the BARS class 'Other area'. The classification accuracy of glasshouses varies strongly and in some areas large contiguous (agricultural) areas are included in the BARS class 'Greenhouses'. In general, classes in urban and rural areas smaller than 1 and 5 ha, respectively, are not included in the BARS Land Use Database. Linear elements, like railways and main roads, are included as line elements, irrespective of their actual width. The BARS database was up-dated in 1993 and 1994, mainly on the basis of topographic maps and city plans. Dependent on the revision dates of the these maps the actuality of different map sheets of the BARS database varies strongly. In future, the BARS database will not be up-dated anymore.

Both the CBS and the BARS Land Use Databases enable discriminination of the main land use types in the LGN2 database (i.e. agricultural area, built-up area, and natural area and forest) and, consequently, could be used for stratification or postclassification sorting. For this purpose the relevant classes should be aggregated. However, the CBS database is more accurate and contains smaller land use units in rural areas. Moreover, the CBS Land Use Database contains some classes, especially greenhouses and main roads, that are also included in the LGN2 database. Further, the discrimimination between wet and dry natural areas in the CBS Land Use Database provides useful additional information (4.2). The BARS database will not be up-dated anymore. The cost of the nation-wide CBS and BARS Land Use Databases amount to Dfl 17 000 and 150 000, respectively.

Table 6 CBS Land Use Database nomenclature

Railways, tramways and metros
Metalled roads (incl. verges)
Unmetalled and half-metalled roads
Water reservoirs
Other water wider than 6 m
Cemeteries
Sports grounds
Airfields and airports
Allotments
Dumping sites
Car wreck sites
Mining areas
Parks and public gardens
Holiday recreation
Recreational objects and areas
Social-cultural facilities
Other public facilities
Industrial areas
Water wih a primarily recreational function
Commercial and trade areas
Residential areas
Building sites for industrial areas
Building sites for other purposes
Woodland
Glasshouses
Other agricultural use
Dry natural areas
Wet natural areas
Other areas
Waddenzee, Eems, Dollard
North sea
IJssel lake
Oosterschelde en Westerschelde

3.5 Topographic maps

The topographic maps of the Netherlands on the scales 1 : 50 000 and 1 : 25 000/10 000 (the topographic maps on the scales 1 : 25 000 and 1 : 10 000 contain the same information) will be available in digital form within a few years. The up-date frequency of the topographic data depends on the population density and varies between 4 and 8 years. The *annual* price (inclusive of up-dates) depends on the up-date frequency and varies between Dfl 840 and 1 680 for one 1 : 50 000 map sheet (i.e. 500 km^2) and between Dfl 550 and 1 110 for one 1 : 10 000 map sheet (i.e. 62.5 km^2). The entire country is covered by 101 1 : 50 000 map sheets and circa 600 1 : 10 000 map sheets.

The topographic databases could be used for discrimination of some main land use types, e.g. urban area, natural area and forest. The topographic databases contain, however, some land *cover* classes, of which the functional *use* can only be determined by using contextual information, e.g. grassland and forest situated in urban area (parks or sport grounds), grassland used for recreational purposes and such-like. These classes cannot be automatically assigned to the correct strata, making the topographic data less suitable for stratification of satellite images or postclassification sorting of classified images. The topographic databases contain some important classes, especially greenhouses, orchards, railways, main roads and buildings in agricultural areas, which are also included in the LGN2 database. Because of the readability of the topographic maps the width of roads and railways on the map do not correspond with their actual width.

The 1 : 10 000 topographic map contains detailled information on linear elements in agricultural area, such as ditches, hedges and (unmetalled) roads, which often coincide with boundaries between agricultural lots. Janssen (1993) showed that the classification result for agricultural crops could be improved considerably by performing a field-based classification. A field contains only one crop type. Lots indicated on the 1 : 10 000 topographic map may comprise several fields. The 1 : 10 000 topographic map may be used as a base map to obtain actual field boundaries by automatic and/or visual interpretation of satellite images. Automatic interpretation techniques include segmentation techniques, such as edge detection, region growing and clustering. Visual interpretation means on screen digitizing of the missing field boundaries. For this purpose the lot boundaries in the topographic database should be projected on the satellite images. The number of field boundaries indicated on the 1 : 10 000 topographic map shows large variation over the country dependent on the occurring landscape.

Some of the lot boundaries in the digital topographic database are not connected with other lot boundaries and form dangling arcs, resulting in polygons which may contain several lots. Algorithmes have to be developed to automatically connect these dangling lines with other lot boundaries.

It can be concluded that the combined use of the 1 : 10 000 topographic database and satellite images may provide field boundaries to be applied in a field-based classification. However, it has to be investigated if the required field boundaries can be obtained in an operational and cost effective way.

3.6 Agricultural Statistics

The General Census of Agriculture, organized by the CBS, is taken annually in May, covering all agricultural holdings with a minimum size of three Netherlands Size Units (NSU). The NSU is a standard expressing the economical size of an agricultural holding. Among other things, the General Census of Agriculture provides information on the areas of the agricultural crops grown (further to be mentioned 'CBS Agricultural Statistics'). The CBS Agricultural Statistics contain cultivated areas, not including roads, ditches and hedges less than 4 m wide that intersect or bound the cultivated area. Fields are assigned to the municipality where the main buildings of the holding are situated, irrespective of real location of the fields. Incidentally, this can lead to large deviations with respect to the real (net) cultivated area in a mumicipality (Central Bureau of Statistics, 1983). The Agricultural Statistics are published per municipality, per province, and per 'agricultural region'. Agricultural regions are more or less homogeneous areas as far as soil type and agricultural land use are concerned. The Netherlands are subdivided in 66 agricultural regions.

In principle, the CBS Agricultural Statistics could be used for the determination of class based a-prori probabilities based on the relative areas of the land cover classes in a CBS agricultural region (3.2.2). Different authors applied such class based a-prori probabilities. Moreover, the effects of applying these class-based a-prori probabilities were assessed in the Drenthe test site. Janssen (1993) applied a-priori probabilities based on the relative areas of the agricultural crops in a test site, situated around the village of Biddinghuizen in the province of Flevoland. The test site consists of large agricultural fields. The maximum likelihood classification was performed, using bands 3, 4 and 5 of a Landsat-TM image that was acquired on 7 July 1989. The crops had a homogeneous spectral appearance on the satellite images and showed a very high spectral seperability. Each crop was represented by only one spectral class. To assess the effect of a-priori probabilities in a situation with poor spectral discrimination, the maximum likelihood classification was also performed for TM band 4 alone. Application of the class-based a-priori probabilities resulted in improvements of the overall classification accuracy of 0.7 and 4.4% for the situations with high and poor spectral seperabilities, respectively, with regard to the clasification accuracies, using equal a-priori probabilities. The 'overall classification accuracy' is calculated by dividing the number of correct classified pixels by the total number of pixels sampled (4.3).

Keeman (1991) applied a-priori probabilities based on the relative areas of the agricultural crops in a test site situated near Nieuw Buinen in the eastern part of the province of Drenthe. The test site is characterized by narrow, elongated fields. The classification was performed, using bands 3, 4 and 5 of a Landsat-TM image that was acquired on 14 July 1987. Classification performance was poor. Because of the narrow fields circa 40% of the pixels consisted of mixed pixels. The poor classification result was largely caused by the misclassification of mixed pixels. Application of the class-based a-priori probabilities resulted in improvement of the overall classification accuracy of only 0.5% with regard to the clasification accuracy, using equal a-priori probabilities.

The Drenthe test site was also classified, using a-priori probabilities based on the relative areas of the agricultural crops. The relative areas were estimated on the basis of the systematic sampling (2.1). The maximum likelihood classification was performed, using bands 3, 4 and 5 of a Landsat-TM image that was acquired on 2 September 1992. As far as not yet harvested, some crops had a heterogeneous spectral appearance on the satellite images caused by drougth damage and withering. Therefore, crops were represented by different spectral classes and showed a moderate spectral seperability. Application of the class-based a-priori probabilities resulted in improvement of the overall classification accuracy of 1.3% with regard to the classification accuracy, using equal a-priori probabilities.

It can be concluded that even for spectral overlapping classes the improvement of classification accuracy by using class-based a-priori probabilities seems small. The effectivity of a-priori probabilities seems not only to be influenced by spectral overlap (3.2.2) but also by spectral variability of the seperate land cover classes. Probably, (large) spectral variability hampers the assignment of representative training areas.

Instead of providing a-priori probabilities, the CBS Agricultural Statistics seem more suitable to be used for stratification and validation purposes. Especially the CBS agricultural regions could be used to divide the area to be classified into smaller, more hogeneous areas. For each agricultural region the areas of the occurring agricultural crops are known. So, the classified areas of agricultural crops can be compared with the areas provided by the CBS Agricultural Statistics.

3.7 Soil maps

A digitized soil map on a scale of 1 : 50 000 is available for the entire country. Drainage classes, which are indicated on the soil map as so-called water table classes, provide information on the depth and seasonal fluctuation of the groundwater table (Van der Sluijs and De Gruijter, 1985). The water table classes are based on the mean highest (MHW) and the mean lowest (MLW) groundwater tables (Table 7), representing the average winter and summer water tables, respectively, in a year with an average precipitation and evaporation. Soil type and water table class may be used as conditioning a-priori probabilities in order to improve land cover classification accuracy. Keeman (1991) studied the relationship between water table classes and land cover in 3 test sites in the eastern part of the Netherlands. The test sites showed only vague relationships between water table classes and land cover. For one test site, situated near Nieuw Buinen (3.6), the data were used to determine conditional a-priori probabilities. Application of these conditional a-priori probabilities resulted in improvement of the overall classification accuracy of 0.9% with regard to the classification accuracy, using equal a-priori probabilities. The field surveys for the soil maps used in the study of Keeman (1991) were performed during the period 1972-1980. From that time changes in water management have effected changes in groundwater level in many areas. By that, information on water table classes on the soil maps of the test sites will probably be outdated. Moreover, the areas studied by Keeman (1991) contained only small areas with water table classes, characterized by high groundwater levels. Especially in areas with high groundwater levels significant relationships between water table class and land cover are to be expected. Eventually, relationships between soil type and land cover were not analysed by Keeman.

Table 7 Water table classes (MHW=mean highest water table, MLW=mean lowest water table; both in cm below surface)

	I	II[1]	III[1]	IV	V[1]	VI	VII[2]
MHW	-	-	<40	>40	<40	40-80	80-140
MLW	<50	50-80	80-120	80-120	>120	>120	>120

[1] A code with * means 'drier part' (MHW deeper than 25 cm).
[2] A code with * means 'very dry part' (MHW deeper than 140 cm).

In order to assess if significant relationships occur between the conditioning variables soil type and water table class and land cover, some recent or recently up-dated soil maps in the provinces of Drenthe and Overijssel were visually compared with the topographic maps on a scale of 1 : 50 000. So far as agricultural crops are concerned only grassland and arable crops were considered. On 'Duin' and 'Vlak' vague soils only forest and dry open natural area occur. In general, arable crops dominate on peat and hystic soils with water table classes above III* (Table 7), while on peat and hystic soils with water table classes III and III* grassland dominates. On very wet soils (water table classes I, II and II*) grassland is mostly the only crop grown. For the remaining soil types and water table classes no distinct relationships with land cover were found. For the above mentioned groups of water table classes situated within peat and hystic soils in the Drenthe test site pixel based priors were derived (Table 8). Only two land cover classes were considered: grassland and arable land. Very wet soil (water table class below III) did hardly occur in the test site. The Drenthe test site was classified using, using bands 3, 4 and 5 of a Landsat-TM image acquired on 2 September 1992. The agricultural crops showed a moderate spectral separability (3.6). Altough application of the conditional a-priori probabilities resulted in minor changes in areas of grassland and arable land, the overall classification accuracy did not change with regard to the classification accuracy, using equal a-priori probabilities. So far as the relationhip between 'Duin' and 'Vlak' vague soils and forest and dry open natural area is concerned it should be remarked that both land cover classes occur in the same stratum (4.2) and show, generally, a high spectral seperability. Therefore, no significant classification improvement is expected by appying a-priori probabilities.

Table 8 A-priori probabilities of grassland and arable land occurring on peat and hystic soils with two groups of water table classes in the Drenthe test site

Water table clases	A-priori probabilities	
	grassland	arable land
III and III*	0.71	0.29
IV and above	0.23	0.77

It seems that, in general, the soil map does not provide a-priori information that can be used for improvement of the classification accuracy. Only, in the case that particular crops require specific physical conditions (e.g. bulbs) or extreme physical conditions allow only specific crops to be grown (e.g. grassland on very wet soils), the soil map could provide useful a-priori information. Especially the relation between grassland and soils with very high groundwater levels may be of importance because of the occurring spectral confusion between (very) wet grassland and winter cereals on TM images obtained in spring. Moreover, the NDVI values of wet grassland may be low which may cause wet grassland to be included in the low NDVI range, representing bare soil. Finally, information on water table class can be used to discriminate between wet and dry natural areas. Both natural areas show different spectral signatures for forest and grassland and, consequently, have to be classified separately (Noordman et al., 1996). In these cases the water table classes should be used for additional stratification, rather than for providing a-priori probabilities. The possibility of using a-priori information from soil maps in order to improve the classification accuracy of bulbs will be investigated in the framework of the so-called CAMOTIUS project (Van der Wel and Gorte, 1995).

3.8 Conclusions

It can be concluded that different digital geographical databases of the Netherlands enable the discriminination of the main land use types in the LGN2 database: agricultural area, built-up area, and natural area and forest. Because of the thematic classes, cost, accuracy and continuity of the data, use of the CBS Land Use Database is preferred. The CBS Agricultural Statistics enable a further subdivision of agricultural area. In the framework of the LGN project it is preferred to use these ancillary digital data for stratification purposes. Stratification is rather simple, it decreases spectral variation and confusion and enables to better focus the discrimination process on problem classes. In special cases ancillary data, not available before the classification, may be applied to correct misclassifications (postclassification sorting). Because it is applied after classification, misclassifications can be corrected by recoding, avoiding a time comsuming new classification. When relevant digital data are available before classification, in general, stratification is prefered to postclassification sorting. Some LGN2 land cover classes, especially greenhouses, orchards, roads and buildings in agricultural areas, are also included in other digital geographical databases. Classifying these classes again by interpretation of satellite images is, generally, waste of time. However, satellite images can sometimes be used for up-dating these classes. Further, the use of satellite images can also be preferred because of financial reasons.

When land cover statistics for strata are available *class-based* a-priori probabilities, estimated by the relative areas of the land cover classes, can simply be included in the classification process. However, the increase in classification accuracy is small. The possibilities of using *conditional* a-priori probabilities, based on soil type and water table classes, are rather poor. Generally, the increase in classification accuracy is small and estimation of a-priori probabilities requires often a considerable additional sampling effort. The cost of acquiring these data does not balance the expected classification improvement. On very wet soils (water table classes I, II and II*) grassland is the only crop grown. Water table classes could also be used to discriminate between wet and dry natural areas. In these cases water table classes could be used for additional stratification, rather than for providing a-priori probabilities.

Combined use of the 1 : 10 000 topographic database and satellite images may provide field boundaries to be applied in a field-based classification. The classification result for agricultural crops could be improved considerably by performing a field-based classification. It has to be investigated if the required field boundaries can be obtained in an operational and cost effective way.

4 Classification methodology

4.1 Introduction

The LGN1 database was produced by automatic classification of manually stratified mono-temporal Landsat TM images from 1986 (Thunnissen et al., 1992b). The classification accuracy showed a large variation over the country. Similarity in spectral reflectances at the image acquisition date impeded consistent identification and mapping of a number of important land cover types. With a view to the up-dating and operational implementation of the LGN database an improved classification method, resulting in a significant improvement of the clasification result, had to be developped. Significant improvements of the classification result could be expected by a reduction of the spectral confusion between the different land cover types. Spectral signatures of a range of cover classes, such as agricultural crops, vary throughout the year. By that, classes which appear very similar in spring, may become separable at other stages of the phenological cycle. It is therefore expected that multi-temporal approaches provide important means to improve classification accuracy. Further classification improvement of spectral overlapping land cover classes could be expected by the use of other digital geographical data. Use of ancillary digital geographical data for stratification of the satellite images seems most suitable (3.8). After classification different postprocessing techniques could be applied to further improve classification accuracy.

