CLOUD CLASSIFICATION ON THE BASIS OF NOAA-APT DATA USING A FUZZY LOGIC APPROACH

Uta Heinzmann
Department of Meteorology and Climatology
University of Basel
Spalenring 145, CH-4055 Basel, Switzerland
Commission VII

Abstract

A combined method for cloud analysis is presented, using remotely sensed satellite data as well as conventional surface observations. Satellite data (NOAA-11) is classified applying a fuzzy logic approach. This algorithm has the following characteristics: Classes are defined by the user directly using fuzzy logic linguistic variables instead of training data. For each pixel the possibility of class membership is given as a fuzzy membership function. The algorithm has therefore the ability to determine cloud cover percentage of mixed pixels, which are frequent at the APT-resolution. Comparison of classification results against a maximum likelihood classification algorithm are demonstrated for one scene. In a first step, four cloud classes, high convective cloud, cirrus, middle cloud and low cloud are classified. Cloud cover fraction can be derived directly via fuzzy logic operators. To compare ground observations of cloud cover to the result of the classification a weighted filter is applied to the image, simulating the ground observer's perspective of the sky.

Keywords: Cloud classification, NOAA, fuzzy logic

INTRODUCTION

Clouds are probably the most important parameter controlling the radiation (and heat) budget, consequently, they have a very crucial impact on climate. A requirement for accurate cloud data does not only exist for climate models on a global scale, but also on a regional scale. One of the objectives of the REKLIP (Regio-Klima-Projekt) climate research project is a better understanding of heat balance in the Upper Rhine Valley. For these purposes it is not only necessary to determine cloud cover, because cloud impact on radiation like reflexion, scattering, emission at different wavelengths are also depending on parameters like cloud type, cloud height, optical thickness.

Retrieval of cloud parameters is based on more than 400 NOAA-AVHRR scenes recorded daily at GIB/AMK (Geographisches Institut Basel, Abt. Meteorologie/Klimaoekologie) since November 1990 between 12:00 and 14:00 GMT.

The research area, the so-called Regio, is located in central Europe covering parts of France, Germany and Switzerland. It is bordered by the mountain ridges of the Vosges in the west, the Black Forest in the east an the Jura in the south. The northern border is at the german town of Pirmasens. The size of the area is about 23 000 km^2.

CLOUD OBSERVATIONS

Surface observations

Apart of their poor spatial resolution, ground observations comprise many sources of errors because visual observation is the only means in operational use. Cloud cover is taken in okta and is defined as the fraction covered by cloud from a central perspective. An observer may for instance under- or overestimate cloud cover significantly due to perspectivic distortions. Usually surface based cloud cover observations overestimate cloud cover by about 15 percent. On the basis of this data it is not possible to determine the spatial distribution of clouds.

Satellite observations

Cloud cover observed by the sensor of a satellite at a certain spectral band could be defined as the fraction of the sky that is covered by clouds from vertical, parallel perspective, with no respect to type, height and density of the clouds. Different cloud cover percentages are observed for different spectral bands.

> It is by no means clear that a unique definition for cloud amount even exists; the effective cloud cover fraction for incoming solar radiation might differ significantly from that for outgoing infrared radiation. (Cess et al., 1982)

Direct comparison between surface observations and satellite cloud statistics is extremely difficult.

> No accepted conversion exists between satellite-derived percentage cloud amount and the okta and tenths scale of surface observations. (Hughes, 1984)

CLOUD CLASSIFICATION

Cloud retrieval algorithms can generally be divided into two classes:

1. Classification algorithms based only on spectral information like simple threshold techniques, unsupervised or supervised clustering techniques; and
2. Classification algorithms based on spectral and spatial textural information. Textural information is extracted by the use of neighbourhood operations (Ebert, 1989; Khazenie and Richardson, 1991).

It is obvious that the classification accuracy increases with the number of input parameters. Application of satellite data in climate research necessarily leads to a conflict between the amount of data that should be taken into consideration and processing time for complicated classification algorithms. Threshold techniques, which only give a crude impression on cloud cover and cloud type, are applied for global cloud climatologies like ISCCP or NIMBUS (Rossow, 1991; Hwang et al., 1988). On the other hand, sophisticated algorithms are applied to single or few NOAA or Landsat-TM scenes (Arking and Childs, 1985; Wielicki and Welch, 1985; Ebert, 1989). These results are not representative in a climatological sense. To take full advantage of satellite data for climate research, it is necessary to devellop classification algorithms which lead to acceptable results without requiring too many processing steps on the basis of data which is available at low cost and in acceptable spatial, temporal and spectral resolution.

Maximum Likelihood Classification

Maximum likelihood (MLK) as a supervised statistical classification technique assumes probability distributions of the form of multivariate normal models for every class defined by training data. The statistics of training data is of essential importance for the resulting classification. There are some particular obstacles to cloud classification of APT-data with maximum likelihood:

- Pixels at every point in multispectral space are classified into the class with the highest probability of class membership. For pixels with very low probabilities, the problem of misclassification is usually solved by application of thresholds. Mixed pixels that are not rejected are classified into one of the classes they represent.
- Normal Distribution of probabilites is rather an assumption than a property of natural spectral classes.
- It is not always possible to determine a sufficiently representative amount for each class when defining classes by digitizing training pixels.
- A deterministic definition of a class is not possible.

