Checking computer classifications on forest decline inventory

Steffen Kuntz

University of Freiburg

Department of Photogrammetry and Remote Sensing
Werderring 6, 7800 Freiburg,
West-Germany

Abstract:

Since the beginning of the 80s large areas in Central Europe have been affected by forest decline - mostly due to air pollution. Thus an urgent need exists by governments of various nations to get updated information on the state of health, and the location of the forest decline for forestial and environmental planning, resp.
Computer aided classification results based on multispectral data were presented by different authors within the last few years. One problem was - and still is - the interpretation of these results.
Main aspects for checking computer-aided classifications are the sample point design for the ground truth and the correlations between scanner data with different ground resolution (pixel size) and ground truth data from terrestrial survey or from CIR aerial photographs.
Methods for the evaluation of classification results must change with scanner data taken from different altitudes according to different pixel sizes. The checking of classification results of single tree crowns need other test strategies than those of tree groups or forest stands.
In this paper computer aided classifications for forest decline inventory from different altitudes are presented and test strategies for checking the results are discussed.

Introduction

Since 1983 multispectral scanner data from altitudes between 300 and 4000 m have been acquired every year over varying test sites in Germany to investigate the usefulness of these data for forest decline inventory. Different authors demonstrated the feasibility of the scanner data using digital image processing techniques and many results have been presented within the last years (e.g. Hildebrandt et al., 1987; Kritikos et al., 1985). But a big problem is still the verification of thematic forest decline mapping with computer-aided classification techniques. Common methods for statistical testing of thematic maps do not suit in this special problem because:

- It is not possible to compare scanner data directly with data from forest decline inventories obtained by terrestial survey or CIR aerial photographs . At one side one has single trees or tree crowns, resp., which represent a particular vitality (damage class) and on the other side one has an area which only in few cases represents exactly a tree and differs in size due to different flight altitudes.
  This also means that different methods are needed to evaluate results of a classifier within single tree crowns or over large areas due to pixel sizes varying e.g. from 0.75 x 0.75 m (scanner data from 300 m altitude) to 30 x 30 m (Landsat TM), resp.
  Another problem is the comparison between reflection mesurements by the scanner and subjective human interpretation of the state of health of single tree crowns. The latter is not based on radiometric values (colour) only, but includes structure and form-elements of crowns, branches, and foliage.
- Stratification in more or less homogenous areas which show uniform damage situation - as is recommended when using multistage sampling - is very difficult due to very inhomogeous situations within the small and heterogenous forests in Europe and due to frequent changes in species and in vitality of crowns standing close together.
- The geometric problems of airborne scanner data make it very hard to exactly locate a specific pixel within a stand or within a tree crown. A problem, whichs grows with increasing pixelsize.

Due to these problems it is not possible to give an exact answer whether or not one particular pixel is classified correctly. To solve this problem it was tried to find a method which allows a judgement of classification quality, when looking at larger groups of pixels (within one crown or within larger sample areas) and to compare them with the real situation.

In the investigation presented here scanner data from 300 m and from 1000 m altitude were evaluated. They were acquired in 1984 from 2 test sites in South-West Germany with a Bendix M^2S-Scanner run by the DFVLR Oberpfaffenhofen.
The evaluation was based on the interpretation of single tree crowns from CIR aerial photographs and on the comparison of the results with the computer aided-classification.
The interpretation methods of forest decline inventory using CIR photos are operational and it is assumend that they represent the real situation better than terrestrial survey because they show the treecrowns like the scanner does. They allow to reveal the current situation as well.

Analysis of scanner data from 300 m altitude

In terrestial and aerial forest decline inventory the objects of interest are single crowns. So at the beginning of the research single tree crowns (spruce) were investigated using scanner data from 300 m altitude.
Due to the high pixel resolution (0.75 x 0.75 m) at this altitude one crown of an old needle tree is represented by up to 100 pixels.
If one compares classification results from this altitude directly with the ground truth (total number of trees per damage class) within a stand obtained by terrestrial research or CIR photointerpretation, the agreement would be poor. In that case one compares parts of the crowns with whole trees because within one tree crown one can find all healthy parts close together with less and severely damaged parts. The human interpreter looks at the whole crown and dedicates it to one damage class. The per point classifier only looks at single pixels which do not normally represent the whole tree (Hildebrandt et.al., 1987).
To see what really happens within the crowns during the classfication process all crowns of a stand in South Germany were interpreted as species and damage classes and were digitized from CIR photos, scale 1:5000. They were used to digital overlay the original scanner data. By the aid of the crown map it was possible to exactly recognize and define more than 400 crowns of that stand and store the coded information in a special annotation channel in addition to the original image channels.
Table 1 shows that only $^{1}/_{3}$ of the crowns could be exactly identified in the scanner data. A total of 404 crowns (spruce and pine) were included in the comparison. 34 decidous tree crowns were also identified but not listed. Pine trees could not be differentiatied into damage classes in the scanner data due to specific characteristics.

