# An Improved Methodology for Detecting Irrigation in Plastic-Covered Crops by Using EO Data

S. Montesinos-Aranda[1] and C. Santos-González

*[1] GEOSYS, S.L., Sector Foresta 23. Locales 7 y 8. Tres Cantos (Madrid). E-mail: smontesinos@geosys.es*

***Abstract***

*During the last two decades Huelva Province (Spain) has experienced a spectacular increase in the surface area occupied by plastic-cover crops (PCC's). This relatively new farming technique is utilised mainly to grow strawberries. As this crop requires irrigation, the water consumption has risen dramatically. Due to the limited water resources of the area, it has become necessary to monitor how much water is used.*

*The present work presents a methodology to detect and quantify the surface area devoted to PCC's by means of remote sensing techniques. An index based on the bands of the sensor Landsat TM and ETM+ was developed to highlight this type of crop. Through a hierarchical unsupervised classification applied over the index image the extent of the PCC's was quantified. This approach proved to be accurate and robust (overall classification accuracy: 97.4%) and more reliable than a supervised classification. However, it was also proved that an appropriate date selection for the image is of paramount importance to apply the index successfully. November, December and January showed to be the best months to discern PCC's with satellite images.*

***Keywords:*** *Plastic-covered crops (PCC's), remote sensing, index, Landsat, classification*

## 1. INTRODUCTION

Plastic-covered crops started proliferating in Spain in the 60's, mainly in the Southern part of the country. The growth of the surface covered by plastic has been increasing ever since. In 1967 there were about 1,300 Has. of plastic-covered crops (PCC's), by the mid-90's that figure had risen to 170,000 Has (90,000 Has. of plastic mulch, 30,000 Has. of greenhouses and 17,000 of plastic tunnels).

In Huelva Province the plastic boom started in the 70's, occupying marginal lands where traditional agricultural practices were not feasible. From the very beginning, plastic-agriculture has been very mono-specific in Huelva, being strawberries the most important crop (only 1 per cent of the surface devoted to this type of agriculture is not used for strawberry) (Moreira *et al.*, 1991).

Irrigation of plastic-covered crops has implied a very important increase in water consumption in this area, as irrigated agriculture did not occur before in the land currently occupied by PCC's. This means a serious threat to the scarce freshwater resources of this territory, where water extraction is not monitored and controlled at the moment.

Controlling the water use in PCC's makes necessary to estimate the extent of the surface area taken up by these crops. Traditional methods such as fieldwork or photo-interpretation are expensive and very time-consuming. Remote sensing techniques can be an invaluable alternative for this task, reducing costs and work time. Among the different available sensors, Landsat TM and ETM+ are thought to be very suitable sensors to monitor PCC's in Huelva due to the broad spatial extent of their scenes, spectral features of their bands and revisit time. These characteristics allow studying wide and remote areas, differentiating PCC's from other land covers and perform multi-temporal analysis.

The present work aims at developing a methodological approach to detect PCC's and estimate the surface area they cover by using Earth Observation techniques. A test site was chosen within the municipalities of Palos de la Frontera and Moguer, where PCC's present a very high density. As mentioned above, Landsat TM and ETM+ were the sensors chosen for this study. Two approaches are presented and compared. The first one is based on the use of an index made up of a combination of bands. The second one uses a supervised classification.

## 2. MATERIALS AND METHODS

### 2.1 Study area

The test site selected for this study is located in Huelva province, within the coastal municipalities of Moguer and Palos de la Frontera, two of the main strawberry producers in Spain. The geographic co-ordinates of the upper left and lower right corners of the study area are (lat. 37º13', long. 0º-54') and (lat. 37º8', long. 0º-46'), respectively.

The climate is temperate Mediterranean, characterised by dry and hot summers, and mild winters with irregular precipitation. Pine tree forests and Mediterranean bushland are the most abundant vegetation types. In this area soils are extremely sandy, which has stopped agriculture from settling for decades. When forced crops were introduced in the 70's, agriculture could take over extensively and natural vegetation was cleared to set PCC's. The availability of so much marginal unoccupied land and abundance of communal woodlands favoured this process.

