THEMATIC INFORMATION EXTRACTION IN A NEURAL NETWORK CLASSIFICATION OF MULTI-SENSOR DATA INCLUDING MICROWAVE PHASE INFORMATION.

Gerrit Huurneman, Rüdiger Gens, Lucas Broekema

International Institute for Aerospace Survey and Earth Sciences (ITC)
P.O.Box 6, 7500 AA Enschede, The Netherlands
Tel. +31-53-4874358, Fax. +31-53-4874335
E-mail: HUURNEMAN@ITC.NL, GENS@ITC.NL

Commission II, Working Group 4

Keywords : ERS-tandem mode, multi source data, coherence map, neural network, land use classification

Abstract

Microwave data (ERS-1 and ERS-2) and optical data (SPOT-XS) were used for the classification of an area with different land use classes. Classifications were executed for the optical data alone and for a combination of the three data sets. Two classifiers, one based on the maximum likelihood algorithm and the other on a neural network approach, were applied. From the ERS tandem mode SAR data a coherence map was created and included in the classifications in the form of an additional dimension in the feature space. The accuracy and reliability of the four classifications are presented and the results discussed.

1. INTRODUCTION

Optical and microwave data can provide complementary information about objects that cover the Earth surface. Optical data contains the information about the reflection of the solar energy in selected parts (bands) of the spectrum. The image elements (pixels) of optical sensors can be seen as vectors of which the components represent the reflection in the different bands. Image elements of microwave data consist of two components, the magnitude and the phase which are stored as complex numbers in two "layers". If optical and microwave data sets are correctly combined, the resulting product will convey more information and could prove to be more useful then either image alone.

The information contained in the multi-sensor data can be extracted visually or by computer supported methods. In computer analysis, a suitable classifier is needed to handle the optical and microwave sensor data. Since the nature of the two data sources is different, which results in different frequency distributions of the data, it makes sense to use the neural network approach to analyze this data. Neural network classifiers are able to handle the multi-nature data from different sources.

Phase information of microwave image data is mainly used in the field of interferometry. Interferometric processing of SAR data from space combines images from two passes of a sensor system or combines the data from two sensor systems in tandem mode. This process derives precise measurements of the differences in path length to the two sensor positions. The main output of interferometry of SAR data is topographic information related to terrain heights or the monitoring of positional changes of the Earth surface. A strong relation exists between the quality of these products and the correlation of the complex data sets. In those areas where high correlations exist an accurate Digital Terrain Model can be realized. On the other hand, height information cannot be extracted in areas with low correlation. So the quality of the products is characterized by the "interferometric correlation", which is a measure of the variance of the interferometric phase estimate. The amount of correlation is a function of the system noise, the volume scattering, baseline configuration and temporal change. Consequently, the interferometric (de)correlation itself contains significant thematic information that can be useful for several other applications.
In SAR images, the magnitude and the phase of each element are the coherent summations of the back scattering and phase of the individual scatterers inside a resolution cell. The variation in the overall phase and the overall magnitude of cells with equal cover type will appear as speckle. However two images taken from the same position at the same moment will be identical (neglecting system noise). If two images are taken from different positions and/or instants of time, variations in pixels representing the same surface cell will appear.
If the structure or chemical composition of the ground-cover changes, then the amount of temporal decorrelation will vary. This variation can contain information about the type of ground-cover and/or the situation in which it exists. Influence on temporal changes, in case of space borne sensor-systems, can be minimized by a high temporal resolution of the system or by a combination of two "identical" systems in a tandem mode. The ERS-1 and ERS-2 systems satisfy this last requirement.

Correlation between the data from a cell is expressed in terms of the summated phase and intensity of the resulting back scatter. A cell is considered to contain a set of individual back scatterers distributed over the cell. The amount of energy back scattered by the individual scatterers can be equal or can vary and the positional

distribution of them are regular or irregular.
If the scatterers are regularly distributed and if they have the same scatter characteristic, the cell is called homogeneous. Small changes in structure or in chemical composition of the ground-cover in such a cell will not change the phase if the imaging geometry did not change.
In a non homogeneous cell the dominant scatterers will have the highest influence on the phase therefore variation in the position of dominance will vary the phase significantly.
Apart from temporal changes, decorrelation can also be a consequence of an improper (too large) distance (base) between the two sensors during the data acquisition. To reduce the baseline decorrelation a careful selection of the orbits of the sensor system(s) is required. For the extraction of topographic and thematic information different constraints are involved. If a data pair is used to extract spatial and/or height information by means of an interferogram, the base-line between the two orbits should range from 200 to1000m to realize an acceptable height resolution. In case of the extraction of thematic information, the decorrelation related to the base line should be minimized. If the decorrelation related to the baseline is zero, the remaining decorrelation is caused by temporal changes in composition and structure of the ground cover within the cells .

