Crop area estimation by data-driven decomposition of mixed pixels in Landsat-TM images

M. S. KLEIN GEBBINCK and T. E. SCHOUTEN

Computing Science Institute, Faculty of Mathematics and Informatics,

University of Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, the Netherlands

Telephone: +31 24 3652710, Fax: +31 24 3553450, E-mail: mauricek@cs.kun.nl

Abstract

Pixel-based classification is a standard approach towards crop area estimation, but it is unsuited to handle mixed pixels. Therefore, we developed the data-driven decomposition method, which combines the linear mixture model with knowledge of the application domain. A quantitative comparison proved that both concepts lead to more accurate estimates when processing simulated images. Experiments on a real Landsat-TM image showed that our method estimated the area of 13 out of 17 lots more accurately than a similar approach based on classification. These results suggest that data-driven decomposition is a more accurate method to estimate the area of an agricultural field.

1 Introduction

In order to manage its agricultural subsidy system, the European Union (EU) needs an automated system that can estimate the area of an agricultural field with sufficient accuracy. Every year the EU spends tens of milliards of dollars on agricultural subsidies, which are allocated based on the type and acreage of the crops a farmer grows. To check a subsidy claim, therefore, an accurate estimation of the area covered by each crop grown by the farmer is required. Furthermore, by summing the area of all fields of a particular crop, the total yield per crop can be predicted. Since the number of fields that are in use throughout the EU is approximately 150 million, even the task of surveying only a small fraction of them is too large to be carried out by human inspectors on-site. With an automated system using satellite imagery, on the other hand, it is possible to check a much larger number of fields, on a more frequent basis, and against lower costs, provided that the area of a field can be estimated accurately enough.

Many crop area estimation methods have been described in the literature. A straightforward way to estimate the crop area per farmer is to locate the pixels representing his or her land and determine the number of pixels covered by each crop type. Based on this direct estimate, more unbiassed estimates can be acquired with techniques using confusion matrices or regression correction (Gallego *et al.* 1993). The traditional approach to determine the crop type of a pixel is classification, which allocates each pixel to only one of a number of selected classes. However, due to the limited resolution of the satellite scanner, many pixels, e.g. those at field boundaries, comprise more than one ground cover type. Classification of these mixed pixels, therefore, leads to inaccurate area estimates because the summed area of the fractions that are wrongly allocated may be considerable. Furthermore, the spectral response of a mixed pixel may be so confusing to the classification algorithm that the pixel is allocated to an incorrect class altogether. A more recent approach to determine a pixel's crop type is decomposition, which tries to divide a pixel between several classes according to its composition (e.g. Drake and Settle 1989, Cross *et al.* 1991, Quarmby *et al.* 1992). The decomposition method is less susceptible to spectral confusion and, in principle, can give more accurate estimates than classification. However, decomposing pure pixels often leads to inaccuracies and, more important, the number of classes that can be discerned is severely limited. In addition to these simple pixel-based methods, several authors have suggested to use the spatial relationships between the image pixels to locate agricultural fields for subsequent classification on a field basis (e.g. Janssen 1994, Schoenmakers 1995). Although this use of domain knowledge offers some improvement over pixel-based classification, it remains unsuited to handle mixed pixels.

The data-driven decomposition method we propose combines the advantages of the decomposition approach and field-based classification. By using contextual information to locate the agricultural fields, a discrimination between pure pixels and mixed pixels can be made. Next, pure pixels are classified and mixed pixels are decomposed to determine their crop type. As will be described in Section 2, the domain knowledge regarding agricultural fields can also be used to improve the accuracy of the decomposition process. Furthermore, it permits to select the classes a mixed pixel is probably composed of in order to circumvent the limitation on the number of classes that can be considered. Application of knowledge about objects other than agricultural fields, e.g. lanes, ditches, and farmhouses, improve the estimation of the field area even more. Thus, by decomposing mixed pixels and utilising knowledge of the application domain, a method was developed that can estimate the area of an agricultural field with relatively high accuracy.

The capabilities and limitations of data-driven decomposition were assessed using both simulated and real Landsat-TM images of agricultural scenes. The aim of our first experiment, see Section 3, was to determine the influence of decomposing mixed pixels and applying domain-specific knowledge on the area estimation accuracy. For this purpose, detailed information about the composition of the individual pixels was needed. Since this ground truth information at sub-pixel level is expensive to gather and difficult to co-register with a satellite image with sufficient accuracy, we decided to simulate a series of Landsat-TM images using increasingly realistic scene models. Based on these images, a comparison between data-driven decomposition and three other approaches was made to determine the merits of the two main concepts behind data-driven decomposition. However, because the scene model that was used to simulate the images was also incorporated in data-driven decomposition, a test on real Landsat-TM data was required as well. Therefore, as described in Section 4, another experiment was devised in which the field area estimates calculated by data-driven decomposition and field-based classification were compared with their assumed true value determined from a large-scale topographical map. Based on these two experiments, a quantitative evaluation of the performance of data-driven decomposition could be made, leading to the conclusions presented in Section 5.

2 Data-driven decomposition

Data-driven decomposition (Klein Gebbinck and Schouten 1996a) attempts to make an accurate estimation of the area of an agricultural field based on two ideas. The first thought is that pixels that are mixed should be split into their composing parts. The method we selected for this task is based on the linear mixture model and will be described in Section 2.1. The second idea is that knowledge of the application domain can be used to improve the accuracy of the crop area estimation and to lift the restriction on the number of crop types. Section 2.2 shows how the contextual information in the image data is used to "drive" the decomposition process; Section 2.3 describes how our knowledge of objects other than agricultural fields is used to obtain a further increase in accuracy.

2.1 *Mixed pixel decomposition*

In order to be able to decompose mixed pixels, the mechanism by which they are created must be modelled. Data-driven decomposition regards each agricultural field of sufficient size as a cluster of pure pixels, which represents the field's interior, surrounded by a layer of mixed pixels that are positioned at the field boundaries. The pure pixels inside the cluster are assumed to have a multivariate normal distribution, which is mainly determined by the crop type of the field. Each mixed pixel of the surrounding layer is thought to represent a spectral response that depends on both the distributions and on the areal fractions of the fields it is composed of. This dependency is linear as long as the amount of multiple scattering between the different crop types is negligible, i.e. each photon reaching the sensor has interacted with only one cover type (Settle and Drake 1993). The linear mixture model has shown to be a valid abstraction for a broad range of applications. For instance, it was used to eliminate the effects of topography and illumination in Viking Lander 1 images (Adams *et al.* 1986); to find porphyry copper deposits (Drake and Settle 1989); to measure tropical forest cover (Cross *et al.* 1991); to estimate the area of agricultural crops (Quarmby *et al.* 1992); and for many other applications estimating the coverage by rangeland vegetation, snow, and urban ground cover types. A number of decomposition methods that assume a non-linear mixture model have also been suggested (e.g. Marsh *et al.* 1980, Foody and Cox 1994, Carlotto 1995), but they all lacked the sensible physical basis of the linear mixture model, and none of them was followed up to any great degree. Because of its solid physical foundation and widely proved applicability, the linear mixture model was chosen as the basis of our data-driven decomposition algorithm.

A mathematical description of the linear mixture model can conveniently be given in matrix-vector notation. Let us assume that our sensor takes samples in n different spectral bands, and that the number of components of a mixed pixel is c. If we follow the notation of Settle and Drake (1993), the spectral response for a mixed pixel $\mathbf{x}$ is given by

$$\mathbf{x} = \mathbf{M}\mathbf{f} + \mathbf{e}. \tag{1}$$

The (nxc) matrix $\mathbf{M}$ contains the spectra of the c so-called endmembers that take part in the mixed pixel. An endmember spectrum is the response that would be received from a ground resolution element containing nothing but the component of interest; since in our case the components of interest are agricultural fields, we take the mean vectors of their pure pixel clusters as the columns of matrix $\mathbf{M}$. The partition of the mixed pixel's area is denoted by the $(cx1)$ vector $\mathbf{f}$, with f_i representing the fraction of the area covered by the ith endmember. The $(nx1)$ error vector $\mathbf{e}$, finally, compensates for the statistical fluctuations in the endmember spectra. According to our model of agricultural fields, the fluctuations around the mean spectrum of endmember i is given by a variance- covariance matrix $\mathbf{N}_i$, which can be estimated from the corresponding pure pixel cluster. Horwitz *et al.* (1971) have shown that as a consequence the spectral response for a mixed pixel will fluctuate according to the variance-covariance matrix $\mathbf{N_f} = \sum f_i \mathbf{N}_i$. Therefore, $\mathbf{e}$ is assumed to have a multi-variate normal distribution with a zero mean and a variance-covariance matrix $\mathbf{N_f}$, with which our short description of the linear mixture model is complete.

