SEDIMENT TYPE UNSUPERVISED CLASSIFICATION OF THE MOLENPLAAT, WESTERSCHELDE ESTUARY, THE NETHERLANDS

Stefanie Adam[1], Ilse Vitse[1], Chris Johannsen[2] and Jaak Monbaliu[1]

1. Katholieke Universiteit Leuven, Laboratory for Hydraulics, Department of Civil Engineering, 3011 Heverlee, Belgium; stefanie.adam(at)bwk.kuleuven.be
2. Purdue University, Department of Agronomy, West-Lafayette, Indiana 47907-2054, USA

ABSTRACT

Sediment stability or erosion resistance of intertidal zones depend on sediment physical characteristics and on biological factors. Obtaining accurate data on the basic biological, chemical and physical processes in sediments is expensive and difficult. Remote sensing methods can produce detailed information on ecological functioning in a cost-effective manner.

A hyperspectral image of the Molenplaat, an intertidal flat in the Westerschelde estuary, the Netherlands, was acquired with the HyMap sensor in June 2004. The goal of this research is to perform, analyse and evaluate unsupervised classification methods for sediment types on the imagery. The unsupervised methods are based on Principal Component Analysis (PCA) or Iterative Self-Organizing Data Analysis Technique (ISODATA), and consist of three steps: (a) classification into spectrally distinct clusters, (b) post-clustering treatment, and (c) assignment of labels to the clusters. The result consists of 13 clusters after the post-clustering treatment, and of 8 or 9 classes after labelling for either the PCA or ISODATA method. A supervised Spectral Angle Mapper (SAM) classification was performed using field data to evaluate the unsupervised classification results. The labelling of the unsupervised clusters was also partly based on the SAM results, due to limited field data.

The comparison of the results reveals that 69% and 73% of the pixels of PCA and ISODATA classification respectively were identically labelled in the supervised classification. Moreover, the mismatches were mainly found in two classes, while the other classes showed high similarities, indicating the plausibility of using unsupervised classification methods for intertidal sediment types. Additional strengths of the unsupervised classification methods are (a) the distinction of classes that were not visited during field work and not classified in the supervised classification, (b) the identification of spectrally distinct areas that should be characterised during field campaigns, and (c) the user-friendliness thanks to limited required field knowledge and short calculation time.

Keywords: Unsupervised classification, sediment type classification, Molenplaat, TideSed.

INTRODUCTION

Estuaries are important from an economical, coastal defence and ecological point of view, because of their particular position at the interface between the marine and terrestrial environment (1). Many of the most important biogeochemical processes occur on the large areas of loose sediments which are exposed at low tide and deposited as a result of the interaction between currents, tides and salinity (2). For example, the distribution and binding of heavy metals are controlled by the fraction of fine sediment. Biological processes on sediments include photosynthesis by benthic microalgae at the sediment surface fuelling primary production and consequently supporting many grazing animals and birds. Sediment stability or erosion resistance depend on grain size and distribution. Biological factors such as accumulations of algal cells in a surface biofilm also cause the sediment to become more stable. In contrast, grazing and bioturbation of macrofauna enhances the erosion rate (3).

However, obtaining accurate data on sediment properties is difficult, due to limited access to the sites and the large spatial heterogeneity of intertidal areas. Modern airborne remote sensors such

as HyMap with hyperspectral capabilities can identify important groups of inorganic and organic materials at a high spatial resolution, improving the understanding of ecological functioning.

In previous research, remotely sensed images were used to characterise intertidal sediments (4). Supervised and unsupervised classification methods of sediment and vegetation types of the Humber estuary gave qualitative results. An accuracy assessment could not be performed due to poor field data on the study site. However, other researchers were able to quantify sand and mud content using a developed empirical model (1). In this study a spectral linear unmixing procedure using collected endmembers, which are spectra of pure materials, was applied on hyperspectral airborne imagery after which an empirical model was calibrated and validated using a significant amount of field data. The accuracy and prediction potential of the model were expressed as a correlation coefficient. The clay distribution could be mapped based on the mud abundance image with a correlation coefficient of 0.79. The correlation was lower in the sand distribution model (r^2 = 0.60). Disadvantages of this kind of studies are the need for field data and the site and image dependency of the empirical model. Therefore, a more automated method with little field knowledge requirement would be helpful in exploring intertidal mudflats.

The goal of this research is to perform and evaluate an unsupervised classification for sediment types on the imagery collected over the study area. The classification methods are based on Principal Component Analysis (PCA) and Iterative Self-Organizing Data Analysis Technique (ISODATA). The results are assessed and the methods evaluated using supervised Spectral Angle Mapper (SAM) classification results. It is believed that automated, unsupervised methods can be valuable for classification of large, inaccessible areas or as an instrument for preparing intensive ground reference measurements in unknown terrain.

This paper reports a part of the work of the TideSed project that aimed at improving the methods used for the characterisation of intertidal sediments (5).

METHODS

Study Area

The study area consists of the Molenplaat, an intertidal flat in the Westerschelde estuary in the Netherlands (Figure 1). With one of the largest wading bird populations in western Europe and several rare habitat types such as freshwater tidal marshes, the Schelde estuary is a site of international recognition and importance for nature conservation. This estuary is also a site of heavy industry and is an important commercial shipping route (6). Coastal zone managers must constantly balance the demands of many conflicting interest groups when making planning decisions which affect this complex system (7).