In practice a zero baseline will not exist, but with a short baseline almost no baseline decorrelation exists. So differences that appear can be considered to be caused by temporal decorrelation. The data can be used as an additional dimension in the feature space for the combined image analysis.

1.1 Neural Network classification

Neural networks are based on a model of the human brain, using certain concepts of its basic structure. The network consists of many simple processing elements (neurons) ordered in layers. These layers are separated into an input, one or more hidden layers and an output layer. The elements in the hidden layer(s) are connected with all or with some elements in the next/previous hidden , input or output layer. In an operational neural network, these connections are weighted in a training stage. The training of the network is based on a set of vectors of which the class membership is known. The neural network classifiers are able to learn from sample patterns. These classifiers do not need a particular frequency distribution as required by some conventional statistical classifiers.

1.2 SAR Coherence image

For the creation of a coherence map, two SAR complex data sets have to be registered and the coherence computed. Coherence is a measure for the relation of the phase information of corresponding signals. To reduce large fluctuations in the map the coherence is computed in a window.
According to Schwäbish and Winter there are several factors which decrease the coherence:

- thermal noise
- temporal changes in atmospheric conditions
- phase errors due to processing
- temporal changes in the object phase
- different viewing positions

2. DATA PREPARATION

The "ground truth" data was collected in the field and their positions indicated on a topographic map. The SPOT image was georeferenced and geocoded to the geometry of that topographic map. Before the classifications were performed, the data sets (SPOT and Coherence Image) were registered. Further a neural net was initiated and trained.

2.1 Data description

For the experiment a data set is selected consisting of an optical image (SPOT-XS) and a tandem of SAR images (ERS-1, ERS2) in single look complex format (SLC). In

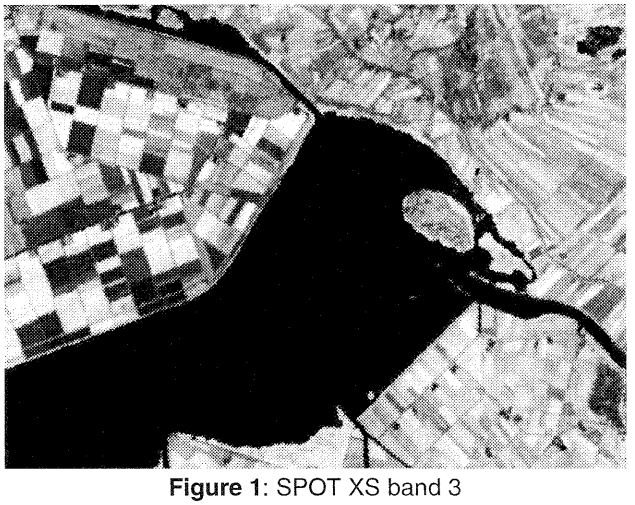

Figure 1: SPOT XS band 3

figure 1 , band 3 of the Spot image is shown. The SPOT image is acquired on 02 August 1995, the ERS-1 image (figure 2) on 19 August 1995 and the ERS-2 (figure 3) on 20 August 1995. Because of the small time interval between the acquisition dates it is expected that the types of ground cover of the sensed area have not changed dramatically. The weather conditions during the data acquisition of the two ERS images were perfect for the experiment; without rain but with different wind force and direction.
The baseline of the two images has a horizontal component of 38 meters and a vertical component of 82 meters

