Prepared for:
Rijkswaterstaat, Dienst Getijdewateren

Bathymetry from landsat–TM images

AFGEHANDELD

May 1989

delft hydraulics
Bathymetry from landsat–TM images

G.H.F.M. Hesselmans
BATHYMETRY FROM LANDSAT-TM IN MAGES

PREFACE

This report contains the results of a research project (DGW 925A) carried out for Rijkswaterstaat on "bathymetry from passive remote sensing observations". The project has been carried out by Dr. G.H.F.M. Hesselmans and ir. G.J. Wensink (project leader Delft Hydraulics). Preceding this project a study has been performed at Rijkswaterstaat Tidal Waters Division. This study was carried out by ir. W. van Hengel and Dr. ir. D. Spitzer (project leader Rijkswaterstaat) and partly funded by the Netherlands Remote Sensing Board. The results have been published in 1988: "Quality analysis of water depth mapping by means of Landsat TM" RWS/MD, and in a report for the Netherlands Remote Sensing Board (bcrs): "Water depth mapping by means of Landsat TM". Although both studies are directed towards mapping of water depth by means of Landsat TM-images and both studies apply the depth algorithm by Lyzenga et al., research methodology differed. In this report a systematic approach is presented to determine uniform regions that form a prerequisite to the practical use of the algorithm by Lyzenga et al. Both methods are complementary.
ABSTRACT

Using the algorithm suggested by Lyzenga et al. (1978) for water depth extraction from remote sensing images, the applicability of Landsat Thematic Mapper imagery has been investigated. The test site was located near the Dutch Wadden island Vlieland. Since both depth information and remote sensing images are contaminated with noise and measuring inaccuracies, a robust regression analysis based on principal components was used. This analysis yields the optimal direction in colour space to extract the depth information.

An additional advantage of the 'principal component' approach is that the main direction of the residual errors in colour space can be determined. In this report the latter direction was used to determine potential sub-regions with a homogeneous bottom and water composition. Uniform regions are a prerequisite to obtain reliable results based on the algorithm by Lyzenga et al.

An iterative approach is proposed to find such uniform regions. In each step better estimates of the optimal depth direction in a certain region are obtained, and at the same time inhomogeneities in regions previously assumed uniform can be detected. According to this analysis several sub-regions could be distinguished in the test region.
1. Introduction

Remote sensing can be used to update charts through the detection and location of uncharted or misplaced reefs and shoals hazardous to navigation. Past studies (Hammack 1977) used data of the Multispectral Scanner System (MSS) on board of the Landsat series of satellites. Unfortunately the 80 m spatial resolution of the MSS limits the charting to relatively large features, while the spectral resolution limits the water depth extraction capability of the instrument to using one or two spectral bands. Gordon and McCluney (1975) report that under the best theoretical conditions, with the sun at zenith, the theoretical penetration depth of the green band is less than 20m while the penetration depth of the red band of the MSS is about 2 m. In this report attention is focused on the Thematic Mapper (TM) a new 7 channel multispectral scanner carried by Landsat 4 and 5. For details see table 1. Its higher 30m spatial resolution is useful for locating smaller hydrographic features, such as rock pinnacles, coral heads, or sea mounts. A repetitive coverage of 16 days permits the identification and separation of permanent bottom features from transient effects, such as water quality and atmospheric effects.

The Thematic Mapper on the Landsat satellite has 7 wavebands. The first three are of direct interest for oceanographic research. Light from these bands may penetrate into the sea. The fourth band however can be used to discriminate between land and sea. For a short description of these four bands see the Appendix. The first three TM bands have been used to make bathymetric maps of a part of the Eastern Wadden Shallows (northern Netherlands). More specific the vliesloot a gully between the Dutch island Vlieland and a sandbank called de richel (see Figure 1). The intensity of the blue band of the image is shown in Figure 2.

The aim of this report is to improve the bottom bathymetry algorithm introduced by Poleyn, Brown and Sattinger (1970) and elaborated by Lyzenga (1978) in two ways. First, a more robust estimation procedure by means of principal component analysis is presented. Second, a method of selecting appropriate uniform regions is presented.
<table>
<thead>
<tr>
<th>frequency band number</th>
<th>frequency range (nm)</th>
<th>signal/noise screen radiance minimum</th>
<th>maximum</th>
<th>colour</th>
<th>resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>450-520</td>
<td>52</td>
<td>143</td>
<td>blue</td>
<td>30 m x 30 m</td>
</tr>
<tr>
<td>2</td>
<td>520-560</td>
<td>60</td>
<td>279</td>
<td>green</td>
<td>30 m x 30 m</td>
</tr>
<tr>
<td>3</td>
<td>630-900</td>
<td>48</td>
<td>248</td>
<td>red</td>
<td>30 m x 30 m</td>
</tr>
<tr>
<td>4</td>
<td>760-900</td>
<td>35</td>
<td>342</td>
<td>infra-red</td>
<td>30 m x 30 m</td>
</tr>
<tr>
<td>5</td>
<td>1550-1750</td>
<td>±17</td>
<td>±110</td>
<td></td>
<td>30 m x 30 m</td>
</tr>
<tr>
<td>7</td>
<td>2080-2350</td>
<td></td>
<td></td>
<td></td>
<td>30 m x 30 m</td>
</tr>
<tr>
<td>6</td>
<td>10400-12500</td>
<td></td>
<td></td>
<td></td>
<td>120 m x 120 m</td>
</tr>
</tbody>
</table>