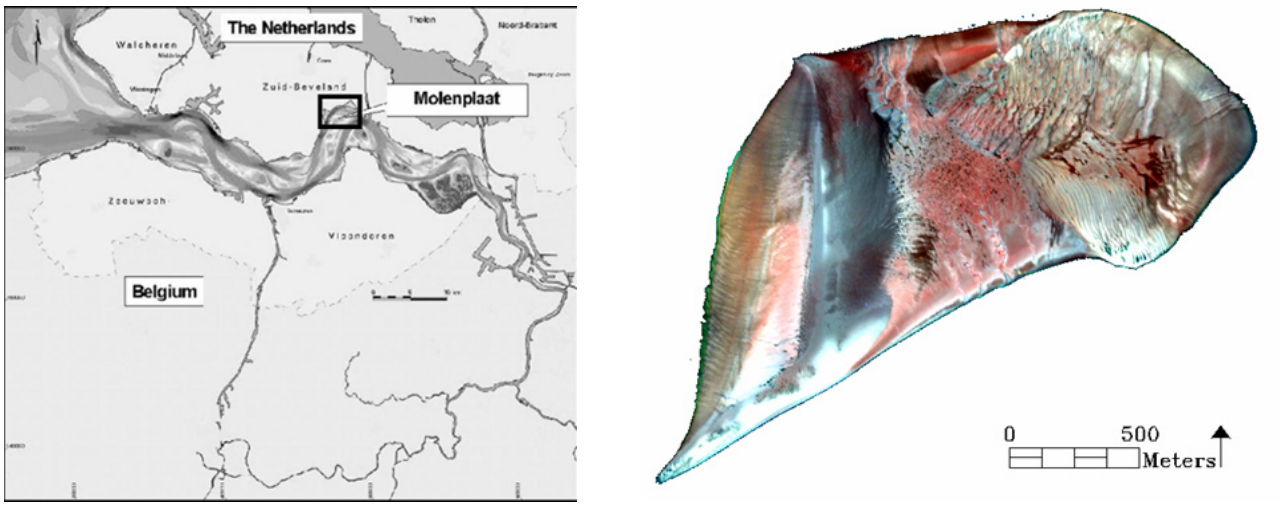

Figure 1: Left: Location of the Westerschelde estuary; Right: False colour image of the Molenplaat (red: 740 nm; green: 635 nm; blue: 543 nm).

The Molenplaat is a predominantly sandy mid-channel intertidal flat with a surface area of approximately 1.5 km^2 (8). The average period of emersion varies between 2 to 4 hours and up to 8 hours for specific locations, per tidal cycle (9).

Data

Images of the Molenplaat were acquired under clear weather conditions on the 8th of June 2004 around noon with the HyMap sensor. The HyMap sensor is an airborne spectrometer which collects radiance measurements of the Earth's surface in the 450 to 2480 nm wavelength region. Typically, the spatial resolution achieved with the HyMap sensor is in the range of 3 to 10 m, depending on the aircraft altitude. The HyMap sensor records 128 spectral bands. However, the data set contains 126 bands because the first and the last band were deleted during the pre-processing. The image was radiometrically, atmospherically and geometrically corrected by VITO (Vlaamse Instelling voor Technologisch Onderzoek; Flemish Institute for Technological Research). To minimise the spectral complexity of the image, the intertidal flat south of the Molenplaat and the water were excluded. For both features, a mask was built and applied on the image.

On the day of overflight, 24 sites were visited to take ground-truth measurements and *in situ* spectral measurements with sensors from Analytical Spectral Devices (ASD) or Ocean Optics. At each site, 9 spectral measurements and three ground samples were taken to account for the pixel size, which is 4 m by 4 m, and the error on the geometric correction of the HyMap imagery (Figure 2). The locations were determined using a Global Positioning System (GPS).

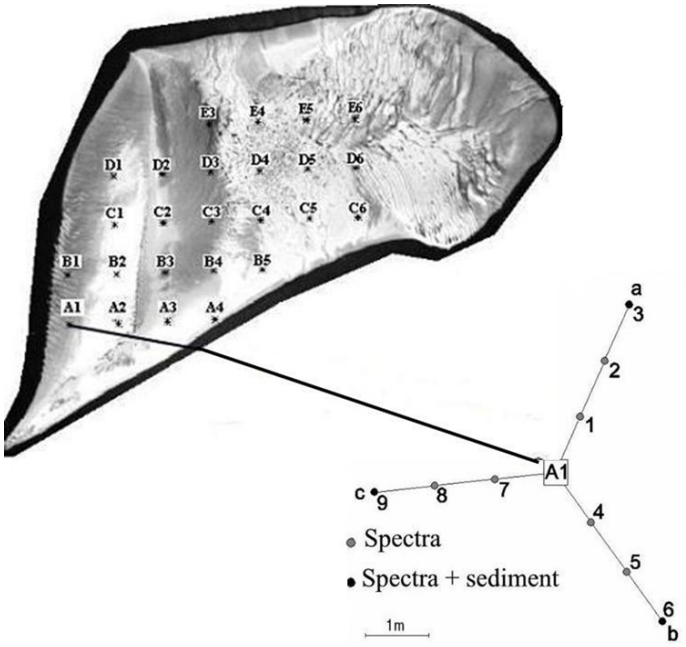

Figure 2: Sampling strategy: 9 spectral measurements and three sediment samples were taken at each site. All the sites indicated in the figure were sampled, except B1; adopted from (10).

Ground samples were analysed at the University of Gent and NIOO (Nederlands Instituut voor Ecologie; Netherlands Institute for Ecology) for grain size distribution, moisture content, chlorophyll *a* (chl a) content and organic matter (o.m.) content. The mean of the three measurements at each site (a, b, c in Figure 2) was calculated.

Classification of the ground reference

The image spectra from the locations of field samples were collected and similar spectra were grouped based on the overall reflectance and absorption features. Since the field spectra and image spectra were very similar (10), and since the Ocean Optics instrument did not acquire the reflectance at wavelengths above 862 nm, it was decided to use the spectra from the image. These spectra integrate the reflectance from an area of 4 m by 4 m. Therefore, the mean value of the sediment properties of the three measurements for each site was considered as being representa-

tive for the spectrum. Then it was checked if the grouped image spectra correspond to similar sediment properties.

For each analysed sediment parameter a classification rule was defined. In previous research, it has been indicated that the fraction of sediments coarser than 250 µm could be determined most accurately using Landsat ETM$^+$ images (11). However, the sensor specifications of Landsat ETM$^+$ and HyMap are different and cohesive properties and erodibility of the sediment are determined by the fraction of particles smaller than 63 µm (mud). In previous research, cohesive properties were observed once the mud fraction was higher than 15 weight% (12). Therefore the classification of grain size distribution was based on this fraction: the sediment type is clayey loam if the fraction of mud is larger than 30%, loamy sand if the fraction of mud is between 15 and 30%, and sand if the fraction of mud is smaller than 15%.

A second classification rule according to the moisture content is based on the sediment stability for different water contents (Bart Deronde, personal communication, 2005). The sediment is dry if the moisture content is lower than 20 weight%, wet if the moisture content is between 20 and 30 weight%, very wet if the moisture content is between 30 and 40 weight%, and saturated if the moisture content is more than 40 weight%.

A third and fourth classification rule according to chlorophyll *a* and organic matter used an *ad hoc* method so that in each group a similar amount of ground samples is present (Bart Deronde, personal communication, 2005). Chlorophyll *a* and organic matter were grouped into three classes. For chlorophyll *a*, the sediment contained a small amount if the content is less than 20 mg/m^2, an intermediate amount if the content is between 20 and 40 mg/m^2 and a large amount if the content is more than 40 mg/m^2.