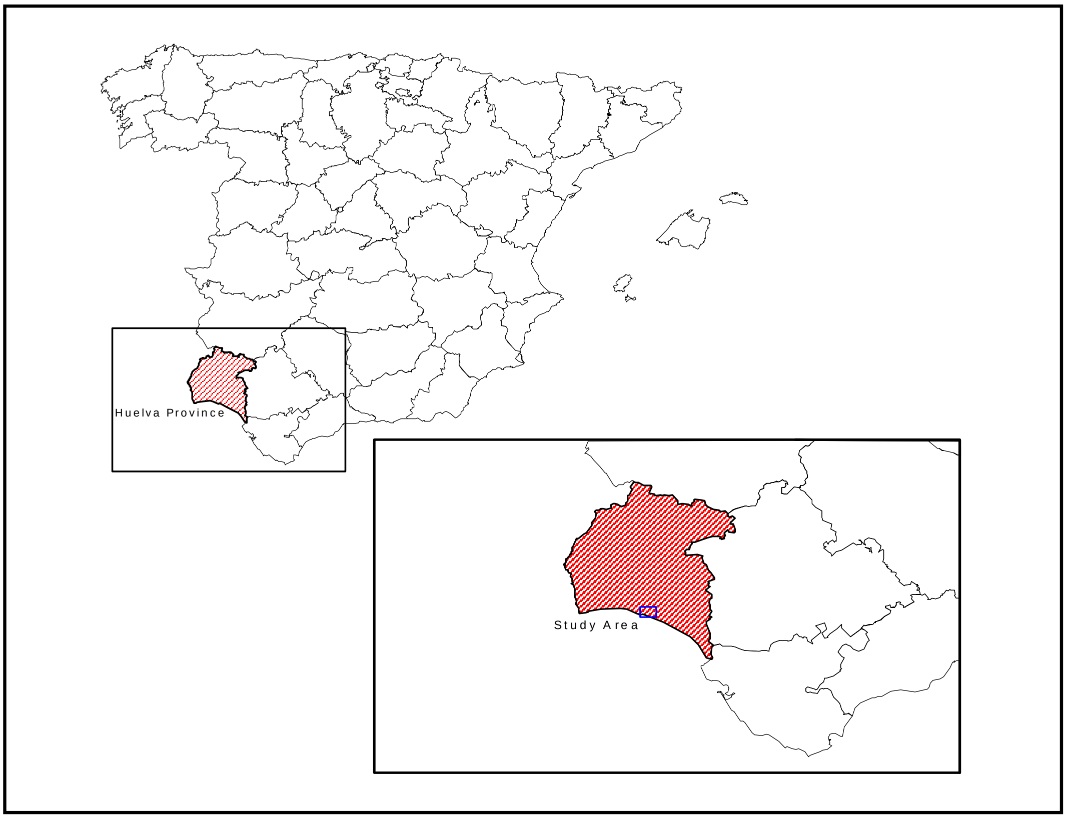


***Figure 1.*** *Study area.*

### 2.2 Plastic-covered Strawberry Farming

In Huelva, strawberry production is totally linked to plastic farming. This technique differs greatly from that used in greenhouses for other crops. Below, plastic-covered strawberry farming is described briefly.

- ***Land preparation and disinfection***

Land preparation aims at giving the plant a proper physical support to vegetate and produce during the campaign. This operation takes place during August and September.

The basic operations are:

- Superficial ploughing to break up the land.
- Application of organic matter, limestone amendments and fertiliser.
- Abundant sprinkling irrigation before disinfecting.
- Disinfectant injection (operation not necessary every year).
- After 3-5 days, the land is ploughed for ventilation and irrigated to eliminate residues from the disinfectant.
- Ridging and mulching (September-early October). Land ridges are built and padded with black plastic mulch (polyethilene). The drip irrigation system is installed along the ridges and the polyethilene is perforated for plantation. The impermeability of this material stops water from evaporating from the soil. Besides, the temperature where the roots will be allocated increases and weeds do not grow.

- ***Transplanting***

Plantation is done between early and mid October, although sometimes it is done in late September. Post-plantation caring is limited to sprinkling or micro-sprinkling irrigation and drip irrigation to reduce the temperature of the ridge.

- ***Plastic Tunnel lying***

To achieve earlier crops, sometimes a complementary protection system that acts as a greenhouse is utilised. This operation takes place between late November and December. There are two techniques:

- Micro-tunnels or tunnels: They have a structure made up of arches that support a layer of transparent plastic or thermal polyethilene. They just cover a single ridge.
- Macro-tunnels: They have the same type of structure as the tunnels, but they are higher (3 m) and cover several ridges.

This plastic layer is removed from March.

## 2.3 Satellite Images

Ten Landsat images were used in this study. After radiometric calibration they were geo-corrected and geo-referenced using a $2^{nd}$ degree polynomial transformation.
Both Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) are multispectral scanning radiometers carried on board of Landsat 4 and Landsat 5, and Landsat 7, respectively. The TM sensors have provided nearly continuous coverage from July 1982 to June 2001, with a 16-day revisit period. The ETM+, the only one operational currently, has been gathering imagery since July 1999 to present with the same re-visit period as TM.