An important characteristic of a decomposition method is the strategy that determines which solution of the linear mixture model is optimal. When decomposing a mixed pixel, the fractions of a pixel's components are estimated from its multi-spectral observation, which in terms of Equation (1) is equal to estimating $\mathbf{f}$ from $\mathbf{x}$. Unfortunately, for every $\mathbf{x}$ there exist infinitely many solutions for $\mathbf{f}$ that all satisfy the linear mixture model. Since the introduction of the linear mixture model by Horwitz *et al.* in 1971, many methods have been suggested to determine a solution that is optimal in some sense (see Foschi (1994) for a brief overview). Basically, however, there exist only three statistically sound techniques, which Pech *et al.* (1986) classify as the maximum likelihood, Bayesian, and empirical Bayesian approach to statistical inference. Of these three alternatives, the maximum likelihood approach has been selected for most studies, and nearly all have reported good results upon its application. Furthermore, as will be described in the remainder of this section, a good approximation of the maximum likelihood solution can be found with relatively few computations. It is for these reasons that data-driven decomposition uses the maximum likelihood approach to determine the optimal solution of the linear mixture model.

The maximum likelihood solution of the linear mixture model selects the combination of fractions and error vector that has the highest probability. In general, if the number of spectral bands is higher than the number of components, all solutions of the linear mixture model require an error vector $\mathbf{e}$ that is non-zero. Since $\mathbf{e}$ has a multi-variate normal distribution with a zero mean and a variance-covariance matrix $\mathbf{N_f}$, the most probable solution is the one that minimises

$$\mathbf{e}^T \mathbf{N_f}^{-1} \mathbf{e} + \ln |\mathbf{N_f}| = (\mathbf{x} - \mathbf{Mf})^T \mathbf{N_f}^{-1} (\mathbf{x} - \mathbf{Mf}) + \ln |\mathbf{N_f}|, \tag{2}$$

where $\ln |\mathbf{N_f}|$ denotes the natural logarithm of the determinant of $\mathbf{N_f}$. Unfortunately, finding the $\mathbf{f}$ which minimises Equation (2) is rather complicated because the variance-covariance matrix is a function of the fractions vector. However, Horwitz *et al.* (1971) have argued that $\mathbf{N_f}$ can be substituted by the mean variance-covariance matrix $\overline{\mathbf{N}} = 1/c \sum \mathbf{N}_i$ without much loss of accuracy. With the dependency of $\mathbf{e}$'s variance-covariance matrix on $\mathbf{f}$ removed, we can minimise Equation (2) by setting to zero its partial derivatives with respect to each f_i. The resulting estimator for $\mathbf{f}$,

$$\hat{\mathbf{f}}(\mathbf{x}) = \mathbf{U}\mathbf{M}^T \overline{\mathbf{N}}^{-1} \mathbf{x} \quad \text{with} \quad \mathbf{U}^{-1} = \mathbf{M}^T \overline{\mathbf{N}}^{-1} \mathbf{M}, \tag{3}$$

consists of a series of matrix-vector operations that can be executed very fast on any computer. Thus, by selecting the most probable solution of the linear mixture model, as the maximum likelihood approach does, an efficient estimator for $\mathbf{f}$ is derived.

Since the linear mixture model is supposed to have a physical basis, the estimated fractions vector has to satisfy two constraints. The first condition that has to be fulfilled is the sum-to-one constraint, which means that the sum of all fractions must add up to one. The second restriction is that all fractions must be non-negative, which is sometimes referred to as the positivity constraint. In general, neither constraint is satisfied by the maximum likelihood estimator defined by Equation (3). However, Settle and Drake (1993) have shown that through standard Lagrangian analysis a slightly different solution can be obtained that does satisfy the sum-to-one constraint. Satisfaction of the positivity constraint, on the other hand, cannot be achieved by some analytic solution, but requires some post-processing of the initial estimate or an alternative approach to solve the linear mixture model. Among the possibilities that have been discussed by Settle and Drake (1993), the simplest post-processing approach is to set any negative fractions to zero and renormalise the fractions vector afterwards. Although this approach may result in a small loss of accuracy, the computational costs are so low that we chose this method for the data-driven decomposition algorithm. By modifying the maximum likelihood estimator in the two ways described above, a solution of the linear mixture model is found that is physically sensible.

2.2 *Data-driven aspects of the method*

The accuracy of an crop area estimation method based on the maximum likelihood solution described in the previous section can be increased by incorporating some knowledge of the application domain. So far, most studies have used no domain-specific knowledge whatsoever, but have employed the estimator to decompose all image pixels, assuming the same endmember distributions for the entire image, and using all available endmembers at the same time. Data-driven decomposition, on the other hand, makes use of known spatial relationships between the pixels to locate the mixed pixels that need to be split, to determine local endmember distributions from the image itself, and to select only those endmembers that are probable components of a mixed pixel.

A first way to get a more accurate area estimation is to separate the pure pixels from the mixed pixels by making a segmentation of the image. As was described in Section 2.1, any agricultural field that is relatively large given the resolution of the scanner can be regarded as an inner cluster of pure pixels surrounded by a layer of mixed pixels at the field boundaries. Since we know that our scenes are a collection of agricultural fields, we can discriminate between pure and mixed pixels by locating the pure pixel clusters. These clusters can sometimes be found with the help of ancillary information sources such as topographical maps (Janssen 1994), but it is also possible to segment the image using standard image processing techniques. Schoenmakers (1995), for instance, developed a hybrid segmentation method that combines edge detection, which looks for local discontinuities in the image, and region growing, which groups neighbouring pixels that are spectrally similar, to find pure pixel regions in optical satellite images. Once the pure pixels belonging to a cluster have been identified, they can be classified either on an individual basis or by some field-based approach that is more accurate (Janssen 1994). Most remaining pixels are probably mixed and should be decomposed rather than classified. Thus, the accuracy of an crop area estimate can be increased by making a segmentation of the image in order to classify the pure pixels and decompose the mixed pixels.

Data-driven decomposition is also more accurate than the classic decomposition approach because of the way it defines the endmember distributions. Traditionally, the statistical distribution of every ground cover type that was to be distinguished was determined using training pixels, laboratory measurements, or feature space analysis. However, since we know that our scene consists of agricultural fields that can be modelled as pure pixel clusters with mixed pixels in between, we can determine more accurate endmember distributions by calculating the statistical distribution of each cluster. The advantage of taking the distributions of fields instead of crops is that the former endmember distributions, which are in fact crop distributions as well, are better adapted to the local variations in ground slope, elevation, soil type, etc; as a result, the estimation of the fractions of a mixed pixel will become more accurate too. However, the size of a field may be too small to estimate its distribution reliably. In this case, data-driven decomposition determines the crop type through maximum likelihood classification of the field's mean vector, and selects the corresponding crop distribution from a database with endmember distributions that were determined in a traditional manner after all. With this mechanism to determine the endmember distributions, data-driven decomposition can split mixed pixels more accurately than the classic decomposition method.

The third way in which data-driven decomposition tries to increase the accuracy of the crop area estimates is to use contextual information for selecting probable endmembers of a mixed pixel. The classic decomposition approach, which decomposes a mixed pixel into all endmembers at the same time, has two drawbacks. The first is that the decomposition is usually sub-optimal because fractions are allocated to incorrect endmembers, and the second is that the number of ground cover types that can be distinguished in the image is severely limited by the number of spectral bands because of the way the linear mixture model is solved. Especially the latter limitation is a serious problem for data-driven decomposition as it regards every agricultural field as a possible endmember. However, by restricting the large set of possible endmembers to a small set of probable endmembers, which can be different for each pixel, an accurate decomposition using a virtually unlimited number of endmembers can be achieved. According to our simple scene model, the components of a mixed pixel are the agricultural fields that surround it; therefore, the endmembers that are needed to decompose a mixed pixel can be found by searching its neighbourhood for pure pixel clusters. The size of the neighbourhood that is searched must be small to ensure that the fields remain contiguous on the one hand, but it must be large enough to contain pure pixels instead of only other mixed pixels on the other hand. Data-driven decomposition solves this dilemma by searching the smallest possible symmetrical neighbourhood—a $3x3$ window centred at the mixed pixel—and following an iterative strategy in which the easy mixed pixels are decomposed first, and the acquired information about their composition is used to decompose the more difficult pixels later. Thus, the spatial context of a mixed pixel can be used to select its probable endmembers, which results in a more accurate decomposition.