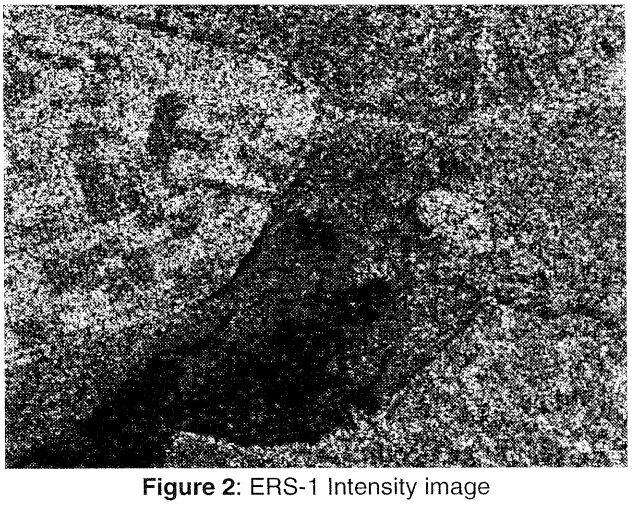

Figure 2: ERS-1 Intensity image

Figure 3: ERS-2 Intensity image

(information provided by ESA). The lengths of these components satisfy the wish to have a limited baseline decorrelation.

The window that is selected from the images reflects an area in the central part of the Netherlands. The area consists of a relative new polder with large agriculture fields, a developed moor land including reed fields, an area with small agriculture fields on the "old" land and parcels with forest. The old and the new land are separated by a lake. In the area there are no relief differences, it is a very flat terrain.

2.2 Field data collection

In order to classify the images, samples were taken that are supposed to be representative of the various types of land cover, and, at the same time, in order to check the accuracy and reliability of the classification another independent set of samples was acquired. Field data was acquired a few weeks later than the image acquisition dates. The eight classes that are selected did not change between the acquisition date and the field check. Only some grass field may have been cut in that period and some stubble fields plowed.

The following classes are recognized :

1. Bare soil.
2. Sugar beets.
3. Stubbles.
4. Forest.
5. Maize.
6. Grass.
7. Water.
8. Reed.

2.3 SPOT georeferencing

The Spot image is georeferenced using the ILWIS image processing software applying an affine transformation including a nearest neighbor "interpolation". The overall accuracy, expressed in RMSE, of the reference points used for the creation of the transformation polygons was less than 1 pixel.

2.4 ERS1/2 intensity images

From the SLC images, intensity images are created. The one with the highest number of clearly visible objects (in this case the ERS-1) is selected to support the registration of the coherence map to the SPOT image.

2.5 Coherence map creation

The coherence map (figure 4) is created using software developed at DLR in Oberpfaffenhofen, Germany. The interferometric software requires some information which is available in the leader file.

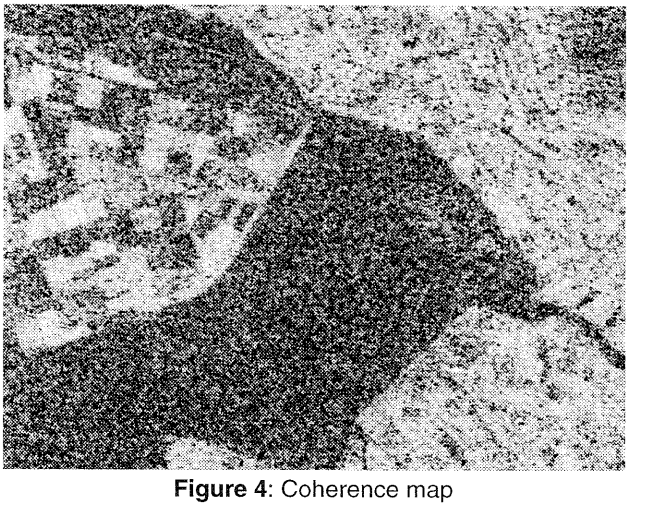

Figure 4: Coherence map

For an accurate registration of two SLC data sets it is necessary to measure the shifts between the images in azimuth and range direction which is done in the intensity images. These shifts are introduced in the data processing for a coarse registration. The WGS84 is chosen as reference system for the resampling of the second scene. To inspect the coarse registration, the fringes and the coherence map are calculated from the coarse registered images. If the registered images cannot produce a sufficient number of fringes, the coarse registration has to be repeated with more accurate shifts. The next step is the fine registration of the images. This is one of the most important steps during the whole data processing because it is the basis for the quality of the later derived interferometric products. The master image remains unchanged while the slave image is fitted to the master. To reduce unwanted variation in the coherence map, a window of 20 azimuth * 2 range elements is used for the computation of the correlation. Based on this filtered map the interferogram is calculated, i.e., the resulting fringes are also corrected under assumption that the Earth is flat.