For organic matter the sediment contained a small amount if the content is less than 2 weight%, an intermediate amount if the content is between 2 and 4 weight% and a large amount if the content is more than 4 weight%.

Supervised SAM classification

The algorithm of Spectral Angle Mapper classification (SAM), which is standard implemented in ENVI (13), determines the spectral similarity between two spectra by calculating the angle between the spectra, treating them as vectors in a space with dimensionality equal to the number of bands (n = 126 in this application). SAM compares the angle between reference vectors and each pixel vector in n-dimensional space. A smaller angle than the specified threshold angle represents a close match to the reference spectrum and the pixel is assigned to that class. Pixels further away than the specified maximum angle threshold for each reference spectrum are not classified (14). In this study, the reference spectra are the image spectra from the sampled sites, and so corresponding to known sediment properties.

Unsupervised PCA classification

The classification method used and described in a previous publication (15) is entirely based on a methodology using Principal Component Analysis (PCA). The principal component transformation (PCT) is a commonly used linear transformation which finds a new set of orthogonal axes with their origin at the data mean. These are rotated so that the data variance is maximised (16). PCA is a powerful technique which decorrelates bands so that maximum information is explained by a few bands such that the classification of hyperspectral images can be based on this reduced dimensionality.

It is suggested that the use of two principal components, PC1 and PC2, is sufficient, since most of the variation in the dataset of intertidal flats is explained by PC1 and PC2. The PCA classification of the Molenplaat supports this statement. PC1 explains 96% of the variation in the dataset and the combination of PC1 and PC2 explains 98% of the variation. In the scatter plot of PC1 versus PC2, each quadrant is representing a cluster of pixels with comparable characteristics. A hierarchical principal component classification is suggested if more than four spectrally distinct clusters are present in the image. The four quadrants defined by PC1 and PC2 are individually considered by separating the pixels of each quadrant in a cluster. Then a PCT is performed on each cluster to

isolate four new clusters using PC1 and PC2 of the second PCT. This procedure leads to 16 (4·4) different clusters.

After this double transformation, separability measures and spectral similarity between mean cluster spectra were used to check if clusters should be merged. The merging procedure consists of two steps. In a first step, the transformed divergence d_{ij}^{T} for each pair of clusters is used as separability measure and is defined as:

$$d_{ij}^{T} = 2\left(1 - \exp\left(-\frac{d_{ij}}{8}\right)\right),$$

with

$$d_{ij} = \int_{x} \left\{p\left(x|\omega_i\right) - p\left(x|\omega_j\right)\right\} \ln\frac{p\left(x|\omega_i\right)}{p\left(x|\omega_j\right)} dx,$$

and where $p\left(x|\omega_i\right)$ and $p\left(x|\omega_j\right)$ are the values of the *i* th and *j* th spectral cluster probability distributions at the position *x*. Because of its exponential character, the transformed divergence has a saturating behaviour with increasing cluster separation. Values range from 0.0 to 2.0 (13). Values lower than 1.9 indicate that the clusters are not clearly separable and are merged.

In a second step the mean spectra of the clusters are compared using spectral angles and spectral feature fitting (SFF). The spectral angle is a value in radian indicating the angle between the spectra that are considered as a *n*-dimensional vector. SFF uses a least-squares technique to compare spectra after continuum removal so that mainly the absorption features are taken into consideration. The similarity between two spectra is expressed as a root mean square error (RMSE). If the RMSE between spectra is smaller than 0.01, then the spectra are very similar. The criterion for merging clusters is:

- the transformed divergence between the clusters is lower than 2.0, AND
- the spectral angle between the mean spectra of the clusters is smaller than 0.06rad, AND
- the RMSE value between the mean spectra of the clusters is smaller than 0.01.

After this merging procedure, the unsupervised clusters are labelled using the available field data.

Unsupervised ISODATA classification

In the ISODATA classification, the number of desired clusters is determined by the user whereupon the classification algorithm selects randomly a number of pixels, equal to the number defined by the user, and calculates the cluster means. The remaining pixels are iteratively clustered using minimum distance techniques. Each iteration recalculates means and reclassifies pixels with respect to the new means. Iterative cluster splitting, merging and deleting is done based on input threshold parameters. The process continues until the number of pixels in each cluster changes by less than the selected pixel change threshold or until the maximum number of iterations is reached (17). The ISODATA unsupervised classification is a standard method implemented in the software package ENVI (Environment for Visualizing Images).

For this application the pixel change threshold was set to 2% and the maximum number of iterations to 1000. The minimum and maximum number of clusters were 15 and 19, taking into consideration the amount of clusters originating from the unsupervised PCA classification which is 16. The merging procedure as described for the PCA unsupervised classification was applied.

RESULTS

Field data

The field spectra of the 24 field sites were grouped based on visual similarity, taking the presence and the depth of the absorption features into consideration. This grouping was straightforward and resulted in eight groups of field sites, each with a unique label (Table 1). The location of the sampled sites are indicated in Figure 2.

In Table 1, the mean values of the sediment properties of the grouped sites are given. Since it is irrelevant to calculate standard deviations on two, three or four numbers, these are not shown. The minimum and maximum values are indicated to show the range of the measured values. As can be seen, these values are for some groups rather different, although most values are situated in the range for a sediment label as indicated in the previous section.

Table 1: Sediment properties for each group of sampled sites (int. = intermediate amount, chl a = chlorophyll a, o.m. = organic matter).

	field sites	**mud fraction (%) (mean) (min – max)**	**moisture (weight%) (mean) (min – max)**	**chl *a* (mg/m^2) (mean) (min – max)**	**o.m. (weight%) (mean) (min – max)**
Group 1	A2 C5 C6 A4	clayey loam (33.5) (20.5 - 48)	very wet (31.7) (29.0 – 34.4)	int. (33.5) (29.4 – 44.3)	int. (3.2) (2.0 – 3.9)
Group 2	A1 E7	loamy sand (25.45) (18.6 – 32.3)	very wet (31.4) (29.5 – 33.3)	small (16.1) (14.6 – 17.6)	int. (2.2) (2.0 – 2.4)
Group 3	D4 D5 E5	clayey loam (46.5) (34.9 – 50.5)	very wet (37.9) (32.5 – 42.5)	large (44.2) (33.9 – 62.7)	large (4.3) (3.5 – 5.4)
Group 4	B2 B5 C1	clayey loam (48.6) (44.6 – 54.4)	saturated (44.0) (40.6 – 49.7)	large (41.3) (36.3 – 44.3)	large (4.9) (4.1 – 6.2)
Group 5	C4 D1 E6	loamy sand (29.8) (25.2 – 38.6)	very wet (32.3) (29.1 – 35.4)	int. (35.0) (20.0 – 45.5)	int. (3.0) (2.3 – 4.3)
Group 6	B3 C3 D3	sand (5.1) (2.1 – 8.2)	wet (23.0) (20.6 – 26.0)	small (11.7) (9.7 – 15.6)	small (1.2) (0.9 – 1.4)
Group 7	A3 C2 D2	sand (0.6) (0.0 – 1.0)	dry (18.6) (16.3 – 20.2)	small (9.5) (8.8 – 10.8)	small (1.1) (1.0 – 1.1)
Group 8	B4 D6 E4	sand (10.0) (6.5 – 14.1)	wet (25.0) (22.0 – 27.0)	int. (31.2) (28.9 – 34.6)	small (1.2) (0.8 – 1.5)