In summary, the data-driven decomposition method overcomes three important weaknesses of the classic decomposition approach. The first improvement is that, instead of decomposing all pixels, a discrimination between pure and mixed pixels is made in order to handle them by classification and decomposition, respectively. The second enhancement is achieved by taking local endmembers, whose distributions are well adapted to local crop variations, instead of endmembers that are global to the entire image. The third weakness that is resolved is the necessity of decomposing a mixed pixel into all endmembers simultaneously; our method is able to predict the probable components of a mixed pixel prior to its decomposition. All these improvements make use of the contextual information included in the image. By formalising some knowledge of the application domain into a simple scene model, the spatial relationships between the pixels can be used to make a more accurate estimation of the area of an agricultural field.

2.3 *Further improvements based on domain knowledge*

The accuracy of the data-driven decomposition method can be increased further by incorporating additional knowledge of the application domain in our scene model. So far, a scene was modelled as a collection of agricultural fields with coinciding boundaries between successive fields. This simple model, however, is not very accurate since in reality agricultural fields are often bounded by lanes, ditches, strips of grass, etc. Furthermore, usually the scene also contains a number of small objects such as farmhouses, ponds, and groves. Therefore, the scene model was extended with two object types, called narrow boundary structures and isolated objects.

Narrow boundary structures require a different approach than agricultural fields in order to process them accurately. Given that the typical ground resolution of a satellite scanner is 20–30 m, a narrow boundary structure will be depicted as an elongated series of mixed pixels that is one or two pixels wide. Since such a structure does not have a pure pixel cluster, it is impossible to derive its statistical distribution directly from the image as described in Section 2.2. Instead, we have to fall back on more traditional techniques, for instance based on training pixels, to determine a few endmember distributions which approximate the distributions of that type of structure. Some examples of endmembers which should be determined because they are components of some well-known boundary structures are: asphalt (roads), soil (uncultivated land and dirt roads), grass (uncultivated land), deciduous forest (tree rows), and water (ditches and canals). Another difficulty due to the lack of a pure pixel cluster is that the presence of a boundary structure cannot be detected, which prevents the endmember selection mechanism introduced in Section 2.2 from finding all components of a mixed pixel. Again, the answer is to revert to the classic approach by including all endmembers typical of narrow boundary structures when decomposing a mixed pixel. The problems that arise from this large number of endmembers—infinitely many solutions of the linear mixture model or an inaccurate decomposition, see Section 2.2—can be solved with some domain-specific knowledge. According to this knowledge, nearly all mixed pixels are formed by one of three possible combinations: two agricultural fields; two agricultural fields and one narrow boundary structure; or one agricultural field and one narrow boundary structure, in case the structure width spans two pixels. Since it can be derived which agricultural fields are probable components of a particular mixed pixel, and a set of endmembers that covers all types of boundary structures is available, a list of every possible composition can be made. To determine the optimal composition, we took the Bayesian approach by choosing the alternative that had the maximum a posteriori probability (MAP). Since none of the possible compositions had our preference a priori, the combination that resulted in lowest value of the expression defined by Equation (2) was selected.[1] By adapting the data-driven decomposition algorithm in the ways described above, narrow boundary structures can be processed accurately.

Isolated objects are processed in roughly the same way as narrow boundary structures. Since isolated objects are small not only in width but also in length, they will usually appear as isolated mixed pixels in the image. Just like narrow boundary structures, an isolated object does not have a pure pixel cluster which is needed to determine its statistical distribution and to mark it as a probable endmember. Therefore, the distributions of those endmembers that are characteristic of this type of object have to be determined with traditional techniques. The endmember selection strategy that was devised to handle boundary structures, however, should not be used for isolated objects because their variety is too large and their number is too small. As a result of their large variety, many distributions are needed to describe all possible objects, which slows down the decomposition process considerably. More important, though, is that many mixed pixels will be decomposed less accurately because fractions can be assigned to a large number of additional endmembers that are characteristic of isolated objects only, while the probability that the mixed pixel indeed represents an isolated object is small. To solve these problems, data-driven decomposition sets apart all mixed pixels that probably represent isolated objects and decomposes them in a separate stage. The criterion that is used to make this decision is based on the probability of the MAP solution described in the previous paragraph: a mixed pixel is regarded as an isolated object only if the maximum a posteriori probability of its decomposition into the usual

[1] Note that $\mathbf{N_f}$ is substituted by $\overline{\mathbf{N}}$, which differs for each combination.

endmembers is too low. The endmember selection strategy that is followed in this case makes use of the knowledge that most of these pixels are a mixture of the isolated object and one of the agricultural fields it is surrounded by. By trying all combinations of one endmember typical of isolated objects and one agricultural field that is located in the pixel's neighbourhood, the optimal decomposition according to Bayes can be found by selecting the solution that has the maximum a posteriori probability. Thus, in a way that is very similar to the way narrow boundary structures are handled, isolated objects are processed efficiently by data-driven decomposition.

In conclusion, the accuracy of the classic decomposition approach can be increased by using the available knowledge of the application domain. Besides our knowledge of agricultural fields, which led to the exploitation of the contextual information in the image, we can also make use of our knowledge of other existing objects. By extending the simple scene model of Section 2.2 with narrow boundary structures and isolated objects, much greater realism is introduced. As a result, the accuracy of the area estimation of an agricultural field by data-driven decomposition will be improved further.

3 Experiments with simulated data

To evaluate the concepts behind data-driven decomposition, the method was applied to several artificial Landsat-TM images of agricultural scenes. Since no real Landsat-TM image with corresponding ground truth information at subpixel level was available, we had to simulate some images in order to make a quantitative analysis possible—see Section 3.1. Besides the data-driven decomposition method, three other well-chosen crop area estimators were applied to these images such that the influence of decomposing mixed pixels and applying domain-specific knowledge could be determined. The results of these experiments are presented in Section 3.2 and are followed by a discussion in Section 3.3.

3.1 *Set-up*

The set-up of the experiment was straightforward. As a first step, several artificial Landsat-TM images were generated based on a detailed ground cover map. Next, these images were processed by data-driven decomposition and three other methods, which represented all four approaches possible when combining the decomposition or classification method with the use of some knowledge of the application domain or no knowledge at all. Afterwards, the results were expressed quantitatively using two accuracy measures such that the merits of each approach could be determined by simple numerical comparison.

To simulate remotely sensed images, we used a method that was based on the scene model presented in the previous section. The method, which was comprehensively described by Schouten and Klein Gebbinck (1995), started from a sketch of the field boundaries that were drawn by hand on a piece of paper. Next, the drawing was scanned and processed to obtain a binary image of high resolution containing 1-pixel wide closed polygons. After labelling the non-boundary pixels with unique field numbers using blob colouring, a mapping was made between the field and the available ground cover types. The boundary pixels could also be related to a ground cover type, thus acting as narrow boundary structures. For each ground cover type that was selected this way, a template was compiled from parts of agricultural fields taken from a classified Landsat-TM image of Portugal. These parts were processed to ensure that they could be connected in a way that preserved both simple point statistics, like mean value and variance-covariance, and texture. The next step was to divide the high resolution image into blocks of nxn pixels, which were used to determine the composition of the low resolution pixels of the target image. This image was then generated by filling each field using the appropriate template according to the previously established map; the spectral response for the mixed pixels was determined by adding the contributions of several templates proportional to the number of matching subpixels in the corresponding nxn blocks. As a last step, isolated objects were introduced by replacing a number of pixels chosen at random by isolated mixed pixels, which were composed of one ground cover type of a nearby field and one totally different ground cover type having a fraction of at least 75%. As a result, an image with Landsat-TM characteristics was generated that consisted of agricultural fields, narrow boundary structures, and isolated objects.

The four crop area estimation methods that were applied are data-driven decomposition, pixel-based decomposition, field-based classification, and pixel-based classification. Data-driven decomposition was already described minutely in Section 2, except for way in which the images were segmented to locate the pure pixel clusters of the fields. Since the accuracy of the area estimator depends on the quality of the segmentation, and an investigation into optimal segmentation methods was considered to be beyond the scope of our present study (but see Schouten and Klein Gebbinck 1995), we decided to use the perfect segmentation, which could easily be derived while simulating the image. This way, the performance of data-driven decomposition could be evaluated without being affected by a possibly bad choice for a segmentation method. Pixel-based decomposition applied the same technique to unmix pixels as its data-driven counterpart, see Section 2.1, but did not use any knowledge of the application domain. As a result, each pixel had to be decomposed using all ground cover types simultaneously. Our field-based classification algorithm, on the other hand, was based on the same domain-specific knowledge as data-driven decomposition, but it allocated a pixel to a single field instead of dividing its area between multiple fields. The object to which the pixel was allocated was determined by the standard maximum likelihood classifier, which looks for the minimum of Equation (2) on the condition that $\mathbf{f}$ is a unit vector.[2] Similar to data-driven decomposition, the probability of the solution thus selected was used to decide whether the pixel under observation represented an isolated object or not. Also, the perfect segmentation was used again to locate the pure pixel clusters. Pixel-based classification, finally, is once again the counterpart of field-based classification that disregards all knowledge of the application domain. As a consequence, the maximum likelihood classifier had to consider all ground cover types during the classification of each pixel. With these four crop area estimators, the potential of both mixed pixel decomposition and the application of domain-specific knowledge to increase the accuracy of the area estimation could be investigated.