Orbit: Circular, near polar
Inclination: 98.2
Altitude: 705 km
Coverage: 81N to 81S
Period: 99 minutes, crossing Equator at 9:45 a.m., local time
Swath width: 185 km
Cycle: 16 days
Quantization levels: 256

Tabel 1 Landsat TM characteristics

Fig. 1 Location of test area near Vlieland
In the next section (Section 2) a theoretical formula is given which relates radiances and water depth. In Section 3 some problems encountered in the realization and implementation of the solution are discussed. In Section 4 the procedure is sketched for processing of remote sensing data to map shallow water bathymetry. The optimal strategy is to combine the 3 TM bands into one algorithm and to calibrate this algorithm by using available seatruth. In Section 5 the results of the method sketched in Section 4 are presented, followed by a discussion in Section 6.

![Vlieland TM-image 16-6-1986 Band 1](image)

**Fig. 2** TM-image of Vlieland test area

A note on the use of principal component analysis by van Hengel (1988) in his related report is desired. His use of principal components is motivated by the need to:
- suppress noise, and
- be able to calculate depth estimates in the absence of sea truth data.
In case of an almost uniform bottom and water composition the theoretical results obtained by Lyzenga indicate that the first principal component corresponds to the optimal depth direction. The higher order principal components describe noise and minor deviations in bottom and water composition. Therefore both objectives can be met in the convenient case of a uniform bottom and water composition.

However in the presence of sub-regions with a distinct bottom and water composition the first principal component no longer is a reliable depth indicator. If, for the sake of the argument, two optically quite different bottom types are present and the range of depth values is moderate, than the first principal component might even be almost orthogonal to the desired direction. This is illustrated in the following simplified Figure:

According to Lyzenga et al. the optimal depth related direction is parallel to the two lines. It is quite obvious that this direction cannot be found by a simple principal component analysis on all data. On the contrary either the data should by split into two groups by means of a cluster analysis or a priori available depth information should be used. The latter is done in this report. Sea truth is used to get a first estimate of the depth direction and then the data are projected in the orthogonal plane. In the orthogonal plane the data can easily be separated into two groups. In successive iteration steps the same analysis can be applied to data corresponding to the separate sub-regions.

Van Hengel (1988) uses the first principal component based on all colour data. No further attempt is made to separate the test site into smaller regions with different optical properties. This may form an extra reason for the poor correlation of sea truth and his computed maps. Alternatively van Hengel and Spitzer (1988) calculated depth maps using standard regression analysis and all three available visible wave bands of the Thematic Mapper. However, neither report attempts to separate the test site under consideration into homogeneous sub-regions. This report presents a flexible and effective approach to detect distinct regions and calibrate the algorithm in each sub-region.
2. Theory

Light coming from below the surface of the sea yields information about all substances in the water, the bottom composition and water depth. The problem is how to get information concerning a specific aspect. Satellite sensors operate in different wave bands. Since each substance in the water has its own spectral signature it is possible, given the concentration of these substances, to estimate the contribution of each group of optically similar substances to each wave band. Unfortunately the number of substances outweighs by far the number of wave bands. Therefore it is impossible to derive the contribution of each substance from the observed radiation. To derive a relation between depth and radiation it is not necessary to know the separate contributions but only to know the effect of all substances lumped together.

A first model for depth processing was formulated by Jerlov (1976). This model relates the radiance in a single wavelength band to water depth. Bottom reflectance and effective water attenuation are the principal ingredients of this model. In summary, the single band model assumes the following (Hallada 1984):

1) bottom reflectance is constant throughout the image, or can be determined at each pixel;
2) water attenuation coefficient is constant throughout the image, and independent of water depth;
3) elastic backscatter by the water can be ignored;
4) atmosphere and sea state are uniform throughout the image;
5) effects of multiple scattering between the water surface and the bottom can be ignored, or assumed constant throughout the image;
6) angular distribution of light does not change with depth; and
7) there is no contribution to the upwelling radiance by the reflected sky radiance from the water surface (i.e. glitter and sun glint).