Tab. 1: Number of digitized tree crowns from the map and the number of the crowns which could exactly be identified within the scanner data (from: Kuntz, 1987)

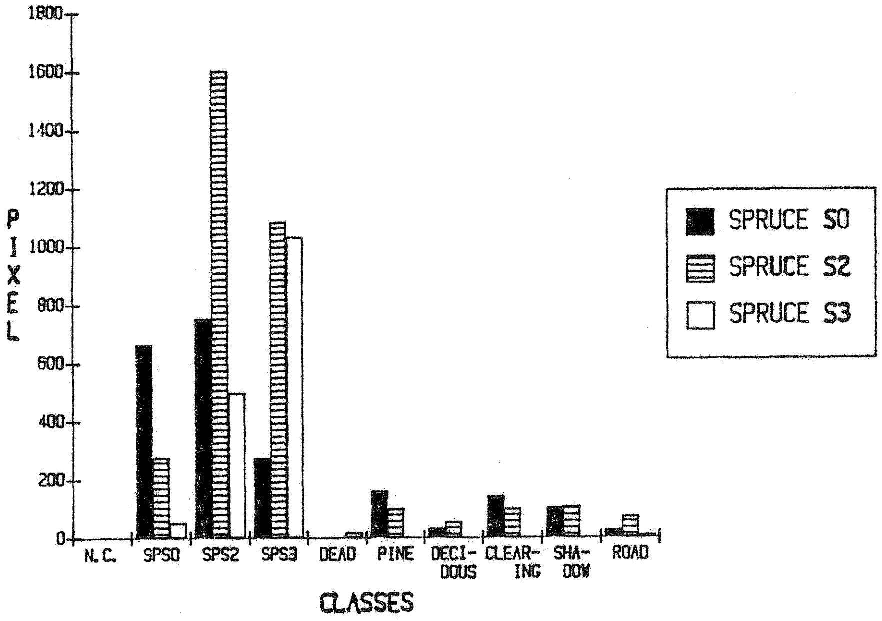
	Damage class				
Species	S 0 healthy	S 1 lightly damaged	S 2 damaged	S 3 severely damaged	S 4 dead
	crown map				
Spruce	125	450	586	111	5
Pine	0	9	45	46	0
	scanner data				
Spruce	25	89	169	74	2
Pine	all crowns: 47				

To check the classification accuracy the digital crown map (lower part of Table 1) in the annotation channel was used. Using special software implemented at the University Computer Centre in Freiburg on a Sperry 1100/82 machine the results of the classfication channel and the annotation channel were transformed into SPSS System files. Using a SPSS-routine (CROSSTABS) it was possible to count the classified pixels within all digitized crowns and show the numbers of pixels for each class in an error matrix (Tab. 2).
In the matrix the classes Spruce S 0 and S 1 obtained by the CIR photointerpretation were combined for the computer aided classification within one class (Sp S0) because it is not possible to seperate the two classes with the applied classifier (supervised maximum likelihood).
Fig. 1 gives an graphic representation of the results of this classification. It is a plot of the column values of the matrix in Tab. 2) for classes "Spruce S0 to S3". It shows the pixel distributions for all crowns of a damage class in a histogram. This allows an easier interpretation of the results than the error matrix.

Tab. 2 : Results of the first classification of a stand (Nr. 143 in Oberschwaben, Forstamt Zeil-Waldenburg, South Germany); scanner data from 300 m altitude.
Sp = Spruce; S0 - S 3 = damage class 0 to 3;
n. c. = not classified;
For better clearness the total numbers of trees (from Tab. 1) are given below the matrix

