Norsk Regnesentral
NORWEGIAN COMPUTING CENTER

Note

Confidence estimation

Report for SatHav Olje

Note no SAMBA/09/08

Authors Anne Solberg

Date February 2008

© Copyright:
Norsk Regnesentral

Norsk Regnesentral

Norsk Regnesentral (Norwegian Computing Center, NR) is a private, independent, non-profit foundation established in 1952. NR carries out contract research and development projects in the areas of information and communication technology and applied statistical modeling. The clients are a broad range of industrial, commercial and public service organizations in the national as well as the international market. Our scientific and technical capabilities are further developed in co-operation with The Research Council of Norway and key customers. The results of our projects may take the form of reports, software, prototypes, and short courses. A proof of the confidence and appreciation our clients have for us is given by the fact that most of our new contracts are signed with previous customers.

Norsk Regnesentral
Norwegian Computing Center
Postboks 114, Blindern
NO-0314 Oslo, Norway

Besøksadresse
Office address
Gaustadalléen 23
NO-0373 Oslo, Norway

Telefon · telephone
(+47) 22 85 25 00
Telefaks · telefax
(+47) 22 69 76 60

Internett · internet
www.nr.no
E-post · e-mail
nr@nr.no

Title	**Oil spill confidence estimation**
Authors	**Anne Solberg**
Date	February
Year	2008
Publication number	SAMBA/09/08

Abstract

This note describes the work aimed at automatic confidence estimation as part of SATHAV.

Keywords	Oil spill detection, earth observation
Target group	KSAT
Availability	Open
Project number	220 345
Research field	Earth Observation
Number of pages	39

© Copyright	Norsk Regnesentral

Contents

1 Introduction

This report describes the work done to improve confidence estimation as part of the oil spill detection service done at KSAT. The existing oil spill service uses manual assignment of confidence high, medium or low to each possible oil spill reported as part of the service. The existing procedure was known to be partly subjective with some inter-operator variance. The main task in this project was to see how this could be improved by incorporating automatic condifence estimation methods.

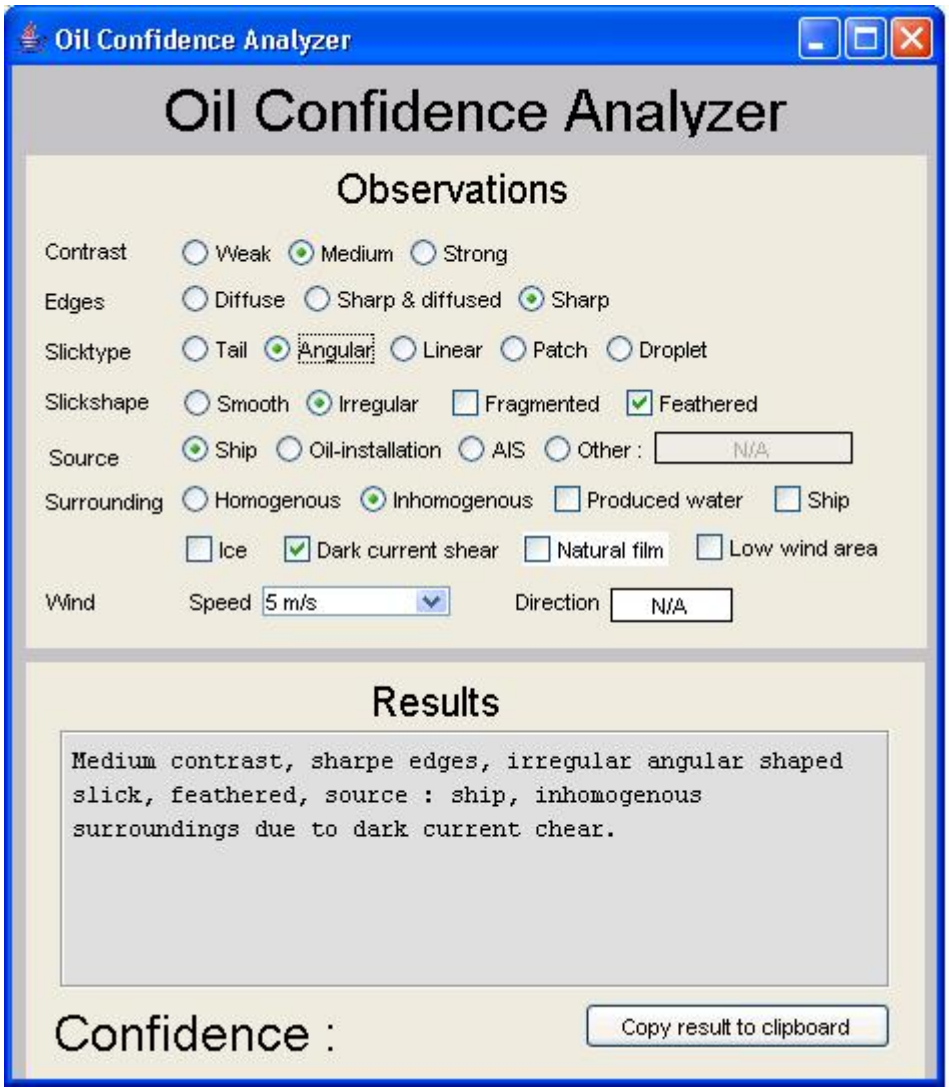
2 Data set and definition of the work

The existing manual assignment of oil spill confidence uses a set of written guidelines, but contains a high amount of subjectivity. To try to reduce the subjectivity, a form called “Oil Spill Confidence Analyzer” was routinely filled out by the operator as part of the oil spill detection service. This form in illustrated in .Figur 1 - Oil spill Confidence Analyser. The first task was to use the information in the Oil Spill Confidence Analyser in combination with oil detection reports to try to define an algorithm for automatic confidence computation based on the parameters set in the Oil Spill Confidence Analyser.

A data set consisting of 16 images with 37 oil spills was prepared by KSAT. NR got access to both the original SAR images and the Oil Spill Detection reports. Available parameters for each oil spill were:

- Slick position
- Confidence (assigned by the operator)
- Region (Land)
- Length
- Width
- Area
- Orientation
- Source
- Contrast
- Edge description
- Slick type
- Slick shape

- Slick surroundings
- Wind direction and wind speed (both from SAR and from forecasts)
- Wave direction and height



Figur 1 - Oil spill Confidence Analyser

The features Contrast, Edges, Slick Type, Shape, and Surroundings were set into a predefined set of categorical classes. The source feature was only set if a source could be directly linked with the spill. **Contrast** was set as either **strong, medium** or **weak. Edges** as **sharp, diffuse** or **sharp&diffuse. Slick ty**pe could be path, droplet, tail, linear or angular. Shape was either smooth or irregular, **and surroundings** either **homogeneous** or **inhomogeneous.**

The 37 oil spills in the data set were assigned the following confidence levels at KSAT:

- High confidence: 3
- Medium confidence: 11
- Low confidence: 23

2.1 Analysis of the oil spills

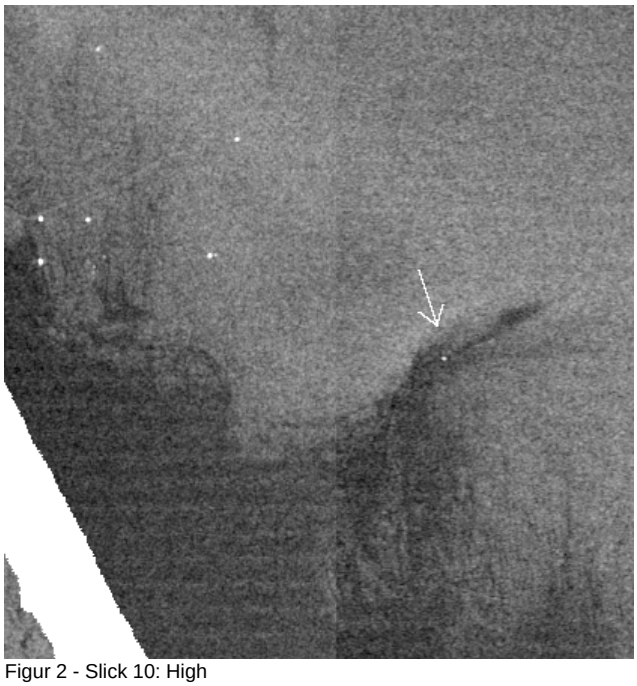
The oil spills are now analysed according to the confidence level assigned at KSAT. Table 1 shows the high confidence slicks and selected feature values. Table 2 shows the medium confidence slicks and their feature values, while Table 3 shows the low confidence slicks.

2.1.1 High confidence slicks

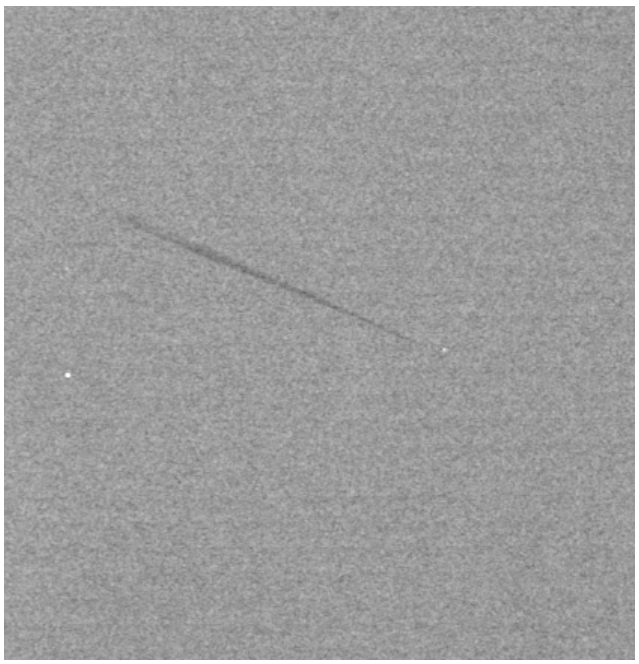
Selected parameters for the high confidence slicks are shown in Table 1.

Tabell 1 –High confidence slicks

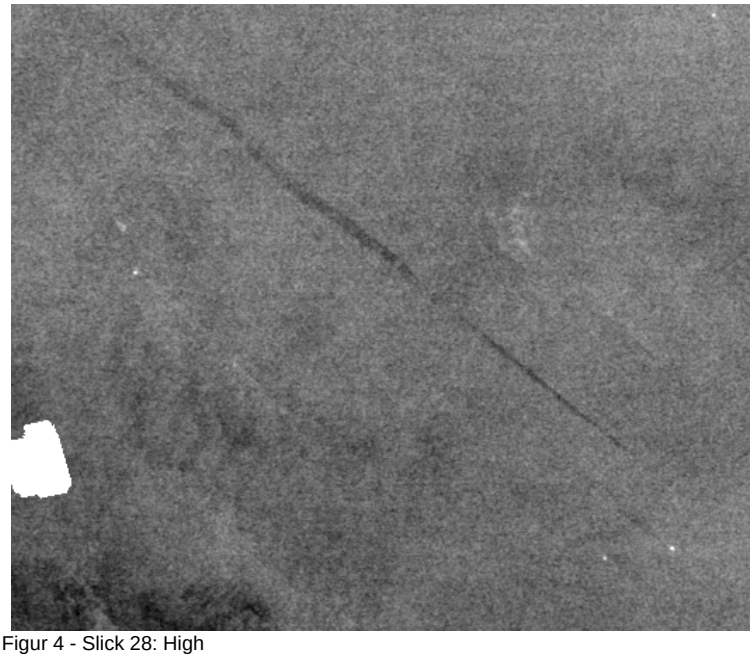
Slick no:	Source	Contrast	Edges	Type	Shape	Surroundings	Model wind speed
10	Ship	Medium	Sharp&diffuse	Tail	Irregular	Homogeneous	3.9 (4.7)
17	Ship	Medium	Sharp&diffuse	Tail	Smooth	Homogeneous	7.2 (9.1)
28		Medium	Sharp&diffuse	Tail	Smooth	Homogeneous	6.8



Figur 2 - Slick 10: High



Figur 3 - Slick 17: High



Figur 4 - Slick 28: High

Figures 2-4 shows the three high confidence slicks. From inspecting the images, we see that they vary. Our interpretation of these figures is that figures 3-4 show long, highly linear slicks. In my opinion, the surroundings varies, they are much more homogeneous in Figure 3 than in Figure 4. Looking at Figure2, it is more difficult to be convinced of the reasons why this has been assigned confidence level High, as the slick marked is closely linked to a low-wind area. What we observe from Table 1 is that almost all of the parameters are equal for the three spills, even if they look somewhat different on these figures.