Dirks and Spitzer (1987) present a two-flow radiative transfer model which handles the effects of backscatter (3) and multiple scattering (5). Variation of the irradiance reflectance \( R \); the ratio of the upwelling (index \( u \)) and downwelling (index \( d \)) irradiance, with the depth \( z \) is expressed by the Ricatti equation:

\[
\frac{dR}{dz} = -b_d + (a_u + b_u + a_d + b_d) R - b_u R^2
\]
where:

\( a_{u,d} \) = absorption coefficient
\( b_{u,d} \) = backscatter coefficient

At the bottom we find a boundary condition: \( R(h) = r \), where \( h \) is the bottom depth and \( r \) is the bottom reflectance. The irradiance reflectance \( R \) just under the water-air interface can be calculated from the bottom depth \( h \) by:

\[
R_O = R_q + \frac{(R_q - R_a) \left[ \frac{r-R}{r-R_a} \right] \exp(-\chi h)}{1 - \left[ \frac{r-R}{r-R_a} \right] \exp(-\chi h)}
\]

where:

\[
R_q = \frac{(c+x)}{2b_u}
\]

\[
R_a = \frac{(c+x)}{2b_u} = \sqrt{c^2 - 4b_u b_d}
\]

\[
c_u = a_u + b_u
\]

\[
c_d = a_d + b_d
\]

\[
c = c_u + c_d
\]

Absorption and scattering is ascribed to suspended and dissolved materials in the water; \([a]\) chlorophyllous material, \([b]\) non-chlorophyllous material and \([c]\) the coloured component of the dissolved organic matter (yellow substance). The bottom reflectance depends on the particle size and composition of the sediment (sand-mud) and whether or not vegetation is present.

In most natural waters \( \frac{b}{a} \ll 1 \) and \( \frac{b}{a} < r \). The reflectance equation can then be approximated (Spitzer and Dirks 1987), resulting in:

\[
R_O = R_\infty + r e^{-\chi h}
\]

where \( R_\infty \) is the mean deep-water radiance. This is effectively the equation derived by Jerlov. The bottom depth can now be calculated from:

\[
h = \frac{\ln(r)}{\chi} - \frac{1}{\chi} \chi_0
\]
where:
\[ X_0 = \ln(R_w - R_o). \]

To reduce the errors in depth calculation due to spatial changes in bottom reflectance and water attenuation spectral wavebands should be combined. A linear algorithm of the type:
\[ h = X_0 + \sum_i x_i X_i \]
is possible if the effective attenuation \( x_i \) in each waveband is constant over the whole scene, and if the ratio of bottom reflectance in each combination of two bands is constant. Paredes and Spero (1983) show that the bottom reflectances can satisfy a more general assumption. In the literature no cases have been reported where the ratio assumption holds. If there are at least as many bands as there are bottom types in the scene, then total independence from bottom type variation can be achieved. Since the Landsat TM has three water penetrating bands, at most total independence of bottom reflectance can be obtained for three bottom types. Independence of bottom types can only by obtained for depths less than the minimum penetration depth of all bands used.
3. Error sources

3.1 Geometric distortions

Geometric distortions are caused by the optics, scan mechanism and detector-array geometry, spacecraft attitude and attitude variations; and the earth's rotation. These distortions can be removed by a transformation using ground-control-points and cubic convolution or nearest-neighbour resampling algorithms. In case of Landsat MSS the earth's rotation causes the largest distortion (see Figure 3). Skewness is largest at the equator and decreases towards the poles. In case of the TM a motor driven scan corrector is present (Slater 1980). The scan corrector produces scans across the swath that are perpendicular to the spacecraft velocity vector - not at an angle. Ground-control-points however are still necessary since the actual image center can vary as much as 20 km from the nominal scene center (Scott Southworth 1985). Furthermore the altitude of the satellite changes slightly causing small variations (≤ 1%) in the size of a single pixel.

Fig. 3 Distortion caused by the relative motion of earth and satellite

3.2 Atmospheric limitations

The radiant flux from the sun is partially absorbed and scattered as it passes down through the atmosphere to the surface of the earth (see Figure 4). The ground scene reflects part of the flux incident on it in the direction of the orbital remote sensor. This reflected flux passing through the atmosphere is again absorbed and scattered, but to it is added scattered light from the atmosphere that has not been reflected from the ground scene. These upward-
scattering effects are the most insidious effects of the atmosphere, as the flux appears to the remote sensor as if it came from the ground scene of interest. Its effect is to reduce the contrast (blur) of the scene and thus make fine detail invisible or at least harder to detect. Furthermore, in adding a uniform flux level to that from the ground scene, it confuses the interpretation of the spectral signatures of the scene. This upward scattering effect, or haze, from the atmosphere is most pronounced in the blue. The signal \( L \) received by a remote sensor can be expressed as follows (Vanouplis 1986):

\[
L = L_{SG} + L_{HG} + L_{w} + L_{p}
\]

where:

\( L_{SG} \) = the sunglitter, i.e. radiance due to direct solar radiance from the water surface

\( L_{HG} \) = the skyglitter, i.e. radiance due to diffuse radiance reflected from the water surface

\( L_{w} \) = water leaving radiance, i.e. radiance of importance for bathymetry

\( L_{p} \) = path radiance, i.e. radiance scattered from molecules (Rayleigh scattering and microscopic particles (aerosol/Mie scattering) in the atmosphere

Fig. 4 Overall picture of interactions between incident and reflected radiance, the atmospheric and the Earth’s surface
The amount of upwelling atmospheric radiances in remotely sensed data is a function of many variables: sensor altitude, atmospheric haze conditions, solar zenith angle, spectral-sensitivity range of sensor, angle of view from nadir and azimuth with respect to the sun, and polarization. Fortunately the addition of a uniform radiances level is automatically compensated for in the calibration of the bottom-depth algorithm mentioned in the previous section.