2.6 Coherence map registration

The coherence map and the ERS-1 image have apart from minor shifts, which are caused by the slightly different orbits of the ERS-1 and the ERS-2, the same geometry. So registration of the coherence map and ERS-1 image will follow the same transformation steps. The following steps are performed, partly with ILWIS and partly with PCI's EASI/PACE.

- resample the intensity image to an azimuth resolution of 20 meter by averaging five columns elements into one.
- select the sub image of the ERS-1 that covers approximately the same area as the SPOT sub image.
- transform the image from slant range into ground range

and resample to 20*20 meter resolution.
- flip-over the image in line direction to create an almost west-east image. This step is required in case of images acquired in descending orbit.
- apply an affine transformation to register the ERS-1 image to the SPOT image.
- extract form the coherence map the same window as is used in ERS-1.
- execute the same steps, with the same parameters for the transformations, on the coherence map as are used for the registration of the ERS-1 image.

3. CLASSIFICATION

3.1 MLH classification

Optical data

For the classification of the optical SPOT-XS data all the bands (3) were used. The classification was performed with the ILWIS software. The positions of the training samples were located in the image by visual inspection, i.e. relating the position of SPOT samples in the topographic map and the corresponding positions on the false color composite map. Since the Maximum Likelihood classifier was applied, the numbers of training samples representing a class were taken into consideration because of the validity of the statistical estimates (mean and variance-covariance matrix). Per class, samples with a size of more than 100 pixels were taken. In the classification process a threshold, in the Mahanalobis distance, of 14 was used. Those pixels that were not classified applying this threshold were considered to belong to the NULL-class. In table 1 the confusion matrix and the corresponding accuracies and reliabilities are shown.

Optical and radar data

The same parameters and sample set are used for the classification of the SPOT image plus the coherence map. Only the feature space dimension is increased by one.
In table 2 the result of the classification is shown.

3.2 NN classification

Mask

Neural network classifications in general do not recognize automatically a NULL class. That means that all pixels are appointed to one of the classes the network was trained for. So also pixels that belong to a not sampled land use will be classified but to a wrong class. To avoid this, a mask was created using the NULL class of the maximum likelihood classification.

Optical data

For the investigation the EASI/PACE software of PCI (version 5.3) is used. A three layer network of fully interconnected processing units is composed. In the input layer one node represents one input channel. For the optical data alone three nodes are used. The output layer comprises 8 units, one for each class. The size of the single hidden layer is varied. Benediktsson et al. (1990) Civco (1991) used both 3 and 4 layer networks. They found that using a four layer network did not improve the classification accuracy. The momentum and learning rate were set at 0.9 and 0.1 respectively. The PCI programs use a back-propagation network which is trained by means of the Generalized Delta Rule. This involves two phases. In the first phase the weights for the inter-unit connections are initialized to random values in the range of -0.5 to 0.5. The input data is presented and propagated forward through the network. Within each processing unit the combined input contained is modified by the sigmoid function before it is passed to other connecting processing units. The second phase is a backward pass through the network, adjusting the weights to reduce the error between the actual and the desired output until it is acceptable or is stabilized. During the classification of the image, each pixel in the output layer is allocated to the class associated with the unit with the highest activation level. After the classification, the previously created mask is used to create a NULL class that is identical to the one of the MLH classification. In table 3 the result of the classification is shown.

Optical and radar data

The same parameters and sample set are used for the classification of the SPOT image plus the coherence map. Only the feature space dimension is increased by one; the first layer contains four nodes.
In table 4 the result of the classification is shown.

4. CONCLUSIONS

The overall accuracy of the four classifications does not vary significantly (about 2%). Also the variance of the average accuracy and reliability is low (about 4%). The main reason for the low variances is the almost perfect classification of the classes 1 (bare soil), 2 (sugar beets), 3 (stubbles), 4 (forest) and 7 (water) in all classifications. Only the accuracy and the reliability of the classes 5 (maize), 6 (grass) and 8 (reed) are improved. In these classes a significant difference exists between the result of the neural network (NN) classifier on the SPOT data plus the coherence map and the result of the maximum likelihood (MLH) classification of the SPOT data alone.