The accuracy of the crop area estimators was measured at the level of individual pixels as well as the entire region. Since all estimators operated on individual pixels, the direct way to determine their performance was to measure how well the composition of a pixel was estimated. A natural metric to express the distance between the true fractions vector $\mathbf{f}_p$ of a pixel p and its estimated value $\hat{\mathbf{f}}_p$ is Minkowski's L1-metric scaled by $\frac{1}{2}$:

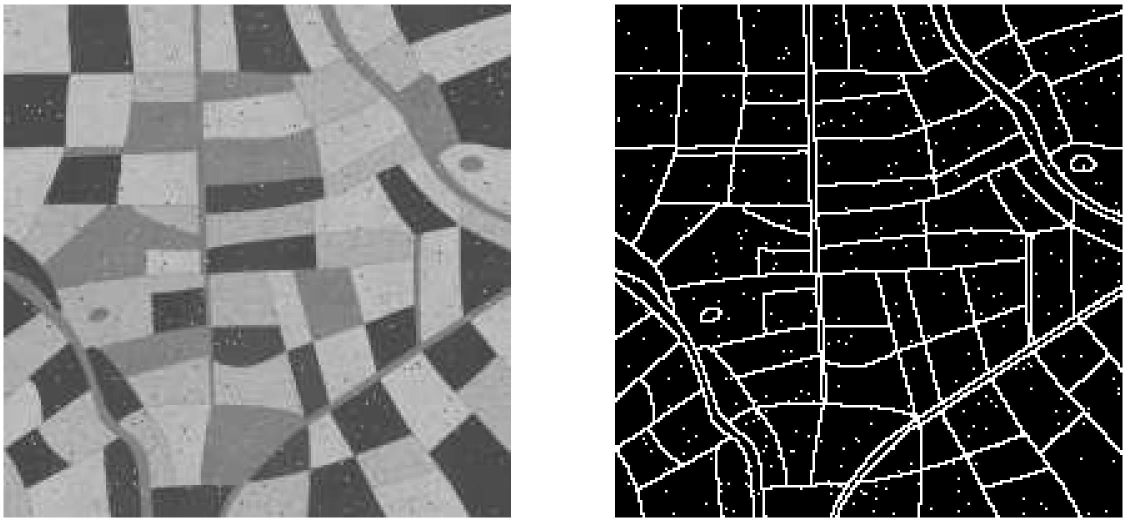
$$e_p = \frac{1}{2} \sum_i \left| \hat{f}_{p,i} - f_{p,i} \right| . \tag{4}$$

In this equation i ranges over all ground cover types, which means that the field area estimates by data-driven decomposition and field-based classification had to be transformed to crop area estimates by relating each field to a crop type. The scalar $\frac{1}{2}$ was introduced to compensate for the effect that the overestimation of one class automatically leads to the underestimation of another class; thus, a pixel composed of 60% maize and 40% barley, for instance, that was classified as maize resulted in an e_p of 0.4 instead of 0.8. From this example it is clear that the methods based on classification will probably give high values of e_p for mixed pixels. However, the crop area estimate for a larger region can still be accurate if opposite errors in pixel area allocation cancel each other out, e.g. overestimating the area of one pixel by 40% is nullified by underestimating the area of two other pixels by 30% and 10%, respectively. Therefore, the accuracy of the crop area estimators was also determined on a larger scale:

$$e_A = \frac{1}{2} \sum_i \left| \hat{A}_i - A_i \right| . \tag{5}$$

Here, $\hat{A}_i = \sum \hat{f}_{p,i}$ represents the estimated area and $A_i = \sum f_{p,i}$ the true area covered by crop i as calculated from all pixels p belonging to the area of interest. Although in a practical situation

[2]Note that $\mathbf{N_f}$ is substituted by a different $\mathbf{N}_i$ for each class.



(a)

(b)

Figure 1: Artificial Landsat-TM image simulated using five ground cover types. (a) Response in band 5; narrow boundary structures of soil can be discerned between high-intensity fields. (b) Positions of the mixed pixels shown in white. The isolated mixed pixels represent isolated objects.

the area of interest probably consists of only one or a couple of agricultural fields, we decided to sum the crop area estimates over the entire region depicted by the image because the same effects could be observed in a more convenient way. Whereas the former approach involved a laborious aggregation of the crop area estimates by the pixel-based estimators in particular, the latter approach required no additional processing at all. With the two accuracy measures presented above, it was possible to determine the crop area estimation capabilities of the different methods on both a small and a large scale.

The procedure consisting of the steps simulation, application, and evaluation as described above is a straightforward way to make a quantitative analysis of the performance of any crop area estimator. Based on the results obtained with the four crop area estimators selected, the impact of the two main concepts of data-driven decomposition—decomposition of mixed pixels and application of domain-specific knowledge—could be investigated. Since the accuracy of the estimates was measured not only per pixel but also over a large region, the potential for crop area estimation at the level of individual farmers could be assessed.

3.2 *Results*

To determine the accuracy of the crop area estimators, a number of different images were used. For the whole experiment—see Klein Gebbinck and Schouten (1996a)—three types of images were generated, which successively comprised only agricultural fields, fields and narrow boundary structures, and fields, structures, and isolated objects. In this paper, however, we only present the results for two images of the third and most realistic type because they suffice to draw the main conclusions.

The characteristics of the first image were chosen such that the crop area estimates for mixed pixels could be evaluated using all four approaches. The image was constructed by merging blocks of $4x4$ pixels of our high resolution ground cover map, which contained narrow boundary structures of soil and a high number—1.0% of all pixels—of isolated objects. The number of ground cover types was limited to five because Landsat-TM data has only four intrinsic dimensions (Mather 1990) and the number of classes pixel-based decomposition can handle simultaneously is at most

	reference	data-driven decomposition	pixel-based decomposition	field-based classification	pixel-based classification
$<e_p>$ (x100%)	–	3.9	13.1	35.3	38.9
e_A	–	12.6	204.7	961.9	1426.8
$\hat{A}_{soil}$	1906.0	1898.4	1760.4	1179	872
$\hat{A}_{barley}$	1047.9	1046.7	1025.9	1279	889
$\hat{A}_{maize}$	1017.3	1013.8	978.8	781	782
$\hat{A}_{vineyards}$	847.9	859.0	847.9	1372	1540
$\hat{A}_{deciduous\ forest}$	480.8	479.0	684.1	686	1214

Table 1: Results obtained by applying different estimators to the five-class image. The column "reference" lists the true area determined from the perfect segmentation; $<e_p>$ represents the error per pixel e_p averaged over all mixed pixels. The values are given in pixel area units.

one more than the true dimensionality of the data (Settle and Drake 1993). Figure 1 shows the resulting image together with the positions of the mixed pixels. This image was processed by each of the four crop area estimators, whose performance is presented in Table 1. The results of data-driven decomposition and field-based classification were obtained with the threshold set at the value that gave the lowest average error per pixel $<e_p>$. With this setting the corresponding value of e_A was close to its optimum as well, while minimising e_A, which as a function of the threshold was less smooth, often resulted in a relatively high $<e_p>$. Both accuracy measures were calculated based on the mixed pixels only. One reason for this decision was that the overwhelming number of pure pixels, which are relatively easy to handle by all methods, obscured the results for the mixed pixels, which are more difficult to process and therefore offer more possibilities to improve the area estimates. Another reason to exclude pure pixels was that we wanted to minimise the influence of using a perfect segmentation, which mainly provides information about pure pixels, e.g. the set of nearby pixels having the same ground cover type and the fact that their fractions vector is a unit vector. The results presented in Table 1 clearly show that data-driven decomposition was the most accurate method, followed by pixel-based decomposition and, at some distance, by field-based and pixel-based classification. This order is quite distinct as the differences in accuracy are large, and is the same regardless of whether the accuracy was measured on the basis of individual pixels or the entire region. Although the accuracy of one approach should be considered only in relation to those of other methods, the crop area estimated by data-driven decomposition was also remarkably accurate in an absolute sense, in contrast to the accuracy of the other three methods.