For land cover classification purposes Landsat TM and SPOT images are already used on an operational basis. However, using these sensors, which operate in the visible and infrared part of the spectrum, regular data acquisition is often hindered by clouds in the Netherlands. Active microwave sensors acquire data independent of clouds. It is therefore investigated if the microwave data supplied by the imaging Synthetic Aperture Radar (SAR) on board of the first European Remote sensing Satellite (ERS-1) can also be used for land cover classification in behalf of the LGN database.

This chapter discusses the improved classification techniques. Attention is paid to stratification (4.2), automatic and interactive mono- and multi-temporal classification (4.4 and 4.5) and postprocessing techniques (4.11). The possibilities of using ERS-1 images in behalf of the LGN database is discussed in 4.10. The classification of greenhouses, orchards and buildings in agricultural areas proved to be troublesome. Specific classification techniques were developed for these classes (4.7, 4.8 and 4.9). The improved classification methods have mainly been developed on the basis of the data gathered in the Drenthe test site and for the production of the LGN2 database (2.1).

4.2 Stratification

Stratification of the satellite images, prior to the classification, is the most suitable method to discriminate the main land use types in the LGN2 database (3.8). Different nation-wide digital geographical databases of the Netherlands enable the discrimininination of the main land use types in the LGN2 database: agricultural area, built-up area (including sport and leisure facilities, dumping sites, mining areas, parks and public gardens and the like), natural area and forest and water. Because of the thematic classes, cost, accuracy and continuity of the data, use of the CBS Land Use Database is preferred (3.4). Because this database was not yet available during the start of LGN2 project the BARS Land Use Database was used for the stratification of the satellite images used in the LGN2 project and for the classification of the Drenthe test site. The individual land use classes were aggregated into main land use classes or strata. Depending on the source data used for the production of the BARS database the actuality of different map sheets of the BARS database varied strongly. The main land use classes were up-dated by visual interpretation of satellite images, supported by simultaneous consultation of topographic maps and aerial photographs. For this purpose the aggregated BARS database was superimposed on the geometrically corrected and contrast enhanced satellite images. The land cover changes were directly digitized on screen. The main land cover changes concerned extensions of towns, sport and leasure facilities and mineral extraction sites. In general, these land cover changes could readily identified on satellite images. First classification results showed that the stratum 'built-up area' had to be subdivided into urban areas and less densely built on areas, such as sports grounds, airports, recreational areas, dumping sites and the like. The spectral signatures of forest and grassland in wet natural areas (i.e. inland marches and peat bogs) differed considerably from the spectral signatures of these classes in dry natural areas. Therefore, a subdivision of natural areas into dry and wet nature area would be very useful. As distinct from the BARS database, this subdivision can be made with the CBS Land Use database (3.4) and the soil map (3.7).

The spectral characteristics of, especially, agricultural crops are likely to vary over distance. The CBS agricultural regions were used to divide the agricultural stratum into smaller, more hogeneous areas. For each agricultural region the areas of the occurring agricultural crops were known (3.6).

4.3 Classification approaches and assessment of classification accuracy

The improved classification procedure consists of the integrated use of satellite images, digital ancillary data, reference data and expert knowledge. The classification method is characterized by a stratified approach (4.2), i.e. every stratum is separately classified. For each stratum the occurring land use is known beforehand and each stratum must be covered by the same satellite images. A stratum is further subdivided when a part of the stratum has to be classified with satellite images from other acquisition dates. Different classification approaches were evaluated: mono-temporal, multi-temporal, automatic and visual classification. Prior to the classification the satellite images were geometrically corrected by identifying ground control points in the original imagery and on the reference topographic maps. In general, first-order order polynomal transformation equations were applied. Some (parts of) satellite images were geometrically corrected using second-order order polynomal transformation equations. The positional accuracy of the corrected satellite images was checked by projecting the BARS vector database on the satellite images. The images were resampled (by applying the nearest neighbour algorithm) to 25 m by 25 m to achieve a close match with standard map sheets of The Netherlands. It was decided to use the maximum likelihood algorithm for the (supervised) automatic classification approach. For specific applications, such as the discrimination between bare and vegetated areas, the Normalized Difference Vegetation Index (NDVI) was applied:

NDVI=(IR-R)/(IR+R)

where R and IR denote the pixel values in the red and near-infrared bands, respectively. Bare soil has a NDVI value near zero. The NDVI value increases as a funtion of the leaf are index.

In order to get an optimal land cover discrimination Landsat TM images were preferred to SPOT images because of the availability of TM bands in the middle-infrared part of the spectrum. Various surveys in agricultural and/or forestry study sites, using TM images, showed TM bands 3, 4 and 5 (Table 1) to be most suitable for land cover classification (Thunnissen et al., 1992a). Moreover, these bands are less affected by haze problems than is the blue and green part of the spectrum (Fuller et al., 1994). Consequently, we used TM bands 3, 4 and 5. When suitable TM images were lacking SPOT images were selected. SPOT bands 1, 2 and 3 (Table 2) were used in the classification process. Field reference data, topographic maps and satellite images were used for selecting training areas representative of the land cover classes to be classified.

The classification approaches were evaluated in the Drenthe test site and in a number of strata of the LGN2 database. Classification results were assessed by visual comparison of the classified images with reference data (i.e. topographic maps and aerial photographs) and by checking the labels of a sample of pixels from the classified image against the reference classes determined in the field. The classification accuracy of the sampled pixels are presented in error matrices (Table 9). From these matrices the percentage of pixels from each class in the image labelled correctly by the classifier can be estimated, along with the proportion of pixels from each class erroneously labelled into every other class. The reference data are usually represented by the columns of the matrix and the classified data are represented by the rows. The 'overall classification accuracy' is calculated by dividing the number of correct classified pixels by the total number of pixels sampled. In addition, accuracies of individual land cover classes can be computed in a similar way. Story and Congalton (1986) distinghuish 'producer's accuracy' and 'user's accuracy' (Table 9). 'User's accuracy' is also called reliability. The 'producer's accuracy', further called accuracy for short, is the probability for a reference sample to be correctly classified. This accuracy measure is calculated by dividing the number of correct classified pixels in a class by the total number of pixels in that class as derived from the reference data (i.e. the column total). The 'user's accuracy' or reliability is defined as the probability that a sample from the classified image actually represents that category on the ground. This accuracy measure is calculated by dividing the number of correct classified pixels in a class by the total number of pixels that were classified in that class (i.e. the row total). The statistical evaluation of the error matrix is discussed in 5.3.

Table 9 The error matrix and the calculated accuracy measures as presented by Congalton, 1991. Numbers in the matrix express numbers of pixels

	Reference data				
Classified data	deciduous forest	coniferous forest	barren	shrub	row total
Deciduous forest	65	4	22	24	115
Coniferous forest	6	81	5	8	100
Barren	0	11	85	19	115
Shrub	4	7	3	90	104
Column total	75	103	115	141	434

Overall accuracy = 321/434 = 74%

	Producer's Accuracy	User's Accuracy
Deciduous	65/75 = 87%	65/115 = 57%
Coniferous	81/103 = 79%	81/100 = 81%
Barren	85/115 = 74%	85/115 = 74%
Shrub	90/141 = 64%	90/104 = 87%

4.4 Mono-temporal classification

At this point, it should be noted that spectral classes may differ from informational classes in the sense that each of the informational classes (in this case LGN2 classes) may comprise several spectral subclasses (Hill, 1993). This appears obvious for the LGN2 classes 'other agricultural crops' and 'other nature area with low vegetation'. The former class includes of course numerous crops and the latter includes a wide range of different species and canopy structures. But it equally holds for most other LGN2 classes. After the classification, the individual spectral classes have to be combined to form the desired informational classes. In general, mono-temporal classification of Landsat TM images, obtained during the period mid-May to late September, provided good results for most LGN2 classes in the strata built-up area (including urban areas and less densely built on areas), natural area and forest and water. However, some of the concerning LGN2 classes could not be distinguished because they possessed similar spectral reflectances. For example, densely forested residential areas could not be distinguished from (patches of) forest situated in urban area. Discrimination of these LGN2 classes could be achieved by recoding the concerning spectral classes to the corresponding LGN2 classes on the basis of the spatial distribution of individual BARS classes. Athough not included in the LGN2 database, larch forms an important tree species because it looses its needle-leaves in autumn, which is of importance in hydrological (groundwater recharge!) and environmental (atmospheric deposition!) studies. Larch proved to be accurately classified when additional images from winter were used.

Tables 10 and 11 show some final results of the mono-temporal classification approach for the Drenthe test site and the LGN2 database. For the calculation of the overall classification accuracy of the Drenthe test site, the reference class mixed forest, classified as one of the spectral classes decidous or coniferous forest, was considered to be correctly classified. In the reference data of the LGN2 database mixed forest was not distinghuished.

Table 10 Error matrix showing the result of the mono-temporal classification approach for the stratum 'forest and natural area' in the Drenthe test site. The classification was performed with a Landsat TM image acquired on 2 September 1991. Numbers in the matrix express numbers of pixels.

Classified data	Reference data					Reliability (%)
	deciduous forest	coniferous forest	mixed forest	open natural area with low vegetation	total	
Deciduous forest	31	6	7	1	45	84.4
Coniferous forest	5	85	14	0	104	95.2
Open natural area with low vegetation	0	7	2	28	37	75.7
Total	36	98	23	29	186	
Accuracy (%)	86.1	86.7	91.3	96.6		

Overall classification accuracy (inclusive of mixed forest): 88.7%

Table 11 Error matrix showing the result of the mono-temporal classification approach for some(aggregated) LGN2 classes. The classification was performed with Landsat TM images acquired on 12 and 24 May 1992. Numbers in the matrix express numbers of pixels (after Noordman et al., 1996)

Classified data	Reference data							Reliability (%)
	deciduous forest	coniferous forest	natural area with low vegetation	urban area (built-up)	green urban area	arable land	total	
Deciduous forest	80	13	7	0	0	0	100	80.0
Coniferous forest	3	94	3	0	0	0	100	94.0
Natural area with low vegetation	5	2	92	1	0	0	100	92.0
Urban area (built-up)	0	0	0	94	5	1	100	94.0
Green urban area	0	0	0	5	92	3	100	92.0
Total	88	109	102	100	97	4	500	
Accuracy (%)	90.9	86.2	91.3	90.2	94.0	94.8		

Overall classification accuracy: 90.4%

4.5 Multi-temporal classification of agricultural crops

For an accurate classification of most agricultural crops the use of multi-temporal satellite data is required. A problem of multi-temporal approaches to automatic image classification is the increasing data volume which results from a simple combination of the spectral bands from each of several acquisitions. In this study the combination of three satellite images would produce a 9-band composite image (only TM bands 3, 4 and 5 were used). Different data reduction methods have been developed, such as optimal band selection, ratios and differences of spectral bands and statistical transformations (e.g. principal component analysis). In numerous remote sensing applications, various combinations of bands in the visible and near-infrared part of the spectrum have proven their suitability to emphasize important plant phenological characteristics. However, these so-called vegetation indices only use a part of the spectral information that is provided by the Landsat Thematic Mapper sensor, and the reduction to one single parameter for each date usually implies too big information losses to be further used for land cover classification. Nevertheless, the use of vegetation indices may be very useful for specific applications such as the discrimination between bare and vegetated areas. Statistical transformations, such as principal component analysis, are data dependent. By that, objects, having similar spectral characteristics, may show different colours in different transformed images (Sheffield, 1985), making it difficult for an interpreter to apply previous experience of colour-surface relationships. As a result, it was decided to use original spectral bands for image classification.

When using a composite image, containing all the selected bands of the obtained satellite images, three bands have to be selected for graphical display. By that, only a part of the spectral information can be used for the selection of training areas. An other problem is related to the number of spectral classes per agricultural crop or field. A particular crop or agricultural field may comprise several spectral classes on *each* of the satellite images, dependent on crop growth conditions (e.g. water supply and soil type), phenological stage (e.g. emergence, ripening and withering) and cultivation practices (e.g. conversion of grassland into arable land, use of different varieties, harvesting and after growth). As a result the composite image will often comprise an excessive number of spectral classes, making the selection of the training areas a time-consuming and difficult task (Hill, 1993 and Fuller et al., 1994). Therefore, it was decided to classify each satellite image seperately (i.e. mono-temporal classification) and to combine the classified images in a GIS to form the final classification result, using conditional 'IF-THEN' statements. Mostly, a NDVI image from spring a was used to mask the satellite images from summer prior to classification.

To enable applications of the LGN2 database on a regional scale, mininimum classification accuracies and reliabilities of 70% are required. However, poor spectral separability and cultivation practices such as harvesting may result in classification accuracies and reliabilities below 70%. In these cases mixed agricultural classes (e.g. maize/sugar beet) were defined. However, when agricultural crops A and B showed large spectral confusion, but crop A covered only a relative small area in comparison with crop B, then class A was assigned to class B and no assignment to a mixed class occurred. The decision to define mixed classes is based on the expertise of the interpreter, data from the CBS Agricultural Statistics and field reference data. In case of partial spectral overlap, resulting in classification accuracies and reliabilities below 70%, the classification process was aimed at the subdivision of the overlapping classes into 'pure' and mixed classes. The inclusion of mixed classes will be demonstrated on the basis of the classification of the Drenthe test site. In general, the definition of mixed classes effects a (considerable) improvement of classification accuracies (inclusive of mixed classes) and reliabilities. However, classification accuracy of the pure classes decreases (Tables 12 and 13).

Table 12 Error matrix showing the classification result for the agricultural crops in the Drenthe test site derived from a Landsat TM image acquired on 2 September 1991. Numbers in the matrix express numbers of pixels

Classified data	Reference data							Reliability (%)
	grass	maize	potatoes	sugar beets	cereals	other agricultural crops	total	
Grass	87	0	8	4	2	0	101	86.1
Maize	0	22	7	7	0	1	37	59.5
Potatoes	11	1	39	3	1	1	56	70.0
Sugar beets	7	7	30	43	0	0	87	49.4
Bare soil	20	3	48	6	47	8	133	0.0
Total	125	33	132	63	50	10	414	
Accuracy (%)	69.6	66.7	29.6	68.3	0.0	0.0		

Overall classification accuracy: 46.3%

Table 13 Error matrix showing the classification result for the agricultural crops in the Drenthe test site derived from a Landsat TM image acquired on 2 September 1991. Mixed classes were included. Numbers in the matrix express numbers of pixels

Classified data	Reference data							reliability (%)
	grass	maize	potatoes	sugar beets	cereals	other agricultural crops	total	
Grass	87	0	8	4	2	0	101	86.1
Maize	0	9	0	4	0	0	13	69.2
Potatoes	8	1	28	1	1	1	40	70.0
Sugar beets	1	0	2	24	0	0	27	88.9
Cereals/potatoes	17	2	37	3	47	6	113	74.3
Maize/sugar beets	0	13	1	19	0	1	39	94.1
Maize/potatoes/sugar beets	12	8	56	8	0	2	86	83.7
Total	125	33	132	63	50	10	414	
Accuracy (%)	69.6	27.3	21.2	38.1	0.0	0.0		
Accuracy inclusive of mixed classes (%)	69.6	90.1	91.7	80.1	94.0	0.0		

Overall classification accuracy: 35.8%
Overall classification accuracy (inclusive of mixed classes): 81.4%

The NDVI image derived from the SPOT image from 3 February 1991 was used to discriminate between grassland and bare soil in the multi-temporal classification of the Drenthe test site. Subsequently, using the NDVI mask, the satellite images from summer were used for the classification of the arable crops. The LGN2 class 'other agricultural crop' formed a heterogeneous class, often consisting of small fields. This class was not included in the classification. The final classification result was obtained by combination of the separately classified satellite images (Table 14 and 15).

As expected the multi-temporal classification approach (overall accuracy exclusive of mixed classes: 55.0%) is superior to the single-date classification (overall accuracy exclusive of mixed classes: 35.8%). Both classification approaches provide comparable overall classification accuracies inclusive of mixed classes (76.0% versus 81.4%). The multi-temporal classification approach without distinction of mixed classes resulted in an overall classification result of 63.4%. In the latter case no additional information is available on the misclassified pixels. In practice, the decrease in classification accuracy of the pure classes has to be balanced with the improvement in classification accuracy (inclusive of mixed classes) and reliability and the additional information resulting from the introduction of mixed classes.