Both instruments provide image data from eight spectral bands. The spatial resolution is 30 meters for the visible and near-infrared (bands 1-5 and 7). Resolution for the panchromatic (band 8) is 15 meters, and the thermal infrared (band 6) is 60 meters for ETM+ and 120 meters for TM . The approximate scene size is 170 x 183 kilometres (106 x 115 miles).

For this study, only the 30m-resolution bands were used. Table 2 shows the acquisition date and type of sensor for each utilised image.

***Table 1.*** *General features of the Landsat TM and ETM+ sensors.*

| Band Number | Spectral Range (nm) | Spatial resolution (m) |
|---|---|---|
| 1 | 0.450 – 0.515 | 30 |
| 2 | 0.525 – 0.605 | 30 |
| 3 | 0.630 – 0.690 | 30 |
| 4 | 0.750 – 0.900 | 30 |
| 5 | 1.550 – 1.750 | 30 |
| 6 | 10.400 – 12.500 | 60 (Landsat 7), 120 (Landsat 5) |
| 7 | 2.090 – 2.350 | 30 |
| Pan | 0.520 – 0.900 | 15 |

***Table 2.*** *List of images utilised in this work.*

| MONTH/YEAR | SENSOR | ACQUISITION DATE |
|---|---|---|
| January | Landsat ETM+ | 07/01/2002 |
| February | Landsat TM | 01/02/1985 |
| March | Landsat TM | 09/03/1998 |
| April | Landsat TM | 04/04/1996 |
| May | Landsat TM | 11/05/1986 |
| June | Landsat TM | 23/06/1996 |
| July | Landsat TM | 09/07/1996 |
| August | Landsat TM | 10/08/0996 |
| September | Landsat TM | 21/09/1988 |
| October | --- | --- |
| November | --- | --- |
| December | Landsat ETM+ | 17/12/1999 |

## 2.4 Date Selection

Since PCC's are seasonal and they are not present throughout the year, it is necessary to chose carefully the time of the year when detection using satellite images should be optimal. From the information given in section 2.2 and visual survey of the above-mentioned satellite images, the strawberry crop calendar can be outlined. There is plastic presence (of one type or another) on the crops from August to March, or even April.

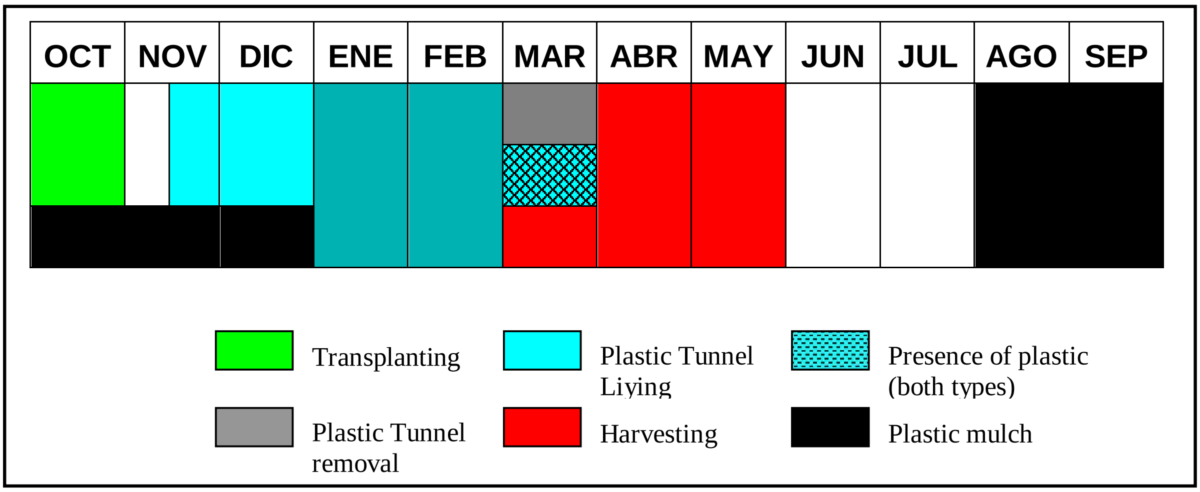


***Figure 2.*** *Phenology and timing of strawberry farming.*

A visual inspection of several images was performed in order to check what the study area looks like throughout the year. To do so, false colour composites (RGB) were created using bands 7-3-1 to best discern plastic, and 4-5-3 to discriminate vegetation. A false colour composite is a combination of three bands where each of them is put through a colour channel (Red, Green and Blue) and then combined to create a colour image.