The second image was generated with the aim of evaluating data-driven decomposition using data that was as realistic as possible. Given the scene model used for simulating data, the only way to increase the level of realism of the previous image was to use a wider variety of ground cover types. However, the number of ten ground cover types present in the second image was too large to be handled by pixel-based decomposition. Therefore, Table 2 only lists the results of data-driven decomposition and both classification methods. Just like in the previous experiment, the threshold used by data-driven decomposition and field-based classification was set at its optimum value. With this setting, both methods needed only a few iterations to determine the probable components of most pixels, but whereas data-driven decomposition thus filtered out most pixels that represented isolated objects, field-based classification was left with a combination of ordinary and isolated mixed pixels. As was to be expected, the doubled number of classes led to less accurate crop area estimates, although the regional crop area estimate of field-based classification showed some improvement. Since the decrease in accuracy was modest for the methods based on classification but even smaller for data-driven decomposition, the order of best crop area estimators remained the same. Furthermore, the crop area estimates by data-driven decomposition calculated for the entire region were still surprisingly accurate. In summary, data-driven decomposition gave the best results of the three estimators when applied to the most realistic image generated.

	reference	data-driven decomposition	field-based classification	pixel-based classification
$<e_p>$ (x100%)	–	5.0	41.1	46.4
e_A	–	22.9	738.7	1743.7
$\hat{A}_{sand}$	216.1	217.8	176	155
$\hat{A}_{soil}$	1446.3	1439.2	1027	378
$\hat{A}_{fresh\ water}$	290.2	291.4	101	87
$\hat{A}_{wheat}$	699.8	700.0	770	813
$\hat{A}_{barley}$	542.0	535.9	717	243
$\hat{A}_{maize}$	521.5	520.8	430	408
$\hat{A}_{rice}$	435.7	433.1	675	726
$\hat{A}_{vineyards}$	244.4	243.1	330	522
$\hat{A}_{grassland}$	594.1	612.6	742	1181
$\hat{A}_{deciduous\ forest}$	309.8	303.1	329	784

Table 2: Performance of the estimators applied to the ten-class image. The true area of each ground cover type is listed under "reference". The error e_p averaged over all mixed pixels is denoted by $<e_p>$. Errors as well as estimates are given in pixels.

3.3 *Discussion*

Because of the big differences in performance between the crop area estimators, it is obvious that both the decomposition of mixed pixels and the application of domain-specific knowledge have considerable influence on the crop area estimation accuracy. However, in order to explain the influence of each concept, first the causes of inaccuracy have to be identified for each approach.

The sources of error were slightly different for each of the four methods applied. Apart from the minor inaccuracies due to the natural variation in the spectral response of each ground cover type, the main source of error for data-driven decomposition was spectral confusion. A good example was the confusion between grass, soil, and barley: since the spectrum of grass was a linear combination of soil and barley, mixed pixels containing the latter two components were often partially allocated to the class grassland—see Table 2. Pixel-based decomposition suffered from the same problem but to a larger extent, because the possibilities for spectral confusion greatly increased when all ground cover types were considered simultaneously. In addition to increased spectral confusion, the larger number of classes also caused an area leakage to unrelated classes because in general at least a small fraction of a pixel's area was allocated to each endmember. The causes of inaccuracy for field-based classification were threefold. First of all, classification of mixed pixels necessarily led to inaccurate results because pixels were allocated to a single class instead of being split. Secondly, a major factor was spectral confusion, which manifested itself when a pixel was classified to an unrelated ground cover type instead of one of the classes it was composed of. The third source of error was that isolated mixed pixels were often regarded as ordinary mixed pixels and vice versa; as a result, incorrect classes were considered. Pixel-based classification, finally, displayed two of the same three problems, including intrinsic misallocation of fractions and spectral confusion. However, it must be noted that spectral confusion in this case as well as in case of field-based classification had a different origin than the spectral confusion occuring during decomposition; whereas linear dependency between the endmember spectra caused spectral confusion in the latter case, the inadequacy of the pure pixel model to represent mixed pixels was the cause of spectral confusion during classification.

The positive effect of decomposition can be observed by comparing either the results of data-driven decomposition and field-based classification, or the results of pixel-based decomposition and pixel-based classification. For instance, the $<e_p>$ of data-driven decomposition for the first image was only 3.9% against 35.3% for field-based classification, while pixel-based decomposition and pixel-based classification had an $<e_p>$ of 13.1% and 38.9%, respectively. Differences in performance of the same magnitude were also found when comparing the crop area estimates for the entire region, which was rather unexpected. As can be inferred from the previous paragraph, the main reasons for these large differences were that classification, unlike decomposition, suffered from intrinsic misallocation of fractions and from spectral confusion due to the inadequate pixel model used. These sources of error also explain why the regional crop area estimates by the classification methods were so inaccurate. Suppose for instance that the natural variation of soil is much higher than that of barley; as a result, mixed pixels composed of soil and barley are much more likely to be allocated to soil than to barley. However, if the spectrum of a third endmember, e.g. grass, closely resembles a mixture of soil and barley, many pixels containing these two ground cover types will be classified as grassland. Since these errors are not random but lead to consistent overestimation of some classes and underestimation of others, the high values of e_A observed are comprehensible. When the number of classes was doubled, which resulted in the second image, the estimates became less accurate because the probability of spectral confusion greatly increased. For data-driven decomposition, however, the error per pixel increased by only 1.1% compared to 5.8% for field-based classification, which indicates that the performance of decomposition is less sensitive to the number of classes than classification. Remarkable, though, was that the regional crop area estimate of field-based classification improved at the same time. The only explanation for this decrease in e_A is that several cases of spectral confusion opposed each other, thus both underestimating and overestimating the area of the same ground cover type. Based on the comparison made in this paragraph, it can be concluded that decomposing instead of classifying mixed pixels increases the crop area estimation accuracy not only on a pixel basis, but also on a larger scale.

The influence of applying domain-specific knowledge can be determined by comparing the results of the knowledge-based methods with the results of their pixel-based counterparts. If this comparison is based on the first image again, we see that the error per pixel decreased with 9.2% when knowledge of the application domain was used by the decomposition approach, whereas $<e_p>$ decreased by 3.6% when field-based instead of pixel-based classification was applied. A comparison based on the crop area estimates for the entire region also shows that incorporation of domain knowledge had a positive effect on the accuracy. The improvement in performance was mainly due to the mechanism to select probable endmembers, which reduced the number of cases of spectral confusion possible. Data-driven decomposition resulted in a further increase in accuracy relative to pixel-based decomposition because the same mechanism also stopped the area leakage inherent to the latter approach. The performance of field-based classification, on the other hand, was somewhat hindered because the domain knowledge about isolated objects could not be applied effectively due to the pure pixel model used. According to this knowledge, isolated objects have a different ground cover type and therefore a different spectral response than the agricultural fields surrounding them. Ordinary mixed pixels, however, have a spectral response which differs from that of neighbouring pure pixels as well and which cannot be discerned easily from the response of isolated mixed pixels based on the pure pixel model used by classification. As a consequence, the threshold of field-based classification could not be set such that a clear separation between ordinary and isolated mixed pixels was made, which invalidated the mechanism to select probable endmembers for a certain number of pixels. It was probably due to the same problem that doubling the number of classes had a considerable effect on the accuracy of field-based classification. Although the application of domain-specific knowledge limited the increase of $<e_p>$ to 5.8% because the possibilities for spectral confusion were reduced, pixel-based classification showed an only slightly higher increase of 7.5%. Therefore, it can be concluded that the use of knowledge of the application domain has a positive influence on the crop area estimation accuracy as long as it can be expressed in the pixel model used.

In this section it has become apparent that decomposition of mixed pixels as well as application of domain-specific knowledge, which are both incorporated in data-driven decomposition, lead to a more accurate crop area estimation per pixel. Apart from eliminating misallocation of fractions inherent to classification, these two concepts also counteract the influence of spectral confusion. Since both sources of error result in consistent over- or underestimation of some ground cover types, the crop area estimates for regions larger than one pixel become more accurate as well. Therefore, data-driven decomposition is expected to be able to estimate the area of an agricultural field with higher accuracy than other methods that are widely used.

4 Application to Landsat-TM images

A problem of using simulated data is that the scene model needed to generate the image will always be a considerable simplification of reality. Therefore, we also carried out an experiment to determine the relative accuracy of data-driven decomposition on a real Landsat-TM image. In Section 4.1 it is described how the experiment was set up such that a quantitative analysis of the accuracy could be made with the limited information at subpixel level available. Sections 4.2 and 4.3 present the results and discuss their implications, respectively.