Table 14 Error matrix showing the multi-temporal classification result for the agricultural crops in the Drenthe test site. Mixed classes were included. Numbers in the matrix express numbers of pixels

Classified data	Reference data							Reliability (%)
	grass	maize	potatoes	sugar beets	cereals	other agricultural crops	total	
Grass	91	2	23	7	3	0	126	72.2
Maize	0	8	0	4	0	0	12	66.6
Potatoes	6	1	73	3	3	1	87	83.9
Sugar beets	1	0	2	22	0	0	25	88.0
Cereals	1	1	1	1	33	2	44	75.0
Cereals/potatoes	11	0	11	3	11	4	40	55.0
Maize/sugar beets	0	13	1	16	0	1	31	93.5
Maize/potatoes/sugar beets	10	8	21	7	0	2	48	75.0
Total	125	33	132	63	50	10	414	
Accuracy (%)	72.8	24.2	55.3	34.9	66.0	0.0		
Accuracy inclusive of mixed classes (%)	72.8	87.9	79.5	71.4	88.0	0.0		

Overall classification accuracy: 55.0%
Overall classification accuracy (inclusive of mixed classes): 76.0%

Table 15 Combination of the separately classified satellite images to form the final classification result for the agricultural crops in the Drenthe test site. The conditional statements were applied in the order in which they are placed

NDVI image (3 Februari 1991)	Classified SPOT image (19 July 1991)	Classified TM image (2 September 1991)	Final classification result
High NDVI		high NDVI	grassland
Low NDVI		sugar beet	sugar beet
		maize	maize
		potatoes	potatoes
	potatoes		potatoes
	cereals	potatoes/cereals	cereals
	maize/potatoes/ sugar beet	potatoes/cereals	potatoes
		potatoes/cereals	potatoes/cereals
		maize/sugar beet	maize/sugar beet
		maize/potatoes/ sugar beet	maize/potatoes/ sugar beet
		other cops	other crops

The overall classification accuracy of the pure classes is well below 70%. This is due to considerable spectral confusion on the satellite images from 29 July (SPOT) and 2 September 1991 (Landsat TM) and harvesting activities, mainly in August. The spectral confusion on the SPOT image from 29 July 1991 was caused by the poor spectral resolution of SPOT, while the spectral confusion on the Landsat TM image from 2 September 1991 was mainly caused by withering of potatoes and water stress. During the month of August 1991 there was hardly any rainfall. Therefore, on 2 september 1991 dry conditions occurred in the field.

The NDVI image from February was expected to accurately distinguish between grassland and arable land. Actually, the result is somewhat disappointing (Table 14). That is largely caused by the early acquisition date of the SPOT image. Evaluation of the reference data and the satellite image showed that about 13% of the fields which showed bare soil in February were sown with grass in spring, while about 27% (!) of the fields which were covered with crops in February were used for growing arable crops. The latter fields concerned both grasslands, converted into arable land, and fields covered with a second crop grown after harvesting of the main crop. (e.g. for the purpose of green manuring). The same cultivation practices were found during the processing of the TM image of 14 Februari 1994, which was used for the multitemporal classification of agricultural crops in the LGN2 database in large areas in the western and southern part of the Netherlands (Noordman et al., 1996 and 4.6).

Spectral confusion between different agricultural crops may have different reasons. An important reason is a deviating crop development due to, for example, shortage of water and/or nutrients or water logging. By that, misclassified pixels often form irregular patterns within a field. Agricultural fields may also show (large) spatial variability in spectral refectance due to, for example, ripening or withering, causing (clusters of) misclassified pixels to occur in a more or less regular distribution. Based on field shape, patterns of misclassified pixels and/or spatial variation in refectance (texture) these fields can often be *visually* recognized as separate fields with a particular crop. Besides shape, pattern and texture the location amidst other fields (context) may also play an important part in the visual interpretation process. Grassland areas, for example, can often readily be visually recognized on the basis of the characteristic variation in reflectance caused by grazing, mowing and regrowth. So in many cases visual interpretation of satellite images is superior to automatic classification, which is solely based on the spectral characteristics of individual pixels.

Prior to the decision to use visual interpretation techniques instead of automatic classification the pro's and cons have te be compared. Factors as time, classification accuracy and the importance of the crop have to be taken into consideration. In practice, visual interpretation often appears to be a valuable tool, complementary to automatic classification. Advanced hardware and software enable the simultaneous interpretation of different satellite images, while the interpretation result can directly be stored in digital form by on screen digitizing. During the production of the LGN2 database visual interpretation proved to be *necessary* in many areas in order to get an acceptable classification result. However, the overall information content of the multi-temporal satellite data is so high, that only automatic approaches can provide classification results for (very) large areas.

4.6 Optimal acquisition dates and availability of satellite images for the clasification of agricultural crops

The classification of agricultural crops in the LGN database is based on a multi-temporal approach (4.5). Landsat TM and SPOT XS are the most suitable satellite images for land cover classification in the framework of the LGN project. Optimal image acquisition dates, required spectral resolution and the availability of satellite images are discussed in this chapter.

In general, Landsat TM images are preferred to SPOT images because of the presence of middle-infrared TM bands required for accurate land cover classification. Moreover, Landsat TM images are considerably cheaper than SPOT images. A SPOT coverage is circa 5 times more expensive per unit area covered. Altough the Landsat 5 satellite, which was launched over eleven years ago in 1984, continues to provide high quality multispectral imagery, the ageing sensors on board Landsat 5 threaten continuity of Landsat imagery. Landsat 6 was lossed at launch and Landsat 7 is scheduled for launch in December 1998. On the other hand SPOT 4, planned for launch in 1996, will be equiped with a middle infrared band. The possibility of SPOT of pre-programmed, off-nadir, imaging will increase the chance of getting suitable images.

Phenological data for the main crops growing in a stratum (e.g. planting/sowing date, ripening) and cultivation practices (e.g. conversion of grassland into arable land and vice versa, harvesting, after growth) should be taken into consideration when selecting optimal image acquisitions. A minimum image data set should include at least one image from spring and one image from summer. For many crops April/May is the optimal image acquisition period in spring. Nevertheless grassland and winter cereals may be confused in early spring (April). Discrimination between grassland and arable land, using a winter image, may be hindered by aftergrowth (e.g. fields covered with crops used for green manuring), conversion of grassland into arable land and vice versa (4.5). The last two practices are especially of importance in areas with mixed land use, where both grasland and arable land cover relative large areas. The growing of a second crop after the harvest of the main crop (e.g. cereals) is strongly increasing last years for the purpose of green manuring, improving soil structure and decreasing the risk of plant diseases. In areas with clay soils the second crop is mostly ploughed before winter. For many arable crops the month of July is the optimal image acquisition period in summer as far as the the minimum image data set is concerned. However, the use of additional images from other periods will, in general, considerably improve the classification result. Experience has shown that the optimal image data set should include images from May, June and August. The increased classification result, in general, counter-balances the additional cost for purchase and processing of these images. In practice, the choice of the images to be used should be based on the occurring land cover, the required classification result and the available satellite images.

When images from suboptimal periods or images which lack spectral bands in the middle-infrared part of the spectrum are used, spectral confusion may effect the inclusion of mixed classes in the LGN database. Mixed classes hamper efficient classification of the satellite images and operational application of the LGN database. Therefore, it is prefered to avoid using suboptimal images as much as possible, even if one would have to wait another growing season for more suitable images.

For training the classifier and validation of the classification result the availability of suitable reference data is of great importance. Reference data of agricultural crops have to be gathered in the field. In practice, the final choice of the images to be used for classification is often made in late summer. For example, a part of the agricultural crops in the LGN2 database of the western part of the Netherlands was classified using satellite images obtained on 16 August 1994. Mostly, the final choice can only be made after visual evaluation of the quick look data for determination of location (especially for SPOT images), quality and information content of the images. Existing catalogues of acquisitions proved unreliable, especially as far as the cloudiness is concerned, and made no distinctions between cloud over sea or land. Moreover, marked differences existed between assessments made at Fucino and Kiruna (TM). However, it can take some weeks after image acquisition before the quick looks themselves are available. Therefore, in practice, often a restricted time period is available for field survey. Many crops are already harvested in late summer. With the modern electronic tools quick look data should be made available on line within 24 hours of acquisition of the image.

4.7 Classification of greenhouses

The LGN1 database was produced by automatic classification of Landsat TM images, using bands 3, 4 and 5. Because of the low spectral separability of greenhouses this class was included in the LGN1 class 'built-up area' (Thunnissen et al., 1992b). However, information on the location of greenhouses is of great importance for physical planning and environmental purposes. Therefore, greenhouses should be included in the LGN2 database.

Greenhouses are included in the topographic databases (3.5) and the CBS Land Use Database (3.4). Both databases were, however, not yet available in digital form during the start of LGN2 project. Greenhouses are also included in the BARS database (3.4). The classification accuracy of greenhouses varies, however, strongly in the BARS database and in some areas large contiguous (agricultural) areas are included in the BARS class 'Greenhouses'. Therefore, it was decided to investigate the possibilities of satellite images to classify greenhouses. Different satellite images (SPOT and Landsat TM) were visually evaluated in several test sites in order to determine the optimal band combinations for the classification of greenhouses (Table 16). Herewith, most attention was paid to the Landsat TM images from 15 and 24 May 1992, which covered the entire country. The visual evaluation of the satellite images was supported by 1 : 25 000 topographic maps and aerial photographs.

Table 16 Satellite images used for determination of the optimal band combinations for the classification of greenhouses

Satellite images	Scene	Acquisition Date
SPOT XS en P	45-245	01-09-1987
SPOT XS	42-244	27-06-1992
Landsat TM	198-24	13-07-1990
Landsat TM	198-24	02-09-1991
Landsat TM	198-24	15-05-1992
Landsat TM	198-24	14-02-1994

For the classification of the multispectral SPOT images bands 1, 2 and 3 (Table 1) were used. Optimal Landsat TM band combinations proved to vary for different areas and acquisition dates. Favourable band combinations found were 1, 2 and 3, 3, 4 and 5 and 2, 6 and 7 (Table 2). In general, the SPOT panchromatic band proved lo be less successfull than the SPOT multitemporal bands. In spite of the low spatial resolution of TM band 6 (120 m), in many areas the band combination 2, 6 and 7 was found to be superior to other band combinations for the discrimination of greenhouses. The greenhouses showed a relatively low radiation temperature on the TM images from 13-07-90, 02-09-1991 and 15-05-1992. The recording of the thermal band of the TM image from 14-02-1994 had failed. The best classification result was achieved by using band combination 4, 5 and 3 (depicted as red, green and blue respectively) of the Landsat TM image from 14-02-1994. For none of the selected band combinations automatic classification of the greenhouses provided sufficient classification accuracy and reliability. In general, the greenhouses showed a relatively large spectral variability and spectral overlap occurred, especially with bare soil and built-up area. However, most greenhouses could visually be distinguished. So, for an optimal classification result an interactive interpretation of the selected band combination(s) should be performed. Therefore, in some test sites the greenhouses were classified by interactive interpretation of band combination 2, 6, 7 (depicted as red, green and blue respectively) of the Landsat TM image from 15 and 24-05-1992. These images covered the entire country. Spectral confusion of greenhouses with bare soil could largely be solved by masking the satellite image with a NDVI image from summer, when most plots were covered with vegetation. Spectral confusion of greenhouses with built-up area could only be solved by using ancillary data, namely topographic maps and aerial photographs. Relatively small greenhouses, consisting of only a few pixels, were often difficult to locate on topographic maps and aerial photographs.

The results of the interactive classification of greenhouses in the test sites were compared with the CBS Agricultural Statistics (3.6) and with the greenhouses on the 1 : 25 000 topographic maps and aerial photographs.The greenhouses in the test sites were digitized from the 1 : 25 000 topographic maps. Small greenhouses could mostly not be distinghuished because of the restricted spatial resolution of the satellite images. The minimum size of the greenhouses required for recognition amounts to circa 0.5 ha. The classification result of large greenhouses or greenhouse complexes often proved to show forms which deviated strongly from the real greenhouses. Moreover, the classified greenhouses and greenhouse complexes were generally considerably larger than in reality. The classified and real greenhouses sometimes overlaped one another less then 50%. The differences beween the classification results and the real greenhouses were caused by the spatial resolution of the Landsat TM images (120 m for band 6 and 30 m for bands 2 and 7) and spectral confusion with neighbouring pixels. Mixed pixels (coinciding only partly with the greenhouses) and houses, buildings and open areas situated among contiguous greenhouses were often classified as 'greenhouses' making the classified greenhouses generally (considerably) larger than the real greenhouses. Consequently, interactive classification of greenhouses can both overestimate and underestimate the area of greenhouses dependent on the occurring situation. Therefore, the classification result cannot be validated by comparing the classified areas with the areas according to the CBS Agricultural Statistics.

Evaluation of the above mentioned results shows that greenhouses cannot be sufficiently accurately classified by visual interpretation of Landsat TM images. Therefore, it was decided to digitize the greenhouses from the most recent 1 : 25 000 topographic maps. To restrict the number of maps to be digitized, only the municipalities, containing more than 10 ha greenhouses, were considered. The concerning municipalities were selected from the 1992 CBS Agricultural Statistics. In principle the boundaries of individual greenhouses were digitized. However, houses, buildings and (small) open areas situated within greenhouse complexes were also included in the LGN2 class 'Greenhouses'. Subsequently, the greenhouse database was up-dated to the situation in 1992. For this purpose the digitized geeenhouses were projected on the satellite images from may 1992. Possible changes, i.e. disappearance or extension of greenhouses, were traced by visual interpretation of the satellite images. Special attention was paid to those municipalities which showed relative large differences between the area of greenhouses digitized from the topographic map and the area of greenhouses according to the CBS Agricultural Statistics of 1992. In a number of these areas the interpretation of the satellite images was supported by use of aerial photographs from 1989. The interactive classification of greenhouses proved to be strongly supported by the simultaneous projection of the greenhouses digitized from the topographic map, providing useful additional information on spectral reflectance, shape, context and location of the greenhouses. The final classification results were semi-quantitatively validated by comparison of the classified areas with the areas according to the CBS Agricultural Statistics and by visual comparison of the classified greenhouses in some test sites with aerial photographs from 1989 and 1992 (Noordman et al., 1996).

As stated above greenhouses are also included in other digital geographical databases. If possible, greenhouses should be copied out of these databases into up-dates of the LGN database. However, in practice, data in the geographical databases will often be outdated compared with the acquisition dates of the satellite images. In these cases, the present LGN2 database forms a reliable starting point for the up-dating of greenhouses in future versions of the LGN database. The up-dating can be restricted to the municipalities where the change in area of greenhouses exceeds a certain minimum area, according to the CBS Agricultural Statistics. Because many greenhouses appear different on different satellite images the simultaneous use of several images is expected to improve the classification result.

4.8 Classification of orchards

The LGN1 database was produced by automatic classification of Landsat TM images, using bands 3, 4 and 5. The classification result of orchards was rather poor because of spectral confusion with grassland and forest (Thunnissen et al., 1992b). Because forests and orchards occur in separate strata (4.2) of the LGN2 database, the spectral confusion between orchards and grassland constitutes the main problem.

Topographic maps contain information on orchards. These maps were, however, not yet available in digital form during the start of LGN2 project. The BARS database contains only large orchard complexes and contains no orchards at all for a large part of the province of Limburg. Therefore, it was decided to investigate the possibilities of satellite images to classify the orchards not present in the BARS database. For this purpose different satellite images (SPOT and Landsat TM) from different acquisition dates were evaluated in two test sites in order to determine the optimal band combinations and acquisition dates for the classification of orchards. The test sites were situated in the 'Betuwe', an area in the middle of the Netherlands densily grown with fruit trees.

Table 17 Satellite images used for determination of the optimal band combinations and acquisition dates for the classification of orchards

Satellite images	Scene	Acquisition Date
SPOT P and XS	42-244	27-06-1986
SPOT XS		42-24427-06-1992
Landsat TM		198-2402-09-1991
Landsat TM		198-2415-05-1992

Due to the high spatial resolution of SPOT panchromatic images (10 m), these images could possibly be suitable to detect textural characteristics of orchards connected with the regular row spacing of the fruit trees. Close visual inspection of the contrast enhanced SPOT panchromatic image from 27 June 1986 (other panchromatic images were not available) showed, however, no spectral variation characteristic of orchards.