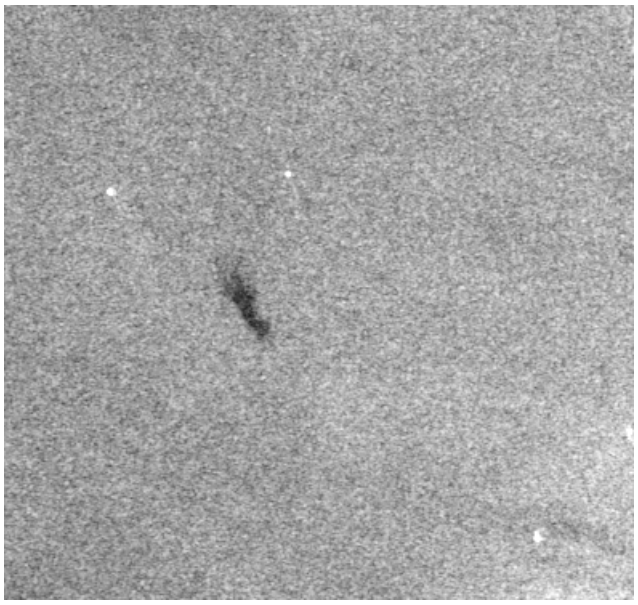
2.1.2 Medium confidence slicks

Table 2 shows the selected feature values for the spills assigned medium confidence at KSAT. The same tendency as for high confidence spill is observed, namely that a majority of the spills are assigned equal feature values.

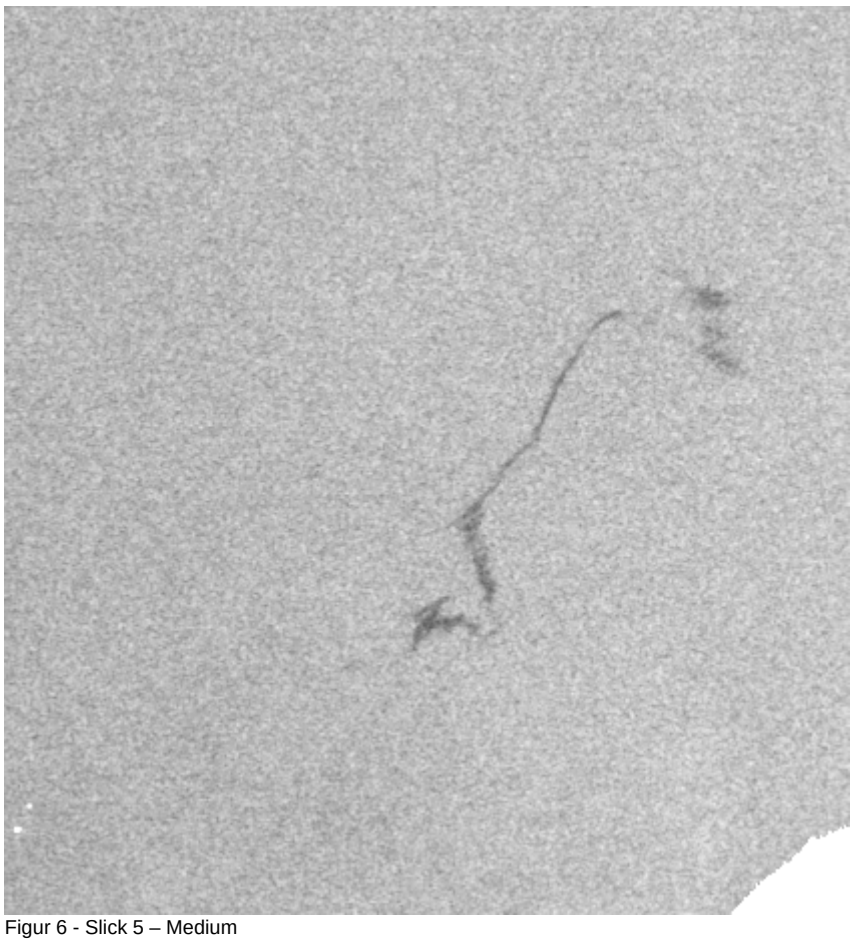
Tabell 2 Medium confidence slicks

Slick no:	Source	Contrast	Edges	Type	Shape	Surroundings	Model wind speed
1		Medium	Sharp&diffuse	Patch	Smooth	Homogeneous	10.1
5		Medium	Diffuse	Tail	Smooth	Homogeneous	12.0 (8.0)
6		Medium	Sharp&diffuse	Linear	Smooth	Homogeneous	11.6 (7.0)
8		Medium	Sharp&diffuse	Linear	Smooth	Homogeneous	2.7 (1.9)
11		Medium	Sharp&diffuse	Linear	Smooth	Homogeneous	4.2 (1.5)
12		Medium	Sharp&diffuse	Linear	Irregular	Homogeneous	3.3 (6.3)
13		Medium	Sharp&diffuse	Linear	Smooth	Homogeneous	4.0 (1.3)
18		Strong	Sharp&diffuse	Linear	Smooth	Homogeneous	4.0 (4.4)
23		Medium	Sharp&diffuse	Linear	Smooth	Homogeneous	6.2
24		Medium	Diffuse	Linear	Smooth	Homogeneous	7.6
32		Medium	Sharp&diffuse	Tail	Smooth	Homogeneous	3.5

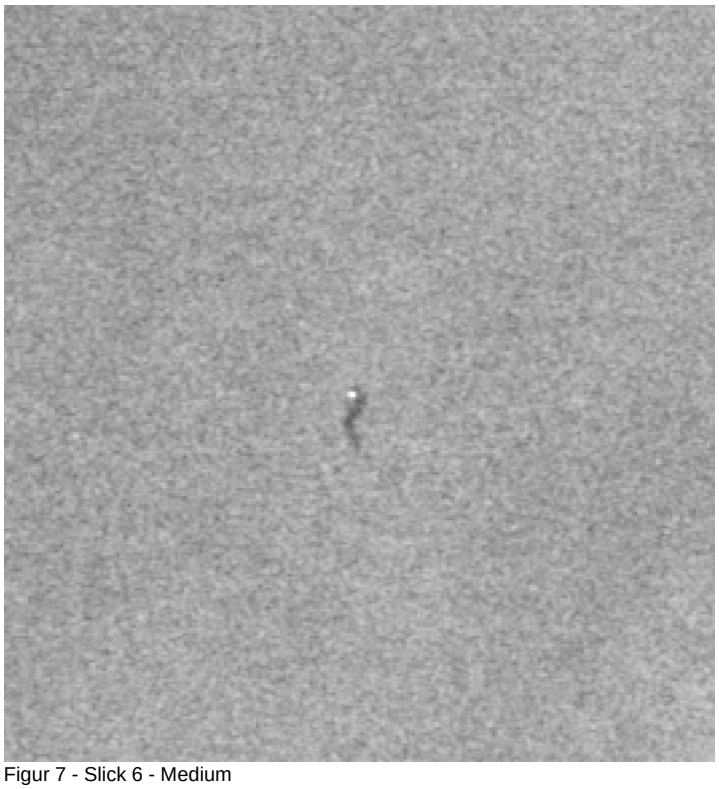
Figures 5-14 show small images for each of the medium confidence spills. All these, including spill number 11, 13 and 18 are said to have homogeneous surroundings. It is also interesting to note that slick 12 has been assigned irregular shape, and all the others smooth.