For some classes, which have a high accuracy in the classification of the SPOT data alone, the accuracy is decreased in the classification of the combined data. This decrease is caused by outlyers in the feature space of the combined data set. The outlyers increase the number of elements appointed to the NULL-class.

From this experiment we can conclude that the input of a coherence map improves the result of the classifications. However it does not give a clear answer which of the two classifiers gives the best result.

REFERENCES

Benediktsson, J.A., Swain, P.H., and Ersoy, O.K., 1990, Neural network approaches versus statistical methods in classification of multi source remote sensing data. IEEE Transactions on Geosciences and Remote Sensing, 28, 540-551.

Civco, D.L., 1991, Landsat TM image classification with an artificial neural network., Proceedings, ASPRS-ACSM

Annual Meeting, Baltimore, MD, Vol 3 pp 67-77.

Schwäbisch, M. And Winter, R., 1995, Erzeugung digitaler Geländemodelle mit Methoden der SAR-Interferometrie. DLR-Nachrichten, Heft 79 (August 1995), pp. 20-24.

Wegmuller, U and Werner, C.L., 1994, Analysis of interferometric land surface signatures. Procedings of PIERS'94, Noordwijk, The Netherlands, Paper Code 039.

Zebker, H.A, Werner, C.L, Rosen, P.A and Hensley, S., 1994, Accuracy of Topographic Maps Derived from ERS-1 Interferometric Radar., IEEE Transactions on Geoscience and Remote Sensing, Vol. 32 No. 4, July 1994.