3.3 Sensor errors

Due to instability and drift in the electronics and bleaching of the absorption filters by exposure to ultraviolet radiation the radiometric sensitivity of the various detectors may change in time (stability 10% over 2 years). Apart from this long term effect a more pronounced feature of the scanning technique is that 16 lines (6 for MSS) are scanned simultaneously by 16 different sensors. Each sensor has its own offset and gain. Since gains differ by more than 10% correction of the image (destriping) is required. The usual approach is to assume that the statistical properties of all 16 sets of lines are equal. This hypothesis can be used to equalize the output of each sensor (Poros and Peterson 1985).

3.4 Environmental conditions

Bottom

As can be understood from the foregoing knowledge of parameters other than bottom depth that influence the satellite observations may be of vital importance. The reflectance changes from 3 to 4% for mud via 4 to 12% for vegetation to 15 to 20% for sand (Muirhead and Cracknell 1986).

Water conditions

Due to suspended material the irradiance may vary by several orders of magnitude (Spitzer and Dirks 1987). A first inspection of the satellite image, which was taken one hour after high tide, shows suspended material in the water which seems to be streaming into the North Sea.

Water depth

Soundings give an estimate of the depth based on the height of the watercolumn near the exploration vessel and the expected tide. An indication of variability
in (measured) bottom depth can be found in the Geomorfologische kaart van de Nederlandse kust wateren. This map shows that away from the gullies changes no larger than a few tenths of a meter are to be expected within a year. Furthermore a comparison with older sounding maps shows that in the gully local changes of a few meters are possible.
4. Method

4.1 Geometric correction

Satellite altitude fluctuations are of minor importance for TM-images and skewness due to the relative earth-satellite motion is suppressed by the scan corrector.

4.2 Atmospheric correction

Correction of the image for atmospheric influences can be done in principle in a spatial uniform atmosphere using only meteorological information about the Rayleigh and ozone optical thickness (van Hengel and Spitzer 1987). The correction consists of a constant added to each pixel in the scene. However, this has neither influence on the outcome of the bottom bathymetry algorithm nor on its parameters, only on the offset of the radiance values. Since only differences are used no atmospheric correction needs to be applied. The atmospheric correction can be applied to uniform regions which can be found by means of clustering techniques (see section 4.3 on the selection of uniform regions). The most striking phenomenon that should be removed by clustering techniques is clouds.

4.3 Image restoration

The periodic structure seen in TM-images (and MSS) is an arte-fact of the measuring procedure and is removed for further analysis by a method called destriping (see Appendix).

Noise can be removed by local averaging of each colour or by datatransformations (principal components, see Appendix) which separates the image in a noisy part and a noise-free component.

4.4 Land-sea segmentation

The next step in any analysis of satellite images for bottom bathymetric purposes is to distinguish between land and water. Such an analysis is a matter of pattern recognition. To label regions in the test site the KMEANS algorithm can be used (Albers et al. 1987, see also Appendix). This algorithm finds data
points with a similar colour composition. Data points on the border of two regions (mixed pixels) are likely to be incorrectly classified and are reassigned the class label of the majority of its neighbours.

4.5 Selection of uniform regions

Since the irradiance depends upon bottom and water composition the theoretical procedure may be applied only to regions where these conditions do not vary. Therefore a kind of pre-processing of the image is necessary to find/classify regions with constant bottom and water properties. Lyzenga (1978, 1981) shows that the bottom reflectance can be estimated if the water optical properties are uniform over fairly large areas. If the logarithmic model mentioned in Section 2 is correct and the bottom reflectance is limited to a few distinct classes, then a plot of $X_i$ versus $X_j$ will form a set of parallel lines (see Figure 5). Each line will correspond to a particular bottom reflectance and the variation along these lines is due to variation in depth. The slope of such a line depends on the ratio of the irradiance attenuation coefficient of the water in bands $i$ and $j$:

$$X_j = \frac{x_i \ln(r_i) - x_j \ln(r_j)}{x_i}$$

The amount of displacement of the lines is a measure of the change in the bottom reflectance. The distance between the two lines is given by:

$$\frac{x_i \ln(r_i) - x_j \ln(r_j)}{\sqrt{x_i^2 + x_j^2}}$$

Fig. 5 Scatter plot of log (radiance) values for mixed bottom

Therefore different clusters can be found in colour-space if different bottom types are present and water composition is uniform. These clusters may be
found by means of the KMEANS clustering technique mentioned in the previous paragraph (4.4). However if the depth profile shows a steep slope then the corresponding line in the colour-space may be interrupted and undesired clusters may be found. This can be avoided by a projection of the data into the sub-space orthogonal to the depth-dependent direction. This projection is meaningless if the colour-space is 1-dimensional, i.e. data in colour-space are grouped along a single line. Whether or not this is the case can be verified by a principal component analysis.