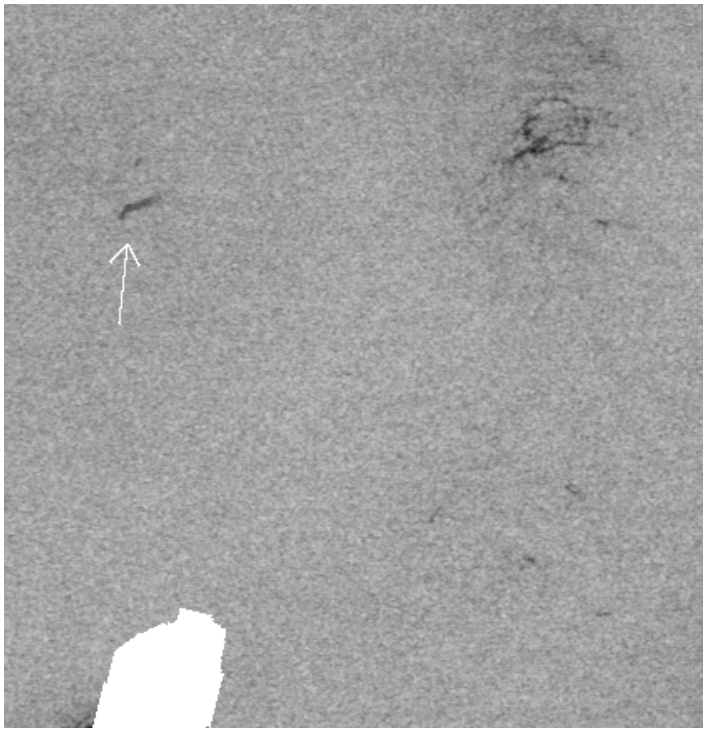
Figur 5- Slick1 Medium



Figur 6 - Slick 5 – Medium

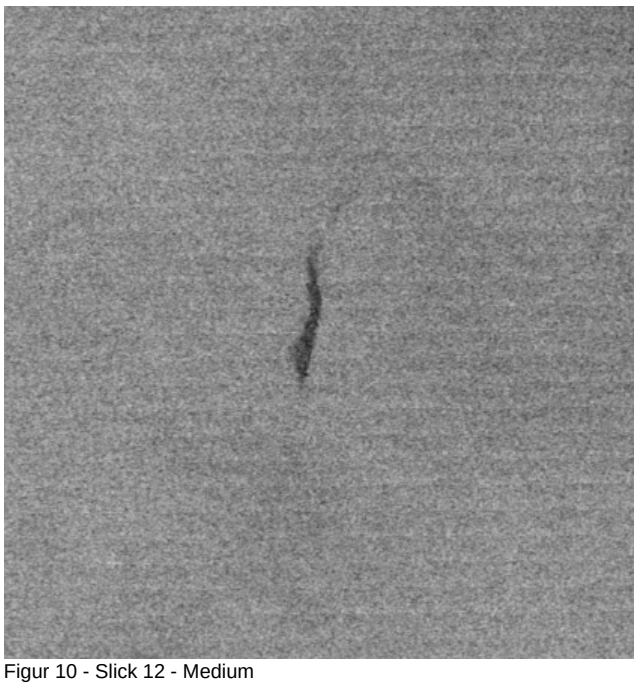


Figur 7 - Slick 6 - Medium

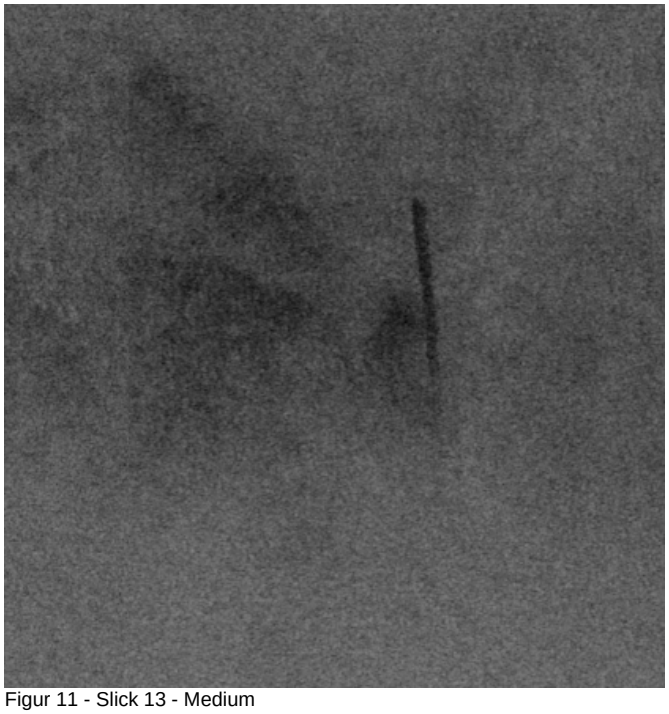


Figur 8 - Slick 8 - Medium

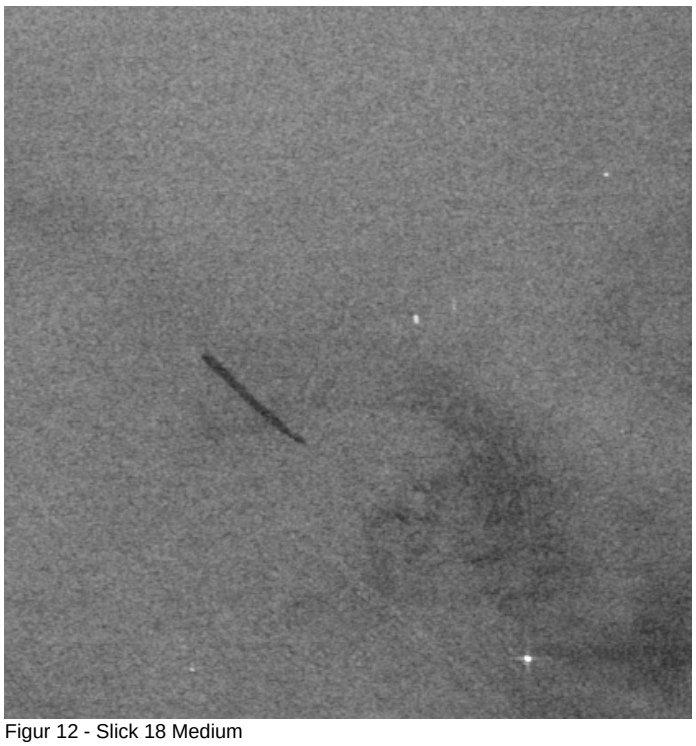
Figur 9 - Slick 11 - Medium



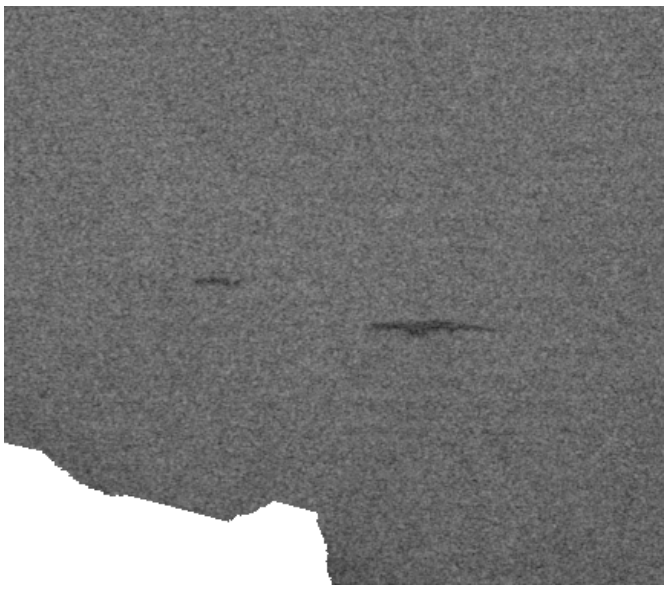
Figur 10 - Slick 12 - Medium



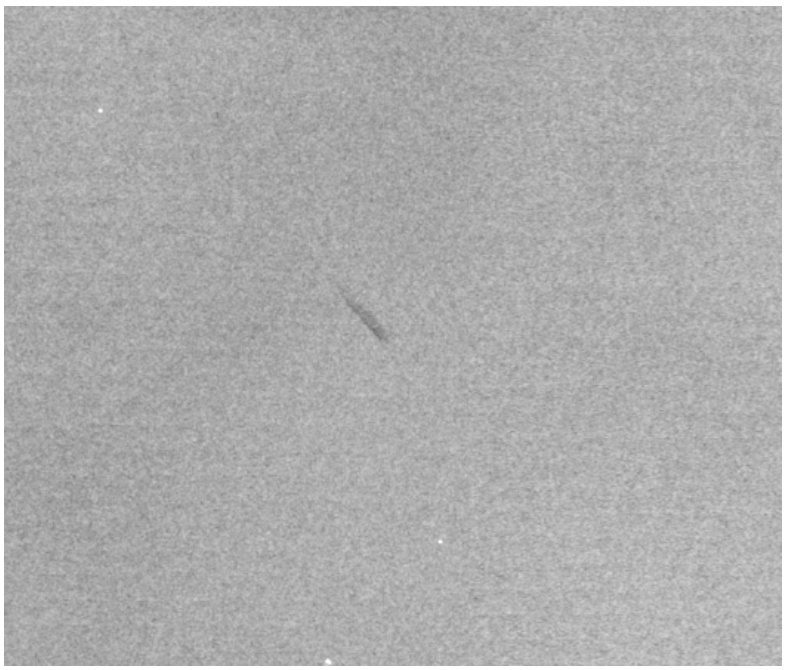
Figur 11 - Slick 13 - Medium



Figur 12 - Slick 18 Medium



Figur 13 - Slicks 23, 24 - Medium



Figur 14 - Slick 32 – Medium

2.1.3 Low confidence slicks

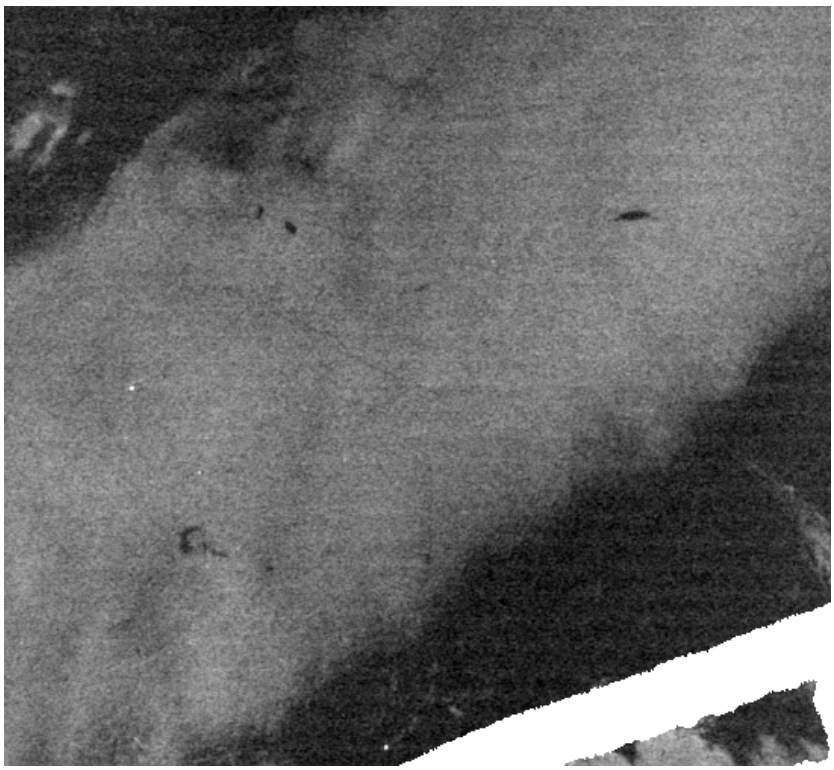
Selected parameters for spills assigned low confidence is given in Table 3.

Tabell 3 - Low confidence slicks

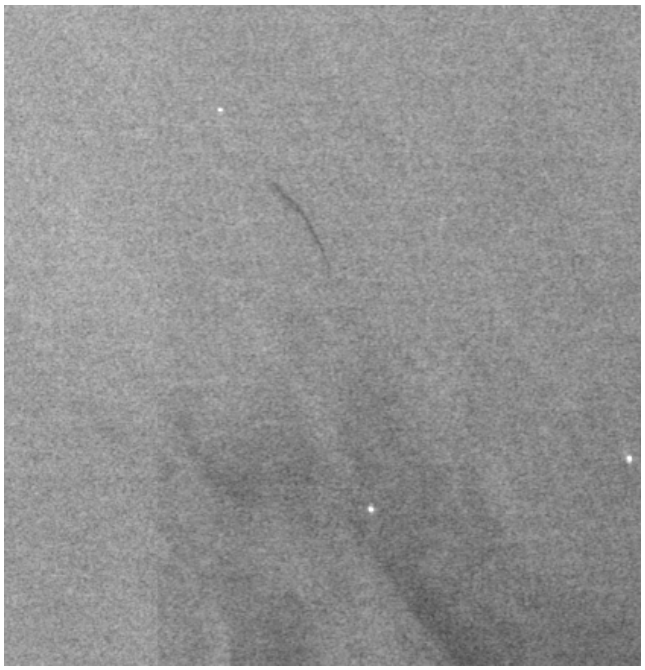
Slick no:	Source	Contrast	Edges	Type	Shape	Surroundings	Model wind speed
2		Medium	Sharp&diffuse	Droplet	Smooth	Homogeneous	7.5
3		Medium	Sharp&diffuse	Droplet	Smooth	Homogeneous	7.9
4		Medium	Sharp&diffuse	Droplet	Smooth	Homogeneous	6.8
7		Medium	Diffuse	Linear	Smooth	Inhomogeneous	3.1 (6.7)
9		Medium	Sharp&diffuse	Angular	Smooth	Homogeneous	4.6 (5.5)
14		Medium	Sharp&diffuse	Linear	Irregular	Homogeneous	5.7 (8.0)
15		Medium	Diffuse	Patch	Smooth	Homogeneous	5.8 (7.0)
16		Medium	Sharp&diffuse	Linear	Smooth	Inhomogeneous	3.6 (3.3)
19		Weak	Diffuse	Droplet	Smooth	Homogeneous	3.8 (5.9)
20		Weak	Diffuse	Droplet	Smooth	Homogeneous	4.5 (5.4)
21		Medium	Diffuse	Tail	Smooth	Inhomogeneous	4.1 (3.2)
22		Strong	Sharp&diffuse	Patch	Smooth	Inhomogeneous	3.4
25		Medium	Sharp&diffuse	Droplet	Smooth	Homogeneous	6.9
26		Medium	Sharp&diffuse	Angular	Smooth	Homogeneous	3.7 (6.8)
27		Medium	Sharp&diffuse	Tail	Irregular	Homogeneous	3.6 (4.9)
29		Weak	Diffuse	Tail	Smooth	Homogeneous	6.4 (8.0)
30		Medium	Diffuse	Patch	Smooth	Homogeneous	6.8 (7.9)
31		Weak	Diffuse	Patch	Smooth	Homogeneous	7.2 (7.9)
33		Medium	Sharp&diffuse	Patch	Smooth	Homogeneous	3.4 (6.6)
34		Medium	Sharp&diffuse	Patch	Smooth	Homogeneous	4.5 (7.4)
35		Medium	Sharp&diffuse	Patch	Smooth	Homogeneous	4.5 (7.1)
36		Weak	Diffuse	Linear	Smooth	Homogeneous	3.8 (5.3)
37		Weak	Diffuse	Linear	Smooth	Homogeneous	3.8 (5.1)

The same tendency as for the medium and high confidence slicks is noted, all parameters are relatively similar, even though the visual inspection of the following figures show larger varations.