For the classification of the multispectral SPOT image bands 1, 2 and 3 (Table 1) were used. The optimal Landsat TM band combination for the classification of orchards consists of the bands 3, 4 and 5 (Table 2). In the beginning of the growing season spectral confusion with grassland dominated, while in the course of the growing season spectral confusion with grassland decreased and spectral confusion with forest increased. The change in spectral confusion is caused by the increase in soil coverage of the orchards during the growing season. Because forest and orchards occur in separate strata the spectral confusion between orchards and forest is a minor problem in the LGN2 database. Therefore, classification of satellite images obtained late in the growing season provided the best clasification results. Nevertheless, for none of the selected satellite images automatic classification of orchards provided sufficient classification accuracy and reliabiliy. Therefore, it was decided to classify the orchards by visual interpretation of satellite images supported by topographic maps. For the LGN2 database TM images obtained on 13 July 1990 (southern half of the Netherlands) and 2 September 1991 (northern half of the Netherlands) were used. In a large part of the province of Limburg orchards are rather small and/or show a low density of trees, hampering an accurate classification. In this area the orchards were digitized from the 1 : 25 000 topographic maps. For practical reasons the digitization of orchards from satellite images and topographic maps was only performed for the municipalities, containing more than 10 ha orchards according to the 1992 CBS Agricultural Statistics. The digitized orchards and the orchards in the BARS database were included in the LGN2 database. Subsequently, the areas of orchards in the LGN2 database were compared with the areas of orchards according to the CBS Agricultural Statistics of 1992. The municipalities of which the areas of orchards in both databases showed relative large differences were selected for further interpretation. For this purpose the orchards in the LGN2 database were projected on the satellite images and misinterpretations and possible changes, e.g. disappearance of orchards, were traced by visual interpretation of the satellite images supported by use of topographic maps.

As stated above orchards are also included in the digital topographic databases. If possible, orchards should be copied out of these databases into future up-dates of the LGN database. In practice, data in the topographic databases will often be outdated compared with the acquisition dates of the satellite images. In these cases, the present LGN2 database forms a reliable starting point for the up-dating of orchards in future versions of the LGN database. The planting of fruit trees will only be perceptible on satellite images after some years. Therefore, some field work may be necessary. The up-dating can be restricted to the municipalities where the change in area of orchards exceeds a certain minimum area, according to the CBS Agricultural Statistics.

4.9 Classification of buildings in agricultural area

Buildings in agricultural areas can in principle be copied out of topographic databases. Most topographic maps were, however, not yet available in digital form during the start of LGN2 project. Automatic classification of different satellite images showed that scattered buildings in agricultural areas could not be sufficiently accurately classified by automatic classification (Thunnissen et al., 1992b). The poor classification results were mainly caused by spectral confusion with bare soil and ripened cereals. However, on many satellite images most buildings in agricultural area could visually be distinguished. Visual interpretation is more succesful than automatic classification because visual interpretation is not only guided by tone but also by size and situation (e.g. with regard to roads) of the buildings. It was investigated if a specific method could be developed for the classification of buildings in agricultural areas, using satellite images and ancillary data.

The classification method was developed in a test site situated between Arnhem and Nijmegen in the eastern part of the fluviatile district of the Netherlands. Five test sites, situated in different characteristic landscapes, were selected for validation of the classification method. In four test sites a qualitative validation was performed. For the test site 'Gelderse Vallei' a semi-quantitative validation was performed. The test site 'Gelderse Vallei' is situated in the centre of the Netherlands in the Pleistocene sandy district. For the test site between Arnhem and Nijmegen 4 optical satellite images were available (Table 18). The roads included in the BARS and the CBS Land Use Databases (3.4) were used as ancillary data in the classification proces. The BARS database contains only the main roads while nearly all roads (inclusive of unmetalled roads) are included in the CBS Land Use Database. Topographic maps and aerial photographs were used for the validation of the classification results. For the test site situated between Arnhem and Nijmegen all buildings in the agricultural stratum were digitized as points. Clusters of buildings were digitized as one point. For the test site 'Gelderse Vallei' all buildings were transfered to transparencies which could be superimposed on hard copies of the classification results. In addition to optical images, the possibility of ERS-1 SAR images for the classification of buildings was investigated. The results of the latter research are discussed in 4.10.

The developed classification method exists of a combination of multi-temporal NDVI images (4.3) and the application of several GIS techniques. All satellite images were converted into NDVI images. Subsequently, a NDVI range was selected containing the buildings in agricultural area. The lower-value of the NDVI range equals zero, while the upper-value was visualy determined by trial and error.

Table 18 Optical satellite images used for the development of a method for the classification of buildings in agricultural area

Satellite images	Scene	Acquisition Date
SPOT	45-245	07-08-1992
SPOT	45-245	08-08-1992
Landsat TM	198- 24	03-08-1986
Landsat TM	198- 24	02-09-1991
Landsat TM	198- 24	15-05-1992

As the upper NDVI value increased not only the number of buildings in the NDVI range increased but also the number of pixels erroneously classified as 'built-up area' increased. Experience showed that the optimal NDVI range should contain at least all large buildings, i.e. farms. The NDVI range may differ for each separate satellite image. After selection of the NDVI range one value was assigned to all pixels within the range. All other pixels were recoded to 0. Besides a large number of buildings, the selected NDVI ranges contained stretches of main roads, fields and patches of bare soil, fields with ripened cereals and mixed pixels. Most fields with bare soil or ripened cereals and some stretches of main roads were considerably larger then the farms and could be separated from the farms and removed from the NDVI range by application of a 'CLUMP' and 'SIEVE' operation. Groups of contiguous pixels in the selected NDVI range (called 'clumps') were identified by their sizes and clumps, to be considered too small for the present application, were filtered out or 'sieved'. Clumps larger than 14 pixels were considered to be fields with bare soil or ripened cereals or stretches of main roads. In this way buildings bordering on fields with bare soil or ripened cereals were erroneously removed from the NDVI range.

Because the classification methodology to be developed applies to buildings in agricultural areas, the test sites were stratified using the agricultural strata as derived from the BARS database (4.2). This stratification has to be performed *after* the 'CLUMP' and 'SIEVE' operation, because stratification may create small clusters of pixels along the strata boundaries which are too small to be sieved. The stratification could also be performed at the end of the classification procedure.

Buildings are supposed to show relatively low NDVI values on every satellite image irrespective of their acquisition date. By combination of NDVI images of different acquisition dates (clusters of) pixels with a permanent low NDVI value could be selected. In some cases (clusters of) pixels, often representing small buildings, did not coincide or overlap on different satellite images because of positional errors due to the geometric correction of the satellite images. This problem could largely be solved by applying a pixel growing technique, i.e. the areas of the original (clusters of) pixels in the NDVI range were expanded by application of a 3 x 3 maximum filter. To prevent large clusters from becoming too large and the number of misclassified pixels from increasing considerably, pixel growing has to be applied to only *one* NDVI image. Application of the pixel growing technique on one image from summer proved to give best results. Pixel growing has to be applied, of course, prior to the combination of the separate NDVI images. All pixels overlapping in at least two NDVI images were stored in the so-called 'building' GIS layer. Besides buildings this 'building' GIS layer contained also some scattered (clusters of) pixels of main roads and bare soil. In general, the classification accuracy of the buildings decreased, the reliability of the classified buildings increased and the number of misclassified pixels decreased as the number of combined NDVI images increased. Most large farms and clusters of small buildings were included in the building GIS layer, while many scattered small buildings (dwelling-houses, sheds and the like) were not included in the building GIS layer. Images from summer, when most fields were covered with vegetation, contributed most to the classification result.

The (clusters of) pixels of main roads, still present in the building GIS layer after combination of the NDVI images, could be removed by overlaying the building GIS layer with the *main* roads included in the CBS Land Use Database. A further improvement of the classification result could be obtained by defining buffer zones around the roads included in the CBS Land Use Databases. The CBS Land Use Database contains nearly all roads (inclusive of unmetalled roads). Most buildings are situated near a road. So the classification result could be improved by creating a buffer of 125 m around all roads and removing all 'buildings' outside the buffer zones from the building GIS layer.

The developed technique for the classification of scattered buildings in agricultural areas was validated in 5 test sites. In four test sites a qualitative validation was performed by a visual comparison of the classification result with topographic maps. For the test site 'Gelderse Vallei' a semi-quantitative validation was performed by superposition of the buildings, transfered to a transparency, on hard copies of the classification result. In all test sites 3 or 4 satellite images were used for the classification. The classification results of the test sites were more or less comparable with the result of test site between Arnhem and Nijmegen where the classification technique was developed. Most large farms and clusters of small buildings were included in the building GIS layer, while many scattered small buildings (dwelling-houses, sheds and the like) were not included in the building GIS layer. In the test site 'Gelderse Vallei' 91% of the large farms and 64% of the (clusters of) small buildings were correctly classified, while the number of (clusters of) pixels erroneously classified as building was less than 2%. In practice the results of the developed classification method will show some variation dependent on the present land cover and the quality and acquisition dates of the used satellite images.

Scattered buildings in agricultural areas are not yet included in the LGN2 database but will be included in up-dated versions of the database. As stated above buildings in rural areas are also included in the digital topographic databases. If possible, these buildings should be copied out of these databases into future up-dates of the LGN database. However, in practice, data in the topographic databases will often be outdated compared with the acquisition dates of the satellite images. In these cases, large buildings could be accurately up-dated by the developed classification methodology, using satellite images.

The possibilities of ERS-1 SAR images to classify buildings in agricultural area is discussed in 4.10.

4.10 Application of ERS-1 images

For land cover classification purposes Landsat TM and SPOT images are already used on an operational basis. However, using these sensors, which operate in the visible and infrared part of the spectrum, regular data acquisition is often hindered by cloudiness or haze in the Netherlands. Especially the multi-temporal classification of agricultural crops in the LGN database requires a regular data acquisition. Active microwave sensors acquire data independent of clouds. With the launch of the first European Remote sensing Satellite (ERS-1) the first long-term spaceborne imaging microwave sensor has become available. The ERS-1 (and its successor the ERS-2) carries an imaging Synthetic Aperture Radar (SAR) which provides a continuous source of radar data for observation of the earth through the years. The use of ERS-1 images increases data reliability. It is, therefore, of importance to investigate if ERS-1 imagery could be used to improve the operational efficiency of the classification of agricultural crops.

Buildings often appear on the SAR images as bright spots. The high backscatter signals are caused by the so-called corner effect of smooth perpendicular surfaces oriented towards the satellite. The classification of scattered buildings in agricultural area, using optical sensors, is troublesome (4.9). Therefore, also the possibility of ERS-1 SAR images to map buildings was investigated.

The applicability of the ERS-1 SAR images was investigated in the Zuid Flevoland test site, situated in the central part of the Netherlands (Fig. 1). The landscape structure is characterized by a division into large, rectangular lots. In general, the lots are 85 ha in size. In total twelve different crops are present in the test site. The main crops grown are potatoes, sugar beets and winter wheat, which cover approximately 53% of the area. The other crops grown are maize, spring barley, winter rape, beans, peas, grass, lucerne and orchards. A crop type map of the test site for the 1992 growing season was established based on a ground-based survey.

The ERS-1 satellite supplies SAR C-band data (VV polarization with an incidence angle of 23° at mid-swath). For the Zuid Flevoland test site 14 ERS-1 SAR images, acquired in 1992, were available (Schotten et al., 1995). On the basis of crop separability indexes optimal sets of ERS-1 images were selected for crop classification. Schotten et al. (1995) found that in the Zuid Flevoland test site application of a field-based multi-temporal classification of in total 12 crops, resulted in an overall classification accuracy of 80% of the fields. This corresponds with 88% of the area (Table 19). For these classification eight images were used from the 1992 growing season starting at 12-05-92 with time intervals varying from 7 to 19 days up to 03-11-92. Using less images, resulted in lower overall classification accuracies. A field-based classification requires the availability of digital field boundaries. These data are, however, not available and digitization of field boundaries for large areas is too expensive. Approaches to obtain field boundaries automatically, such as edge detection and region growing techniques, cannot be applied on an operational basis. Therefore, also the possibilities of a pixel based classification of ERS-1 images was investigated. A pixel-based classification is complicated by the presence of speckle. In order to reduce the speckle a Gamma-Gamma MAP (Maximum A Posteriori) filter (Lopes et al., 1993) was applied. This filter reduces speckle while preserving spatial resolution and structural features. After application of the Gamma-Gamma MAP filter an extra 5 x 5 majority filter was applied. The latter filter was used to simplify the selection of training areas because the variances within the fields were still quite large. To compare the results of the pixel based classification with those of the field-based classification, the pixel based (maximum likelihood) classification was carried out with the same eight-date multi-temporal data set for the Zuid Flevoland test site.

Table 19 Error matrix showing the field-based crop classification result of the Zuid Flevoland test site for the eigth-date SAR data set. The figures are given in hectares (adapted after Schotten et al., 1995)

Classified data	Reference data													Reliability (%)
	potatoes	sugar beets	winter wheat	grass	maize	winter rape	spring barley	orchards	onions	beans	peas	lucerne	total	
Potatoes	1751	189	8	22	0	0	0	0	10	0	10	0	1989	88
Sugar beets	281	1429	27	22	3	0	0	5	0	0	18	0	1784	80
Winter wheat	0	24	1722	10	9	0	0	0	30	2	0	0	1797	96
Grass	0	4	14	1287	26	0	0	0	5	12	0	0	1347	96
Maize	23	5	5	39	259	0	0	19	0	0	0	0	349	74
Winter rape	0	0	0	0	0	565	0	0	4	0	0	0	570	99
Spring barley	0	0	0	0	0	0	824	0	0	0	7	0	831	99
Orchards	19	0	2	18	0	0	0	306	0	1	10	0	357	86
Onions	17	2	29	17	0	0	0	12	477	7	6	0	566	84
Beans	6	0	19	16	24	0	0	0	6	371	6	0	449	83
Peas	49	41	12	6	10	0	0	11	7	14	302	0	453	67
Lucerne	0	0	0	0	0	0	0	0	0	0	6	484	490	99
Total	2146	1693	1839	1436	330	565	824	353	539	407	365	484	10981	
Accuracy (%)	82	84	94	90	78	100	100	87	88	91	83	100		

Overall classification accuracy: 88%

Table 20 Error matrix showing the pixel-based crop classification result of the Zuid Flevoland test site for the eigth-date SAR data set. The figures are given in hectares

Classified data	Reference data													Reliability (%)
	potatoes	sugar beets	winter wheat	grass	maize	winter rape	spring barley	orchards	onions	beans	peas	lucerne	total	
Potatoes	927	277	20	13	8	3	0	0	5	4	10	0	1268	73
Sugar beets	445	611	16	14	14	3	0	5	7	2	28	0	1146	53
Winter wheat	271	146	1651	133	21	17	223	13	288	75	81	12	2929	56
Grass	45	24	37	1044	60	10	15	44	21	14	4	22	1339	78
Maize	1	4	2	1	12	0	1	1	2	2	7	0	31	38
Winter rape	207	204	13	20	14	490	5	17	11	5	65	2	1054	47
Spring barley	3	3	6	2	1	6	525	1	10	5	6	1	570	92
Orchards	29	17	3	39	17	4	1	222	10	8	16	18	382	58
Onions	10	10	12	8	4	1	1	1	114	4	24	0	186	61
Beans	160	353	52	76	166	22	24	33	55	241	80	35	1296	19
Peas	0	0	0	0	0	0	0	0	0	0	0	0	1	31
Lucerne	49	45	28	88	12	9	30	17	18	46	46	393	780	50
Total	2146	1693	1839	1436	330	565	824	353	539	407	365	484	10981	
Accuracy (%)	43	36	90	73	4	87	64	63	21	59	0	81		

Overall classification accuracy: 55%

The pixel-based crop classification leads for all crops to significant lower classification accuracies and reliabilities than the field-based classification (Tables 19 and 20). The pixel-based classification resulted in an the overall classification accuracy of 55% (Table 20). The variances in backscatter still present within the individual fields lead to small misclassified clusters of pixels throughout the result of the pixel-based classification.