The same procedure can be applied if bottom conditions are uniform and water properties are different.

The result of the clustering procedure consists of a number of uniform regions. Whether or not these areas differ with respect to bottom and water composition is still to be evaluated. In both cases, the clustering procedure does not yield the correct result if the bottom reflectance and/or the water composition varies with depth within the area selected.

In this report the orthogonal projection is used in an iterative procedure. First the overall depth dependent direction is colour space is determined, then the data are projected in the orthogonal plane. This leads to candidate uniform regions. Then the procedure is repeated and applied to each candidate region, till no more sub regions can be distinguished.

4.4 Bathymetric maps

Based on training data from the sounding maps the algorithm from Section 2 is optimized for each uniform water area. Deep-water radiances are obtained by averaging over part of a uniform region. If deep-water regions are not present a suitable deep-water radiance value is chosen. Deep-water radiance values are subtracted from the corresponding waveband. Since the latter subtraction may lead to negative values, these data are truncated before the logarithmic transformation is performed which leads to the \( X_i \).

Since the non-linear logarithmic transformation may introduce relatively extreme \( X_i \) values, a robust linear regression method based on principal component analysis is used. A further advantage of the method is that both errors in satellite observations and bathymetric chart are taken into account. Ones the parameters have been determined the algorithm can be applied to each region of the scene.
5. Results

5.1 Environmental conditions

During this study relevant information such as bottom and water composition of the test site was gathered. In spite of much interest in the Wadden Sea, bottom classification maps appear to be scarce. Recent information concerning the Danish Wadden Sea can be found in a paper by Bartholdy and Folving (1986), however the most recent study from the Dutch part of the Wadden Sea by Glopper stems from 1967 and is only concerned with high lying sand-banks. Therefore this information can only be indicative.

As yet no information concerning surface slicks and/or the composition of the water masses in the Vlieland region and its optical properties has been found. In the period November 1987 - March 1988 depth soundings were made by Rijkswaterstaat in the Wadden Sea area. These measurements have been used to obtain a digitised map of the region. The satellite images were recorded on June 16, 1986 and June 23, 1986. So a period of more than a year has elapsed between the moment of satellite image recording and the acquisition of depth measurements. Therefore the actual depth at the moment the satellite picture is taken is not known due to possible changes in gully position.

5.2 Land-sea segmentation

To separate land from sea clusters were searched and found with the KMEANS algorithm. The result is shown in Figure 6. Four distinct regions can be distinguished: dunes, forest, land and sea.

5.3 Information content of sea data

The three TM-wavebands in both images of the Vlieland area are highly correlated, as can be concluded from a principal component analysis of the water part of these image.

<table>
<thead>
<tr>
<th>eigenvalues</th>
<th>16-6-1986</th>
<th>23-6-1986</th>
</tr>
</thead>
<tbody>
<tr>
<td>variance component 1</td>
<td>26.4 (91.8%)</td>
<td>26.2 (90.8%)</td>
</tr>
<tr>
<td>variance component 2</td>
<td>1.8 (6.1%)</td>
<td>2.0 (6.7%)</td>
</tr>
<tr>
<td>variance component 3</td>
<td>0.6 (2.0%)</td>
<td>0.7 (2.5%)</td>
</tr>
</tbody>
</table>
Fig. 6 Land-sea segmentation for the Vlieland test site

Because of this high correlation it is expected that it will be difficult to distinguish the separate influences of depth, bottom composition and water composition to the radiance levels in the 3 bands. A way to remove noise from an image is by using the spatial correlation. One might expect that the spatial filtering of the images will result in larger eigenvalues for the higher order components. These higher order components provide the necessary information to select uniform regions in a later stage. Linear filtering by a Gaussian filter on a 3 x 3 square shows that this is not the case.

<table>
<thead>
<tr>
<th>eigenvalues</th>
<th>16-6-1986</th>
<th>23-6-1986</th>
<th>filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>variance component 1</td>
<td>23.3 (96.4%)</td>
<td>22.7 (94.5%)</td>
<td>1 2 1</td>
</tr>
<tr>
<td>variance component 2</td>
<td>0.7 (2.8%)</td>
<td>1.0 (4.1%)</td>
<td>2 4 2</td>
</tr>
<tr>
<td>variance component 3</td>
<td>0.2 (0.8%)</td>
<td>0.3 (1.4%)</td>
<td>1 2 1</td>
</tr>
</tbody>
</table>
The image recorded on June 23 seems to be most promising. The first component is very smooth compared to the second one (see Figure 7). One might say that in both images the first component contains all the information, while the higher components represent noise. Note that due to the discrete nature of the satellite images the variance of the higher components is expected to be of order of magnitude 1.