In August (beginning of the strawberry farming season), it was appreciated that a very small part of the area devoted to PCC's is covered with plastic. The image from September was taken in 1988. This part of Huelva was not holding PCC's extensively at that time. The image does not show much plastic presence, but it hard to give an opinion whether this lack of plastic cover is due to the time of the year or the year itself. October and November could not be assessed as images were not available, but it is thought plastic should be present and visible in those months. In December and January plastic is really evident in the Landsat images, showing a clear contrast with the surroundings. The image from February was acquired in a very early year (1985) and it is not possible to appreciate a wide plastic extent. However, it is also thought to be a good month to detect PCC's. March and April still show plastic on the land, but in the 4-5-3 composites it is possible to appreciate vegetation presence in the parcels with plastic. That implies that the strawberry plants have developed and both spectral signatures (plastic and strawberry plants) are mixing. This can be proved with a spectral analysis (see section 2.5). Those months would not be optimal for plastic detection as maximum pixel purity is being sought. Therefore it can be concluded that the best months to detect PCC's by means of Landsat images are November, December and January, although October and February may be suitable as well. For the completion of this study it was selected an ETM+ image from December 1999

## 2.5 Spectral Analysis and index creation

To generate a index, first it is necessary to study the spectral signatures of the different land covers present in the study area in order to select the most appropriate and discriminating band combination. The spectral signatures of six land covers (PCC's, sandy soils, ocean water, inland water, vegetation and built-up areas). were extracted from the images of December 1999, January 2002, and March 1998. Figure 3 shows these spectra.