Results of the Classification	**Digitized Crowns**						
	Sp S0	Sp S1	Sp S2	Sp S3	Sp dead	Decidous	pine
	declaration of values in %						
Sp S0	41.5	28.3	8.1	2.9	0.0	6.4	18.9
Sp S2	23.4	37.3	47.1	27.0	3.1	3.9	25.8
Sp S3	7.6	13.9	31.8	56.0	53.1	2.3	22.5
Sp tot	0.0	0.0	0.0	0.1	29.7	0.0	0.0
Decidous	2.8	1.2	1.6	0.2	0.0	56.5	2.3
Pine	13.5	6.2	3.0	1.4	0.0	1.4	16.6
Glade	3.6	7.3	3.0	0.3	0.0	22.3	7.8
Shadow	6.1	4.6	3.2	1.5	0.0	4.6	3.2
Road	1.5	1.2	2.2	5.6	0.1	1.4	2.9
n. c.	0.0	0.0	0.0	0.0	0.0	1.1	0.0
Sum	100	100	100	100	100	100	100
	Numbers of Digitized Crowns per Class						
	25	89	169	74	2	34	47

Fig. 1: Graphical representation of the first classification using airborne scanner data from 300 m altitude.



It is obvious that one can find pixels representing nearly all classes in each crown but that there are more or less typical distributions of classified pixels within the different damage classes.
Especially the pines don't show any satisfying results, but even for the skilled interpreter it was very difficult to identify the pines in the original scanner image visually. Only with the help of the crown map it was possible, to acrely seperate pine from spruce trees. The statistics didn't show any significance either so it was impossible in this case to seperate pine from spruce using the classifier mentioned above.

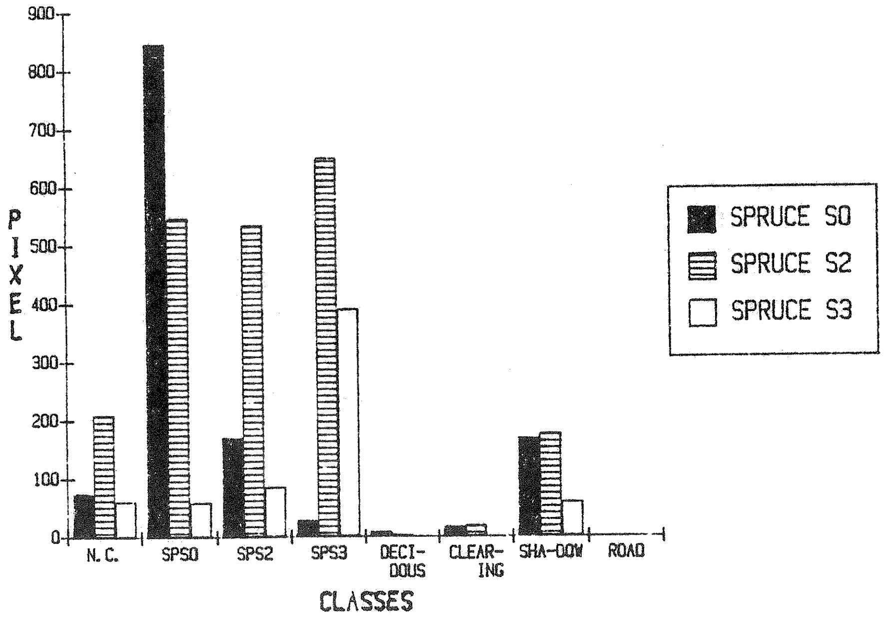
Employing a second approach the effects of the scan angle were considered when the training fields for the classification were defined. With the software system FIPS (Freiburgs Image Processing System) it is possible to take special conditions into consideration during the actual classification process. For example the classifier uses only a special trainings area set when processing a predefined region (rows or colums or grey values in a special channel) (Lösche, 1984).
In this case four colum regions were defined depending on the scan angle, and for each region separate training fields were taken.
The results of this classification for the first region (scan angle < 20 °) are shown in Tab. 2. For the other regions the results are quite similar.
Because it is not possible to give a per point pixel accuracy with this method but only the distribution of the pixels in specific damaged tree crowns it is difficult to decide if a classification is "good" or becomes "better" than another. But as a result to the limitation on a relativly small scan angle

the classification accuracy seems to increase in the classes Spruce S0 and S3. The class Spruce S 2 shows less accuracy.
Damage class S 2 has spectral signatures lying exactly between classes S 1 and S 3, and it is not predictible how the classifier will react in a specific crown. But the graphic representation of the classification results in Fig. 2 clearly show that the distribution of the classified pixels within the crowns are typical for each damage class (Fig. 2).
This also means that the use of texture filtering within the crown or a classifier which uses neighbourhood information may provide better results than a per point classifier. Further investigation of that matter will be done in the near future in our department.