Fig. 7 First and second principal colour component

5.4 Sea segmentation

A search for clusters in the sea area did not lead to more distinct areas with different types of bottom or water composition. Clusters found by the KMEANS algorithm appeared to be spatially mingled. The clusters seem to be more related to depth than to bottom and water composition, i.e. variation in depth is larger than variations due to bottom and water composition or depth is strongly correlated with bottom and water composition. Without further in situ measurements it cannot be concluded whether or not the differences are real.
Fig. 8 Digitized bathymetric chart of the Vlieland test site

5.5 Depth calibration

Bottom depth values from the sounding map (see Figure 8) were used to calibrate the bottom bathymetry algorithm. After the Gaussian filter had been applied the following estimates for the deep water radiances were made.

<table>
<thead>
<tr>
<th>deepwater radiance</th>
<th>16-6-1986</th>
<th>23-6-1986</th>
</tr>
</thead>
<tbody>
<tr>
<td>radiance band 1</td>
<td>86</td>
<td>68</td>
</tr>
<tr>
<td>radiance band 2</td>
<td>30</td>
<td>22</td>
</tr>
<tr>
<td>radiance band 3</td>
<td>24</td>
<td>18</td>
</tr>
</tbody>
</table>

Before the logarithmic transformation was performed, data were truncated at 0.2. This resulted in the following algorithms:

\[
\text{depth} = -25.6 - 76.9 \times X_1 + 750 \times X_2 - 509 \times X_3 \quad \text{[dm]} \quad \text{image 16-6-1986}
\]

\[
\text{depth} = 138 - 184 \times X_1 + 244 \times X_2 - 11.8 \times X_3 \quad \text{[dm]} \quad \text{image 23-6-1986}
\]
The calibrated algorithm was applied to the whole sea area (see Figure 9). Comparison with the sounding map shows a qualitative resemblance. However, a more detailed inspection shows that the estimated position and depth of specific gullies are incorrect. The rms errors are 100 dm and 46 dm for the images of June 16 and June 23 respectively.

![Bathymetric chart of the Vlieland test site](image)

Fig. 9 Bathymetric chart of the Vlieland test site

Finally, the data of both images were projected into the plane orthogonal to the depth dependent direction. This allows a guided search for uniform areas. The component showing the largest variance for the image of June 23 is shown in Figure 10.

The result is a smooth image which shows no distinct boundaries and no distinct uniform regions. A gray level distribution histogram of this image shows one broad Gaussian shaped bump and a small peak representing part of the North Sea. It appears that this projection is strongly correlated \((r = 0.99934)\) with the first principal component of the logarithmic X-data. In spite of the poor sepa-
rability of the area into uniform regions an attempt was made to distinguish 7 areas for the image of June 23. The classification is shown in Figure 11. The algorithm has been calibrated for each region with the following result:

<table>
<thead>
<tr>
<th>area</th>
<th>$x_0$</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>rms error [in dm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 North sea</td>
<td>186</td>
<td>-85</td>
<td>101</td>
<td>1.3</td>
<td>36</td>
</tr>
<tr>
<td>2 North sea</td>
<td>49</td>
<td>-134</td>
<td>257</td>
<td>19</td>
<td>35</td>
</tr>
<tr>
<td>3 Wadden sea (shallow)</td>
<td>-314</td>
<td>-237</td>
<td>746</td>
<td>-249</td>
<td>52</td>
</tr>
<tr>
<td>4 Wadden sea (shallow)</td>
<td>-389</td>
<td>-373</td>
<td>1183</td>
<td>-536</td>
<td>60</td>
</tr>
<tr>
<td>5 near beach</td>
<td>-32</td>
<td>-194</td>
<td>492</td>
<td>-183</td>
<td>43</td>
</tr>
<tr>
<td>6 Wadden sea (deep)</td>
<td>22</td>
<td>-4.8</td>
<td>169</td>
<td>-4.6</td>
<td>17</td>
</tr>
<tr>
<td>7 Wadden sea (deep)</td>
<td>-307</td>
<td>-7.1</td>
<td>587</td>
<td>-329</td>
<td>47</td>
</tr>
</tbody>
</table>

**Fig. 10** Projection of the data into the plane orthogonal to the depth direction. The dominant direction in this plane is shown.
Fig. 11 Segmentation of sea area based on orthogonal projection
6. Discussion, conclusions and recommendations

Segmentation of a satellite image into land-sea areas is rather easy and the criterion is mainly a matter of the total radiance measured in all three wavebands. However the difference between deep-sea and the forest on Vlieland is rather small and depends on the different ratio of green and blue. Separation of land and sea areas can be done with the fourth band (infra red) as well. Segmentation into land and sea uses the whole data range of about 100 digitization levels. Whereas a further division of the water-masses into areas with different bottom and water composition can be based upon no more than 10 levels. Selection of regions with constant properties is therefore difficult without an adequate model and every interpretation should be checked by in situ measurements.

The strong correlation between the separate wavebands indicates that a separation of the data in a depth dependent and a depth independent part might be difficult. Theory (chapter 2) suggests that the logarithm of the radiances should be linearly correlated and not the radiance levels. This strong correlation indicates that if bottom and water composition differences are present, they are strongly correlated with depth. This limits the applicability of the theory (chapter 2) and prohibits the selection of uniform subregions.