It would be interesting to investigate the influence of sediment properties on the in situ spectral measurements and the spatial variability of the sediment properties since it was noted that for some sites the measured values for chlorophyll *a* and mud fraction were very different, although the locations of sampling were less than six meters apart. However, these topics were not part of the paper's objectives.

The average spectrum for each group is shown in Figure 3.

Comparing the field and spectral properties, the following observations are made:

- In theory the reflectance increases with decreasing grain size, since the smaller grains provide a higher proportion of surface area available as reflectors and larger grains can absorb more radiation (18). Generally, the image spectra show the highest reflectances for the clayey loam sites, especially in the near infrared (0.7-1.1 µm). As the mud content increases, the reflectance increases: group 4 with an average mud content of 48.6% shows higher reflectance than group 1 with an average mud content of 33.5% and higher reflectance than the loamy sand and sand sites. However, this relation is not always observed. Possible reasons are the effects of surface roughness, shadowing, and moisture and organic matter content.
- The reflectance decreases with increasing moisture content since more radiation is absorbed. This is demonstrated in the sand spectra where the driest sand of group 7 shows higher reflectance values than the wet sand of group 6 and 8. In previous research, it has been observed that this negative relation is valid as long as the moisture content is below a critical point that is determined by soil properties. Beyond this critical point, the reflectance increases with increasing moisture content (20). This is observed in the spectrum of group

4. Although the moisture content is very high, the reflectance is also very high. The moisture content is probably higher than the threshold value, thereby increasing the reflectance and enhancing the effect of high reflectance with small grain size.

- There is a positive relation between the amount of chlorophyll *a* and the depth of its absorption feature around 670…680 nm (Figure 4). The depth of this absorption feature has been used in previous research to quantify the amount of chlorophyll *a* (19,20).

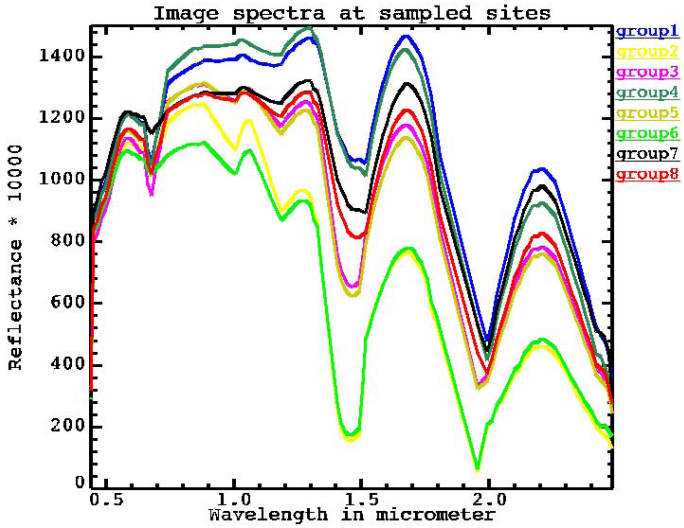

Figure 3: Average spectrum of each group of sampled sites.

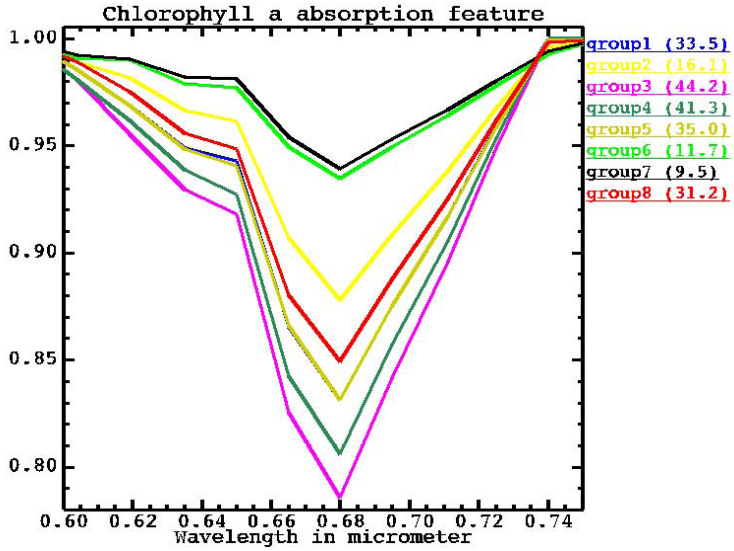

Figure 4: Absorption feature of chlorophyll a for the sampled sites. The continuum of the spectra is removed to emphasise the absorption feature. The amount of chlorophyll a in mg/m^2 is indicated between brackets. The group 1 spectrum is covered by the group 5 spectrum.

Supervised SAM classification

Although the number of field observations is rather limited, a SAM classification was performed using the image spectra at the sampled sites, as shown in Figure 3. A default threshold angle of 0.10 rad was used for all spectra. The result is shown in Figure 5.

To have an assessment of the accuracy of the classification results, it was checked if the sampled sites were classified correctly. It appeared that 20 out of 24 used field sites were correctly classified. The four misclassified sites are D4, D5, E4 and E5. Three out of the four misclassified field sites are members of group 3 (D4, D5 and E5) but classified in group 1 and group 4. However, group 1, group 4 and group 3 are characterised by similar sediment properties (Table 1). There-

fore, these misclassifications will not result in large errors for the end-user who is interested in sediment properties. Additionally, those three field sites are located in a region with large spectral differences over small areas, indicated by the speckle of blue (group 1), magenta (group 3), sea green (group 4) and green (group 6) pixels and leading to a difficult assignment of the final class to these field sites. Since less than 0.1% of the pixels of the Molenplaat were classified as group 3 and since there exist large similarities between sediment properties of group 3 and group 4, the classes of group 3 and group 4 were merged. The fourth sampled site that was not correctly classified, E4, is located in an unclassified region.