4.1 *Set-up*

For the site that was studied in this experiment—see Klein Gebbinck and Schouten (1996b) for a more elaborate description—the only source of subpixel information was a topographical map containing parcel boundaries. Therefore, we tried to estimate the area of each parcel as accurately as possible in order to make a quantitative evaluation of the different approaches. The methods that were applied were data-driven decomposition and field-based classification. Since both methods made an estimation of the crop area per field instead of the area per parcel, the results had to be transformed before a quantitative analysis could be made.

The test area that was selected was located near the village of Zeewolde in the south of the Flevoland polder. The land, which has been reclaimed from the IJsselmeer, was flat with an altitude of 3.30 m below sea level and was mainly used for the cultivation of arable crops. A network of roads, canals, and ditches divided the area into lots, which are basic units of approximately $500x1600$ m^2 that can be rented or bought. Most lots were subdivided into several agricultural fields, whose boundaries were determined by the farmers themselves and whose locations were subject to change every year. In most cases, however, the fields were still large enough to calculate their statistical distribution with sufficient accuracy. It were these characteristics that made the area ideally suited for our experiment.

In order to make a quantitative analysis possible, the Landsat-TM image and the corresponding topographical map were co-registered following Janssen (1994). First the boundaries of 17 agricultural lots were digitised from a map of scale 1:10,000. Next, several ground control points were identified in both the image and the digitised map, which were used to calculate the parameters of an affine transformation. With this transformation, finally, the digitised boundaries were mapped onto the image, which gave a root mean square error of only 0.283 pixel. Furthermore, it was calculated that the factor by which the boundaries were scaled was approximately 29.99, which corresponded with the 30 m resolution of Landsat-TM images very well. Based on these figures as well as on careful visual inspection of the lot boundaries superimposed on the satellite image, it was assumed that co-registration with an error in accuracy of less than one pixel was achieved.

Apart from data-driven decomposition, the only crop area estimation method that was applied was field-based classification. The pixel-based approaches were excluded for three reasons. First of all, the number of endmembers that were present in the image was to large to be handled by pixel-based decomposition. Secondly, since both pixel-based methods allocated area to ground cover types instead of to agricultural fields like data-driven decomposition and field-based classification did, calculation of the area per lot was much more complicated. The third reason was that according to the topographical data, which was used to evaluate the crop area estimators, all pixels inside the boundaries of a lot belonged entirely to that lot; because this view is shared by data-driven decomposition and field-based classification with regard to agricultural fields, these methods were favoured during evaluation, while the pixel-based approaches were not. Since both methods that were applied needed the image to be segmented, a perfect segmentation could be used without biassing their relative accuracy. Therefore, the image was segmented by extending the lot boundaries with field boundaries by means of on-screen digitising. After vector-to-raster conversion using a sufficient margin to account for co-registration errors and positional inaccuracy of the field boundaries, a segmentation of high quality was available. With the segmentation thus acquired, application of data-driven decomposition and field-based classification was straightforward.

To determine the performance of the crop area estimators, some post-processing of their results was required. Since the only supervised information at our disposal was the area of the 17 lots that were digitised, the crop area estimates per field had to be aggregated to lot area estimates before a comparison could be made. The assumption behind this comparison was that the outer boundaries of the agricultural fields within a lot coincided precisely with the boundaries of the lot itself. In that case, the area of a lot could be found by summing the area of all the fields situated within the lot. However, care had to be taken that the estimated lot area was compensated for the area allocated to narrow boundary structures inside a lot, as these structures were not present on the topographical map. The accuracy of the estimates obtained in this way was measured using the same metric already presented in Section 3.1:

$$e_A = \frac{1}{2} \sum_i \left| \hat{A}_i - A_i \right| . \tag{6}$$

Although this accuracy measure appears to be similar, it expresses an error that is slightly different due to the different meaning of A_i and $\hat{A}_i$: whereas in Equation (5) A_i and $\hat{A}_i$ represented the true and estimated area covered by the i-th crop, respectively, in Equation (6) they denote the true and estimated area of the i-th lot. So, after post-processing the results, the performance of the crop area estimators could be determined quantitatively using a simple accuracy measure.

The experimental set-up described above is a way to assess the accuracy of the crop area estimators on a real Landsat-TM image with the limited supervised information at subpixel level that was available. However, because our topographical map displayed land ownership instead of land cover, the accuracy could be determined only indirectly. Based on this supervised information, it was also impossible to determine the accuracy of the pixel-based approaches without bias. Therefore, only data-driven decomposition and field-based classification were applied so that a comparison of their performance would indicate the influence of decomposing mixed pixels.

4.2 *Results*

The results presented in this section were obtained on a real Landsat-TM image of the Zeewolde test area. The image, which is displayed in Figure 2, was acquired on 24 May 1990 and was free of clouds. However, because the date of acquisition was rather early in the growing season, only fields covered by cole-seed, alfalfa, winter wheat, and grass were found, while the others appeared to lie fallow. The narrow boundary structures present were assumed to be composed of water, forest, grass, soil, and asphalt; therefore, each of these classes was characterised by upto three different distributions extracted from the image by a remote sensing expert in order to serve as endmembers for data-driven decomposition and field-based classification. Application of these methods gave the results shown in the second and third super-column of Table 3. The threshold setting with which these results were obtained was chosen such that e_A was minimal. While the optimal setting could be determined unambiguously in case of data-driven decomposition, for field-based classification the choice was less clear because e_A as a function of the threshold had many local minima that were close to the global minimum. However, even the lowest value of e_A achieved by field-based classification was still three times as high as the best value achievable by data-driven decomposition. More important, though, was that the decomposition approach gave a more accurate area estimate than classification for 13 of the 17 lots, while in one case (lot 4) the estimates by the two methods differed by less than one pixel. Overall, data-driven decomposition appeared to be better suited for estimation of the lot area than field-based classification.

The experiment was repeated under the assumption that soil was not one of the possible components of the narrow boundary structures. We took this step because field-based classification in particular obviously underestimated the area of nearly all lots that were largely covered by bare soil. Apparently this approach allocated considerable fractions of these lots to narrow boundary

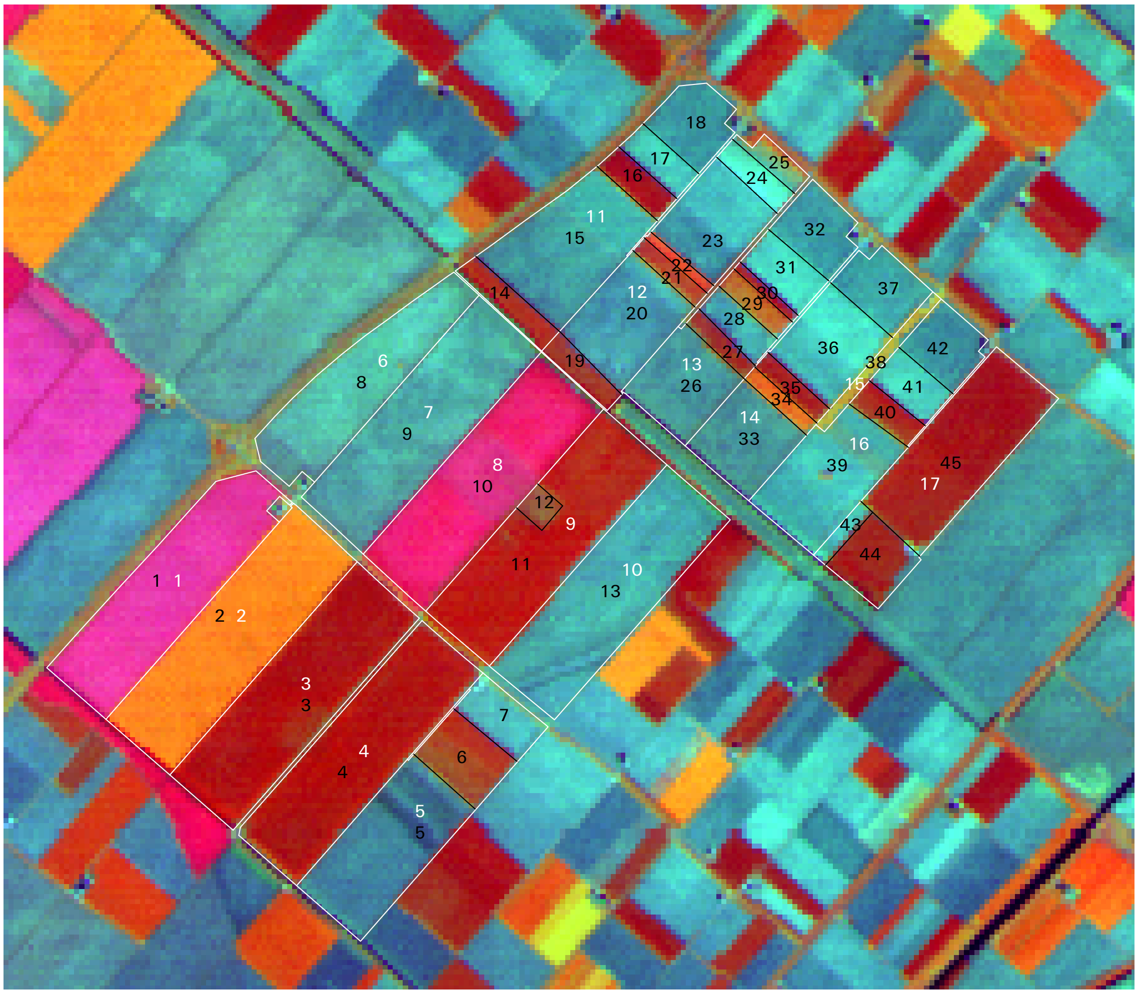


Figure 2: False colour composite (R/G/B: TM-band 4/5/3) of the Zeewolde study site with superimposed boundaries. Lot boundaries (numbers) are given in white, field boundaries (numbers) are shown in black. Visual inspection suggests a co-registration error of less than one pixel.