Tab. 3 : Results of the second classification of stand nr. 143 (in Oberschwaben, Forstamt Zeil-Waldenburg, South Germany), scanner data from 300 m altitude
Sp = Spruce; S0 - S 3 = damage class 0 to 3; n. c. = not classified

Results-of the-classification	Digitized Crowns declaration of values in %						
	SpS0	SpS1	SS2	SpS3	Sp dead	Decidous	Pine
Sp S0	65.4	63.9	25.6	9.6	0.0	11.2	13.6
Sp S2	7.7	13.6	25.0	13.8	0.0	0.4	22.8
Sp S3	3.1	2.2	30.3	62.8	81.8	0.3	49.4
Decidous	2.3	0.5	0.1	0.3	0.0	62.5	1.6
Glade	0.0	1.5	0.9	0.0	0.0	8.3	2.7
Shadow	17.7	12.3	8.3	3.0	0.0	4.6	1.2
n. c.	3.8	6.0	9.8	10.0	18.2	12.7	8.7
Sum	100	100	100	100	100	100	100

Fig. 2: Graphical representation of the second classification using airborne scanner data from 300 m altitude.



Analysis of scanner data from 1000 m altitude

For the development of inventory methods for large areas using scanner data from high altitudes one must consider that only few pixels will represent a crown o a group of trees standing close together. Problems increase when checking the classification accuracy because larger areas have to be considered. Also, it is impossible to select one specific pixel within a stand as a representative for the actual damage situation.
In the original scanner image the intersection of every 100th line and every 100th column was taken as centres of 15 x 15 pixel boxes. Every box contains 225 pixels which approximatly represent an area of 1400 m^2.
The areas so defined were marked with a specific grey value and were stored in an annotation channel in the original image as an overlay of systematic sample areas.
After plotting the sample areas together with the original scanner image on an ink jet plotter it was possible to transfer them with enough accuracy onto the CIR-photos (1:5000 scale) with a Bausch & Lomb Stereo Zoom Transferscope.
Every marked area in the photos was magnified and then analyzed using a dot grid. The raster size of 3 x 3 mm^2 was chosen to approximately represent the pixel size in the scanner data.
With this method it is even possible to interpret parts of trees and to compare the digital classification results with human interpretation methods. Simultaneously the damage classes for the whole tree crowns were interpreted in the CIR aerial photographs with an Aviopret and a verbal description of the area was given.
The areas were interpreted using the same classes as were used for the computer aided classification (damage classes S 0 to S 4 for spruce and fir, young plantation, decidous, glades, roads, shadow, non-classifiable). So for every sample area the exact statistics (number of grid points per class) from the CIR photo were available.
Using the SPSS-programm CROSSTABS after the computer aided classification for every sample area a distribution of classified pixels for all classes could be obtained.

This method makes it possible to compare results of the computer aided classification with those of the CIR photointerpretation in a direct manner. Even statistical methods for the comparison of the two distributions are possible. When using a greater amount of pixels per sample area the very exact localization at the ground or in CIR photos is not so essential.
The approach allows a statement for a group of pixels only, not for single pixels.
The method also allows to judge why, at a special sample area, the agreement between CIR-photointerpretation and scanner is poor.

A quick and easy method comparing the two datasets directly, is the plotting of the pixel histograms (Fig. 3 to 5), showing the numbers of pixels per class of a sample area in comparison with the results of the CIR photointerpretation (number of dot grid points per class) .

The sample areas are systematically distributed over the scanner image and not stratified into homogeneous parts or stands. They represent partially very inhomogeneous areas including different stands, roads etc. This allows a judgement what happens during the classification process when heterogeneous areas are close together and influence each other (problem of mixed pixels). Especially when having many mixed pixels within a large area - a situation which occurs very often in the forests in Middle Europe - one must take into account that a stratification only in homogeneous stands may result in overestimation of the accuracy. On the other hand the per point information of the damage situation over a large area is so interesting - as was stated by foresters in the Black-Wood-Forest (Hildebrandt et.al) - that classification errors over inhomogenous areas may not so essential.

Besides the visual interpretation statistical methods were used to compare the two distributions (classification versus CIR-photointerpretation) of the sample areas. A Kolmogoroff-Smirnoff test (K-S-Test) was used to test if there are significant differences between the two distribution. Spearman's Rank Correlation Coefficient (R_s) was determinded as well. Table 3 shows the test results for the 47 sample areas.

Tab. 4 : Statistical test results of the class distributions within the 47 sample areas (for the K-S-Test only the tree-classes were selected, the R_s used all classes within the sample areas).