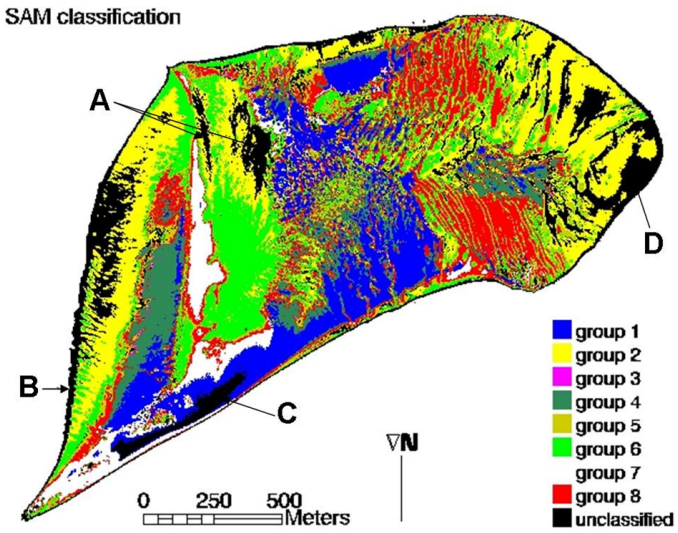

Figure 5: Supervised classification results. The sediment properties of each class are listed in Table 1.

The regions that were left unclassified show spectral properties that are very different from the image spectra at the sampled sites (Figure 6). Region A shows spectra with low reflectance in the VIS and NIR, indicating that water absorbing the NIR radiation is present. Even lower NIR reflectance is observed in region B, or pixels near the waterline, probably sediment in suspension. In region C the spectrum has got very high overall reflectance values (more than 20%) and doesn't show an absorption feature of chlorophyll *a*. This region is very dry sand without microphytobenthos. The spectral properties of pixels in region D are similar to the spectral properties of the group 2 class, but the unclassified regions show lower NIR reflectance, probably due to higher moisture content or finer grain size. However, the determination of the sediment properties in this region would require sediment sampling and analysis.

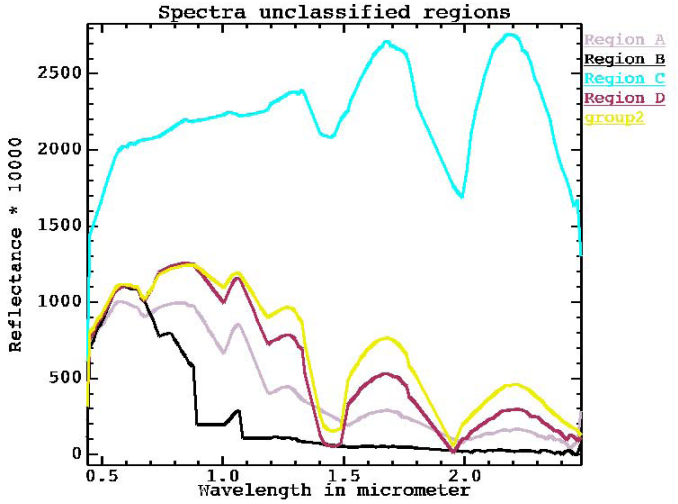

Figure 6: Spectral properties at the unclassified regions as indicated in Figure 5.

Apparently a threshold angle of 0.10 rad is appropriate, since regions with very distinct spectral properties are left unclassified. If the sediment and spectral properties at these regions would have been determined during the field work, those regions would probably have been distinguished as separate classes.

Unsupervised PCA classification

The unsupervised classification was performed as described above. After the merging procedure there were 13 clusters out of the 16 original clusters (Figure 7). Visual comparison with the supervised classification result shows large similarities, indicating a qualitative good result for the unsupervised classification procedure.

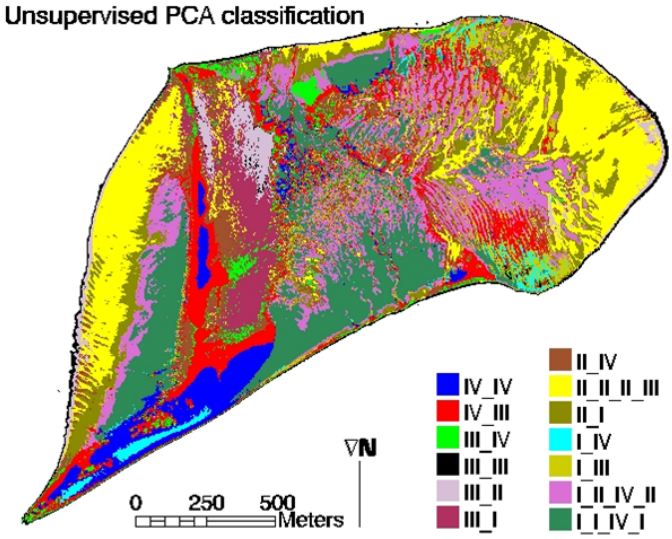

Figure 7: Unsupervised PCA classification result after the merging procedure. The legend indicates the cluster of pixels: the first Roman figure is the quadrant in the scatter plot of PC1 and PC2 after the first PCT, and the second Roman figure is the quadrant after the second PCT. If two clusters were merged, the first two Roman figures indicate the first and the following two Roman figures the second cluster.

Labels were assigned to the clusters using the available field data. If all the field data of one group were located in a cluster, the cluster was labelled with the sediment properties of that group. The cluster I_I_IV_I contained all the sampled sites of group 1 and group 4, very wet clayey loam with intermediate contents of chlorophyll *a* and organic matter, and saturated clayey loam with large contents of chlorophyll a and organic matter respectively. Apparently the distinction between different moisture contents and between small differences in chlorophyll *a* or organic matter was problematic. This cluster received a label that indicates the sediment properties of group 1. The cluster II_II_II_III contained all the sites of group 2, very wet loamy sand with a small amount of chlorophyll *a* and an intermediate amount of organic matter, and was labelled accordingly.

There were three clusters with distinctive unclassified spectra that are similar to the spectra plotted in Figure 6. In the supervised classification these regions were classified as a single unclassified class, since no sediment samples were taken. However, the unsupervised classification could separate these regions. Some sediment characteristics could be derived based on spectral properties and field knowledge. The black pixels, III_III, is sediment in suspension located at the waterline. The cluster III_II is sediment saturated with water, and the cluster I_IV in the south-western part of the Molenplaat is very dry sand without chlorophyll *a*.