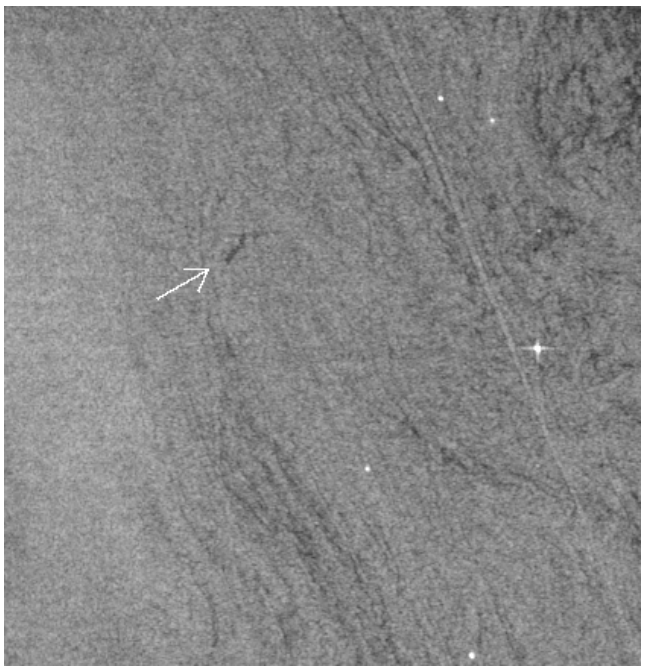
The images corresponding to these slicks are found in Figures 15-30.