The result of the multi-temporal pixel-based classification of ERS-1 SAR images is rather poor. However, sometimes the available optical satellite images do not allow some crops to be classified with sufficient accuracy or reliabilty. In these cases mixed classes are included in the LGN database (4.5). These mixed classes could possibly be used for masking ERS-1 SAR images prior to a pixel-based classification. Since only a restricted number of classes have to be distinguished, the classification result is likely to increase. Two examples of mixed classes which occur in the LGN2 database are grass-winterwheat and potato-sugarbeet-maize. These two classes were used to mask the filtered ERS-1 images of the Zuid Flevoland test site. These areas were classified again, but now only for the crops present in the mixed class. The results are presented in Tables 21 and 22.

Table 21 Influence (%) of crop segmentation on the pixel based classification result of the eigth-date ERS-1 SAR data set for potato, sugar beet and maize

Crop type	Classification of all crops		Classification of potato, sugar beet and maize	
	accuracy	reliability	accuracy	reliability
Potato	43	73	56	73
Sugar beet	36	53	70	52
Maize	4	38	25	33

Table 22 Influence (%) of crop segmentation on the pixel based classification result of the eigth-date ERS-1 SAR data set and of the ERS-1 SAR image from 16-8-1992 for winter wheat and grass

Crop type	Eight date data set				Image from 16-8-1992	
	classification of all crops		classification of winter wheat and grass		classification of winter wheat and maize	
	accuracy	reliability	accuracy	reliability	accuracy	reliability
Winter wheat	90	56	91	93	98	83
Grass	73	78	91	89	75	96

The classification result of the masked ERS-1 SAR images for the mixed class potato-sugar beet-maize is rather poor. The backscatter signatures of these crops remain similar throughout the growing season in contrast with the backscatter signature of most of the other crops. Masking of the ERS-1 images for the mixed class grass-winter wheat, however, shows a significant improvement of the classification result. Classification accuracy and reliability for grass increase from 73% to 91% and from 78 to 89%, respectively. For winter wheat classification accuracy and reliability increase from 90% to 91% and from 56 to 93%, respectively. Winter wheat is also one of the few crop types which can be discriminated from the other crops with only one ERS-1 image (Schotten et al., 1995). Therefore, an extra classification was done with one image, acquired on 16-08-1992, for the mixed class grass-winter wheat. This resulted in an reasonable classification result (Table 22). However, this good result may not be expected for all ERS-1 acquisition dates during the growing season. Optimal image acquisition dates for the discrimination between specific crops can be calculated with so called separability indices (Schotten et al., 1995). To determine these optimal acquisition dates, backscatter time series from different years and for different crops have to be analysed.

It can be concluded that ERS-1 imagery cannot be used to improve the operational efficiency of the classification of agricultural crops in the framework of the LGN database. Only in specific cases ERS-1 SAR images could be used complementary to images from optical satellites to improve the classification result of agricultural crops which show spectral confusion in the optical part of the spectrum. However, further research is required to determine the optimal acquisition dates for the discrimination between specific crops. Possible application of ERS-1 images in the production process of the LGN database will strongly depend on the occurring mixed classes, the number of required images and the availability and (processing) cost of the images. In this framework also other available SAR images (JERS-1 and Radarsat) and SAR images which are foreseen with planned missions (Envisat) have to be included in further research.

The classification of buildings situated in the agricultural stratum of the LGN database was troublesome (4.9). Therefore, also the possibility of ERS-1 SAR images to map buildings was investigated. Buildings appear on the SAR images as bright spots. The high backscatter signals are caused by the so-called corner effect of smooth perpendicular surfaces oriented towards the satellite. The incoming radar signal is reflected by corners formed by soil-wall surfaces. The interpretation of ERS-1 images from three different tracks over the Zuid Flevoland test site learned that only a limited number of farms could be distinguished. Only on images of one of the tracks (track 29) a number of buildings could be classified by a simple slicing procedure. On images of the other two tracks (tracks 423 and 151) buildings did not have a higher backscatter than several agricultural fields and thus making a differentiation difficult. The different results of the tracks can be explained by the looking direction of the SAR system and the orientation of most buildings in the test site. Many roads in the test site are Northwest-Southeast or Southwest-Northeast oriented and most of the farms are build with the same orientation along these roads. The looking direction of the SAR in track 29 appears to be perpendicular towards the east-oriented walls of the farms. An implication of this feature for the rest of the Netherlands is that only buildings with a specific orientation (perpendicular towards the looking direction of the radar) can be discriminated. Thus, multi-temporal SAR data coming from different tracks will increase the maximum number of buildings which can be distinguished.

Images of the phase difference between two SAR images acquired from two (nearly) repeated SAR coverages are called interferograms (Van der Kooij et al., 1995). The coherence of the phase of the interferogram is a quality measure for the phase preservation, with time and surface type. Temporal decorrelation can be caused by small changes of location of scatterers within the resolution cell between the passes.

Temporal decorrelation depends strongly on the nature of the scattering object. Water will decorrelate very fast (within approximately 0.1 second). Vegetated areas decorrelate fast especially during the growing season, while bare agricultural fields may show complete loss of coherence by management practices like plowing. On the contrary, solid scatterers like houses do hardly decorrelate for long time intervals. When used in a proper way the coherence can provide valuable additional information on land use (Van der Kooij et al., 1995). It is therefore, interesting to investigate the possibilities of multi-pass inferometry for the classification of buildings in agricultural area.

4.11 Postprocessing

Following the classification, different postprocessing techniques could be applied to further improve classification accuracy.

Classified data often manifest a salt-and-pepper appearance due to the inherent spectral variability encountered by a classifier when applied on a pixel-by-pixel basis (Lillesand and Kiefer, 1994). Majority filtering is often applied to remove this salt-and-pepper appearance, thereby producing a more map-like product. Majority filtering assigns the most frequently occurring class in a N x N pixel (moving) window to the central pixel of the window. Gurney and Townshend (1983), Kenk et al. (1988) and Thunnissen et al. (1992a) showed that simple majority filtering results in significant classification accuracy increase when the pixel is significantly smaller than the size of the objects being classified. However, the majority filter has a number of disadvantages, resulting in a less desirable final product. These undesirable characteristics stem from the lack of control over minimum polygon size and pixel class conversion (Kenk et al., 1988). This results in pixels being converted to very dissimilar classes, the positional shifting of hard natural boundaries and the disappearance of small objects.

To remove noise and to improve the overall classification accuracy it was decided to apply a 3 x 3 pixel majority filter on the output from the automatic classifier. The disappearance of small objects was not considered to be a problem for applications on a regional scale. If required, important classes can be excluded from filtering. In order to prevent pixels from being converted to very dissimilar classes, each stratum was filtered separately.

The influence of mixed classes on the classification result could possibly be decreased by application of a *selective* majority filter. A selective majority filtering involves the assignment of the most frequently occurring pure class in a N x N pixel window to the central pixel of the window labeled with a mixed (agricultural) class in the original classification result. The mixed class code of the 2central pixel of the window can only be changed to the code of one of the constituent classes. This approach is based on the assumption that pixels of a mixed class are likely to be surrounded by correctly classified pixels of pure classes. A selective majority filtering was applied to the classification result of the agricultural crops in the Drenthe test site.

A preliminary data set was used which deviated from the final classification result as presented in Table 14. The reliabilities of the mixed classes in the preliminary dataset were relatively low compared with those in Table 14. The data were filtered three times and different window sizes were applied. Selective majority filtering was found to improve the accuracy of the pure classes (between 4 and 30%). The accuracy improved more as window size increased and the filter was applied more times in succession. The reliability of the pure classes, however, increased only sligthly (up to 2%) or (in most cases) decreased (between 2 and 38%). The decrease in reliability of pure classes was caused by erroneous assignment of mixed classes. The risk of erroneous assignment of mixed classes increased as window size increased, the reliability of the mixed classes decreased and the number of the constituent classes increased. In the case of relative large window sizes, mixed classes were erroneosly assigned to pure classes in neighbouring fields. The reliability of the major mixed class in the dataset used in the selective majority filtering approach was relatively low, which frequently resulted in wrong class assignment. When the number of constituent classes increased, the chance of assigning wrong classes also increased. In fact only one pure pixel in the window could change the class of the central pixel. Therefore, the effect of selective majority filtering could possibly be improved by applying different tresholds, rather than using a simple majority rule.

Evaluation of the results of selective majority filtering in the Drenthe test site showed that in the case of mixed classes, which contain only two crop types and have a relative high reliability (at least 70%), application of a 3 x 3 selective majority filter could effect a considerable improvement of the classification result. The filter should be applied three times in succession. In other cases the result of application of a selective majority filter is doubtful. Because the reliability of the mixed class is of decisive importance for the succes of selective majority filtering, the filter should only be applied after evaluation of the classification result. When a mixed class consists of crops which were harvested on the image acquisition date, the pixels in this class will not be surrounded by correctly classified pixels of a pure class. So in this case no selective majority filter may be applied.

Experience gained during the production of the LGN2 database showed that under specific conditions (e.g. occurrence of large fields which contain clusters of pure classes as well as clusters of mixed classes with high reliability) other window sizes may be applied and mixed classes which contain more than two crop types may be filtered. Application of selective majority filtering in these cases requires competent and experienced inpreters which have knowledge of the area to be classified.

During the production of the LGN2 database in some areas misclassification of small clusters of pixels occurred which did not disappear by application of a simple 3 x 3 majority filter. Some frequently occurring misclassifications were:

- In specific cases mixed pixels were continually erroneously classified as a particular class. For example, mixed pixels, containing deciduous forest and water, were continually classified as coniferous forest.
- Locally, geometric deviations beween different satellite images covering the same area were larger than 1 pixel. Combination of the separately classified images

could result in classification errors, especially along transitions between different land cover types.
- Locally spectral confusion between two agricultural crops resulted in small scattered clusters of wrongly classified pixels.

Most of the above-mentioned type of misclassifications were corrected by application of a 'CLUMP' and 'SIEVE' operation and/or existing or ad-hoc developed filters (Noordman et al., 1996).

In special cases ancillary data, not available before the classification, may be applied to correct misclassifications (postclassification sorting, 3.2.3). Because it is applied after classification, misclassifications can be corrected by recoding, avoiding a time comsuming new classification. In order to improve the classification of the main roads in the LGN2 database, postclassification sorting was applied (Noordman et al., 1996). The polygons of the main roads, including verges and vegetated areas within roundabouts, were derived from the CBS Land Use Database (3.4). The recoding of the classified pixels was not unambiguous and resulted, locally, in a poor discrimination between the road surfaces and the contiguous vegetated areas.

For specific LGN2 classes other data sources were superior to satellite images (e.g. greenhouses, 4.7) or visual interpretation of satellite images was superior to automatic classification (e.g. orchards, 4.8). Classes copied out of external databases or obtained by visual interpretation have to be excluded from majority filtering.

As a final step in the post-classification processing of the LGN2 database a visual check of the classification result was performed. For this purpose the classification result was visually compared with the 1 : 50 000 topographical map and obvious misclassifications were interactively corrected.

4.12 Conclusions

An improved classification methodology for the LGN database has been developed, consisting of the integrated use of satellite images, digital ancillary data, reference data and expert knowledge. The classification method is characterized by a stratified approach, i.e. every stratum is separately classified.

For an optimal classification result the following strata have to be distinguished: agricultural area, urban area, less densely built on area, dry natural area (including forest), wet natural area (including forest) and water. If the stratification is outdated with respect to the acquisition dates of the satellite images the strata can be up-dated by visual interpretation of the satellite images, supported by simultaneous consultation of topographic maps and aerial photographs. The CBS agricultural regions should be used for a further subdivision of the agricultural strata.

In general, mono-temporal classification of Landsat TM images, obtained during the period mid-May to late September, provide good results for most LGN2 classes in the strata urban area, less densely built on area, natural area and forest and water. Although not included in the LGN2 database, larch proves to be accurately classified when an additional image from winter is used. For an accurate classification of most agricultural crops the use of multi-temporal satellite data is required. In order to maximize spectral information, to ensure an optimal selection of training areas and to avoid an excessive number of spectral classes, it is preferred to use (a selection of) the original spectral bands for image classification and to classify each satellite image separately. Subsequently, the classified images can be combined in a GIS to form the final classification result, using conditional 'IF-THEN' statements. The use of a vegetation index (e.g. NDVI) may be useful for the discrimination between bare and vegetated fields, especially in spring. The extent to which use of multi-temporal data improves the classification result is dependent on the cover types involved, crop growth conditions and the spectral resolution, number and acquisition dates of the used satellite images.

When poor spectral separability and cultivation practices such as harvesting may result in classification accuracies and reliabilities below 70%, mixed agricultural classes (e.g. maize/sugar beet) have to be defined to ensure the required minimum classification result. In general, the definition of mixed classes effects a (considerable) improvement of the (overall) classification result.

Due to deviating crop development or the occurrence of particular crop phenological stages considerable spectral confusion may occur. By that automatic classification, which is solely based on the spectral characteristics of individual pixels, will effect poor classification results. Based on field shape, patterns and/or spatial variation in reflectance and the the location amidst other fields, these fields can often be *visually* recognized as separate fields with a particular crop. So in many cases visual interpretation of satellite images is superior to automatic classification. Prior to the decision to use visual interpretation techniques the pro's and cons have te be compared. Factors as time, classification accuracy and the importance of the crop have to be taken into consideration. In practice (e.g. LGN2), visual interpretation often appears to be a valuable tool, complementary to automatic classification. Advanced hardware and software enable the simultaneous interpretation of different satellite images, while the interpretation result can directly be stored in digital form by on screen digitizing.

The classification of greenhouses, orchards and buildings in agricultural area prove to be troublesome. Therefore, special classification techniques were developed for these classes. Greenhouses show a relatively large spectral variability and spectral overlap occurs, especially with bare soil and built-up area. Both automatic and visual classification of greenhouses result, generally, in insufficient classification results. Visual classification is, however, superior to automatic classification. In order to get accurate information on the location of greenhouses, they should be digitized from maps or copied out of existing databases.

Orchards show spectral confusion with grassland and forest. In the beginning of the growing season spectral confusion with grassland dominates, while in the course of the growing season spectral confusion with grassland decreases and spectral confusion with forest increases. The change in spectral confusion is caused by the increase in soil coverage of the orchards during the growing season. Because forest and orchards occur in different strata of the LGN database, the spectral confusion between orchards and grassland constitutes the main problem. Therefore, classification of satellite images obtained late in the growing season provides best classification results. Nevertheless, automatic classification of these images provides insufficient classification result. Visual interpretation of these satellite images, supported by topographic maps, provides sufficient classification results in most areas.

Buildings in agricultural areas can not be sufficiently accurately classified by automatic classification of individual optical satellite images. The poor classification results are mainly caused by spectral confusion with bare soil and ripened cereals. Visual interpretation proves more succesful. Visual interpretation is not only guided by tone but also by size and situation (e.g. with regard to roads) of the buildings. A specific classification method has been developed, using these specific characteristics of buildings. The method exists of the combined use of multi-temporal NDVI images and ancillary data and the application of specific GIS techniques. By applying this method most (large) farms and clusters of small buildings are correctly classified, while scattered small buildings (dwelling-houses, sheds and the like) are only partly correctly classified. Images from summer, when most fields are covered with vegetation, contribute most to the classification result. The backscatter values of ERS-1 SAR images are not suitable for the classification of buildings in agricultural area. It is, however, interesting to investigate the possibilities of ERS multi-pass SAR inferometry for the classification of buildings in agricultural area.

From 1997 LGN classes like greenhouses, orchards and buildings in agricultural area, can simply be copied out of other available geographical databases. However, there may be financial or copyright contraints. Moreover, in practice, data in these geographical databases will often be outdated compared with the acquisition dates of the satellite images. In the latter case, satellite images could be used for up-dating of the concerning LGN classes. Interactive on screen up-dating of the LGN classes greenhouses and orchard can be strongly supported by the simultaneous projection of these classes already present in the LGN database, providing useful additional information on spectral reflectance, shape, context and location of the concerning LGN classes. The up-dating of the greenhouses can be improved by masking the satellite image with a NDVI image from summer, when most plots are covered with vegetation. By that the spectral confusion of greenhouses with bare soil decreases considerably. For purpose of efficiency the up-dating of greenhouses and orchards can be restricted to the municipalities where the change in area exceeds a certain minimum area, according to the CBS Agricultural Statistics.