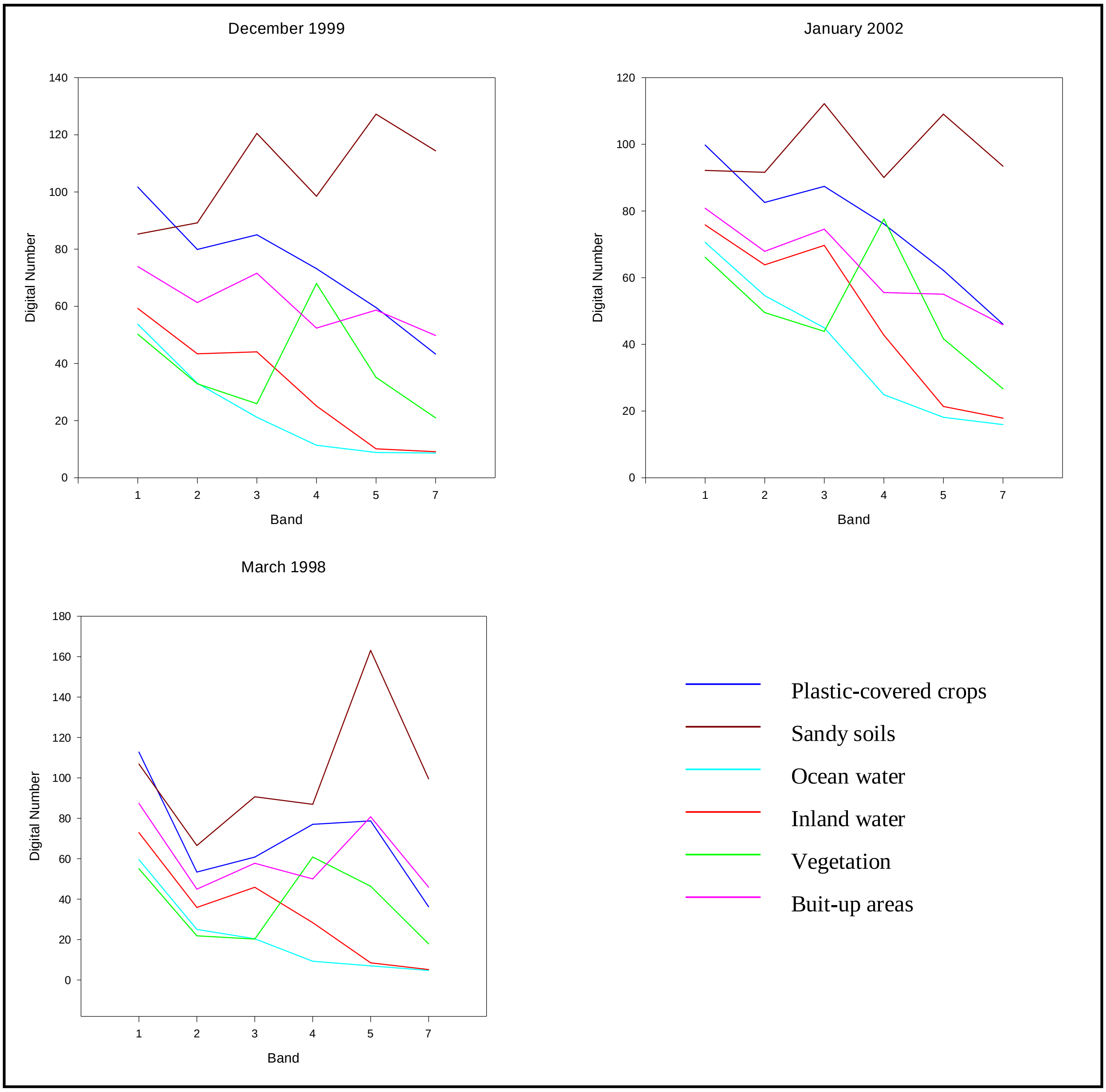


***Figure 3.*** *Spectra of different land covers from December 1999, January 2002 and March 1998.*

It is possible to observe that the plastic spectra from December 1999 and January 2002 are quite similar, whereas the spectrum from March 1998 differs from them in several features. It is specially interesting to notice the change of values in bands 3 and 4. During the winter months the values of band 3 are higher than those of band 4 in PCC's. On the contrary, in March the trend shifts and values in band 4 are higher, which is a very characteristic feature of vegetation. This confirms that vegetation influences plastic spectral response and March is not a good month to assess plastic, as it was mentioned in section 2.4. Therefore, the index was developed studying the spectra from December and January.

An index aims to highlight a parameter or quality with the highest values, consequently a combination of bands that yields the highest values for PCC's is sought.

Band 1 shows the highest values for PCC's, but sand is also very reflective in this wavelength and can behave in the same way as plastic in certain cases. By subtracting band 7 from band 1 (band 1 – band 7), sand does not mix with plastic any more, but water does. To overcome this, (band 1 – band 7) is multiplied by band 1 and band 4, which is the final band combination. Finally, the index image is converted to a 8-bit image to compress the wide range of values into a range of 256 values, so that it is possible to have a better idea of what high, medium or low values are.

Plastic-Covered Crops Index (PCI):

$$PCI = (band1 - band7) \times band1 \times band4$$

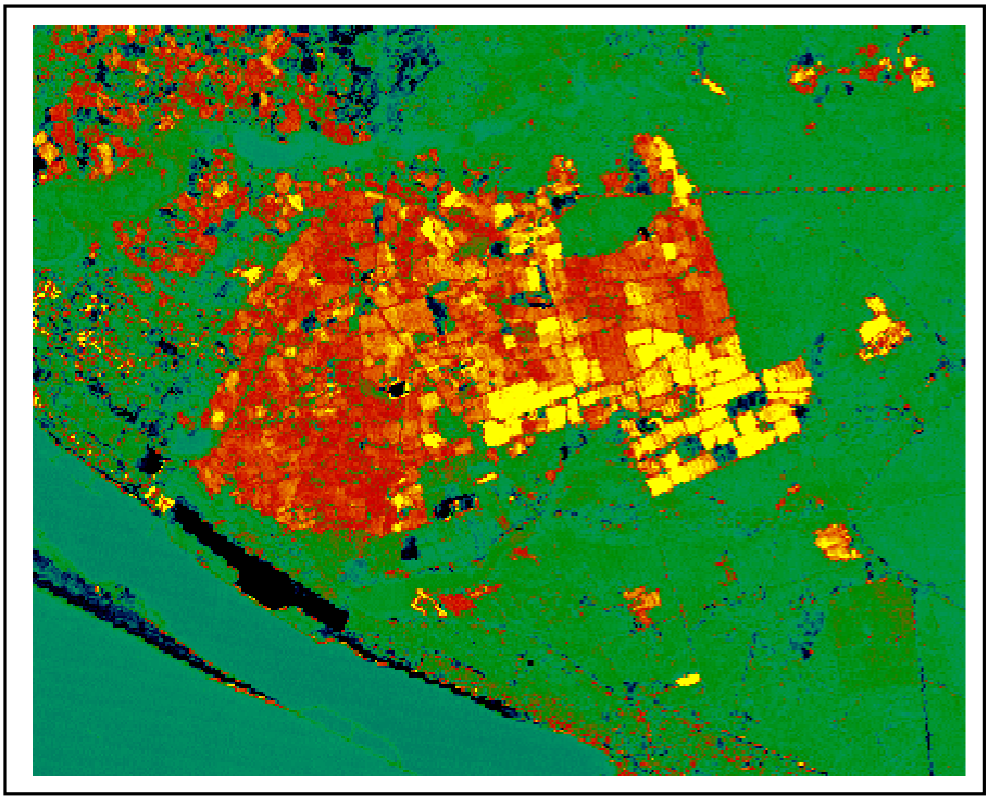

***Figure 4.*** *Coloured image derived from the PCI. Red-yellow colours represent high values, PCC's in this case, whereas green-blue colours represent other land covers.*