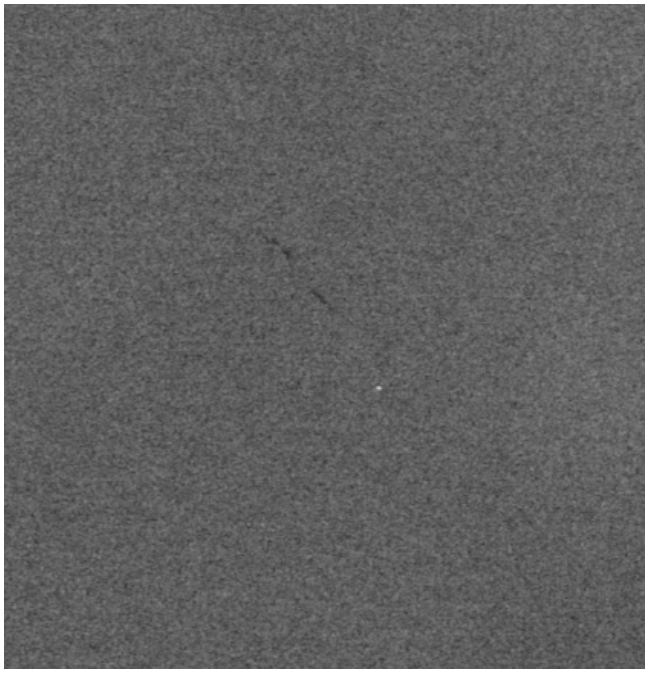
Figur 15 - Slicks 2,3,4: Low



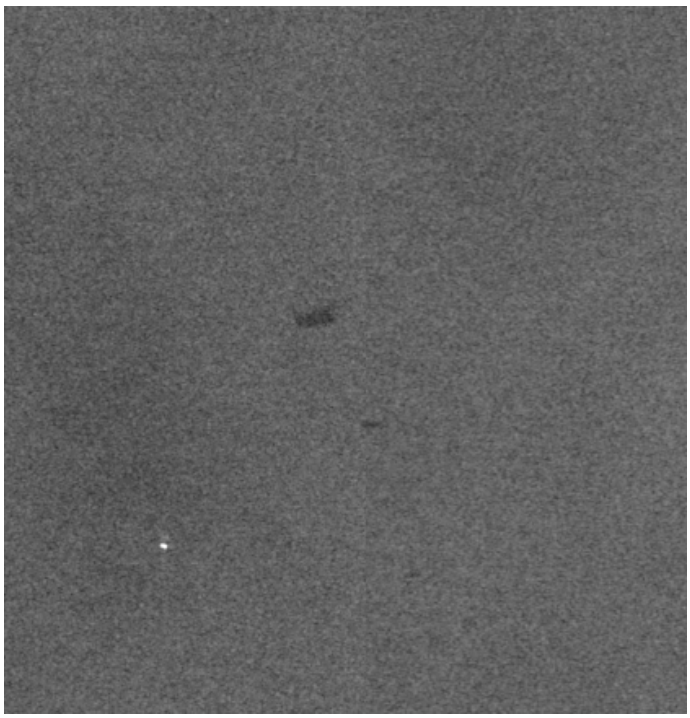
Figur 16 - Slick 7: Low



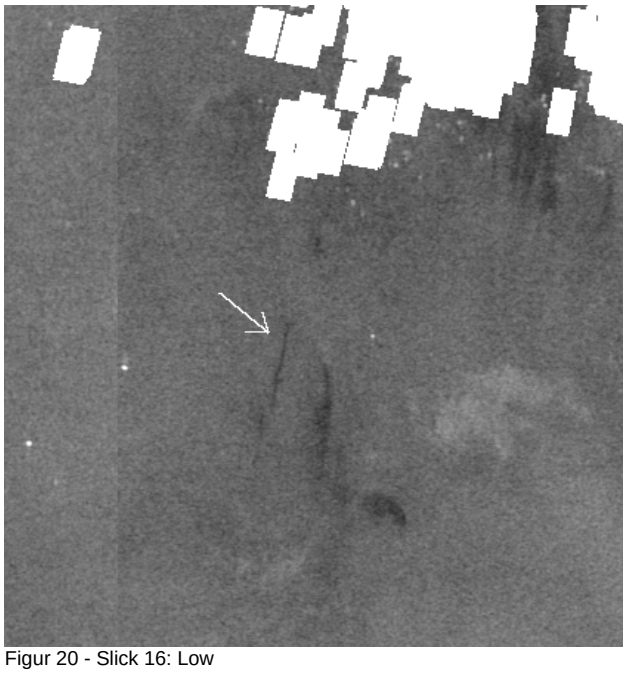
Figur 17 - Slick 9: Low



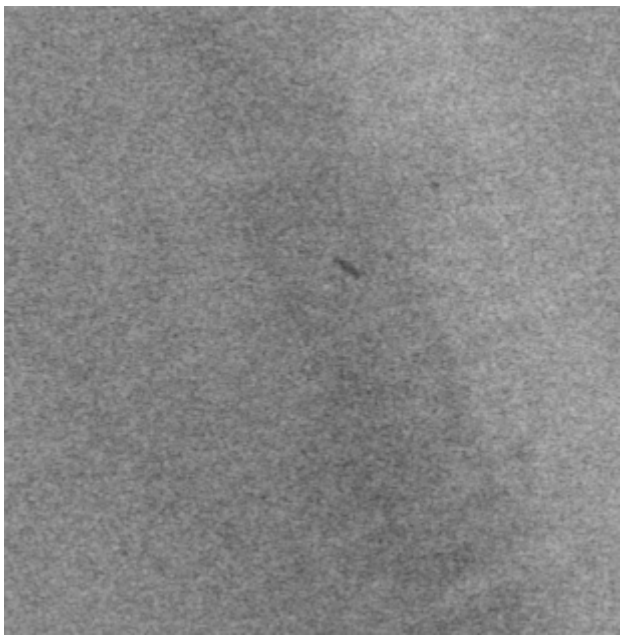
Figur 18- Slick 14: Low



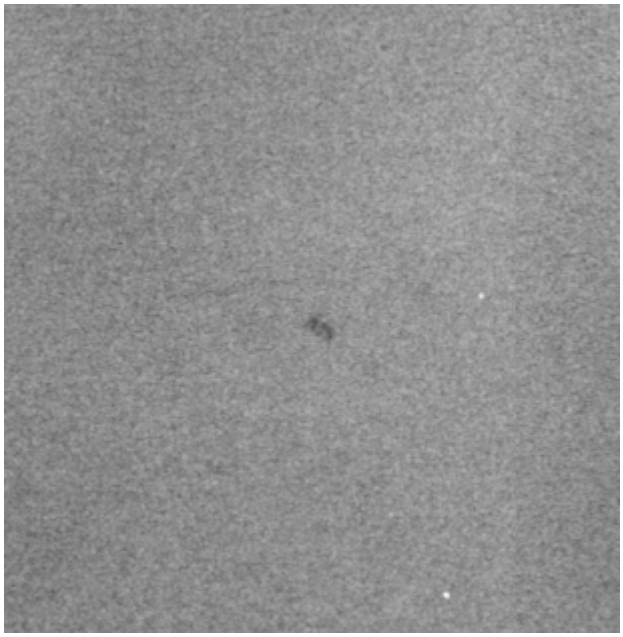
Figur 19 - Slick 15 : Low



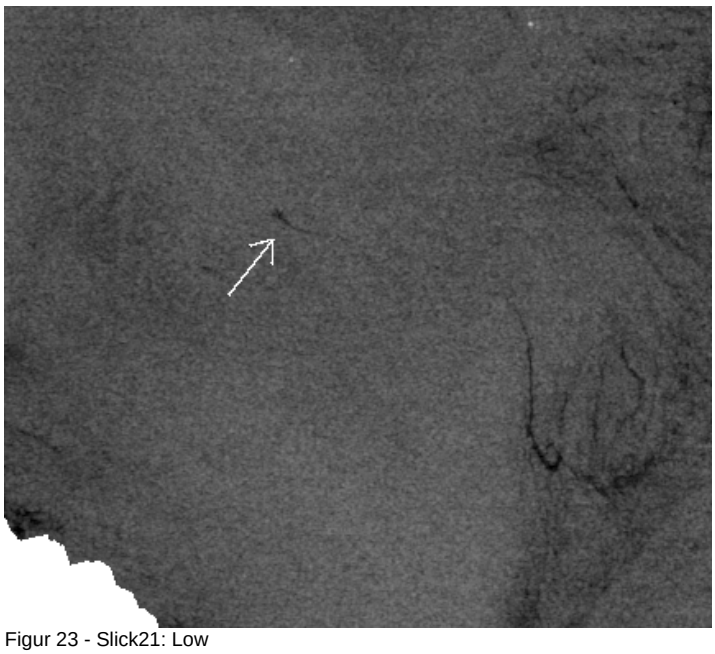
Figur 20 - Slick 16: Low



Figur 21 - Slick 19: Low

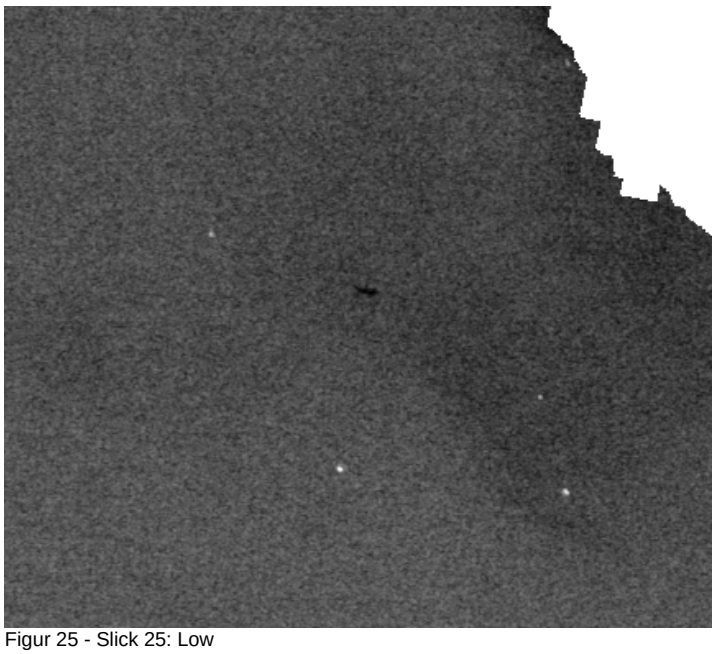


Figur 22 - Slick 20: Low

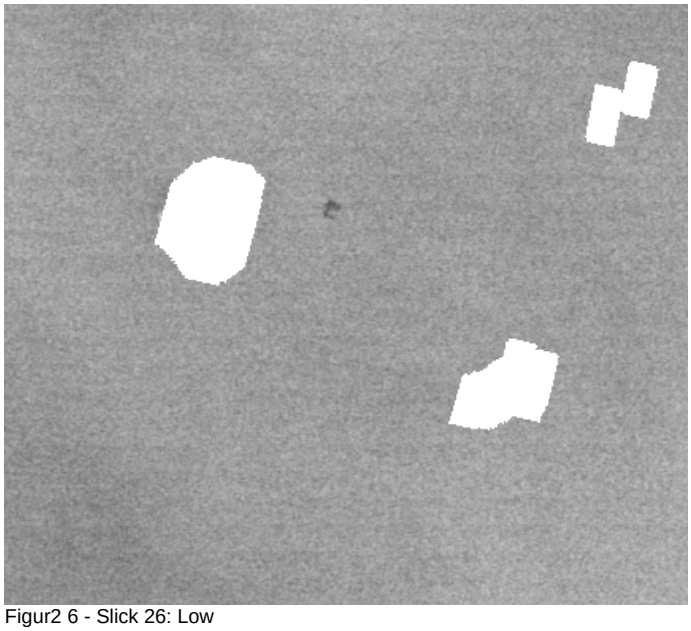


Figur 23 - Slick21: Low

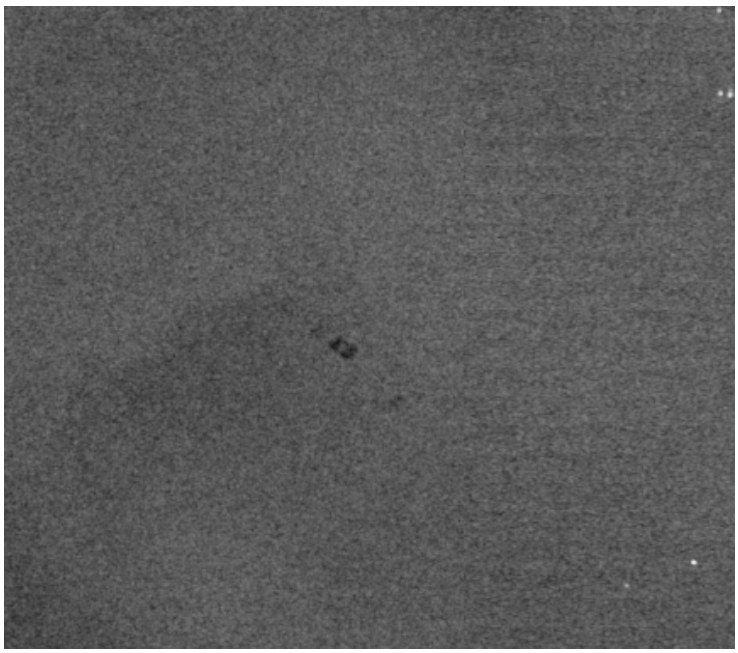
Figur 24 - Slick 22: Low



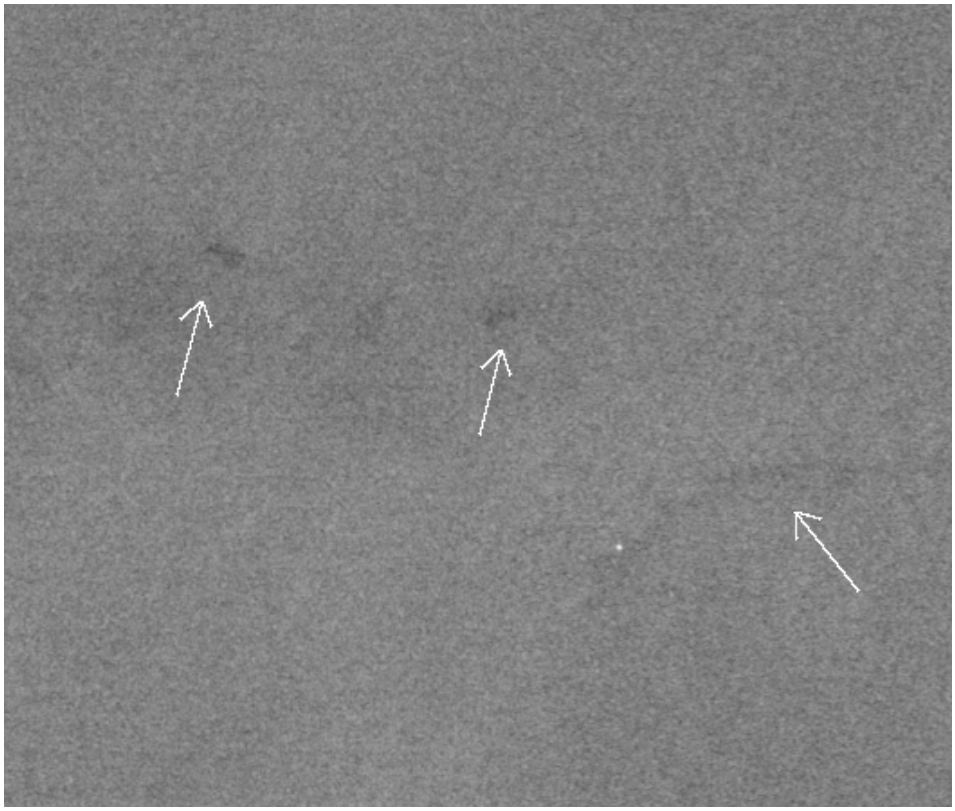
Figur 25 - Slick 25: Low



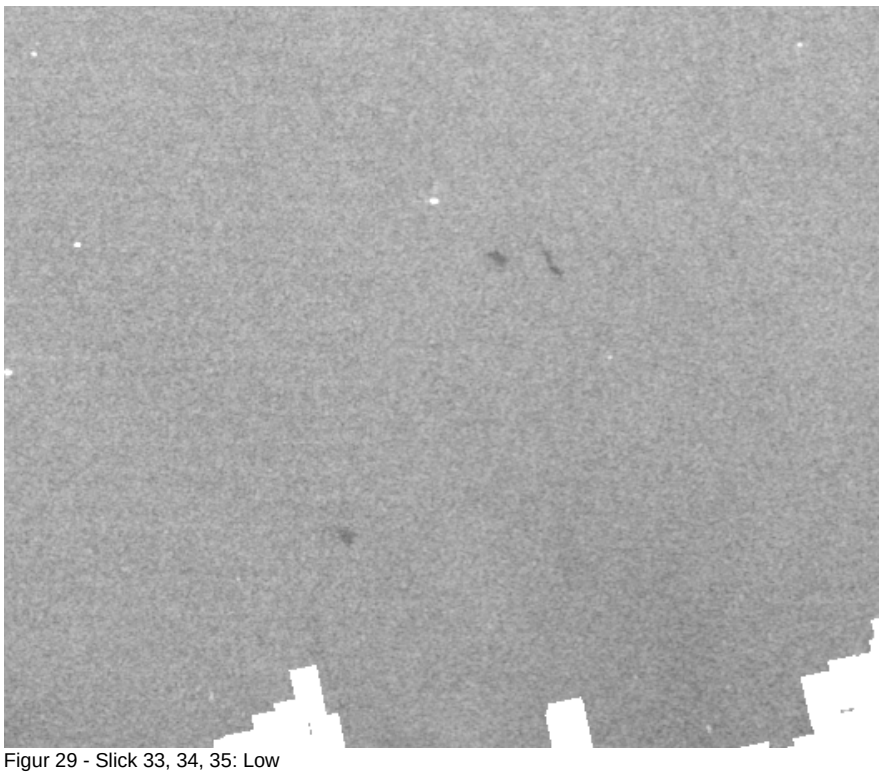
Figur2 6 - Slick 26: Low



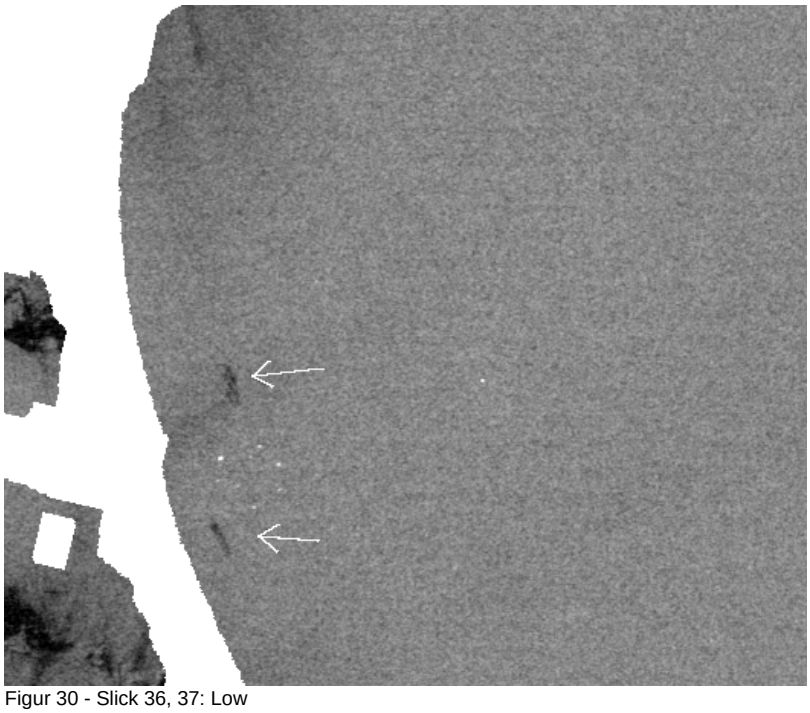
Figur 27 - Slick 27: Low



Figur 28 - Slick 29, 30, 31: Low



Figur 29 - Slick 33, 34, 35: Low



Figur 30 - Slick 36, 37: Low

2.2 Estimating confidence from the KSAT features

The first goal of the project was to investigate if the feature values filled in at KSAT could be used as a fundament for an automatic procedure to compute confidence levels automatically. This would yield a more objective procedure (if the features values are so simple that filling them in is objective).

To investigate this, further data analysis was done. Based on pure inspection of the images and the tables with the feature values as discussed in the previous section, it seems that

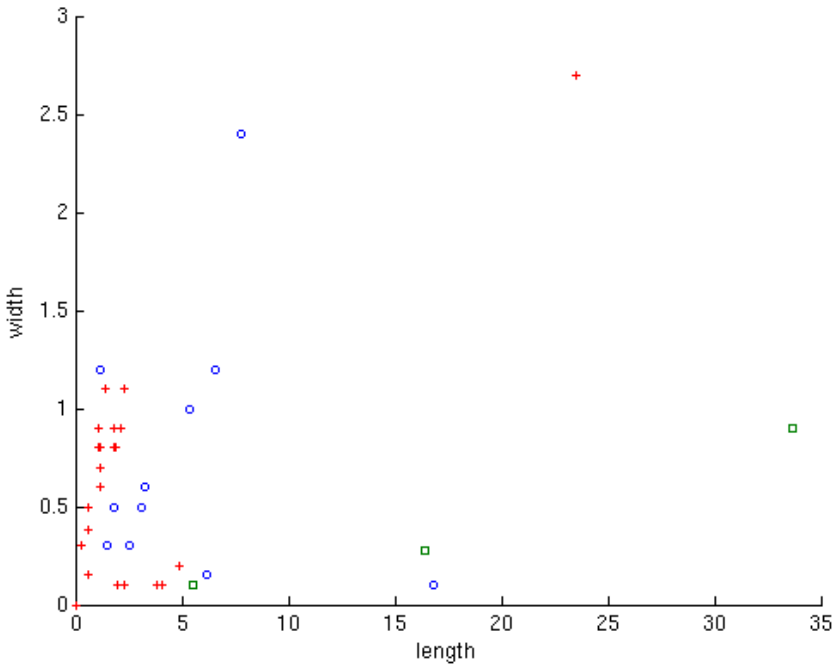
- The feature values are not always objectively filled in as we after additional visual inspection cannot always understand why a particular spill has been assigned a certain confidence level.