	Kolmogoroff-Smirnoff Test:	Spearman's Rank Correlation Coefficant	
	p = 0,05	p = 0,05	p = 0,1
differences	41	32	30
no differences	6	15	17

It is surprising that the distributions of 6 sample areas show no significant differences when using the K-S-Test at the 95% confidence level because this test is very conservative. It is questionable if such a strong test is adequate for the comparison of CIR-photointerpretation (with an interpretation error of approximatly 10 %) and airborne scanner data.
The R_s-test is less conservative and it seems more resonable to use it at a lower confidence level. Otherwise reasonably good classifications will not be accepted.

In 18 sample areas significant differences can be explained by the special terrain situation (mountainous) and the flight direction (north-south). They are located at steep slopes (east and west exponation) and at the lower parts of the mountains.
Therefore different illumination conditions in combination with the directional reflectance properties of the forest canopies explain the classification errors. It is expected that incorpo-

rating scene radiation models will provide better results. Further research will be done on this topic.
The differences in the remaining sample areas can be attributed to the very inhomogeneous conditions within the areas (especially in partially failed or discolored plantations, young copse or pole wood, clearings, and when roads are included). The large amount of mixed pixels contributes in classification errors depending on the classifier being used.

Summary

Results of computer aided classifications of single tree crowns and stands (using multispectral scanner data accquired from 300 and 1000 m altitude, resp.) were presented. Methods for checking the results were shown.
It could be demonstrated that spruce crowns (scanner data from 300 m altitude) show characteristic pixel distributions within the specific damage classes.
To check the results of classifications from 1000 m altitude a net of systematically distributed sample areas was designed. A direct comparison of CIR-photointerpretation results with the results of the computer aided classification of the scanner data within the sample areas was possible. Even statistical methods could be applied.
It could be shown that the classification accuracy decreases at steep slopes depending of different reflectance properties and on heterogenous. Homogeneous areas showed good classification accuracy.

Fig. 3 to 5: The figures show the plots of the CIR-photointerpretation (CIR) and the computer classification (COM) for three sample areas. The sample areas were selected to show agreement or disagreement, resp. for the three possibilities using the statistical tests mentioned above: Fig. 3 agreement between CIR and COM using K-S-Test and R_s, Fig. 4 agreement only with R_s, Fig. 5: no agreement).
The numbers above the plots show the number of tree crowns per class.
The lines connecting the data point are only used to guide the eye.

plant. = plantation, Sp = Spruce, F = Fir, S0 - S3 = damage classes (healthy - damaged - severly damaged)