The other clusters could not be labelled using the field data, since one cluster could include field data of more than one group, or a sampled site was located at the border between clusters. To label the other clusters it was decided to overlay the supervised and unsupervised classification results, so that unsupervised clusters are labelled according to the supervised class with the most pixels in common. Some clusters have pixels distributed over several supervised classes, e.g.

cluster IV_III with 13478 pixels has 7130 pixels with the group 8 class and 3920 pixels with the group 7 class in common. The majority rule was always applied, so that in this example the sediment properties of group 8 were assigned as label to cluster IV_III, although a different labelling might have been expected by visual comparison.

No unsupervised cluster exists that has the majority of pixels in common with the group 4 or the group 7 classes, so that these labels were not assigned to any cluster. It should be noted that most pixels of the group 4 class are in common with the unsupervised cluster that received the label of group 1. This label assignment is not a radical change, since the labels of group 1 and group 4, very wet clayey loam with intermediate contents of chlorophyll *a* and organic matter, and saturated clayey loam with large contents of chlorophyll *a* and organic matter respectively, are similar. This is also observed for the labels of group 7 and group 8, dry sand with small amounts of chlorophyll *a* and organic matter, and wet sand with intermediate amount of chlorophyll *a* and small amount of organic matter respectively.

The final classification result has 8 classes (Figure 8). Three of these classes are unclassified regions in the supervised classification.

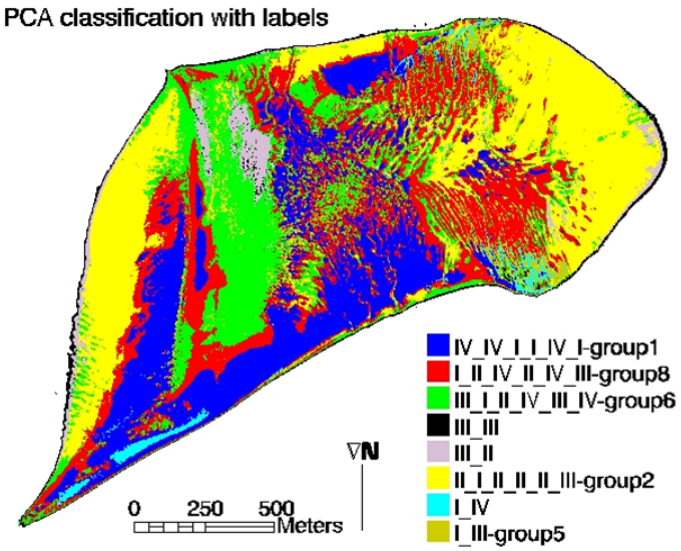

Figure 8: Unsupervised PCA classification result after labelling. The colours in the legend correspond to the colour and label of the supervised classification. III_III is sediment in suspension, II_II is saturated sediment, I_IV is dry sand.

Unsupervised ISODATA classification

The maximum number of iterations was needed in the ISODATA classification procedure with the defined parameters, leading to the conclusion that in each iteration more than 2% of the pixels still changed cluster. No clusters were merged so that 19 clusters, the maximum number, were obtained. After the ISODATA clustering procedure, the merging steps as described in the PCA unsupervised classification were applied and resulted in 13 clusters.

If a cluster contained all the sampled sites of one group, the cluster received the sediment properties of the group as label. The other clusters were labelled using the supervised classification results. This labelling procedure led to 9 classes, of which 4 were unclassified regions in the supervised classification results, namely cluster_1, cluster_2, cluster_18 and cluster_19 (Figure 9). As in the PCA classification the labels of group 4 and group 7 were not used.

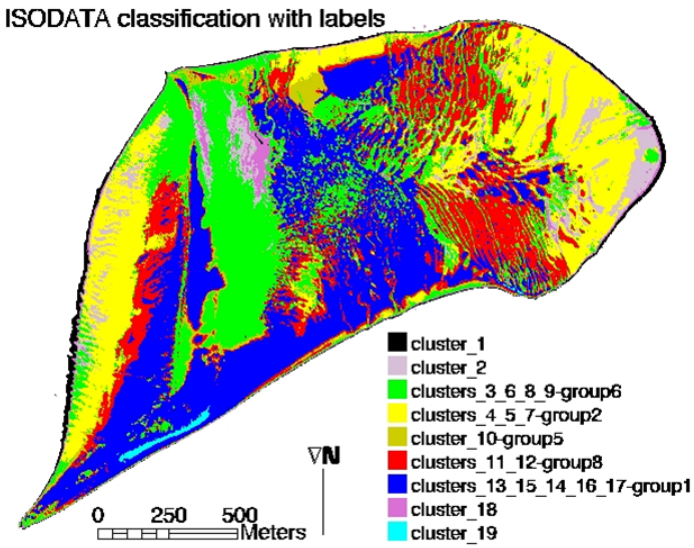

Figure 9: ISODATA classification results after merging and labelling. The colours in the legend correspond to the colour and label of the supervised classification. Cluster_1 is sediment in suspension, cluster_2 is saturated sediment, cluster_18 is very saturated sediment and cluster_19 is dry sand.

Comparison of supervised and unsupervised classifications

In order to compare the classification results on a pixel-by-pixel basis, confusion matrices were calculated. Although we realise that this methodology in which firstly the labelling is done based on the supervised classification and secondly the similarity is assessed using the same supervised classification, is biased, there was no other objective manner for label assignment or similarity assessment, since field knowledge is limited. The unsupervised classes corresponding to the unclassified region in the supervised classification are merged into one class. The PCA unsupervised classification (Table 2) shows an overall similarity of 69% with the supervised classification. This value is calculated by dividing the total number of pixels in the diagonal (77763) by the total number of pixels of the supervised classes (113060).

This means that the majority of the pixels are assigned identically in both classification results. The largest differences are found in the unclassified region and in the I_III-group5 class. The numbers of pixels in the corresponding supervised classes are much higher: 18826 versus 8943 for the unclassified region, and 12497 versus 5396 for the I_III-group5 class. The similarity between the supervised and unsupervised results is higher for the other classes.

The similarity assessment for the ISODATA classification results is shown in Table 3. The overall similarity is 73% (82098/113060). The same observations are made for the unclassified and cluster_10-group5 classes as for the PCA classification.