- The confidence assigned by the KSAT operator sometimes seems a bit difficult to explain by inspecting the images again.

- Many of the oil spills that look visually different in the images presented here have equal parameter values.

Scatter plots of the feature values with respect to the KSAT confidence levels assigned will give a good indication of if the features values can be used to divide the data into the given confidence levels. A full set of pairwaise scatter plots (two features on each plot) was produced. Most of them showed low discriminatory power. Two of the best plots with respect to discrimination of confidence levels are shown in Figures 31-32. In all figures, high confidence slicks are given as squares, medium confidence as circles, and low confidence using the + symbol.

Figur 31 - Scatter plot of Slick Width vs. Slick Length



Figur 32 - Scatter plot of Slick type vs. Wind

None of the other scatter plots indicated that the KSAT features were very useful for confidence discrimination. Building an automatic procedure for confidence estimation based on them was thus not further investigated. Instead, we used the remaining resources to check if features computed in the existing NR automatic algorithm for oil spill detection could be used.

2.3 Features from NRs automatic algorithm

The automatic algorithm for oil spill detection consists of dark spot segmentation, followed by feature extraction for each dark spot, and then classification of each spot as either oil spill or look-alike. Because the original SAR images were available to us, we could run the software on the images and save the result after the feature extraction step. The result after feature extraction is for each dark spot in the image, a list of feature values computed. The following features were believed to be useful for confidence estimation:

Slick surroundings:

- Homogeneity of surroundings
- Number of neighboring slicks

Slick contrast:

- Slick local contrast
- Smoothness contrast

Slick shape:

- Shape Complexity
- Slick planar moment
- Slick width
- Slick area

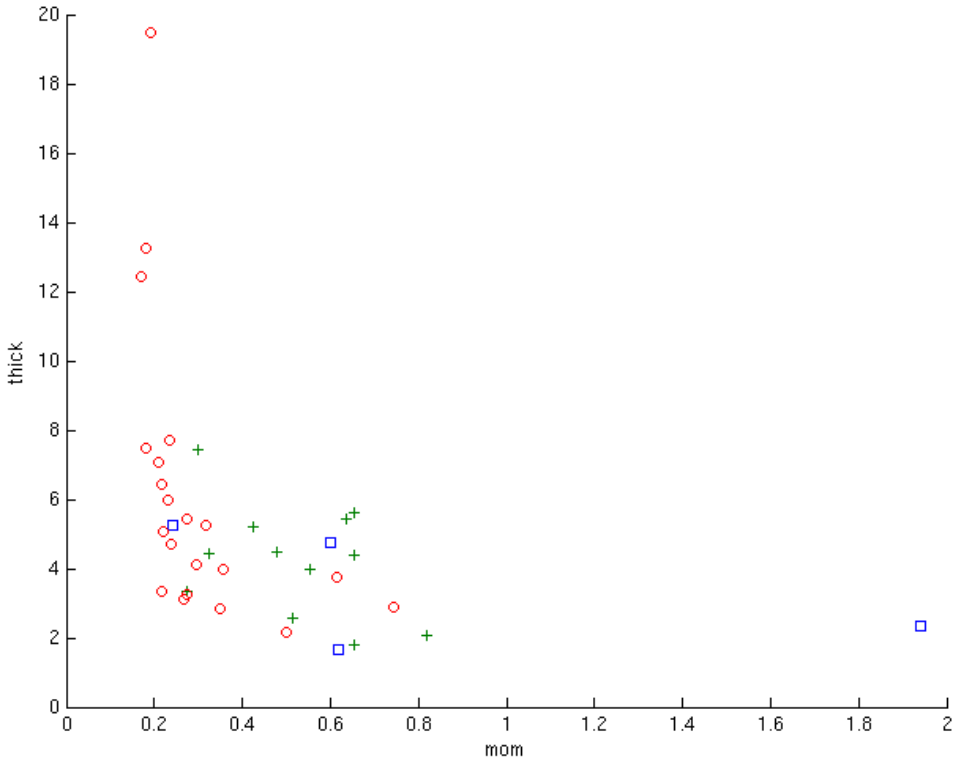
Source:

- Distance to possible source

Wind:

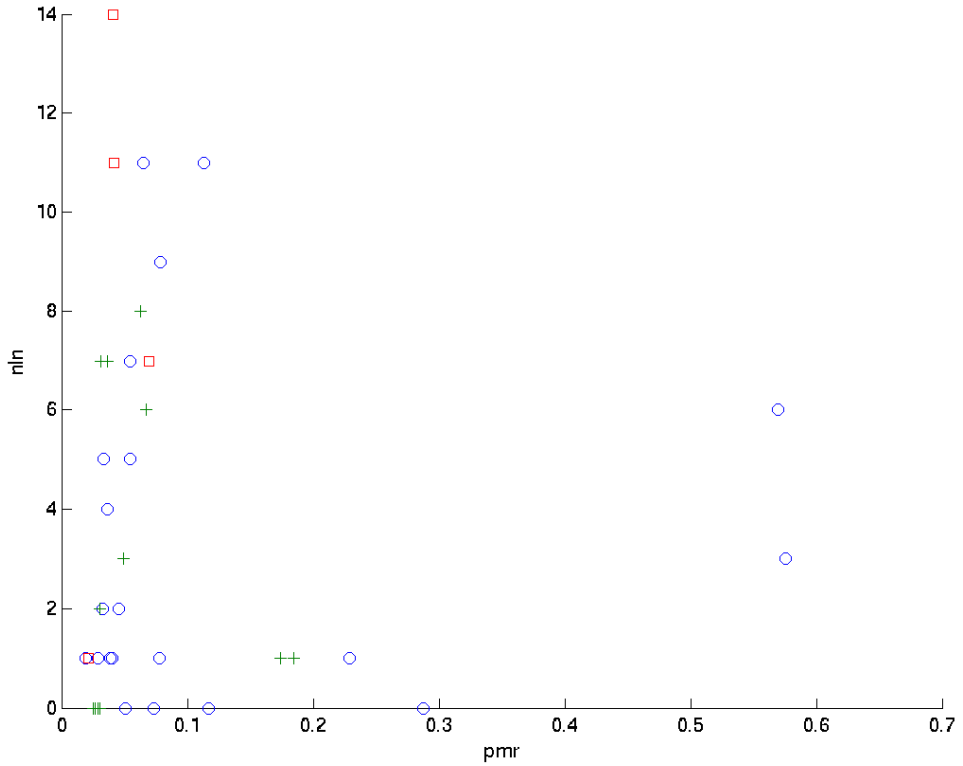
- Model wind (exported from the KSAT sheet)

Scatter plots were also produced for these features with respect to the confidence levels assigned by KSAT. Examples are given in Figures 33-34.



Figur 33 - Scatter plot of automatic features Slick moment vs. Slick width

Figur 34 - Scatter plot of automatic features Nof. Neighboring slicks vs. Homogeneity of surroundings



A new automatic procedure for confidence estimation will of course use more than two features at a time (this is just convenient for visualization). The automatic features span out a larger set of values. However, we still believe that the subjectivity involved in the confidence assignment could be a problem. To elaborate this, if we design an algorithm to produce the same confidence assignements as the KSAT confidence labels, these should be unquestionable. A repeated analysis by another operator should result in the same confidence assignment. To avoid restricting the automatic confidence procedure to this, we start off by developing an unsupervised algorithm for confidence assignment, and later compare the results to the KSAT assignments.

A number of factors can be claimed to be relevant for confidence assignment. We believe that confidence should increase with the observation of the following factors:

- Source. If a source is close, or even directly linked to the slick, the confidence should increase.

- Wind speed. In low wind, look-alikes are frequent. As the wind level increases, the likelihood of observing look-alikes decreases and the likelihood that a given dark spot is oil will increase.

- Local contrast. The relation between contrast and the likelihood that a slick with given contrast is oil will depend on wind level. In low wind, high contrasts are often observed both for look-alikes and for oil spills. As the wind level increases, the contrast will decrease. At high winds, even slicks with low contrast are likely to be oil.

- Homogeneity of surroundings: A single slick on a homogeneous background has higher likelihood of being oil than the same slick in the middle of a low-wind area.

- Slick shape. Slicks with regular shape, e.g. wide/circular with low edge complexity are likely to be fresh outlets from a stationary source. Slicks with linear or piecewise linear are likely to come from a moving ship.

With this in mind, we developed a set of score values. One score factor was developed for each of the factors wind, source, contrast, homogeneity, and linearity. The scores will then be combined to yield to total score that should be used to assign a confidence level automatically. All the score factors will be normalized to [0,1} to avoid scale-sensitive measures in the combination step.

2.3.1 Score factor for wind

The basic assumption if that higher speed speed gives higher probability that a dark spot is oil. The following score function is used:

```
IF WIND<WindMean
        WindScore = 0
ELSE
        WindScore = (WIND-WindMean)/(WindMax-WindMean)
```

2.3.2 Score factor for source

A very short distance between the slick and a source (detected bright spot) will increase the confidence. The present version uses the distance from the source to the closest point in the dark spot (because this was already implemented in the automatic oil spill detection algorithm), but later versions can include the distance from the ends of the spot (if linear). The following score is used:

```
IF DISTANCE>40
        SourceScore = 0
ELSE
        SourceScore = (50-DISTANCE)/50;
```

2.3.3 Score factor for contrast

High contrast will in general increase the probability that a given dark spot is oil, but the expected value for contrast will depend on the wind speed. The following score is used:

```
IF WIND<7 m/s
        IF Contrast<1.5
                ContrastScore = 0
        ELSE
                ContrastScore = (Contrast-1.5)/(ContrastMax-1.5)
```

ELSE (WIND>7)
    IF Contrast<1.0
        ContrastScore = 0
    ELSE
        ContrastScore = (Contrast-1.0)/(ContrastMax – 1.0)

2.3.4 Score factor for surroundings

If the surroundings are homogeneous, the confidence should increase. As a feature for homogeneity of the surroundings, the number of neighboring dark spots is used (called N) Another homogeneity feature is the power-to-mean ratio (PMR)of the surroundings.. The following score factor is used:

IF N>4
    NofNeighbScore = 0
ELSE
    NofNeighbScore = (4-N)/4

IF PMR>0.06
    PMRScore = 0
ELSE
    PMRScore = (0.06-PMR)/(0.06-PMRMin)

2.3.5 Score factor for linearity

The more linear, or piecewise linear, a slick is, the more likely it is to be man made. There is still a possibility that a linear slick might be legal discharges from a moving ship, but as the SAR sensor alone cannot verify this 100%, suspect slicks should be assigned high confidence. As a measure of linearity we use Slick first planar moment , which will show higher values the more linear and elongated the slick is. The score factor used is:

IF (MOMENT<MomentMean)
    MomentScore = 0
ELSE
    MomentScore = MOMENT/(MomentMax-MomentMean)

2.3.6 Total score

The total score should increase as the presence of several of the other factors are found. The more of the different score types set, the more likely is the slick to be oil. Thus, a simple weighted sum of the individual scores can be used:

$$\text{Totalscore} = w_1\text{ScoreWind} + w_2\text{ScoreSource} + w_3\text{ContrastScore} + w_4\text{NofNeigbhScore} + w_5\text{PMRScore} + w_6\text{MomentScore}$$

The weights are not optimized in this part of the testing as the data set used is somewhat limited, they can be optimized at a later point.
Based on the total score, the following confidence assignment gave reasonably good results based on experiments:

IF TotalScore>2:
    Confidence = High
ELSE IF TotalScore > 1
    Confidence = Medium
ELSE IF TotalScore is close to 1
    Confidence = Medium/Low
ELSE
    Confidence = Low

2.3.7 Evaluating confidence assigned by the score factors

The score scheme was tested on all images and all oil spills in the available data set. This score scheme resulted in confidence assignments as shown in Tables 4-67 and Figures 35-39. In the tables, the main score factors are the largest factors that contribute to the total score. We see that most of the slicks assigned Medium confidence by KSAT are here assigned either High or Medium confidence. The 3 slicks assigned High confidence by KSAT are here assigned either High, Medium or Low, respectively. These 3 slicks were discussed in more depth earlier.