Fig. 3

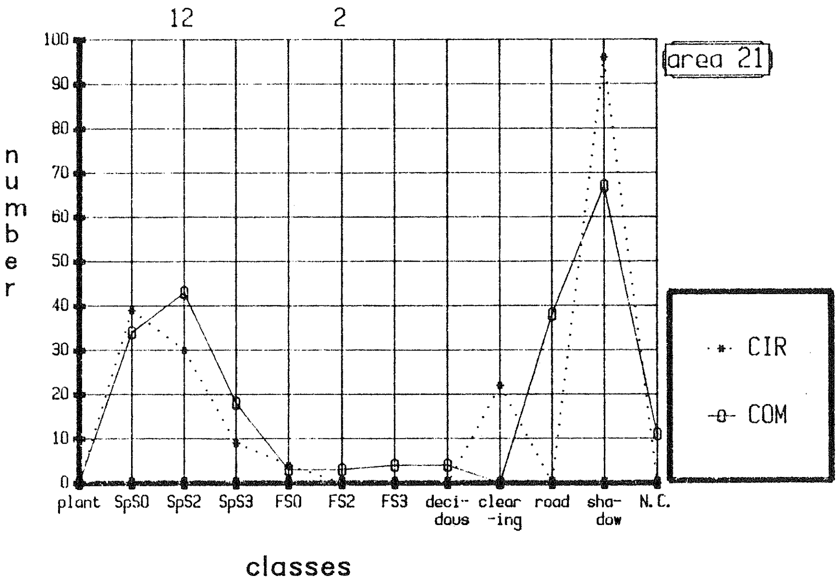


Fig. 4

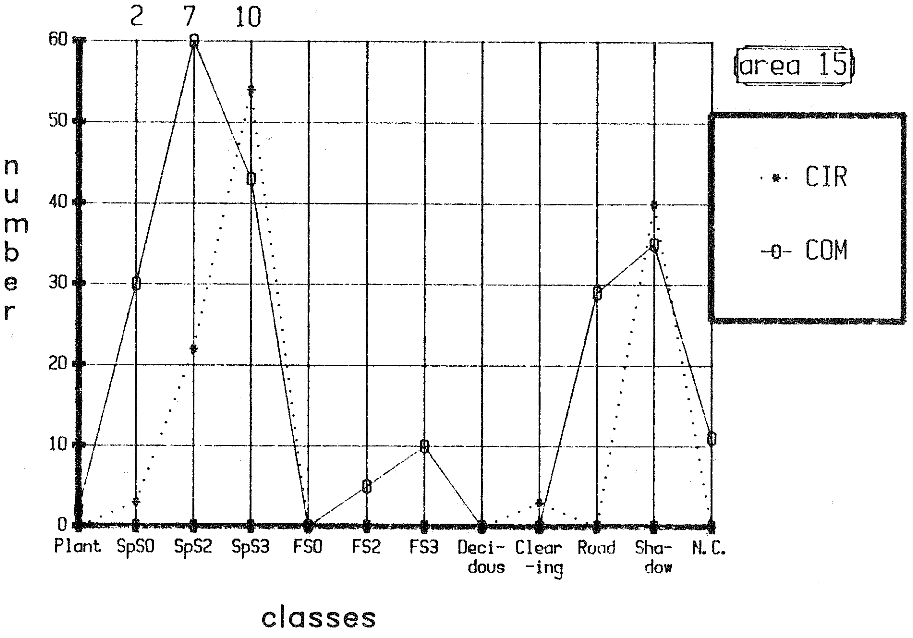
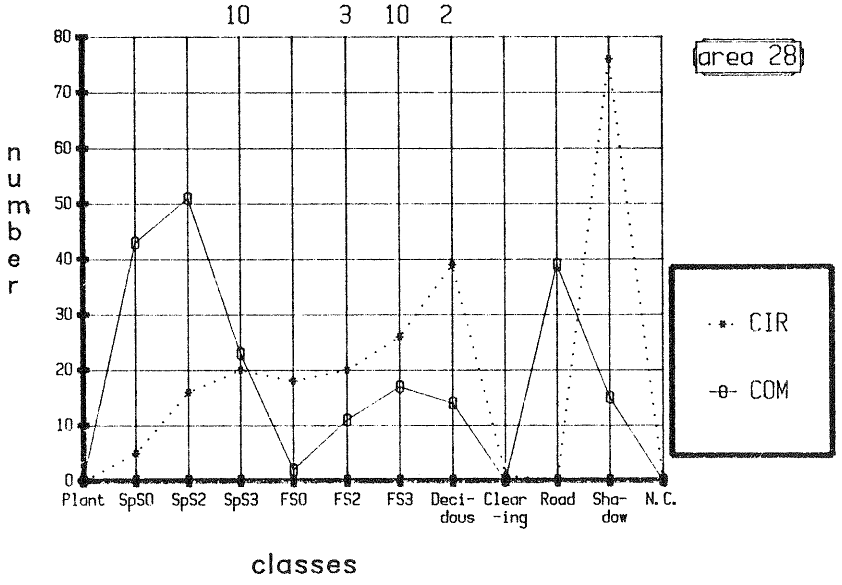


Fig. 5



Literature

G. Hildebrandt, A. Kadro, S. Kuntz, C. Kim: 1987
Entwicklung eines Verfahrens zur Waldschadensinventur durch multispektrale Fernerkundung. KfK-PEF 25, März 1987

G. Kritikos, D. Kübler, K. Herrmann: 1985
Vergleichende Waldschadensklassifizierungen mit Scannerdaten aus verschiedenen Flughöhen. DFVLR Interner Bericht

S. Kuntz: 1987
Analyse computergestützter Klassifizierungen geschädigter Nadelbäume. Vortrag gehalten beim 2. Statuskolloquium Waldschäden bei der DFVLR Oberpfaffenhofen vom 30.11.-2.12.1987; vorgesehene Veröffentlichung in einem Berichtsband.

G. Lösche: 1984
QUMLOI: ein Werkzeug zur Klassifizierung spektraler Signaturen unter Einbezug von Zusatzinformationen. Interner Bericht, Freiburg Forstfakultät.

L. Sachs: 1984
Angewandte Statistik. Springer Verlag Berlin, Heidelberg, New York, Tokyo, Sechste Auflage 1984