### *2.5.1 Unsupervised Classification*

An index just assigns values to pixels. In this case it has been attempted to give high values to PCC's and lower values to the other land covers. But the index does not say what threshold value should be used to consider pixels as PCC's. Therefore, a further step must be taken to threshold the index image. This action could be done manually by an experienced operator, but an automatic hierarchical unsupervised classification is considered to be a more accurate way to do it, as it is based on a statistical approach.

Unsupervised classifications are used to separate spectral classes when there is insufficient observational or documentary information about the nature of the land cover types of the study area. In this situation, it is not possible to estimate the mean centres of the classes (like in the case of a supervised classification), or even know the number of classes. In an unsupervised classification the operator sets the number of classes s/he wants as classification output. The members of each class are drawn from separate bivariate-normal distributions. It is assumed that there are n groups (classes) of pixels which centres are not known in the feature space. In this study the *K-means* clustering algorithm was used (Tou & González, 1974). It generates the class centres (means) arbitrarily. Then, iteratively, pixels are grouped within the closest class by using a *minimum distance* technique. In each iteration, the class means are recalculated and pixelas are re-classified respect to the new means. All the pixels are classified unless a standard deviation threshold distance is specified. In that case some pixels may not be classified if they do not satisfy the selected criteria. This process goes on until the number of pixels in each class changes by less than an established threshold or the maximum iteration number set is reached.

To best understand the classification process and class generation, a hierarchical approach was applied. It consists in generating an initial two-class classification, from which the class that do not contain the land cover/s of interest is masked. Next, the class/es of interest undergo a further n-class classification. This process continues until the operator considers the land cover/s subject of study have been fully isolated.

In the present work, it was only necessary a two-class classification to discriminate PCC's in the image from December 1999.

### *2.5.2 Supervised Classification*

An alternative approach to discern PCC's is using a supervised classification. Unlike an unsupervised classification, this method is based on external knowledge about the land covers of the study area. The supervised methods require the input of this information before the chosen classification algorithm is applied. This input is derived from fieldwork, photo-interpretation, or any other source of information. There are several classification algorithms, and all of them require the specification of the number of classes and certain statistical parameters in advance. These estimates are extracted from samples of pixels, called *training areas*. Each *training area* is a user-defined group of pixels that falls into a known land cover type.

In this work, the algorithm *Maximum Likelihood* was used (Richards, 1994). It assumes that the statistics for each class in each band are normally distributed and calculates the probability that a pixel belongs to a given class. Unless a probability threshold is selected, all pixels are assigned to the class with the highest probability (i.e. maximum likelihood).

Training areas were generated from seven different land uses: transparent or white plastic, plastic mulch (black), ocean water, inland water, sandy bare soil, vegetation and built-up areas. They were used as input in the supervised classification and no threshold value was selected. Therefore, all the pixels were classified into one of the seven classes.

## 2.7 Post-Classification

The post-classification process consists in applying two different filters to the resulting classification image in order to give more spatial coherency eliminating speckle or holes in the classified areas (*clumping*) and isolated pixels (*sieving*).

a) Sieving filter: This filter eliminates all those clumps of similarly classified pixels smaller than a specified size (3 pixels in this case), considering adjacent pixels in eight directions (two vertical, two horizontal and four diagonal).

b) Clumping filter: This filter clumps together adjacent similar classified areas using morphological operators. The selected classes are clumped together by first performing a dilate operation and then an erode operation on the classification results using a kernel of a specified size (2 rows and 2 columns in this case).

### 2.8 Results Validation

To validate the results obtained with both methodologies, a confusion matrix was performed. It compares a classification result with ground truth information. For detailed information on confusion matrices, see Jensen 1986. Ground truth data was extracted from photo-interpretation of digital orthophotos. The photogrammetric flight was carried out in October 1997. Since this date belongs to the previous farming season, photo-interpretation of a false colour composite from the satellite image was used as additional support for this task.