Many of the slick assigned High confidence by the automatic procedure as linear slicks on a homogeneous background. Three large contributions to the score were always found for these (as seen in Table 3).

The Medium confidence slicks, found in Figure 34, are also high regular in shape. The homogeneity varies, but common for all these slicks is that they has at least two large score factors.

The Medium/Low category was assigned a category of its own because this set of slicks had very similar scores, close to 1.0. They have either one strong score factor or two medium high score factors.

The slicks assigned to the Low confidence category typically have only one medium high score factor present.

Slick number	Score	KSAT confidence	Main score factors
17	3.35	High	Homog, Source, Linear
6	2.64	Medium	Wind, Homog, Source
5	2.54	Medium	Wind, Linear, Homog
1	2.54	Medium	Wind, Homog, Contrast
12	2.37	Medium	Contrast, Homog, Linear

Tabell 4 High confidence slicks after automatic confidence assignment

Slick number	Score	KSAT confidence	Main score factors
2	1.55	Low	Contrast, Homog
7	1.64	Low	Wind, Source, Homog., Linear
13	1.18	Medium	Contrast, Linear
18	1.48	Medium	Homog, Linear
23	1.45	Medium	Homog, Linear
24	1.49	Medium	Homog, Linear
28	1.30	High	Contrast, Linear, Homog
32	1.28	Medium	Contrast, Homog., Wind
34	1.19	Low	Homog, Source

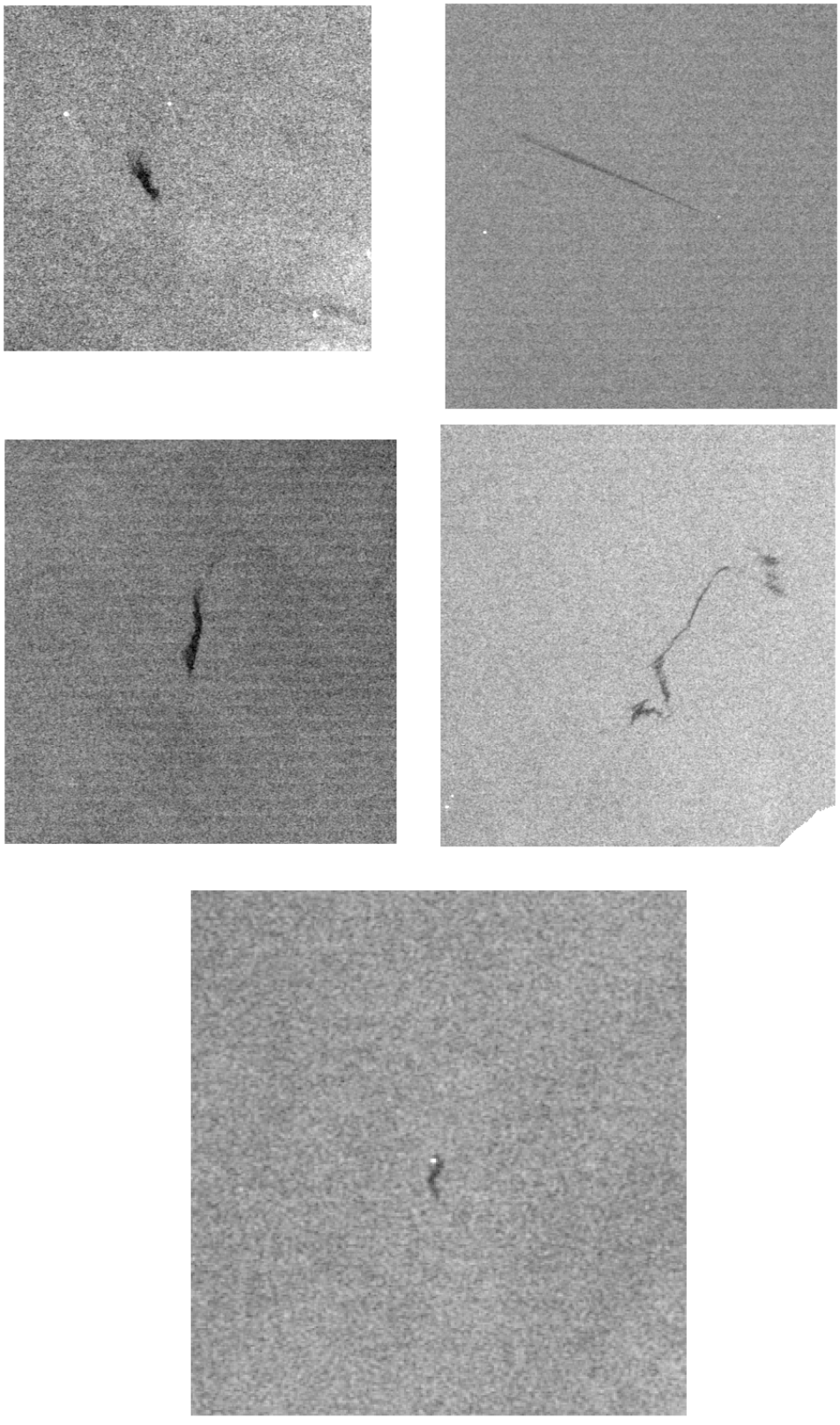
Tabell 5 Medium confidence slick assigned by the automatic procedure

Slick number	Score	KSAT confidence	Main score factors
8	1.006	Medium	Contrast, Homog
14	1.06	Low	Source, Homog,
16	1.05	Low	Source, Linear
22	1.00	Low	Contrast
26	1.10	Low	Contrast, Homog,
27	1.02	Low	Contrast, Homog

Tabell 6 Medium/Low confidence slicks assigned by the automatic procedure

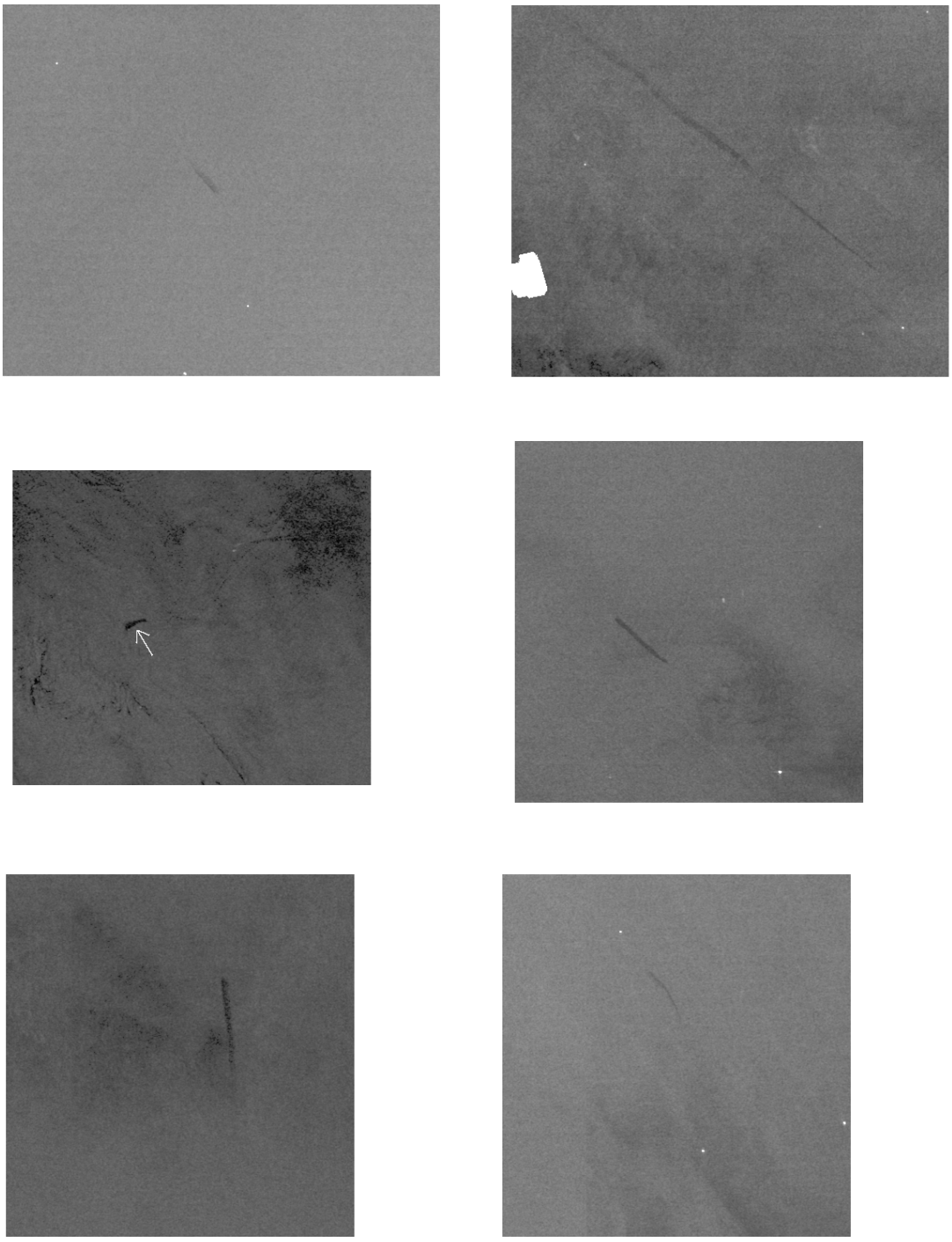
Tabell 7 - Low confidence slicks assigned by the automatic procedure