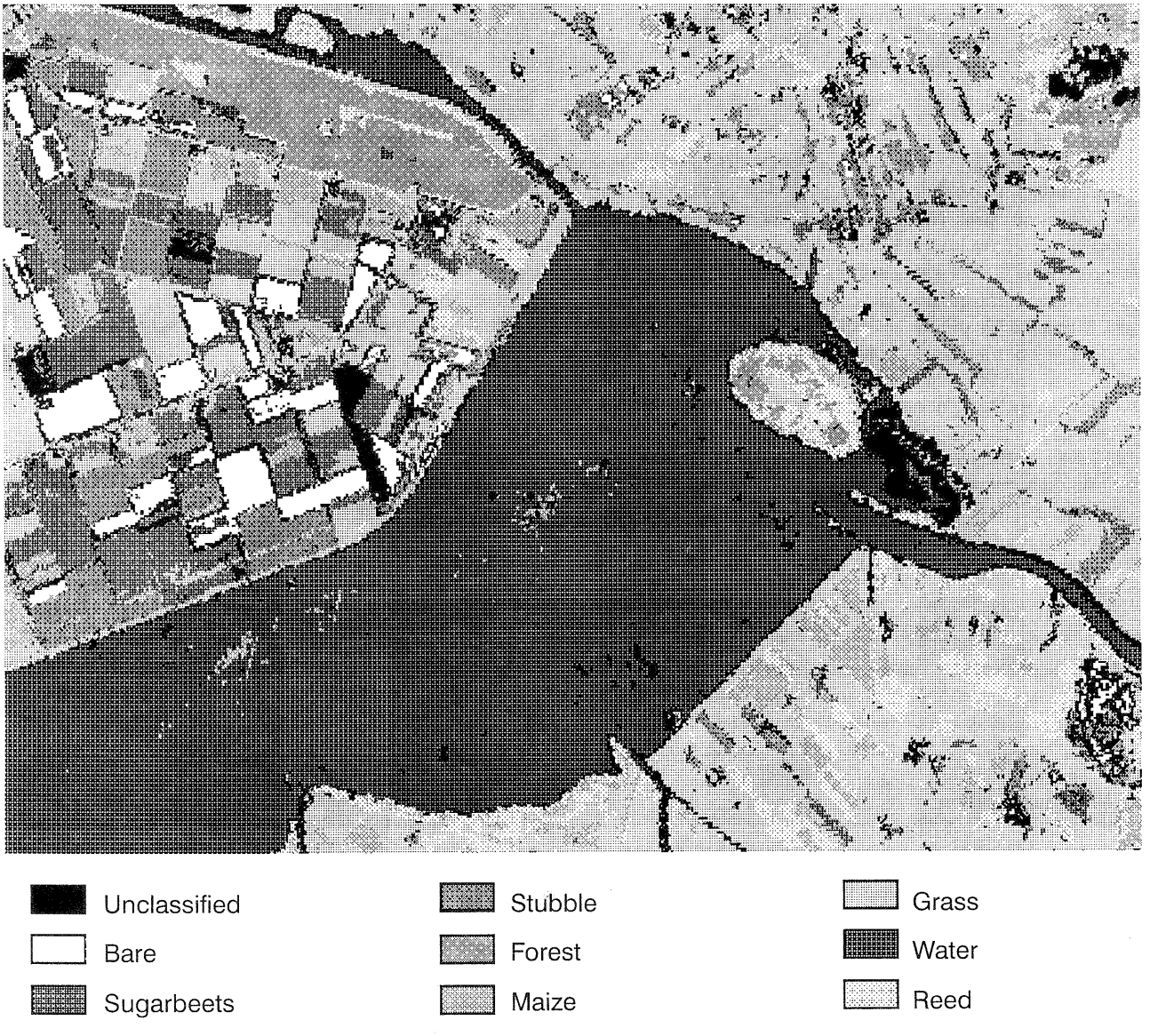

Figure 5: Neural Network classification optical and radar data

	1	2	3	4	5	6	7	8	uncl	ACC
1	189	0	0	0	0	0	0	0	0	1.00
2	0	365	0	0	0	3	0	0	3	0.98
3	0	0	251	0	0	0	0	0	0	1.00
4	0	0	0	221	0	0	0	29	3	0.87
5	0	0	0	0	46	1	0	128	0	0.26
6	0	1	0	0	114	326	0	47	0	0.67
7	0	0	0	0	0	0	414	0	0	1.00
8	0	0	0	0	57	1	0	29	0	0.33
REL	1.00	1.00	1.00	1.00	0.21	0.98	1.00	0.12		

average accuracy = 76.52 %
average reliability = 78.98 %
overall accuracy = 82.63 %

Table 1: Confusion matrix Maximum likelihood optical data

	1	2	3	4	5	6	7	8	uncl	ACC
1	168	0	0	0	0	0	0	0	21	0.89
2	0	364	0	0	0	3	0	0	4	0.98
3	0	0	250	0	0	0	0	0	1	1.00
4	0	0	0	206	0	0	0	25	22	0.81
5	0	0	0	0	82	17	0	76	0	0.47
6	0	0	0	0	88	362	0	34	4	0.74
7	0	0	0	0	0	0	411	0	3	0.99
8	0	0	0	0	45	4	0	38	0	0.44
REL	1.00	1.00	1.00	1.00	0.38	0.94	1.00	0.22		

average accuracy = 79.00 %
average reliability = 81.74 %
overall accuracy = 84.43 %

Table 2: Confusion matrix Maximum likelihood optical and radar data

	1	2	3	4	5	6	7	8	uncl	ACC
1	126	0	0	0	0	0	63	0	0	0.67
2	0	364	0	0	0	4	0	0	3	0.98
3	0	0	251	0	0	0	0	0	0	1.00
4	0	0	0	227	1	1	5	16	3	0.90
5	1	0	0	0	32	4	1	137	0	0.18
6	0	0	0	0	81	393	0	14	0	0.81
7	0	0	0	0	0	0	414	0	0	1.00
8	1	0	0	0	28	23	0	35	0	0.40
REL	0.98	1.00	1.00	1.00	0.23	0.92	0.86	0.17		

average accuracy = 74.19 %
average reliability = 77.06 %
overall accuracy = 82.68 %

Table 3: Confuson matrix Neural Network optical data

	1	2	3	4	5	6	7	8	uncl	ACC
1	168	0	0	0	0	0	0	0	21	0.89
2	0	299	0	0	22	46	0	0	4	0.81
3	0	0	250	0	0	0	0	0	1	1.00
4	0	0	0	216	0	0	0	15	22	0.85
5	0	0	0	0	73	18	0	84	0	0.42
6	0	0	1	1	70	382	0	30	4	0.78
7	0	0	0	2	0	0	409	0	3	0.99
8	0	0	0	2	35	10	0	40	0	0.46
REL	1.00	1.00	1.00	0.98	0.37	0.84	1.00	0.24		

average accuracy = 77.40 %
average reliability = 80.16 %
overall accuracy = 82.45 %

Table 4: Confusion matrix Neural Network optical and radar data