Two classes were considered: PCC's and Others. Polygons were digitised over land plots occupied by PCC's to create class "PCC's" and over other land uses to create class "Others". Those polygons were rasterised and crossed with the final output of both methodological approaches to generate a confusion matrix for each method.

## 3. RESULTS AND DISCUSSION

The surface area classified as PCC's using the index-unsupervised classification and a supervised classification was 2,560.6 and 1,718.6 Has., respectively. The resulting confusion matrices for both methods showed very similar overall accuracy (index: 97.37%, supervised classification: 96.31%). However, the index classification showed to identify PCC's better, as 97.8% of the PCC sample was correctly classified, whereas that that figure dropped to 90.6% in the case of the supervised classification. On the other hand, the supervised classification performed better to identify the generic class "Others" (100% of pixels correctly classified). In the index classification it was 97.1%.

***Table 3.*** Confusion matrix of the index-unsupervised classification approach.

Overall accuracy = (5959/6120) 97.37%

Kappa Coefficient = 0.94

| | **Ground Truth (Pixels)** | | |
|---|---|---|---|
| **Class** | PCC's | Others | Total |
| PCC's | 2362 | 107 | 2469 |
| Others | 54 | 3597 | 3651 |
| Total | 2416 | 3704 | 6120 |

| | **Ground Truth (Percent)** | | |
|---|---|---|---|
| **Class** | PCC's | Others | Total |
| PCC's | 97.76 | 2.89 | 40.34 |
| Others | 2.24 | 97.11 | 59.66 |
| Total | 100.00 | 100.00 | 100.00 |

***Table 4.*** Confusion matrix of the supervised classification approach.

Overall accuracy = (5894/6120) 96.31%

Kappa Coefficient = 0.92

| | **Ground Truth (Pixels)** | | |
|---|---|---|---|
| **Class** | PCC’s | Others | Total |
| PCC’s | 2190 | 0 | 2190 |
| Others | 226 | 3704 | 3930 |
| Total | 2416 | 3704 | 6120 |

| | **Ground Truth (Percent)** | | |
|---|---|---|---|
| **Class** | PCC’s | Others | Total |
| PCC’s | 90.65 | 0.00 | 35.78 |
| Others | 9.35 | 100.00 | 64.22 |
| Total | 100.00 | 100.00 | 100.00 |

A visual comparison of the results against a false colour composite suggests that the supervised classification greatly underestimates the surface extent of PCC’s. This could be due to the fact that parcels are very small in general, which results in numerous border pixels with a mixed spectral behaviour. Those mixed pixels are then classified into a class other than PCC’s, as they are spectrally more similar to it.

On the contrary, the index classification seems to delineate the area occupied by PCC’s rather accurately. Still, some isolated pixels from highly reflective land covers such as built-up areas (especially sheet metal roofs), breaking waves, sand from the beach, and roads are classified as PCC. Many of them can be eliminated with the *Sieving* filter as they do not form big clusters. Others remain, especially in built-up areas. To overcome this, it is advisable to mask all those known built-up areas beforehand, as it is the only way to eliminate those intruding pixels. However, those mis-classified pixels only represent 1.8% of the total area classified as PCC. The supervised classification do not mix those pixels with “plastic” ones.

It must be kept in mind that most mixed pixels covering both PCC’s and any other land use, after index classification and post-classification, are considered as belonging to PCC’s. This will result in a slight overestimation of the surface occupied by plastic, which is something that has to be assumed as it is not possible to go further with a 30-meter spatial resolution.

These results also confirm that images from this time of the year are very appropriate for PCC’s detection.