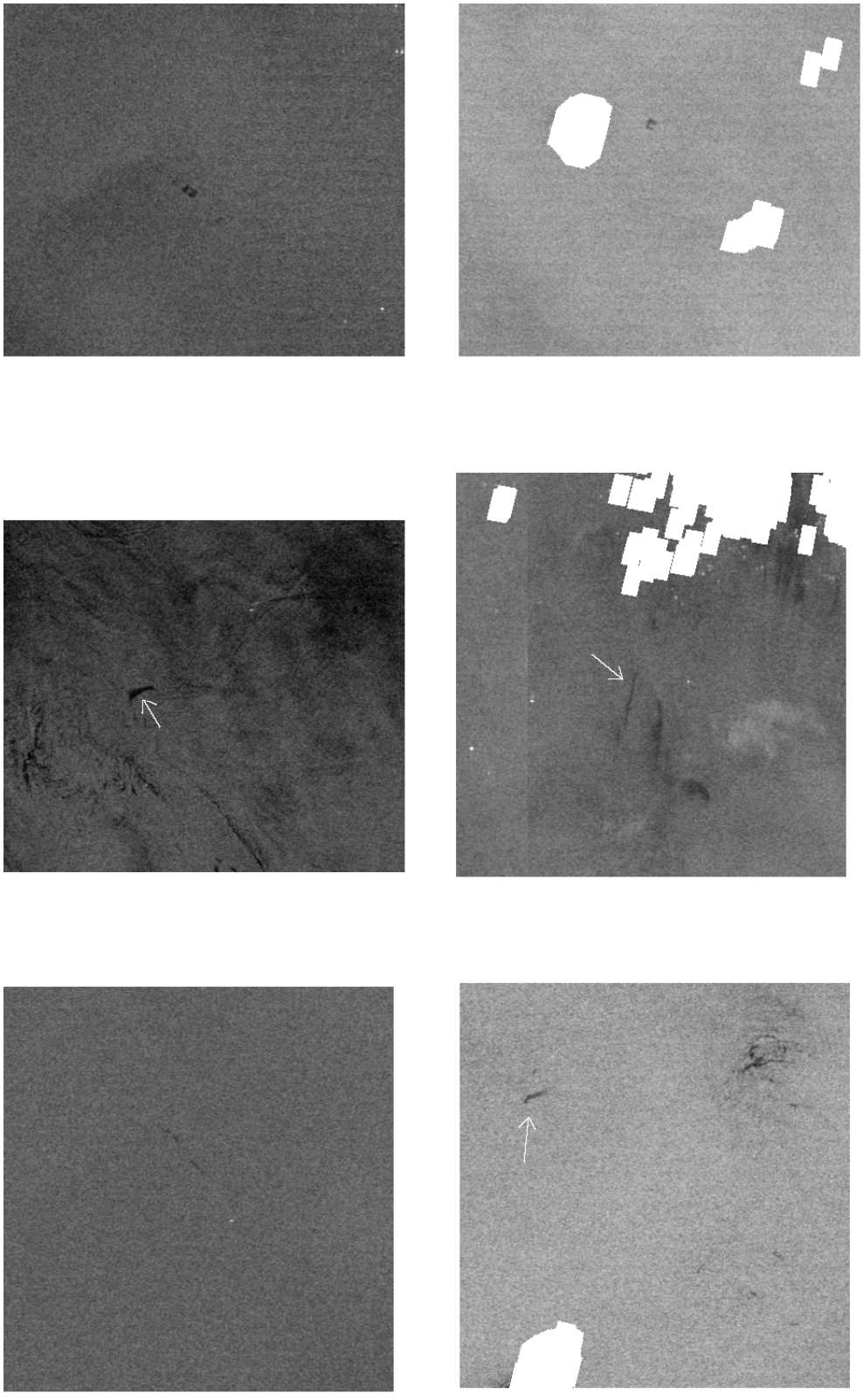
Slick number	Score	KSAT confidence	Main score factors
3	0.95	Low	Contrast
4	0.55	Low	Contrast
9	0.15	Low	
10	0.67	High	Source
11	0.69	Medium	Linear
15	0.95	Low	Homog.
19	0.54	Low	Homog.
20	0.81	Low	Contrast, Homog.
21	0.38	Low	
25	0.85	Low	Wind, Contrast, Homog.
29	0.65	Low	Homog.
33	0.63	Low	Homog.
34	0.85	Low	Homog
36	0.52	Low	Homog, Linear
37	0.82	Low	Homog, Linear



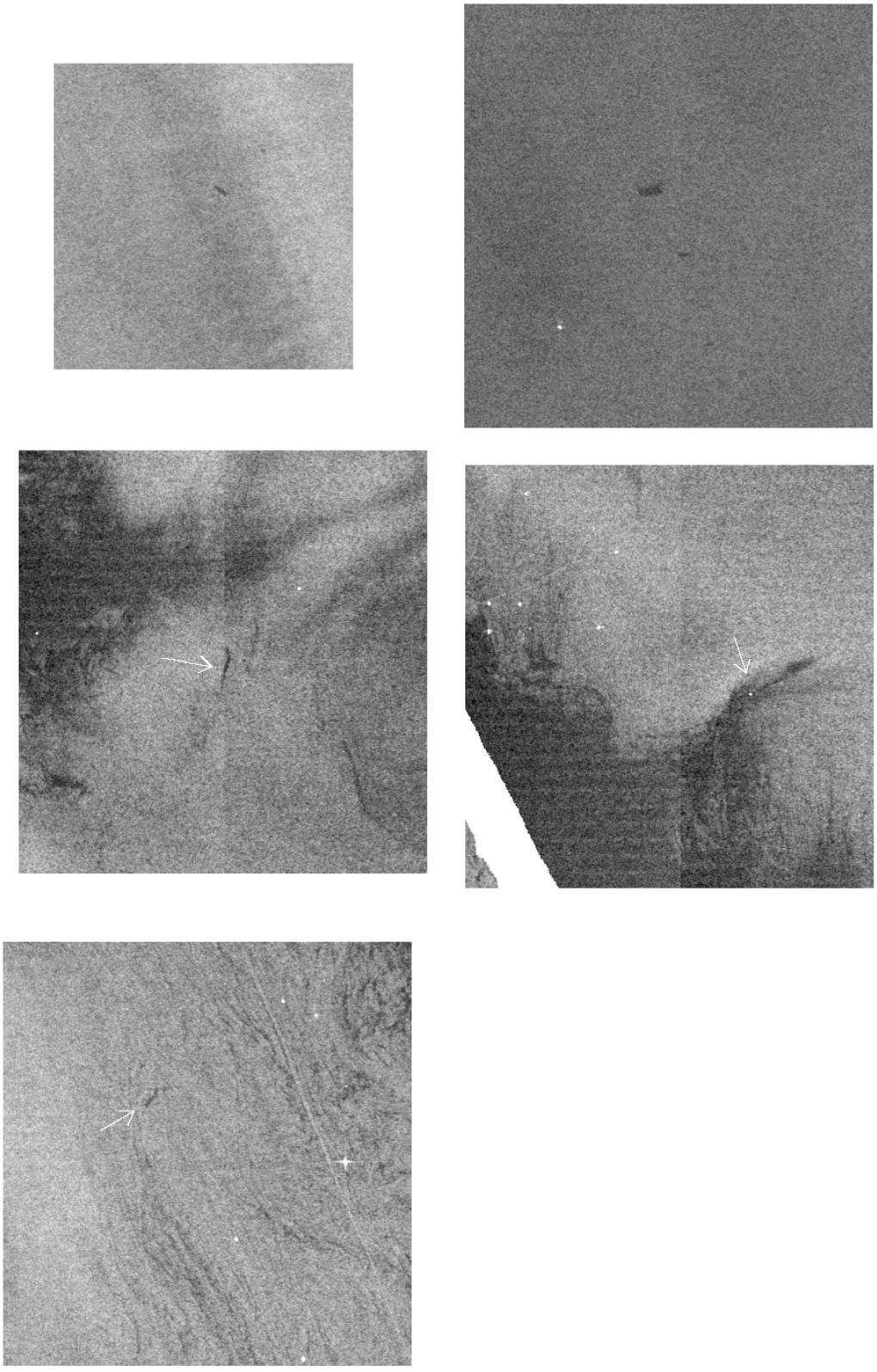
Figur 35- Oils spills assigned high confidence by the automatic procedure

Figur 36 - Medium confidence slicks from the automatic procedure

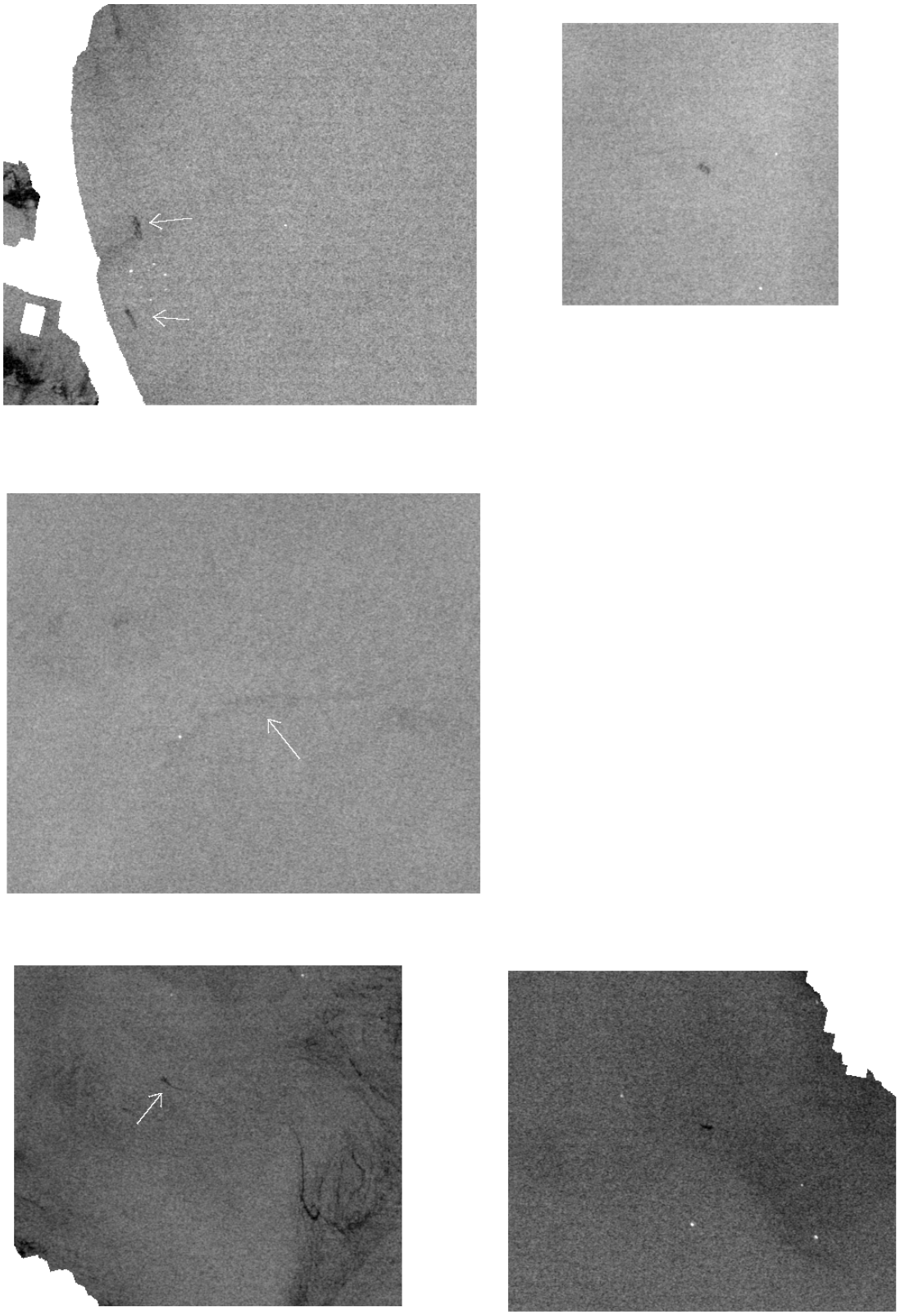




Figur 37 - Medium/Low confidence slicks from the automatic procedure



Figur 38 -Low confidence slicks from the automatic procedure



Figur 39- More Low confidence slicks from the automatic procedure

3 Discussion of the results and guidelines for further work

Automatic confidence assignment based on both input from the form filled in by the KSAT operator, and features derived from the automatic oil spill algorithm have been tested for use in a new procedure for automatic confidence assignment. The factors filled in at KSAT as part of the oil spill detection report were found to have little discriminatory power. An visual inspection of all images showed that sometimes the confidence level assigned by the KSAT operator was difficult to explain and the objectivity involved could be questioned.

A new procedure for automatic confidence assignment based on computing score factors based on features describing source, homogeneity, contrast, linearity, wind, and slick surroundings was defined. The results looks promising, but more testing on a larger data set is necessary before the procedure could be used at KSAT.

With respect to further work on defining an operational automatic procedure for use at KSAT, we recommend:

- Test the simple scheme developed on a larger data set. In particular, the data set should contain a higher percentage of medium and high confidence slicks, in addition to verified oil slicks.
- Try to develop new features that are designed with the score factors homogeneity, source, linearity, contrast, and slick surroundings.
- Create a prototype program that can be tested further at KSAT.