The ERS-1 SAR, which operates in the microwave part of the spectrum, is not hindered by cloudiness or haze. Especially the multi-temporal classification of agricultural crops in the LGN database requires a regular data acquisition. A *field-based* multi-temporal classification of ERS-1 images provides, in principle, good classification results. A field-based classification requires the availability of digital field boundaries. These data are, however, not available and digitization of field boundaries for large areas is too expensive. A *pixel-based* multi-temporal classification of ERS-1 SAR images leads for all crops to significant lower classification results than the field-based classification. In some cases a pixel-based classification of ERS-1 SAR images could be used complementary to images from optical satellites to improve the classification result of agricultural crops which show spectral confusion in the optical part of the spectrum (i.e. mixed classes). Further research is required to determine the optimal acquisition dates for the discrimination between specific crops. Possible operational application of ERS-1 images in the production process of the LGN database will strongly depend on the occurring mixed agricultural classes, the number of required images and the availability and (processing) cost of the images. In this framework also other available SAR images (JERS-1) and SAR images which are foreseen with planned missions (Radarsat and Envisat) have to be included in further research.

Following the classification, different postprocessing techniques could be applied to further improve the classification result. To remove noise and to improve the overall classification accuracy, a 3 x 3 pixel majority filter has to applied on the output from the automatic classifier. The disappearance of small objects is not considered to be a problem for applications on a regional scale. In order to prevent pixels from being converted to very dissimilar classes, each stratum has to be filtered separately. The influence of mixed classes on the classification result could possibly be decreased by application of a *selective* majority filter. Evaluation of the results of selective majority filtering shows that in the case of mixed classes, which contain only two crop types and have a relative high reliability (at least 70%), application of a 3 x 3 selective majority filter could effect a considerable improvement of the classification result. The filter should be applied three times in succession. Only under specific conditions other window sizes may be applied and also mixed classes which contain more than two crop types may be filtered. Application of selective majority filtering in these cases requires competent and experienced inpreters which have knowledge of the area to be classified.

During the production of the LGN2 database in some areas misclassification of small clusters of pixels occurred which did not disappear by application of a simple 3 x 3 majority filter. Most of these misclassifications could be corrected by application of a 'CLUMP' and 'SIEVE' operation and/or existing or ad hoc developed filters.

In special cases ancillary data, not available before the classification, may be applied to correct misclassifications (postclassification sorting).

Classes copied out of external databases or obtained by visual interpretation of satellite images have to be included in the final LGN database without application of the 3 x 3 majority filter.

As a final step in the post-classification processing it is advised to perform a visual check of the classification result. In this way obvious classification errors can be interactively corrected.

5 Collection of reference data and classification accuracy assessment procedure

With a view to the validity of the decisions based on the information in the LGN database it is of great importance that sufficient reference data are obtained for assessing the accuracy of the results obtained. This chapter discusses some considerations and techniques for reference data collection and classification accuracy asessment, especially sampling scheme (5.1), sample size (5.2) and sample evaluation procedures (5.3). Because of practical considerations adapted reference data collection and accuracy assessment procedures have been set up for the LGN database (5.4).

Congalton (1991) distinguishes 'site-specific' and 'non-site-specific accuracy assessment'. Site-specific accuracy assessment means that the locational and classification accuracy are both assessed together. Non-site-specific accuracy assessment applies only to the classification accuracy while ignoring locational accuracy. In this report 'classification accuracy' means 'site-specific accuracy'. When 'non-site-specific accuracy assessment' is discussed it will be explictely mentioned. The non-site-specific accuracy assessment applied in the LGN project will be discussed in Section 5.5.

5.1 Sample scheme

The most usual way of assessing the accuracy of a per-pixel classified remote sensing images is by selection of a sample of pixels from the classified image and checking their labels against classes determined from reference data. Sampling may be selected on the basis of a variety of sampling schemes. To ensure that the results are an accurate reflection of the performance of the classifier, a random distribution of sampling points over the whole stratum must be sought (Van Genderen et al., 1978). A difficulty that can arise with simple random sampling is that it is area-weighted (Richards, 1986 and Congalton, 1991). That is, large classes tend to be represented by a larger number of sample points than the smaller classes (e.g. Table 4); indeed some very small classes may not be represented at all. To avoid this it is necessary to ensure small classes are represented adequately. For this reason stratified random sampling is often applied where a minimum number of samples are selected from each strata. The most appropriate stratification to use is the actual thematic classes themselves. Consequently, the user should choose a random sample within each thematic class to asses the classification accuracy of that class. That means that the reference data to be used for the accuracy assessment can only be gathered after the classification has been performed in stead of in conjunction with the training data collection. Moreover, the gathering of reference data can cause temporal problems for classes that change quickly in time (e.g. agricultural crops). In these cases, only a simple random sampling can be applied.

To make the ground collection effort more efficient, a systematic sampling could be performed. Systematic sampling is a method of sampling in which the sampling units (pixels) are selected at some equal interval in space. The advantage of systematic sampling is that it is easier to locate the samples on the ground. A major disadvantage of systematic sampling is that if the population contains some periodicity, then the regular spacing of the sample units might result in unrepresentative samples (Congalton, 1988a). By application of spatial autocorrelation analysis of the patterns of error in the classification results of different test sites Congalton (1988b) showed that for the agriculural and range test sites systematic sampling overestimated the population statistics (i.e. mean and variance), while in the forest test site systematic sampling adequately estimated population statistics. These differences have to do with the complexity of the areas. The pattern of error in the forest classification result was more complex and linear than the more simple, blocky pattern found in the range and especially the agricultural classification results. This was caused by the periodicity in the error patterns, which was pronounced in the agricultural area, slightly less in the range area and hardly present in the forest area. Congalton concludes that 'depending on the complexity of the area as determined by spatial autocorrelation analysis, systematic sampling may yield adequate results. However, periodicity in remotely sensed data due to positive correlation between errors could result in a poor estimate of the population parameters ...'. Therefore, 'systematic sampling should be used only with extreme caution'. However, it must be mentioned that the overestimation of the mean (estimate of classification accuracy) as found by Congalton in the agricultural study area amounted 2.5% maximum.

In many circumstances, simple random and even stratified random sampling and systematic sampling pose serious problems as the selected pixels are not easily accessible for field verification, making the sampling expensive and time consuming. Under such conditions, (random) cluster sampling (i.e. sampling groups of neighbouring pixels) has been frequently selected as the sampling strategy (Congalton, 1988a), especially to collect information on many pixels very quickly. It is much easier and cheaper to visit a few large areas than to visit many smaller areas. A problem related to cluster sampling lies in the spatial autocorrelation of neighbouring pixels because each pixel is not independent of the other. From correlation analysis, Congalton (1988a) concluded that, in order to maximize the information derived from cluster sampling, small custers should be taken, using no more than 10, or at most 25 pixels per cluster.

5.2 Sample size

To ensure that the results of the accuracy assessment are an accurate reflection of the performance of the classifier, in addition to a random distribution of sampling points over the whole stratum a minimum sample size for each class is required. Van Genderen et al. (1978) and Rosenfield et al. (1982) have adressed this problem, using binomial statistics. Commencing from different view points their approaches lead to different minimum sample sizes. Rosenfield et al. are interested in ensuring that the accuracy indicated from the samples (i.e. sample mean) is a reasonable (constant) approximation of the actual map accuracy. In contrast, Van Genderen et al. base their approach on ensuring that the set of samples is representative. Both have their merits and in practice one may wish to choose a compromise of between 30 and 60 samples per category (Richards, 1986). However, Congalton (1991) states that 'because of the large number of pixels in a remotely sensed image, traditional thinking about sampling does not often apply. (......) Therefore, practical considerations more often dictate the sample size selection. A balance between what is statistically sound and what is practically attainable must be found'. From experience Congalton recommends as a 'good rule of thumb' the collection of a minimum of 50 samples for each land cover class in the error matrix. For especially large areas (i.e. more than 4000 km^2 or a large number of land cover classes (i.e. more than 12 classes) he recommends to increase the minimum number of samples to 75 or 100 samples per land cover class. It may be useful to adjust the number of samples for each land cover class based on the relative importance of that class within the objectives of the classification or by the inherent variability within each of the land cover classes.

5.3 Sample evaluation

Once the sampling has been performed the classification accuracy has to be evaluated. The most common way to represent the classification accuracy is in the form of an error matrix, as described in 4.3. The error matrix can be statistically evaluated. If appropriate sampling is performed confidence limits can be determined, based on the binomial distribution. Binomial probabilities may be calculated for the classification as a whole or for individual classes. Hord and Brooner (1976) and Richards (1986) describe a procedure by which the 95 percent confidence interval can be derived, given the sample size and the number of correctly classified pixels in a sample. The most common method to derive sample estimates from systematic sampling, is to assume that the sample is random and use the appropriate equations from simple random sampling (Congalton, 1988). Congalton (1988) showed that, if the situation dictates it, (random) cluster sampling can also be used for estimation of the classification accuracy. However, small clusters should be taken using no more than 10 pixels per cluster or 25 pixels per cluster, maximum.

5.4 Sampling schemes and classification accuracy assessment procedure in the LGN project

For the assessment of the classification accuracy of the LGN database a distinction was made between the non-agricultural classes and the agricultural classes. In order to create un unbiased and statistical valid sample of pixels, simple random sampling or stratified random sampling schemes are preferred (5.1). However, due to time and budget constraints it has been decided to base the accuracy assessment of the *non-agricultural classes* on a systematic sampling scheme. The (potential) sampling points coincide with the points of intersection of the 1 km grid lines on the Dutch topographic maps. The use of these points facilitate the location of the sampling points. To avoid oversampling a stratified systematic sampling was performed. The stratification is based on (groups of) the actual thematic classes themselves. The sampling is concentrated on the most important (groups of) LGN2 land cover classes, namely (the LGN2 class codes are indicated between brackets): deciduous forest (2.1), coniferous forest (2.2), open vegetated nature area (3.1 and 3.2), built-up area (5.1 and 5.2) and green urban area (5.3 up to and including 5.6). From all the points of intersection of the 1 km grid lines situated in the different (groups of) LGN2 classes 100 samples were selected randomly. The reference classes for the selected sampling points was derived from topographic maps and, if necessary, from aerial photographs. Because of the positional accuracy of the pixels in a geometrically corrected image (the RMS error can be as large as 1 à 2 pixel, section 4.3), it is difficult to determine the reference class if a selected pixel is situated in a heterogeneous area. In these cases the reference class represents the class that covers the majority of the area in a 3x3 pixel window. This is allowable because the LGN2 classification result was spatially smoothed using a 3x3 majority filter.

The (groups of) non-agricultural LGN2 classes to be sampled are rather complex and it is assumed that no periodicity in the data occurs. Therefore, the population estimates were obtained by assuming simple random sampling and using the appropriate equations.

The accuracy assessment of the land cover classes which are less important and/or show little variability (i.e. bare soil in nature area (3.3), water (4.1 and 4.2) and bare soil in rural built-up areas (5.7)) and the land cover classes that are (partly) copied out of other high quality databases (i.e. greenhouses (1.8), orchards (1.9), heath land (3.1) and main roads and railways (5.8)) was only performed qualitatively and/or in a non-site-specific way (section 5.5).

For the assessment of the classification accuracy of the *agricultural classes* reference data should be gathered in the field. In practice there is often a time lapse between acquisition date and classification date, hampering the execution of a stratified sampling as it can not be based on the classification result. So, in these cases only a simple random or systematic sampling has to be performed, requiring a large sample size to ensure an adequate sample size for all the classes concerned. Another problem, related to the time lapse between acquisiton date and classification date concerns the dynamic land use change in agricultural areas. After the selection of suitable satellite images, often a restricted time period is available for gathering reference data of agricultural crops. Because a stratified classification of the agricultural area is performed (4.2), every stratum has to be classified seperately. Totally, circa 70 agricultural strata were distinguished. The sampling effort can be adjusted based on the homogeinity of the strata. Only heterogeneous strata require a comprehensive sampling. Nevertheless, the gathering of a statistical valid sample of pixels for the agricultural crops in the entire Netherlands is estimated to take 150 à 200 working days. A random cluster sampling may only slightly reduce the sampling effort. So due to time and money constraints a simple random or systematic sampling and even a random cluster sampling per stratum is not feasible. Therefore, in the framework of the LGN database a 'controlled' cluster sampling is proposed.

The 'controlled' cluster sampling is a sampling method in which all (agricultural) plots bordering on a number of selected sections of roads are sampled. For all agricultural strata a number of outwardly representative sections of roads are selected on the basis of topographic maps and satellite images. Topographic maps provide useful information on land cover (grasland or arable land), the spatial distribution of land cover and plot size, while satellite images provide more detailed in formation on the occurring agricultural crops. For the main agricultural crops in an stratum at least 10 plots have to be sampled. If the mininimum number of plots are not present along the selected sections of roads, additional samples have to be collected for the concerning crops. The sampled pixels used for the accuracy assessment consist of clusters of 3x3 pixels in the centre of the plots, minimizing the intracluster correlation and maximizing the information derived from cluster sampling (Congalton, 1988). For small plots only the centre pixel is sampled. In addition to accuracy assessment a part of the reference data is also used for training of the classifier. The proposed sampling procedure has some main disadvantages:

- The sampling procedure does not guarantee the samples to be truly representative of the classification result. Therefore, the resulting error matrices provide no statistically valid estimates of classification performance and no confidence limits can be calculated;
- A part of the sampled fields is often used to get information on the reflection characteristics of the different agricultural crops. Subsequently, large homogeneous fields are selected on the screen by the interpreter for training the classifier. In practice, mostly only a few sampled fields prove to be suitable training areas. Moreover, many satellite images cover several strata. In these cases the selected training are often transferred to other strata, requiring no or only minor adaptions. Therefore, it is believed that the use of a small part of the reference data for training purposes effects no significant overestimation of classification performance. Nevertheless, it is thought desirably to sample in future some large, homogeneous plots exclusive for training of the classifier, complementary to the controlled cluster sampling;
- The clusters of pixels or individual pixels selected for the assessment of the classification accuracy of agricultural crops represent areas of homogeneous land cover. This leads to the problem that those areas which represent mixtures of different agricultural crops will not be represented by the sample. In consequence, accuracy values based on such samples must be expected to be significantly higher than the true accuracy of the classified image (Corves and Place, 1994). However, accuracy assessment of the LGN1 database by comparison of the classification result with gridded reference areas on a pixel by pixel basis showed that (mixed) pixels along field boundaries may wrongly be considered in error (Thunnissen et al., 1992a en 1992b). This is caused by positional errors occurring from inaccuracies in the position of boundary pixels in the classified image and the gridded reference areas. These inaccuracies in the position of boundary pixels are caused by geometric correction of satellite images, delineation of boundaries during field survey or by interpretation of aerial photographs, digitization of reference maps and vector to raster conversion of the reference maps.

  Mixed (boundary) pixels may be classified as one of the constituent classes or as one of the remaining classes. When mixed pixels are classified as one of the constituent classes the positional inaccuracies only cause a slight shift in boundary position. These location errors are of minor importance for applications on a regional scale. Therefore, boundary pixels classified as one of the constituent classes are considered to be correctly classified. Boundary pixels classified as one of the remaining classes are erroneously classified. Separate classification accuracy assessment of boundary pixels in the LGN1 database showed that in most reference areas most of the boundary pixels were correctly classified as one of the constituents classes (Thunnissen et al., 1992a en 1992b). Visual inspection of the LGN2 database showed that also in mostly all agricultural strata most boundary pixels were correctly classified. Therefore, omitting boundary pixels from the sampled clusters is not expected to cause significant overestimation of classification performance. However, when locally many boundary pixels are misclassified, omitting boundary pixels from the sampled clusters must indeed be expected to significantly overestimate classification performance.

- In the controlled cluster sampling approach only agricultural fields are sampled. However, the agricultural strata contain also minor roads, ditches, farms and farm yards. These classes are classified as one of the agricultural crops. By that, the reliability of the classified agricultural crops is underestimated.