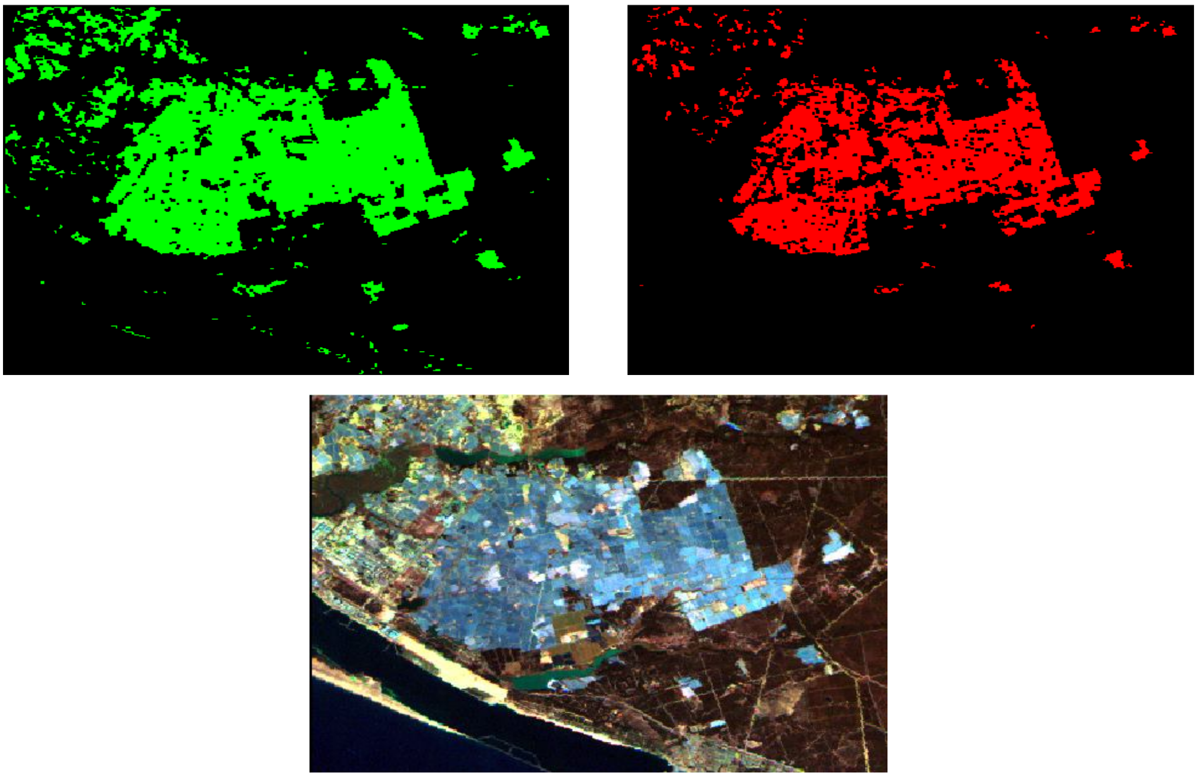

***Figure 5.*** *a) Final classification from the index-unsupervised classification approach. b) Final classification from the supervised classification approach. c) False colour (RGB) composite made with bands ETM 7 (R),ETM 3(G) and ETM 1(B). Bluish colours correspond to PCC's.*

## 4. CONCLUSIONS

The present work has presented a methodology to detect PCC's in the province of Huelva. Two approaches have been tested: a band combination (index) with a hierarchical unsupervised classification, and a supervised classification.

By visual and spectral inspection is was proved that it is extremely important to use an image acquired in one of the following months: November, December and January. During these months pixels that cover PCC's are the purest as there is no vegetation in the crops that could interfere and generate a mixed spectral signature. Furthermore, the index would not perform properly using other months as it was developed to be used with the spectral response of PCC's at this time of the year.

The index and later hierarchical unsupervised classification yielded very good results, although some isolated pixels with similar spectral behaviour from other land uses were classified as PCC's (e.g. built-up areas, beach, breaking waves). To improve the accuracy of the classification it is advisable to mask known built-up areas.

The supervised classification showed worse results as it mis-classified many PCC pixels. Due to the fact that PCC land plots are small in general, the farming technique leaves some uncovered soil, and a very extensive road network crosses this area, a lot of mixed pixels appear. Those pixels are considered spectrally closer to other land covers by the classification algorithm and therefore, they are included into other categories. The index proved to be less strict and most mixed pixels were classified as PCC's.

It must be mentioned that Landsat ETM+ (launched in July 1999) offers better instrument calibration than Landsat TM, which could result in a poorer index performance when historic images from Landsat TM are used. In addition, when there is a high load of suspended solids in the water, some of those water pixels can behave spectrally like plastic. To overcome this problem (if it happens), it is strongly advised to create a mask for water using band 4.

For a further validation, it would be interesting to implement this method in other areas were PCC's exist.

## 5. REFERENCES


J.R. Jensen, (1986), *Introductory Digital Image Processing*, Prentice-Hall, Englewood Cliffs, New Jersey.

J.M. Moreira, A. Ramos, A. Lobato and A. Fernández, (1991), *Evaluación de Superficies de cultivos de fresón mediante imágenes Landsat TM. Su uso en un sistema de pronósticos de cosecha, III Reunión Nacional del Grupo de Trabajo en Teledetección Espacial,* 1989, Madrid (Spain).

J.A. Richards, (1994), *Remote Sensing Digital Image Analysis*, Springer-Verlag, Berlin.

J.T. Tou and R.C. González, (1974), *Pattern Recognition Principles,*Addison-Wesley Publishing Company, Reading, Massachusets .