	topographical map		decomposition		classification incl.		classification excl.	
lot	main class	area	area	deviation	area	deviation	area	deviation
1	cole-seed	825.6	840.7	15.1	830	4.4	830	4.4
2	alfalfa	945.0	953.6	8.6	904	-41.0	904	-41.0
3	winter wheat	928.8	933.4	4.6	954	25.2	954	25.2
4	winter wheat	888.7	886.8	-1.9	887	-1.7	887	-1.7
5	bare soil	932.8	932.7	-0.1	904	-28.8	930	-2.8
6	bare soil	617.8	624.2	6.4	590	-27.8	631	13.2
7	bare soil	886.1	866.8	-19.3	810	-76.1	845	-41.1
8	cole-seed	890.7	862.6	-28.2	895	4.3	895	4.3
9	winter wheat	886.4	898.7	12.3	894	7.6	894	7.6
10	bare soil	879.8	884.0	4.2	864	-15.8	914	34.2
11	bare soil	1055.9	1055.5	-0.4	1013	-42.9	1037	-18.9
12	bare soil	1004.6	1009.8	5.2	969	-35.6	999	-5.6
13	bare soil	880.8	879.2	-1.6	840	-40.8	871	-9.8
14	bare soil	901.2	887.4	-13.8	847	-54.2	877	-24.2
15	grassland 3	76.9	75.1	-1.9	84	7.1	94	17.1
16	bare soil	826.1	835.9	9.8	795	-31.1	814	-12.1
17	winter wheat	883.6	883.8	0.2	882	-1.6	897	13.4
e_A		–		66.7		222.9		138.3

Table 3: Lot area estimates by data-driven decomposition and field-based classification for the Zeewolde study site. The sub-column "main class" shows the dominant cover type according to a remote sensing expert. The super-column suffixes "incl." and "excl." indicate whether soil was included in or excluded from the set of possible components of boundary structures during classification. The unit of all values is the area of a single pixel.

structures composed of soil, which probably did not even exist in reality. The results of field-based classification after excluding soil as a possible boundary component are presented in the last columns of Table 3; the results of data-driven decomposition are not shown as they deteriorated over the entire line. As expected, field-based classification underestimated the area of most lots covered by soil less severely or even overestimated them; unfortunately, the area of two lots (numbers 15 and 17) whose main ground cover types were not soil was (further) overestimated as well. As a result, e_A fell from 222.9 to 138.3, which was only twice the value achieved by data-driven decomposition. A qualitative comparison of the error per lot, however, showed virtually no improvements: the decomposition approach was still more accurate in 12 cases, approximately as accurate (± 1 pixel) in two cases, and less accurate in only three cases compared to the classification method. To summarise, excluding soil as a possible component of narrow boundary structures did improve the performance of field-based classification, but in general it did not lead to more accurate area estimates than those of data-driven decomposition.

4.3 *Discussion*

Application of data-driven decomposition and field-based classification to the Zeewolde image showed a big difference in performance. However, since the way in which the results were evaluated was not straightforward, the factors that caused these differences have to be investigated first. Only after this, it may be possible to draw conclusions about the accuracy of the crop area estimators.

In addition to the three sources of error already identified in the experiment with simulated images, there appeared to be one other cause of inaccuracy for field-based classification. The first source of error found in Section 3.3 was the misallocation of fractions that is inherent to the classification approach. Although the lack of supervised information about the composition of individual pixels prevented its verification, it probably did decrease the accuracy of the area estimates. Spectral confusion, on the other hand, was a major source of error that could be confirmed with the help of a topographical map. Since field-based classification found many boundary structures composed of asphalt where, according to the map, canals, ditches, or tree rows were situated, the occurrence of spectral confusion was evident (Klein Gebbinck and Schouten 1996b). The third source of error, which was caused by the difficulty to distinguish ordinary from isolated mixed pixels, was once more impossible to verify due to the nature of the supervised information. However, since visual inspection of the image did not reveal any isolated objects in the area of interest, the effect on the estimation accuracy was probably small. The last cause of inaccuracy, which was already discussed during the presentation of the results of this experiment, concerned the status of the ground cover type soil. If soil was regarded as a possible component of narrow boundary structures, the area of lots covered by soil was underestimated, while many boundary structures of soil were found that did not exist in reality. The reason for this problem was that the classification approach was not sensitive enough to the slightly different distributions of soil to allocate many pixels composed of soil to the correct object. Although excluding soil as a possible boundary component improved the performance, it also resulted in overestimation of a few lots and the underestimation of several narrow boundary structures of soil. Therefore, it was concluded that soil was a component of some boundary structures after all and that the underestimation of lots of soil simply was another case of spectral confusion. To summarise, the inaccuracy of field-based classification was caused by the same three factors that were identified when processing simulated images.

Apart from the two sources of error found in the simulation experiment, the accuracy of data-driven decomposition may be affected by other factors, which may become apparent by examining the lots where the decomposition approach leads to less accurate estimates than field-based classification. The first cause of minor inaccuracy, which was the natural variation of a ground cover type, could not be confirmed with the supervised information available. However, the occurrence of spectral confusion, which was the second cause, was verified using a topographical map. Although the estimated composition of the narrow boundary structures agreed with the map much better than the estimates by field-based classification, data-driven decomposition did experience some problems with structures composed of water in particular (Klein Gebbinck and Schouten 1996b). To determine whether there were any other sources of error, we investigated lots 1, 8, and 9 to find out why the decomposition approach gave less accurate results than the classification method. The problem with lot 1, whose area was overestimated by both methods, was that the pixels at the south-west boundary partially covered an undigitised field of the same crop. As a result, the crop area of lot 1 was overestimated during decomposition; the classification approach, on the other hand, allocated these pixels to an incorrect class due to spectral confusion, thus decreasing the already overestimated area. The area of lot 8 was severely underestimated by data-driven decomposition and slightly overestimated by field-based classification. A close look at its north-east boundary, however, revealed an uncharted strip of bare soil, which invalidated the requirement that field and lot boundaries should coincide. Consequently, underestimation of the lot area indicated a more accurate crop area estimate. The area of lot 9 was overestimated for about the same reason as lot 1. Because of spectral confusion, a number of pixels situated north-east of the lot's boundary were incorrectly allocated to this lot. But whereas data-driven decomposition assigned fractions of the boundary pixels themselves to lot 9 as well, field-based classification allocated them to an incorrect class altogether, which compensated for the overestimation mentioned before. Therefore, after investigating the three cases in which decomposition was less accurate than classification, we conclude that spectral confusion is still the only major source of inaccuracy of data-driven decomposition, although the way in which the results were evaluated may obscure this observation.

Based on this experiment it is not possible to draw any conclusions about the absolute accuracy of data-driven decomposition and field-based classification, but their relative accuracy can be determined. A problem for the evaluation of the accuracy was that the available supervised data concerned the area owned or rented by a farmer, whereas the crop area estimators tried to calculate the area covered by each crop. These two types of information could only be compared by assuming that the crops cultivated by a farmer grew precisely to the boundaries of his or her land, i.e. the outer field boundaries coincided exactly with the lot boundaries. Since this assumption was not always true, e.g. lot 8, the absolute accuracy of the estimators could not be determined. However, as both area estimation methods suffered from the same mismatch between the supervised information and the area estimates, the accuracy of their estimates could be compared. According to this comparison, data-driven decomposition was more accurate in 13 (12 excl. soil) of the 17 cases, while the three cases in which field-based classification seemed to be more accurate could all be explained. Overall, this difference in performance led to an e_A of 66.7 for data-driven decomposition, while the same error measure gave a value of 222.3 (138.9 excl. soil) for field-based classification. Based on these figures, it is reasonable to conclude that data-driven decomposition can estimate the crop area of the agricultural fields of the Zeewolde study site more accurately than field-based classification.