The pixels of the supervised group 5 class are located in classes labelled with the sediment properties of group 5, group 6 and group 8 in the unsupervised classifications. However, the sediment properties of group 5, very wet loamy sand with intermediate amounts of chlorophyll *a* and organic matter, and of group 6 and 8, wet sand with small amounts of chlorophyll *a* and organic matter, and wet sand with intermediate amount of chlorophyll *a* and small amount of organic matter, show large differences (Table 1). Especially the mud content, which is 30 weight% for group 5 and 5 weight% and 10 weight% for group 6 and 8 respectively, is very different. This might indicate that the influence of mud content on the spectrum is not very high, or, more likely, that other factors are affecting the relation between mud content and reflectance.

Table 2: Similarity assessment between unsupervised PCA and supervised classification results.

	Group 1	Group 2	Group 5	Group 6	Group 8	unclassified	**Total**
IV_IV_I_I_IV_I-group1	**15879**	0	7	0	923	782	17591
II_I_II_II_II_III-group2	0	**20702**	156	4279	0	8685	33822
I_III-group5	0	2	**3198**	300	1822	74	5396
III_I_II_IV_III_IV-group6	0	3955	3995	**14461**	170	1584	24165
I_II_IV_II_IV_III-group8	1556	0	5133	0	**16138**	316	23143
III_III_III_II_I_IV-unclassified	186	666	8	0	698	**7385**	8943
Total	17621	25325	12497	19040	19751	18826	113060

Table 3: Similarity assessment between unsupervised ISODATA and supervised classification results.

	Group 1	Group 2	Group 5	Group 6	Group 8	unclassified	**Total**
clusters_13_15_14-group1	**17589**	0	2	0	2920	905	21416
clusters_4_5_7-group2	0	**17827**	0	2876	0	4300	25003
Cluster_10-group5	0	0	**3885**	0	1515	680	6080
clusters_3_6_8_9-group6	0	6054	4771	**16163**	0	1492	28480
clusters_11_12-group8	0	0	3839	1	**15316**	131	19287
clusters_1_2_18_19-unclassified	32	1444	0	0	0	**11318**	12794
Total	17621	25325	12497	19040	19751	18826	113060

It should be noted that in these confusion matrices the pixels of the group 4 and group 7 supervised classes were not taken into account, since their sediment properties were not used as a label for the unsupervised clusters. However, all the pixels should be considered. Therefore, group 4 class was combined with group 1 class, and group 7 class was combined with group 8 class. The resulting confusion matrices are shown in Tables 4 and 5. Lower similarities resulted, 67% and 68% for the PCA and ISODATA results respectively.

The effect of several parameters like mud, moisture and chlorophyll *a* content of intertidal sediments on their spectral characteristics should be further investigated to improve unsupervised and supervised classification methods, e.g. by focussing on the wavelengths which allow to discriminate these classes.

As indicated in previous research (21), the unsupervised ISODATA classification can produce very different final clusters for the same dataset depending on how certain parameters, e.g. number of clusters, allowable dispersion around cluster means and pixel change threshold, are controlled. To solve this ambiguity partly, a larger number of clusters than expected is calculated and a procedure to merge similar clusters followed. An ISODATA classification with a maximum number of clusters of 35 was performed and the resulting clusters merged as described before. The result is a classification image with 14 clusters. These clusters were labelled using the supervised classification result. The similarity between the unsupervised classes and the supervised classes was 73%. It should be noted that in this unsupervised classification the labels of the group 4 and group 7 classes of the supervised classification were not used. The observation that nearly the same number of classes resulted from an ISODATA classification with a different number of clusters is encouraging the further use of the merging method.

An error source in the unsupervised PCA classification is the method of cluster formation. In this paper the quadrants of the scatter plot of PC1 and PC2 are the cluster boundaries, since the zero values of PC1 and PC2 represent the data mean. However, patterns that don't correspond to the quadrants are often seen in the scatter plot. The extraction of quadrants might not be the appropriate method of clustering. Therefore, other clustering methods have to be tested, leading to a quick,

consistent and correct unsupervised classification of intertidal sediments. Even more basically, information about the spectral separability of the classes is needed.

Table 4: Similarity assessment between unsupervised PCA and supervised classification results for all the pixels.

	Group1+3+4	Group 2	Group 5	Group 6	Group 7+8	unclassified	**Total**
IV_IV_I_I_IV_I-group1	**22742**	0	7	0	3710	782	27241
II_I_II_II_II_III-group2	0	**20702**	156	4279	0	8685	33822
I_III-group5	32	2	**3198**	300	1822	74	5428
III_I_II_IV_III_IV-group6	15	3955	3995	**14461**	171	1584	24181
I_II_IV_II_IV_III-group8	6329	0	5133	0	**20041**	316	31819
III_III_III_II_I_IV-unclassified	207	666	8	0	1047	**7385**	9313
Total	29325	25325	12497	19040	26791	18826	131804

Table 5: Similarity assessment between unsupervised ISODATA and supervised classification results for all the pixels

	Group 1+3+4	Group 2	Group 5	Group 6	Group 7+8	unclassified	**Total**
clusters_13_15_14-group 1	**25037**	0	2	0	9829	905	35773
clusters_4_5_7-group 2	0	**17827**	0	2876	0	4300	25003
cluster_10-group 5	46	0	**3885**	0	1543	680	6154
clusters_3_6_8_9-group 6	42	6054	4771	**16163**	0	1492	28522
clusters_11_12-group 8	4168	0	3839	1	**15419**	131	23558
clusters_1_2_18_19-unclassified	32	1444	0	0	0	**11318**	12794
Total	29325	25325	12497	19040	26791	18826	131804

CONCLUSIONS

The goal of this research was to perform and interpret an unsupervised classification for sediment types on the Molenplaat imagery. Automated unsupervised methods are valuable for the classification of large, inaccessible areas or as an instrument for preparing intensive ground reference measurements in unknown terrain. Unsupervised classification results were compared with supervised ones.

A supervised classification was performed using the available field data. Twenty-four sediment samples were analysed for grain size, moisture content, organic matter and chlorophyll *a* content, and spectrally and geographically characterised. Eight groups of spectra were formed and received a unique label based on the sediment properties. The mean spectrum of each group was applied in the supervised Spectral Angle Mapper classification with a threshold angle of 0.10rad. Of the 24 sampled sites 20 were classified correctly. Some regions with very distinctive spectral characteristics remained unclassified.