In the Drenthe test site a controlled cluster sampling was performed (2.1). The results of the classification accuracy assessment, using the reference data from the controlled cluster sampling, are presented in Table 23. Comparison with the results of the accuracy assessment performed with the reference data from the systematic sampling (Table 14) shows that individual classes show some deviations. However, the overall classification accuracies of both approaches are comparable. To better found the reliability of the controlled cluster sampling, it is advised to compare the results of the controlled cluster sampling with the results of the random or systematic sampling in some other test sites. However, using the LGN land cover database in practice, we must realize that the results of the classification accuracy assessments are only statistical valid as the error matrices are truly representative of the entire classification. Because, especially the sampling of the agricultural crops is not based on a (stratified) random sampling approach, the results must be used with caution. Nevertheless, we think that the results of the accuracy assessments provide a fair indication of the quality of the classified images.

For LGN2 database the controlled cluster sampling was only performed in a restricted number of strata (Noordman et al., 1996). This is mainly caused by the restricted time and man-power available for the field survey. The actual collection of field reference data should be incorporated in the operational implementation of the LGN database (6.2).

Finally, it should be remarked that when performing a classification accuracy assessment, it is often implicitely assumed that the reference data are 100% correct. In practice this assumption is not true.

In addition to the accuracy assessment on the basis of the site-specific controlled cluster sampling also a non-site-specific accuracy assessment of the agricultural crops was performed. The latter method will be discussed in 5.5.

Table 23 Error matrix showing the multi-temporal classification result for the agricultural crops in the Drenthe test site. The classification accuracy assessment was performed by using the reference data from the controlled cluster sampling. Numbers in the matrix express numbers of pixels

Classified data	Reference data							Reliability (%)
	grass	maize	potatoes	sugar beets	cereals	other agricultu-ral crops	total	
Grass	968	77	168	44	7	11	1275	75.9
Maize	4	85	6	14	0	6	115	74.0
Potatoes	71	20	387	52	19	16	565	68.5
Sugar beets	0	12	10	159	4	2	187	85.0
Cereals	59	8	23	2	185	31	308	60.1
Cereals/pota-toes	77	6	49	4	49	35	220	44.5
Maize/sugar beets	2	100	3	75	0	4	184	95.1
Maize/pota-toes/sugar beets	75	97	66	29	10	25	302	63.6
Total	1256	405	712	379	274	130	3156	
Accuracy (%)	77.1	21.0	54.4	42.0	67.5	0.0		
Accuracy inclusive of mixed classes (%)	77.1	69.6	70.5	69.4	85.4	0.0		

Overall classification accuracy: 56.5%
Overall classification accuracy (inclusive of mixed classes): 71.3%

5.5 Non-site-specific accuracy assessment

The CBS Agricultural Statistics contain information on the areas of the agricultural crops for each of the 66 'agricultural regions' in the Netherlands (3.6). Agricultural regions are more or less homogeneous areas as far as soil type and agricultural land use are concerned. By comparing the classified areas in the LGN2 database with the areas provided by the CBS Agricultural Statistics a non-site-specific accuracy assessment of the LGN2 database was performed. The CBS Agricultural Statistics, however, contain net cultivated areas, while the agricultural stratum in the LGN2 database contains the total agricultutal area inclusive of ditches, (minor) roads, hedges, farm yards, farms and other buildings. A part of this agricultural infrastructure was included in the neighbouring fields during the classification or by majority filtering of the classification result. The remaining part of the agricultural infrastructure was mainly classified as grassland or bare soil. This implies that especially the grassland and bare soil areas are overestimated in the LGN2 database. The part of the agricultural infrastructure, classified as bare soil, was assigned to the arable class(es) which contained a considerable area bare soil on the image acquisition date. The accuracy of the LGN database could be improved considerably in future by copying the agricultural infrastructure out of other available digital geographical databases.

5.6 Conclusions

In order to create un unbiased and statistical valid sample of pixels, simple random sampling or stratified random sampling schemes are preferred. However, due to time and budget constraints it has been decided to base the accuracy assessment of the non-agricultural classes in the LGN database on a stratified systematic sampling scheme. For the concerning classes the results of this sampling scheme are comparable with the results of a random sampling scheme. By selecting the points of intersection of the 1 km grid lines on the Dutch topographic maps as (potential) sampling points, the location of the sampling points is facilitated considerably. By reason of efficiency, the sampling is concentrated on the most important (groups of) LGN2 land cover classes. From all the points of intersection of the 1 km grid lines situated in the different (groups of) LGN2 classes 100 samples are selected randomly. The reference classes for the selected sampling points can be derived from topographic maps and, if necessary, from aerial photographs. The accuracy assessment of the land cover classes that are less important and/or show little variability or are (partly) copied out of other high quality databases, is only performed qualitatively and/or in a non-site-specific way.

The gathering of a statistical valid sample of pixels for the agricultural crops in the entire Netherlands is estimated to take 150 à 200 working days. So due to time and money constraints a simple random or systematic sampling and even a random cluster sampling per stratum is not feasible. Therefore, in the framework of the LGN database a 'controlled' cluster sampling is proposed. The 'controlled' cluster sampling is a sampling method in which all (agricultural) plots bordering on a number of selected sections of roads are sampled. For all agricultural strata a number of outwardly representative sections of roads are selected on the basis of topographic maps and satellite images. For the main agricultural crops in an stratum at least 10 plots have to be sampled. The sampled pixels used for the accuracy assessment consist of clusters of 3x3 pixels in the centre of the plots in order to maximize the information derived from cluster sampling. For small plots only the centre pixel is sampled. Comparison of the accuracy assessments of the classification result of the Drenthe test site, performed with reference data from both a controlled cluster sampling and a systematic sampling, shows that individual classes may show some deviations. However, the overall classification accuracies of both approaches are comparable. To better found the reliability of the controlled cluster sampling, it is advised to compare the results of the controlled cluster sampling with the results of the random or systematic sampling in some other test sites.

Using the LGN land cover database in practice, we must realize that the results of the classification accuracy assessments are only statistical valid as the error matrices are truly representative of the entire classification. Because, especially the sampling of the agricultural crops is not based on a (stratified) random sampling approach, the results must be used with caution. Nevertheless, we think that the results of the accuracy assessments provide a fair indication of the quality of the classified images.

The CBS Agricultural Statistics are suitable for a non-site-specific accuracy assessment of the LGN2 database. One should, however, take into consideration that the

CBS Agricultural Statistics contain net cultivated areas, while the agricultural stratum in the LGN2 database contains the total agricultural area inclusive of ditches, (minor) roads, hedges, farm yards, farms and other buildings. The accuracy of the LGN database could be improved considerably in future by copying the agricultural infrastructure out of other available digital geographical databases.

6 Application, cost and benefit analysis and operational implementation of the LGN database

6.1 Applications

The main users of the LGN(2) database are national and regional governmental agencies. Because of its digital format the LGN database can be easily combined with other digital information. Is has frequently been used for different purposes in the fields of environmental protection, water management, nature conservation and physical planning on regional and national scales. Mostly, the LGN data are combined with other geographical information, such as soil type, water-table, the occurrence of seepage, meteorological data, and the application of animal manure, fertilizers and pesticides. Underneath some main applications of the LGN database are summarized.

Physical planning
The LGN database can provide an overview of the land cover in a specific area, in contrast to the CBS Agricultural Statistics, which are available for administrative units only and therefore have a low spatial resolution. The LGN database was used to allocate soil and groundwater protection areas, to design monitoring networks for soil and groundwater quality and to determine the relation between land use/cover and soil and groundwater quality (e.g. TAUW, 1994 and Van Drecht et al., 1996). LGN has also be used for an inventory of land use in nature development and groundwater extraction areas.

The DLO Winand Staring Centre (SC-DLO) has developed a spatial decision-support system for physical planning, called WATRO (Steenvoorden et al., 1993). This system contains information on land use (LGN) and soil and groundwater characteristics. Using WATRO, land use planners can assess the effects of different land use scenario's on the environment or investigate the suitability of an area for a specific type of land use.

Extension of built-up areas threatens the open space in rural areas (e.g. Ministery of Agriculture, Nature Management and Fisheries, 1995). To support physical planning policy, it is therefore of great importance to have up-to-date information on urbanization. This information can be derived from the LGN database.

Environmental protection
The effect of (proposed) measures on the quantity and quality of groundwater and surface water are studied regularly, using regional water and solute transport models. In these models, evapotranspiration, groundwater recharge and irrigation are often related to land cover/use, just like nutrient loads and uptake, and the distribution of nutrient applications and uptake over the year. Atmospheric deposition and the application of pesticides are also often related to land cover. The required information

on land cover/use for these models can be derived from the LGN database (e.g. Schouwmans et al., 1990; Thunnissen et al., 1992a, Reijerink and Breeuwsma, 1992; Vermulst, 1993; Querner, 1993; Van Walsum et.al., 1996; Van Walsum and Veldhuizen, 1996; Van der Bolt et al, 1994; Meinardi, 1994 and Gehrels, 1995).

Wopereis (1991) studied the suitability of soils for low-emission manure application on grassland in sandy areas.

For national and regional policies in the fields of water management and the environment, information on land use is indispensable. For the underlying environmental and water system studies, the LGN database was used (e.g. Nationale Milieuverkenningen (National Institute of Public Health and environmental Protection, 1993), de Derde Nota Waterhuishouding (Ministery of Transport and Public Works, 1989) en provinciale Milieu- en waterhuishoudingsplannen (e.g. Provincie Gelderland, 1994)

In the early eighties, pesticides were found in drinking-water extracted in the Drentsche Aa catchment area (province of Drenthe). The main causes of these contaminations were spray drift, surface run-off, and spilling when water was being pumped for spraying. To prevent this, alternative filling places were constructed so that farmers need not use streams and ditches for their water supply. Because pesticides are mainly used on arable land, the sites of these filling places were selected, using the LGN database in order to minimize the distance between the filling places and the fields.

Water management

SC-DLO has made a cost-benefit analysis of several promising combinations of measures to prevent soil erosion and flooding in the southern part of the Province of Limburg. Information on slope, land cover (LGN) and soil type has been taken into account (Van Eck et al., 1995).

The Directorate-General of Public Works and Water Management and the Provinces of Gelderland and Overijssel are developing inundation models for the major rivers of the Netherlands. An important input parameter in these models is the roughness coefficient. The spatial distribution of the roughness coefficient can be calculated, using data from the LGN database.

Mobile telecommunication

The planning of a national mobile telecommunication network is based on predictions of signal propagation. The signal propagation is influenced by surface roughness, which can be estimated on the basis of land cover. The LGN database is used for the planning of a telephone network.

6.2 Cost and benefit analysis and implementation of the LGN database

6.2.1 Cost and earnings

The cost of production of the LGN2 database amounted to Dfl 1 120 000 (i.e. circa Dfl 0.35 per ha). The selling price of the LGN2 database has been determined on the basis of the cost, the expected number of users and negotiations with (potential) users. The selling price of the LGN2 database depends on the area required, the number of classes, the spatial resolution and the number of applications. The earnings from sale of the LGN2 database amount to circa Dfl 1 170 000, inclusive of warrants at Dfl 350 000 (situation June 1996).

6.2.2 Benefits

The LGN database is used for many applications (6.1). According to some users, the benefits of using the LGN database exceed the cost. In practice, the LGN database, once being available, appears often to be used for all kinds of unintended applications. An accurate estimation of the benefits, however, is troublesome. Information from the LGN database is mainly used by governmental agencies which are responsible for policies in the fields of environmental protection, water management, nature conservation and physical planning on regional and national scales. The effects of using the LGN database on the quality of the pursued policy are difficult to assess.

In addition to the above-mentioned users, different drinking water companies and water boards are interested in the data. However, purchase of the data is often hampered by the high selling price. Finally, there is much interest in digital land cover data for research and educational purposes. In practice, however, hardly any budget is available for purchase of data. Therefore, many subareas of the LGN2 database have been supplied free of charge for research and educational purposes.

6.2.3 Operational and commercial implementation

Operational implementation of the LGN database has been achieved when the database or parts of the database are up-dated at regular intervals and the up-dating is largely paid by the users of the database. Commercial implementation implies that the up-dating is completely paid by the users. The cost of up-dating the whole LGN2 database is estimated at Dfl 900 000. At this moment operational implementation of the LGN database has been achieved. Meanwhile the production of the LGN3 (!) database has already been started. Commercial implementation of the LGN database has not yet been achieved.

The chance of commercial implementation of the LGN database in near future is determined by the need of land cover data and the cost of gathering these data. In this framework the LGN database has to compete with other available nation-wide digital land cover/use databases, especially the topographic databases (3.5) and the CBS Land Use Database (3.4). The LGN database can be competitive when the database distinguishes itself from the other databases by cost, thematic classes, classification accuracy and reliabilitiy, geometric accuracy, up-date frequency, timeliness or data processing. These factors will be discussed separately:

Cost (all prices mentioned are exclusive of VAT)
Both the LGN database and the 1 : 50 000 and 1 : 10 000 topographic databases are sold against market prices, while the CBS Land Use Database is sold below market price. The *annual* price of the topographic databases (inclusive of up-dates) depends on the up-date frequency and varies between Dfl 840 and Dfl 1 680 for one 1 : 50 000 map sheet (i.e. 500 km^2) and between Dfl 550 and Dfl 1 110 for one 1 : 10 000 map sheet (i.e. 62.5 km^2). The prices of the 1 : 50 000 and 1 : 10 000 topographic databases for the entire Netherlands amount to circa Dfl 112 500 and Dfl 455 500 per annum, respectively. These prices are inclusive of up-dates. The annual prices are based on the most frequent applications of the databases. For more or less intensive applications higher or lower prices may hold. The prices are constant per unit of area.

The price of the LGN2 database for the entire Netherlands amounts to Dfl 200 000. The price per unit of area increases as the required area decreases. For example, the average price of the LGN2 database for one province amounts to circa Dfl 75 000. The prices of up-dated versions of the LGN2 database are estimated at 70 à 80 per cent of the price of the LGN2 database. However, when optimal satellite images are available and no mixed classes have to be included in the LGN database, the prices can be reduced further. The price of the LGN database decreases as the number of required classes decreases and/or raster size increases. Also the number of applications may influence the price. By applying this flexible price policy, it is aimed as much as possible that application of the database is not hampered by too high cost. Nevertheless, for many (potential) users the cost are too high. Although the database, once being available, is used for all kinds of unintended applications, the cost have, mostly, to be paid from the budget of only one or two projects. In practice, this often results in long, internal decision-making. Moreover, in practice, many (potential) users prove to be unacquainted with the LGN database.

The price of the CBS Land Use Database for the entire country amounts to circa Dfl 17 000. The price is likely to rise in future.

Thematic classes
Underneath, it is briefly discussed to what extent the LGN2 classes differ from or correspond with classes in the CBS Land Use Database and the topographic databases.

The CBS Land Use Database distinghuishes 33 land use classes, especially for *artificial, non-agricultural areas* (Table 6). The classification is largely based on functional land *use* and the separate classes may contain several land *cover* types. For instance 'Parks and public gardens' consist mainly of grass and forest, and gardens and public greens, situated in residential areas, are assigned to the class 'Residential areas'. The topographic databases contain topographic point, line and area information. Especially some area elements in these databases, for example grassland and forest, are characterized by land *cover* and the functional *use* of these classes can only be determined by using contextual information, e.g. grassland and forest situated in urban area (parks or sport grounds), grassland used for recreational purposes and such-like. For the main land *use* types of the LGN2 database (i.e. 'built-up area', 'natural area', 'forest' and 'agricultural area'), the database contains information on land *cover*. So, for example, grassland used for recreational or sport purposes is discriminated from pastures.

In the CBS Land Use Database forests are labeled as woodland. Forested areas in the topographic databases are subdivided in polygons bounded by linear elements, mostly footpaths. Deciduous and coniferous forest are distinghuished. When both deciduous and coniferous forest occur within the same polygon it is labeled as mixed forest. Because the LGN2 database contains information per *pixel*, it can distinguish deciduous and coniferous forest within separate polygons in the CBS Land Use Database and the topographic databases.