Figures 8 and 9 show clearly that the in situ measurements provided by the soundings map and the satellite image based map do not match. Gullies seem to have been moved over distances of more than 100 meters. Since this is of the same order of magnitude as the width of a single gully, this is fatal in our approach. If the position of the gullies is questionable than its depth is unreliable as well. The latter prevents an adequate calibration of the algorithm and explains the large differences between 'sea truth' and calculated bathymetric map.

The two satellite images have been obtained at different tides, i.e. outgoing tide for the image of June 23 and ingoing tide for the image of June 16. Since different tides result in different sediment currents this may account for some of the differences found in the two images. Away from these regions with suspended matter results should be more accurate. Unfortunately sediments currents seem to be predominant in the gullies, which form the most interesting part.
The logarithmic transformation of the irradiance values allows a linear relation between the $X_i$ and depth. However two types of errors are introduced in this procedure. First, too low radiance levels are truncated to avoid non-positive values and small positive radiance levels yield large logarithmic outputs. Second, the relative error is changed. If errors were equal before the transformation, this is not so afterwards. Therefore different weights should be given to the various measurements depending on the radiance level.

The regression procedure based on principal component analysis is less sensitive to outliers than ordinary linear regression based on least square techniques. However, it still assigns equal weights to all data points and the effect of truncation is not properly handled. To avoid all these problems it may be better to fit the original model relating irradiance and depth and not to perform the logarithmic transformation. This will lead to increased computational efforts, but does more justice to the physics.

The projection of the data into the space orthogonal to the main depth dependent direction offers the possibility to check whether or not the algorithm is applied to uniform regions. This feature of the analysis procedure provides a sensitive control mechanism and may be of great importance in areas where in situ measurements are hard to get, since it allows a more selective use of measuring efforts. A prerequisite for this projection is that the data are not too strongly correlated.

In this report we find that the first principal component of the logarithmic X-data almost completely matches the dominating direction in the residual-space. Therefore we conclude that the optimal depth direction is described by a combination of second and third principal component (this was verified by a direct analysis). This finding cannot be reconciled with the approach by van Hengel (1988) and puts serious doubt on the reliability of the available sea truth and/or on the assumption of a uniform bottom and water composition in the test area. We must admit that the rate of convergence of the presented iteration approach depends on the amount of homogeneity of the water composition. A uniform water composition results in an unambiguous estimate of the optimal depth direction in the first iteration step, while large differences in water composition may result in poor estimates of the optimal direction and make sub-regions indistinguishable.

In short techniques have been developed to make bathymetric maps. The necessary procedures have been implemented, e.g. clustering technique, robust regression, orthogonal projection. However the procedure could not be tested because of a
lack of reliable data. The drift over more than 100 meters found appears to be quite plausible. Therefore further effort should be primarily directed towards the acquisition of reliable data. An experiment in this line is in progress at this moment. In situ data will be collected near Saba by the Dutch Royal Marine.
REFERENCES


REFERENCES (continued)


APPENDIX A
A.1 Principal component analysis

Principal component analysis tries to explain the variance of a n-dimensional set of data. The first component gives the direction in which the variance is largest. The second component explains most of the variance of the data in the sub-space orthogonal to the first component. The third component is orthogonal to the first and second component and explains most of the variance of the data projected into the sub-space orthogonal to the first two components. In this way the procedure can be extended to at most n components. Each component explains less variance than the previous one.

The rational behind this procedure is that the first components will not only explain most of the variance but also express most of the information contained in the data. Two algorithms are provided by the literature. Although mathematically equivalent their are some practical differences.

The first algorithm is known as singular value decomposition (SVD) and starts from the original data. SVD requires a matrix of size n x m. With n the dimensionality of the data (the number of wavebands), and m the number of data points (the number of pixels in the satellite image). In case of a reasonably sized image this matrix becomes to large to handle on a personal computer. Therefore this method was abandoned.

The second algorithm uses the covariance matrix of the data. This n x n sized matrix is much more easy to handle. Two routines (TRED2 and TQL1) from the Numerical Recipes book have been used. The combination of these two routines yields the most efficient known technique for finding all the eigenvalues and eigenvectors of a real, symmetric matrix, according to the authors of this book.

A.2 Robust estimation

Principal component analysis can be used in a more robust estimation of a straight line through a number of data points. If all variables show some kind of measuring error (e.g. both bottom depth and radiances are noisy) then both errors should be dealt with in the linear regression. This is done by the principal component method. A further advantage is that the method is less sensitive to outliers (see Figure 12).
Fig. 12 Examples where robust regression analysis is desirable

Suppose we want to find an optimal linear relation between \( y, x_1, \ldots, x_n \), i.e. we like to find good estimates for \( a_i \) in:

\[
y = a_0 + a_1 x_1 + \ldots + a_n x_n
\]