Two unsupervised classifications, using PCA or ISODATA, were performed. The spectral separability of the clusters and the similarity between the mean spectra of each cluster was used as criterion to merge clusters. The resulting unsupervised classifications consisted of 13 clusters. Since the Molenplaat is mainly a muddy and sandy area with small spectral differences, it is difficult to label the obtained clusters using the limited field data and spectral information. Therefore, information was extracted from the supervised classification to label the clusters with sediment properties. Clusters that received the same label were merged. This led to 8 and 9 classes for the PCA and ISODATA classifications respectively. The labels of two supervised classes, group 4 and group 7, were not used in the unsupervised classifications. However, most pixels of these supervised classes were labelled in the unsupervised classifications with similar sediment properties, e.g. the majority of pixels in the supervised group 4 class, saturated clayey loam with large amounts of

chlorophyll *a* and organic matter, were identically labelled or as very wet clayey loam with intermediate amounts of chlorophyll *a* and organic matter in the unsupervised classifications. In other words, these dissimilarities lead to small errors for the user who is interested in sediment properties.

The similarities between the supervised and unsupervised classifications were assessed using confusion matrices. 73% and 69% of the pixels in the ISODATA and PCA classifications respectively, were identically labelled as in the supervised SAM classification. Supervised and unsupervised classifications lead to similar results. Better field knowledge is needed to estimate which method is the most accurate.

In the final PCA and ISODATA classification three and four classes respectively were unclassified in the SAM classification, indicating that unsupervised classification methods are valuable in less accessible areas where sediment samples and spectral information are difficult to gather. A fast unsupervised classification can also be used to plan field campaigns, so that spectrally different regions are characterised for their sediment properties.

In conclusion, PCA and ISODATA are fast and objective classification methods that require little field knowledge and that lead to similar classification results as supervised SAM classification.

For future research, more ground samples should be taken to label more classes and to perform an accuracy assessment. Additionally, the applied merging procedure should be further studied. Spectral separability measures are valuable, but the used threshold value is a subjective choice, as is the choice of the threshold spectral angle. Other spectral matching algorithms, e.g. the spectral correlation measure, can be tested. Another suggestion is the investigation of the effect of sediment properties on the reflectance characteristics to improve supervised and unsupervised classification methods.

ACKNOWLEDGEMENTS

The TideSed project is a joint effort of VITO, the University of Gent (Laboratory of Protistology and Aquatic Ecology; Marine Biology Section), NIOO-CEME (the Netherlands Institute of Ecology – Centre for Estuarine and Marine Ecology), and the Katholieke Universiteit Leuven (Laboratory for Hydraulics). It is supported by the Belgian Science Policy office in the framework of the research programme for earth observation *STEREO* (Support to the Exploitation and Research of Earth Observation data). The authors are grateful to Bart Deronde (VITO) to assist with grouping the sediment properties, and to two anonymous reviewers for their comments and suggestions.

REFERENCES

1 Rainey M P, A N Tyler, D J Gilvear, R G Bryant & P McDonald, 2003. Mapping intertidal estuarine grain size distributions through airborne remote sensing. Remote Sensing of Environment, 86: 480-490

2 Black K S & D M Paterson, 1997. Measurement of the erosion potential of cohesive marine sediments: A review of current in situ technology. Journal of Marine Environment Engineering, 26: 43-83

3 Paterson D M, 1997. Biological mediation of sediment erodibility: Ecology and physical dynamics. In: Cohesive sediments, edited by N Burt, R Parker & J Watts (Wiley), 215-229

4 Thomson A G, J A Eastwood, M G Yates, R M Fuller, R A Wadsworth & R Cox, 1998. Airborne remote sensing of intertidal biotopes: BIOTA I. Marine Pollution Bulletin, 37: 164-172

5 http://www.kuleuven.be/hydr

6 http://www.pml.ac.uk/biomare/sites/Westerschelde.htm

7 http://www.proses.be

8 Smith G, A Thomson, I Möller & J Kromkamp, 2003. Hyperspectral imaging for mapping sediment characteristics. In: 3rd EARSeL Workshop on Imaging Spectroscopy, edited by M Habermeyer, A Müller & S Holzwarth (EARSeL, Paris) 439-446

9 Herman P M J, J J Middelburg & C H R Heip, 2001. Review article. Benthic community structure and sediment processes on an intertidal flat: results from the ECOFLAT project. Continental Shelf Research, 21: 2055-2071

10 Van Engeland T, 2005. Using field data and hyperspectral remote sensing to model microalgal distribution and primary production on an intertidal mudflat. MSc thesis, Ghent University, 33 pp.

11 Ryu J-H, Y-H Na, J-S Won & R Doerffer, 2004. A critical grain size for Landsat ETM+ investigations into intertidal sediments: a case study of the Gomso tidal flats, Korea. Estuarine, Coastal and Shelf Science, 60: 491-502

12 Mitchener H & H Torfs, 1996. Erosion of mud/sand mixtures. Coastal Engineering, 29: 1-25

13 ENVI User's Guide, 2000, ENVI Version 4.0, Research Systems, Inc.

14 Kruse F A, A B Lefkoff, J B Boardman, K B Heidebrecht, A T Shapiro, P J Barloon & A F H Goetz, 1993. The Spectral Image Processing System (SIPS) - Interactive Visualization and Analysis of Imaging spectrometer Data. Remote Sensing of Environment, 44: 145 - 163

15 Adam S, S Salama & J Monbaliu, 2004. Sediment characterization in 'De IJzermonding' using an empirical orthogonal function: application to CASI. In: Airborne Imaging Spectroscopy Workshop, Bruges, Belgium 8 October 2004 (TELSAT, Brussels, Belgium)

16 Richards J A, 1993. Remote sensing digital image analysis: an introduction (2nd ed.) (Berlin: Springer-Verlag) 363 pp.

17 Tou J T & R C Gonzalez, 1974. Pattern Recognition Principles (Addison-Wesley Publishing Company, Reading, Massachusetts) 377pp.

18 Weidung L, F Baret, G Xingfa, T Qingxi, Z Lanfen & Z Bing, 2002. Relating soil surface moisture to reflectance. Remote Sensing of Environment, 81: 238-246

19 Hakvoort H, K Heymann, C Stein & D Murphy, 1997. In-situ optical measurements of sediment type and phytobenthos of tidal flats: a basis for imaging remote sensing spectroscopy. German Journal of Hydrography, 49: 367-373

20 Carrere V, 2003. Mapping microphytobenthos in the intertidal zone of Northern France using high spectral resolution field and airborne data. In: 3rd EARSeL Workshop on Imaging Spectroscopy, edited by M Habermeyer, A Müller & S Holzwarth (EARSeL, Herrsching) 395-410

21 Jiang, H, J R Strittholt, P A Frost & N C Slosser, 2004. The classification of late seral forests in the Pacific Northwest, USA using Landsat ETM+ imagery. Remote Sensing of Environment, 91: 320-331