In the CBS Land Use Database two agricultural classes are distinghuished: 'glasshouses and 'other agricultural use'. The topographic databases distinguish five agricultural classes: arable land, grassland, orchards, tree nurseries and glasshouses. Finally, the LGN2 database contains 10 agricultural classes: the main agricultural crops (Table 5) and glasshouses.

The LGN2 database as well as the CBS Land Use Database and the topographic databases contain little information on land cover in open natural areas.

Scattered buildings in agricultural areas are only included in the topographic databases. This class will, however, be included in up-dated versions of the LGN2 database, using a newly developed method to derive this class from satellite images (4.9).

Railways and the main roads and watercourses are included in the CBS Land Use (3.4) with their real width. The width of roads, railways and watercourses in the topographic databases do not correspond with their actual width. The railways and the main roads and watercourses in the LGN2 database were for the greater part copied out of the CBS Land Use database and up-dated by visual interpretation of satellite images.

Classification accuracy and reliabilitiy

Altough no exact figures are known on the classification accuracy and reliability of the topographic databases and the CBS Land Use Database they are supposed to be high. Nevertheless, locally edge matching problems and label errors were found in the CBS Land Use Database. So far little experience has been gained by using the topographic databases. As far as the LGN2 database is concerned the classification accuracy and reliability of the non-agricultural classes vary between 80 and 95%, while most values exceed 90% (4.4 and Noordman et al., 1996). Spectral confusion occurs only between different land cover classes within the same stratum (e.g. deciduous and coniferous forest). The classification accuracy and reliability of the agricultural crops, inclusive of mixed classes, are, in general, above 70% (Noordman et al., 1996). The classification accuracy of greenhouses and orchards in the LGN2 database, finally, are estimated at 90 and 70%, respectively (Noordman et al., 1996).

Geometric accuracy

Both the topographic databases and the CBS Land Use Database are based on the interpretation of aerial photographs (on scales 1 : 17 000 and 1 : 10 000, respectively), while the LGN database is based on the interpretation of satellite images. Therefore, the topographic databases and the CBS Land Use Database are superior to the LGN database as far as geometric accuracy is concerned. On the average, the geometric accuracy of the LGN database amounts to circa ±25m, which is sufficient for applications on regional scale.

Up-date frequency

The up-date interval of the topographic data varies between 4 and 8 years. From 1-1-1998 the up-date interval will amount to 4 years for all map sheets. The CBS Land Use Database will be up-dated every 3 years. Most users of the LGN database prefer an up-date interval of 5 years. The up-dating of the LGN database is complicated by the need of satellite images acquired in specific periods (4.6). Because of cloudiness, optimal satellite images will not be available every year. An analysis of available Landsat images (Landsat 1 to 5) of the Netherlands during the period 1975-1988 showed that the average frequency of obtaining a cloud-free summer image of the entire country is about once per two years (Van der Laan, 1989). A cloud-free summer images can exist of imagery from different acquisition dates. Moreover, some of the analysed satellites had overlapping acquisition periods. The simultaneous availability of spring and summer images was not analysed. By experience gained during the period 1986-1995 the average frequency of obtaining suitable Landsat 5 images from spring and summer is estimated at once every 3 à 4 years. For the classification of particular classes (e.g. extension of towns and mineral extraction sites) monotemporal images are sufficient and the acquisition period is less critical. Therefore, these classes can be up-dated more frequently. Future satellites, especially SPOT 4 and Landsat 7, are likely to increase the availability of images of suitable spatial and spectral resolution (4.6).

Timeliness

Timeliness refers to the period required for delivering of products. When sufficient equipment and trained staff are available, the classification result can, in principle, be delivered within a few months after the acquisition date of the satellite images. A fast on line availability of high quality quick look data is a prerequisite for a timely delivery of classification results (4.6). On the contrary, the interpretation of aerial photographs can take a few years. Both the topographic databases and the CBS Land Use Database are based on the interpretation of aerial photographs.

Data processing

The topographic databases and CBS Land Use database are stored in vector format, while the LGN database is stored in raster format. The 1 : 10 000 topographic database, the CBS Land Use database and the LGN database occupy circa 20 Gb, 500 Mb and 20 Mb of disk space, respectively, for the entire Netherlands. Especially the processing of the topographic databases and, to a less extent, the CBS Land Use database for relatively large areas may be very time consuming and requires large amounts of disk space.

6.2.4 Conclusions

It can be concluded that (a part of) the LGN database may distinghuish itself favourably from the CBS Land Use and/or the topographic databases, especially concerning cost, thematic classes, timeliness and dataprocessing. That means that the mentioned databases do not replace each other. Altough there may be some overlap, the databases largely supplement each other. In practice, combined use of different databases may give a surplusvalue to the separate databases. In order to achieve commercial implementation or continue operational implementation of the LGN database, the advantages of the LGN database with respect to other digital geographical databases should be exploited as much as possible. Further, in order to meet customers' needs more satisfactorily, the LGN database should become a more flexible product and the classification accuracy and reliabilitiy of some classes should be improved. Finally, more users should be found. In order to reach these objectives, the following activities have to be performed in near future:

- Setting up a subscription system. Participants will be provided with an up-dated version of (some areas or classes of) the LGN database at more or less regular intervals on payment of a fixed price each year. A subsciption system has advantages for both users and producers of the database. The users pay a rather small amount each year, facilitating the financing. The producers are assured of continuity of revenues, facilitating the work planning, especially the acquisition of satellite images (including SPOT Programming Requests), the gathering of field reference data, the appointment of skilled employees and the delivery of timely products;
- Defining a more flexible product by variation of up-date intervals for different classes and/or areas, dependent on the needs of the users, the rate of change and the importance of the concerning classes and the availability of suitable satellite images;
- Improving the classification results by using only images from optimal acquisition periods and with the required spectral resolution. Mixed classes have to be avoided;
- The classification result of agricultural crops could be improved considerably by

performing a field-based classification. By that, the usefulness of the LGN database in projects on local to regional scale would increase. In principle, combined use of the 1 : 10 000 topographic database and satellite images may provide field boundaries to be applied in a field-based classification. It is advised to perform further research into algorithmes to obtain the required information for a field-based classification in an operational and cost effective way;
- The recently available LGN2 database will be used for a comprehensive marketing campaign;
- The production cost should be reduced further.

References

Bolt, F.J.E. van der, P. Groenendijk en H.P. Oosterom, 1994. Nutriëntenbelasting van grond- en oppervlaktewater in de stroomgebieden van de Beerze, de Reuzel en de Rosep. Wageningen, DLO-Staring Centrum. Rapport 306 (in Dutch).

Central Bureau of Statistics, 1983. De oppervlakte cultuurgrond per gemeente; in: Statistiek van de landbouw 1981. 's-Gravenhage 5-8.

Congalton, R.G., 1988a. A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data. Photogrammetric Engineering and Remote Sensing 54: 593-600.

Congalton, R.G., 1988b. Using spatial autocorrelation analysis to explore the errors in maps generated from remotely sensed data. Photogrammetric Engineering and Remote Sensing 54: 587-592.

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment 37: 35-46.

Corves, C. and C.J. Place, 1994. Mapping the reliability of satellite-derived landcover maps; an example from the Central Brazilian Amazon Basin. International Journal of Remote Sensing 15: 1283-1294.

Drecht, G. van, H.F.R. Reijnders, L.J.M. Boumans en W. van Duijvenbooden, 1996. De kwaliteit van het grondwater op een diepte tussen 5 en 30 meter in Nederland in het jaar 1992 en de verandering daarvan in de periode 1984-1993. Bilthoven, RIVM. Rapport 714801005 (in Dutch).

Eck, W. van, D. Slothouwer, J.B. Strik en G.F.P. IJkelenstam, 1995. Erosienormeringsonderzoek Zuid-Limburg. Kosten en baten van erosiebestrijdingsmaatregelen in Zuid-Limburg. Wageningen, DLO-Staring Centrum. Rapport 364.2 (in Dutch).

Fuller, R.M., G.B. Broom and A.R. Jones, 1994. The land cover map of Great Britain: An automated classification of Landsat Thematic Mapper data. Photogrammetric Engineering and Remote Sensing 60: 553-562.

Gehrels, J.C., 1995. Niet-stationaire grondwatermodellering van de Veluwe. Een studie naar de invloed van grondwaterwinning, inpoldering en verloofing op de grondwaterstand sinds 1951 (2 rapporten: tekst en figuren). Amsterdam, Vrije Universiteit (in Dutch).

Genderen, J.L. van, B.F. Lock and P.A. Vass, 1978. Remote sensing: statistical testing of thematic map accuracy. Remote Sensing of Environment 7: 3-14.

Gurney, C.M. and J.R.G. Townshend, 1983. The use of contextual information in the classification of remotely sensed data. Photogrammetric Engineering and Remote Sensing 49: 55-64.

Hill, J., 1993. High precision land cover mapping and inventory with multi-temporal earth observation satellite data. The Ardèche Experiment. Commission of the European Communities. Joint Research Centre. Luxembourg, Office for official publications of the European Communities.

Hutchinson, C.F., 1982. Techniques for combining Landsat and ancillary data for digital classification improvement. Photogrammetric Engineering and Remote Sensing 48: 123-130.

Janssen, L., 1993. Methodology for updating terrain object data from remote sensing data. The application of Landsat TM data with respect to agricultural fields. Wageningen, Thesis, Agricultural University Wageningen.

Kenk, E., M. Sondheim and B. Yee, 1988. Methods for improving accuracy of thematic mapper ground cover classifications. Canadian Journal of Remote Sensing 14: 17-31.

Kooij, M.W.A. van der, D. van Halsema, W. Groenewoud, G.J. Metz, B. Overgaauw and P.N.A.M. Visser, 1995. SAR land subsidence monitoring. Delft, Beleids-commissie Remote Sensing. Rapport 95-13.

Laan, F.B. van der, 1988. Integratie van topografische kaarten en satelliet- gegevens voor aktualisatie van landgebruik. Kartografische Tijdschrift XIV: 69-78.

Laan, F.B. van der, 1989. The chance on cloudfree imagery of the Netherlands. In: Products and services. Detailed guide to imagery available. Annex to Remote Sensing Newsletter no. 38. Delft, Programmabureau Beleidscommissie Remote Sensing.

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

Lopes, A., E. Nezry, R. Touzi and H. Laur, 1993. Structure detection and statistical adaptive speckle filtering in SAR images. International Journal of Remote Sensing 14: 1735-1758.

Meinardi, C.R., 1994. Groundwater recharge and travel times in the sandy regions of the Netherlands. Vrije Universiteit, Amsterdam. Docteral thesis. Also published as Report 715501004, RIVM, Bilthoven.

Ministery of Agriculture, Nature Management and Fisheries, 1995. Discussienota Visie Stadslandschappen (Hoofdrapport en 6 thema rapporten). Den Haag, Ministerie LNV, Directie Natuurbeheer (in Dutch).

Ministery of Transport and Public Works, 1989. Derde nota waterhuishouding; water voor nu en later. Tweede kamer, 1988-1989, 21250 nrs. 1-2. 's-Gravenhage, SDU uitgeverij (in Dutch).

National Institute of Public Health and Environmental Protection, 1993. Nationale Milieuverkenningen 1993-2015. Alphen aan de Rijn, Samson H.A., Tjeenk Willink (in Dutch).

Noordman, E., H.A.M. Thunnissen en H. Kramer, 1996. Gebruikershandleiding voor het LGN2-bestand. Grondgebruiksbestand vervaardigd met satellietbeelden en andere geografische informatie. Wageningen, DLO-Staring Centrum. Technisch Document 43 (in Dutch).

Provincie Gelderland, 1994. Milieu/waterverkenning 1994. Arnhem.

Querner, E.P., 1993. Aquatic weed control within an integrated water management framework. Wageningen Agricultural University, Wageningen. Docteral thesis. Also published as Report 76, DLO Winand Staring Centre, Wageningen.

Reijerinck, J.G.A. en A. Breeuwsma, 1992. Ruimtelijk beeld van de fosfaatverzadiging in mestoverschotgebieden. Wageningen, DLO-Staring Centrum. Rapport 222 (in Dutch).

Richards, J.A., 1986. Remote sensing digital image analysis; An introduction. Berlin, Springer-Verlag.

Schotten, C.G.J., W.W.L. van Rooij and L.L.F. Janssen, 1995. Assessment of the capabilities of multi-temporal ERS-1 SAR data to discriminate between agricultural crops. International Journal of Remote Sensing 16: 2619-2637.

Schoumans, O.F., A. Breeuwsma and H.A.M. Thunnissen, 1989. Use of remote sensing and geographic information systems in soil and groundwater vulnerability assessment. Proceedings of the 9th symposium of the European Association of Remote Sensing Laboratories (EARSeL), held in Espoo, Finland, 27 June-1 July 1989. Luxembourg, Office for official publications of the European Communities, 54-60.

Sheffield, C., 1985. Selecting band combinations from multispectral data. Photogrammetric Engineering and Remote Sensing 50: 607-617.

Sluijs, P. van der and J.J. De Gruijter, 1985. Watertable classes: A method used to indicate seasonal fluctuation and duration of water tables on Dutch soil maps. Agricultural Water Management 10: 109-125.

Steenvoorden, J.H.A.M., A.K. Bregt and B.J. van Bleek, 1993. WATRO: beleidsondersteunend instrument op het gebied van water en milieu voor de ruimtelijke ordening. Wageningen, DLO-Staring Centrum. Rapport 266 (in Dutch).

Story, M. and R. Congalton, 1986. Accuracy assessment: a user's perspective. Photogrammetric Engineering & Remote Sensing 52:397-399.

Strahler, A.H., 1980. The use of prior probabilities in maximum likelihood classification of remotely sensed data. Remote Sensing of Environment 10: 135-163.

TAUW, 1994. Bodemkwaliteitsmeetnet Stroomgebied Drenthse Aa en Elperstroom. Ontwerp. Deventer, Tauw Milieu bv. Rapport 3313964 (in Dutch).

Thunnissen, H.A.M., M.N. Jaarsma and O.F. Schouwmans, 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., R. Olthof, P. Getz en L. Vels, 1992b. Grondgebruiksdatabank van Nederland vervaardigd met behulp van Landsat Thematic Mapper opnamen. Wageningen, DLO Winand Staring Centre. Rapport 168 (in Dutch).

Thunnissen, H.A.M. and H. J. van Middelaar, 1995. The CORINE Land Cover database of the Netherlands. Final report of the CORINE Land Cover project in the Netherlands. Wageningen, DLO Winand Staring Centre. Report 78.

Vermulst, J.A.P.H., 1993. Toetsing van de waterkwaliteismodellen NITSOL en PHOSOL op stikstof- en fosfaatvrachten in de Schuitenbeek. Lelystad, Rijkswaterstaat, Rijksinstituut voor Integraal Zoetwaterbeheer en Afvalwaterbehandeling. Nota 93.002 (in Dutch).

Walsum, P.E.V. van en A.A. Veldhuizen, 1996. Modelstudie waterhuishouding Fochteloërveen en omgeving. Simulatie van scenario's voor het waterbeheer met SIMGRO. Wageningen, DLO Winand Staring Centre. Rapport 399 (in Dutch).

Walsum, P.E.V. van. J.M.P.M. Peerboom und J.H.A.M. Steenvoorden, 1996. Hydrologische Untersuchung von Hochmoorresten und ihrer Umgebung. Felduntersuchung und Modellierung. Zeitschrift für Kulturtechniek und Wasserbau.

Wel, F.J.M. van der, 1993. Camotius Vooronderzoek. Universiteit Utrecht. Delft, Beleidscommissie Remote Sensing. Rapport 93-05 (in Dutch).

Wel, F.J.M. van der en B.G.H. Gorte, 1995. Camotius Hoofdfase. Universiteit Utrecht and ITC. Delft, Beleidscommissie Remote Sensing. Rapport 95-06 (in Dutch).

Wopereis, F.A., 1991. Bodemgeschiktheid voor emissie-arme mesttoedieningstechnieken van grasland in zandgebieden. Wageningen, DLO-Staring Centrum. Rapport 159 (in Dutch).

Unpublished sources

Keeman, D.M. 1991. Using Ground water table classes for improving image classification. Dept. of Land Surveying and Remote Sensing. Wageningen, Agricultural University. MSc-thesis (in Dutch).