The above conclusion probably still holds when surveying agricultural regions other than the Zeewolde area. The main reason for this conjecture is that both the order of best crop area estimators and their underlying sources of error were the same in the experiments with the simulated images as well as the Zeewolde image, which may suggest that this trend will also be found when processing other images. However, the characteristics of these scenes should not differ too much from those of the Zeewolde area: whereas a smaller average field size will probably still be acceptable, large variations in terrain altitude may well be not. Under this condition, therefore, we believe that the decomposition of mixed pixels in combination with the application of domain knowledge increases the accuracy of the crop area estimates derived from a real Landsat-TM image.

5 Conclusions

The purpose of our study was to develop a method to estimate the area of an agricultural field with higher accuracy than traditional methods such as pixel- and field-based classification and pixel-based decomposition. The main problem of all classification approaches is that they are unsuited to handle mixed pixels that occur, for instance, at the field boundaries, because they assume that every pixel represents only one ground cover type. Therefore, we developed the data-driven decomposition method, which applies the same linear mixture model used by pixel-based decomposition to unmix pixels composed of more than one ground cover type such that each pixel can be allocated to several classes. The other main concept of data-driven decomposition is the application of domain-specific knowledge based on a scene model comprising agricultural fields, narrow boundary structures, and isolated objects. One aspect of this knowledge is that the spatial relationships between the pixels of an agricultural field can be used to locate mixed pixels that should be decomposed, to determine local endmember distributions from the image, and to select endmembers that are probable components of a mixed pixel. An important advantage of this endmember selection mechanism is that it permits the decomposition of a scene into a far greater number of endmembers than can be considered by pixel-based decomposition. Further aspects incorporated in our approach are that fields are often bounded by narrow structures with well-known compositions, e.g. roads and ditches, and that isolated objects, which are composed of ground cover types not found in their surroundings, can be detected because their spectral response is not a mixture of the responses of neighbouring pixels. Since all these knowledge rules help to improve the correct allocation of mixed pixels, the crop area estimates by data-driven decomposition were expected to be more accurate than those of traditional crop area estimators.

In order to evaluate the two concepts behind data-driven decomposition, it was applied to simulated Landsat-TM images. After a quantitative comparison of the results of our approach and the estimates by three other well-chosen traditional methods, it appeared that both decomposition of mixed pixels and the use of domain-specific knowledge had a positive influence on the accuracy of the area estimates. The principal merits of decomposing mixed pixels were that misallocation of fractions, which is inherent to the classification approach, was resolved and that spectral confusion was considerably reduced. The main advantage of using domain knowledge was that it led to a further reduction in spectral confusion. The resulting increase in accuracy was found not only for individual pixels but also for regional crop area estimates. The explanation for this somewhat surprising result was that the errors due to spectral confusion as well as intrinsic misallocation were not random but caused consistent overestimation of some ground cover types and underestimation of others. Since data-driven decomposition counteracted both sources of error, its crop area estimates were more accurate than those of the other methods for individual pixels, agricultural fields, and entire regions.

The data-driven decomposition method was also evaluated in an experiment with real Landsat-TM data. Based on the supervised information available, it was impossible to measure the accuracy of the estimator in an absolute sense. However, the accuracy of our approach could be compared quantitatively with the performance of other methods in an indirect way. A comparison between data-driven decomposition and a similar method using classification showed that the former approach estimated the area of agricultural lots more accurately in 13 of the 17 cases, while the latter approach seemed to be more accurate in three cases due to the incompleteness of the supervised data or coincidentally opposite errors. The main source of error that could be determined with the available supervised information was spectral confusion, although misallocation of fractions inherent to classifying mixed pixels probably still was a major cause of inaccuracy. Based on these results, which confirmed the findings of the experiment with simulated data, it can be concluded that data-driven decomposition is able to estimate the area of agricultural fields with higher accuracy than many other crop area estimation methods that are used today.

6 Acknowledgments

We would like to thank Gerard Nieuwenhuis, head of the Remote Sensing Department of the DLO Winand Staring Centre for Integrated Land, Soil and Water Research, for the opportunity to carry out the experiments on the Zeewolde image, and Edwin Noordman and Henk Kramer of the same department for their help with the preparation of the data set.

References

ADAMS, J. B., SMITH, M. O., and JOHNSON, P. E., 1986, Spectral mixture modelling: a new analysis of rock and soil types at the Viking Lander 1 site. *Journal of Geophysical Research*, 91, 8098–8112.

CARLOTTO, M. J., 1995, Non-linear mixture model and application for enhanced resolution multispectral classification. *Proceedings of the 1995 International Geoscience and Remote Sensing Symposium* (Piscataway, New Jersey: I.E.E.E.), pp. 1168–1170.

CROSS, A. M., SETTLE, J. J., DRAKE, N. A., and PAIVINEN, R. T. M., 1991, Subpixel measurement of tropical forest cover using AVHRR data. *International Journal of Remote Sensing*, 12, 1119–1129.

DRAKE, N. A., and SETTLE, J. J., 1989, Linear mixture modelling of Thematic Mapper data of the Peruvian Andes. *Proceedings of the EARSeL symposium held in Helsinki, Finland, on 27 June–1 July 1989* (Luxembourg: Commission of the European Communities), pp. 490–495.

FOODY, G. M., and COX, D. P., 1994, Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions. *International Journal of Remote Sensing*, 15, 619–631.

FOSCHI, P. G., 1994, A geometric approach to a mixed pixel problem: detecting subpixel woody vegetation. *Remote Sensing of Environment*, 50, 317–327.

GALLEGO, F. J., DELINCE, J., and RUEDA, C., 1993, Crop area estimates through remote sensing: stability of the regression correction. *International Journal of Remote Sensing*, 14, 3433-3445.

HORWITZ, H. M., NALEPKA, R. F., HYDE, P. D., and MORGENSTERN, J. P., 1971, Estimating the proportions of objects within a single resolution element of a multispectral scanner. *Proceedings of the 7th International Symposium on Remote Sensing of Environment* (Ann Arbor, Michigan: Environmental Research Institute of Michigan), pp. 1307–1320.

JANSSEN, L. L. F., 1994, *Methodology for Updating Terrain Object Data from Remote Sensing Data*. Ph.D. Dissertation (Wageningen, the Netherlands: Wageningen Agricultural University).

KLEIN GEBBINCK, M. S., and SCHOUTEN, T. E., 1996a, *Accurate Area Estimation by Data-Driven Decomposition of Mixed Pixels*. CSI Report No. CSI-R9622 (Nijmegen, the Netherlands: Computing Science Institute, University of Nijmegen).

KLEIN GEBBINCK, M. S., and SCHOUTEN, T. E., 1996b, *Application of Data-Driven Decomposition to Landsat-TM Images for Crop Area Estimation*. CSI Report No. CSI-R9626 (Nijmegen, the Netherlands: Computing Science Institute, University of Nijmegen).

MARSH, S. E., SWITZER, P., KOWALIK, W. S., 1980, Resolving the percentage of component terrains within single resolution elements. *Photogrammetric Engineering and Remote Sensing*, 48, 1079–1086.

MATHER, P. M., 1990, Theoretical problems in image classification. In *Application of Remote Sensing in Agriculture*, edited by J. A. Clark and M. D. Stevens (London: Butterworths), pp. 127–135.

PECH, R. P., DAVIES, A. W., LAMCROFT, R. R., and GRAETZ, R. D., 1986, Calibration of Landsat data for sparsely vegetated semi-arid rangelands. *International Journal of Remote Sensing*, 7, 1729–1750.

QUARMBY, N. A., TOWNSHEND, J. R. G., SETTLE, J. J., WHITE, K. H., MILNES, M., HINDLE, T. L., and SILLEOS, N., 1992, Linear mixture modelling applied to AVHRR data for crop area estimation. *International Journal of Remote Sensing*, 13, 415–425.

SCHOENMAKERS, R. P. H. M., 1995, *Integrated Methodology for Segmentation of Large Optical Satellite Images in Land Applications of Remote Sensing*. Ph.D. Dissertation (Nijmegen, the Netherlands: University of Nijmegen).

SCHOUTEN, T. E., and KLEIN GEBBINCK, M. S., 1995, Quality measures for image segmentation using generated images. *Proceedings Image and Signal Processing for Remote Sensing II* (Washington: S.P.I.E.), pp. 411–422.

SETTLE, J. J., and DRAKE, N. A., 1993, Linear mixing and the estimation of ground cover proportions. *International Journal of Remote Sensing*, 14, 1159–1177.