The procedure is implemented as follows. First, the average values \( \bar{y}, \bar{x}_1, \ldots, \bar{x}_n \) are subtracted from respectively \( y, x_1, \ldots, x_n \). Then the \( n+1 \) principal components in the \( n+1 \) dimensional data space are calculated and the component having the smallest eigenvalue is selected. Let its components be: \( e_0, e_1, \ldots, e_n \). Then, the optimal \( a_i \) are given by:

\[
a_0 = \frac{-\bar{y}}{e_0} - \frac{e_1 + \ldots + e_n}{e_0}
\]

\[
a_i = -\frac{e_i}{e_0} \quad \forall i \in \{1, \ldots n\}
\]

A.3 KMEANS algorithm

The KMEANS algorithm minimizes the total square distance to a predefined number of clusters. The algorithm works as follows:

1) cluster centers are defined (by the user)
2) each data point is assigned to the nearest cluster center
3) new cluster centers are calculated by taking the average of the data points assigned to the particular cluster
4) data points are reassigned from one cluster \( (i) \) to another cluster \( (j) \) if this decreases the total square distance. This is the case if:

\[
\frac{n_i}{n_j + 1} || \hat{x} - \bar{m}_j ||^2 < \frac{n_i}{n_i - 1} \frac{1}{n_i} || \hat{x} - \bar{m}_i ||^2
\]
where \( n_i \) is the size of cluster \( i \) and \( n_j \) is the size of cluster \( j \), and \( \mathbf{m}_i \) and \( \mathbf{m}_j \) are the respective cluster centers. If both clusters are rather large this amounts to assigning a point to the nearest cluster center. The new cluster centers become:

\[
\mathbf{m}_{i,\text{new}} = \frac{n_i \mathbf{m}_{i,\text{old}} - x}{n_i - 1} \quad \mathbf{m}_{j,\text{new}} = \frac{n_j \mathbf{m}_{j,\text{old}} + x}{n_j + 1}
\]

Step 4) is repeated for all data points until no more changes occur. Although this routine is sure to converge to a minimum, a global optimum is not guaranteed.

A.4 Destriping algorithm

The 16-line period in the TM-landsat images was removed as follows. Of all 16 sets of lines the average \( m_i \) and variance \( V_i \) was calculated. The average \( \bar{m} \) and \( \bar{V} \) of respectively all \( m_i \) and \( V_i \) was determined. Then new values \( x_{\text{new}} \) were calculated for each colour component of each pixel (Richards 1986).

\[
x_{\text{new},i} = \bar{m} + \sqrt{\frac{V}{V_i}} (x_{\text{old},i} - \bar{m})
\]

Another way to destripe a satellite image is by means of Fast Fourier techniques. This method calculates the average value of each row and then the high frequency components of this sequence are removed. Finally the values in each row are adjusted to fit the new average value. A further refinement might be to apply this procedure to the variation of each row as well. Van Hengel (1988) used this technique to destripe images. Drawbacks of the FFT approach are:

- the number of lines is in general not equal to \( 2^N \) and
- some genuine frequency components above a certain threshold may be removed.

In order to extend the length of the array of means/variances to \( 2^N \) assumptions have to be made that may result in the introduction of unintended frequencies. The removal of all frequency components above a threshold is not supported by any theoretical model: only the removal of frequencies corresponding to the number of detectors and higher harmonics has a physical basis. Note that on visual inspection the FFT approach may yield a better (smoother) result.
A.5 Description of TM-bands

The first four TM bands can be characterized by the following qualitative properties (Slater 1980).

**Band 1:** In the blue-band the peak transmittance of clear water is found. A maximum penetration depth of 55m (at 475 nm) in very clear water has been reported by Gordon and Cluney (1975). In coastal waters the maximum penetration shifts to the green due to dissolved organic material (yellow substance) and may decrease considerably. Under ordinary conditions a depth of 25m is more realistic. Some of the radiance is absorbed by blue chlorophyll. Also atmospheric scattering reduces ground radiances.

**Band 2:** The green-band encompasses the region between the two chlorophyll absorption regions and therefore corresponds to the green of healthy vegetation. It has been shown that the ratio of the blue to green radiance of a water body is a measure of the dissolved organic materials and plankton present.

**Band 3:** The red-band encompasses the red chlorophyll absorption region and is used for plant-vigour determinations. In addition this band is the most useful visible spectrum region for soil-boundary and geological boundary discriminations. Surface features often exhibit high contrast in this region, and atmospheric haze is lower than in the rest of the visible spectrum; thus the contrast and therefore the resolution of the imagery is high in this band.

**Band 4:** Ratios of bands 2 and 4 are sensitive to the amounts of green biomass and moisture in vegetation, and as this band corresponds to a region of peak reflection from vegetation, it is useful for vegetation detection and assessment.
Delft Hydraulics

Main Office
Rotterdamseweg 185
P.O. Box 177
2600 MH Delft
The Netherlands
Telephone (31) 15 - 56 93 53
Telefax (31) 15 - 61 96 74
Telex 38176 hydel-nl

Location 'De Voorst'
Voorsterweg 28, Marknesse
P.O. Box 152
8300 AD Emmeloord
The Netherlands
Telephone (31) 5274 - 29 22
Telefax (31) 5274 - 35 73
Telex 42290 hylvo-nl