Fuzzy Logic Classification

According to fuzzy logic theory, for every pixel value x and every class A, there is a membership function of x to A. In binary logic $mA(x)$ can only be equal to 0 or 1, while $mA(x)$ in fuzzy logic can have all real values from 0 to 1. The membership function can be considered as a measure for the possibility of x belonging to A.

$$A = \{(x, mA(x))\}$$

Using this formalism, two basic membership functions describing high and low spectral response were defined by Blonda (1991) for the purpose of land-use classification on the basis of Landsat-TM data. These functions can be modified by fuzzy logic operators as follows.

$$VERY(A(x)) = (mA(x))^2$$
$$QUASI(A(x)) = sqrt(mA(x))$$
$$NOT(A(x)) = (1 - mA(x))$$

Fuzzy Logic operators AND and OR can be either minimum operators or product operators depending on the degree of compensation that is to be expressed.

$$(A(x))AND(B(x)) = MIN(mA(x), mB(x))$$
$$(A(x))AND(B(x)) = mA(x) * mB(x)$$
$$(A(x))OR(B(x)) = MAX(mA(x), mB(x))$$
$$(A(x))OR(B(x)) = MIN(1, mA(x) + mB(x))$$

Combinations of the functions HIGH(x) and LOW(x) with these fuzzy logic operators enable the user to define classes by describing the spectral response in every channel. It is not necessary to determine the statistical parameters of training data, which have the above mentioned uncertainties.

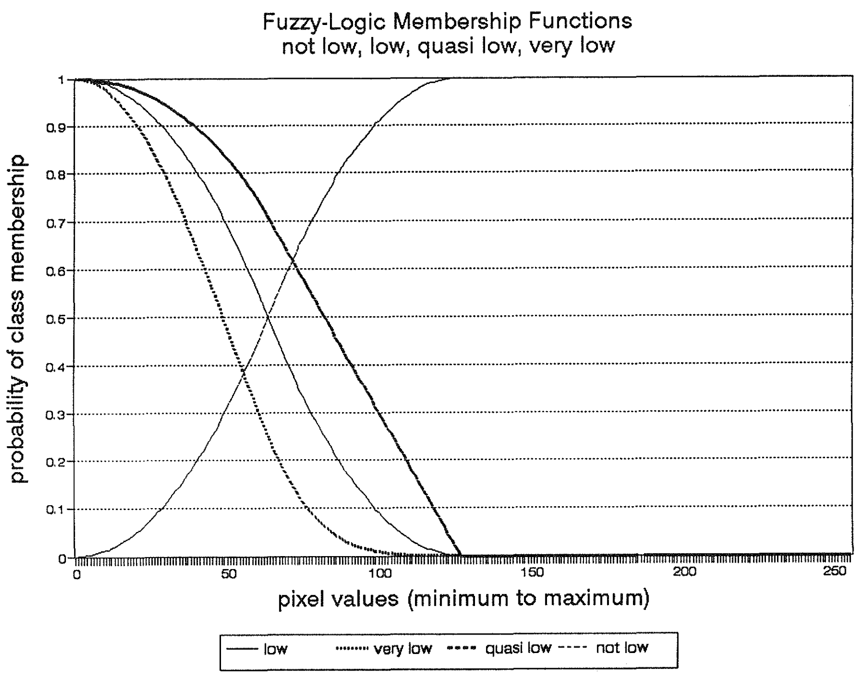

Figure 1: Membership function HIGH modified with fuzzy logic operators very, quasi, not

THE FUZZY LOGIC CLASSIFICATION ALGORITHM

On the basis of this approach, a classification method (FLOP) was develloped for cloud classification with NOAA-APT. NOAA-APT data format has a spatial resolution of 3.3 km^2 and a spectral resolution of two channels:

1. channel 1 (0.58-0.68 μm)
2. channel 4 (10.3 - 11.3 μm)

A simple geometric correction of APT data is done before transmission of the data as an analog signal. At GIB/AMK, geometric distortion is corrected by polynomials of third order. Coefficients are estimated by identification of at least 12 ground control points. Due to auto-contrasting during reception of the APT data, the pixel values have different scales and offsets for every scene. For maximum likelihood classification it would be necessary to apply a calibration or to determine training data for every scene seperately. The FLOP classification algorithm therefore defines the classes by giving the position of the values relative to minimum and maximum of every scene. One percent of the pixel values are cut off at the upper and lower part of the histogram, because these values often are due to noise. Instead of applying a fixed, deterministic threshold value seperating cloud from background for each pixel in multispectral space, the result of the fuzzy logic classification algorithm is the value of the membership function for each class at every spatial location. For every cloud type, there is a resulting image, which can be converted to cloud cover.

Using the fuzzy logic language, 4 cloud classes are characterized spectrally by the linguistic variables listed in table 1. Class definitions have to be adapted to changes in temperature and radiation during the year.

Using the minimum operator for AND, the cloud cover percentage can then be defined as follows:

$$c = (mA(x_1) + mA(x_2) + \cdots + mA(x_n)) \; AND \; 1$$

VALIDATION, RESULTS AND DISCUSSION

The application of the FLOP algorithm to the images shown in figure 2 gives the image of cloud cover percentage shown in figure 3. To validate results, the same scene was classified using a supervised maximum likelihood technique. A set of 6 classes has been selected for this purpose.

As mentioned before, there is no uniform conversion between surface based cloud observation and cloud cover derived from satellite data. To compare classification results, the cloud cover images were filtered with a template weighting each pixel with a value

class	channel 1	channel 4
Low cloud	high	not very high
middle cloud	not high and not very low	not high and not very very low
high convective cloud	high	high
cirrus	not high	very high

Table 1: Fuzzy logic membership functions for 4 cloud classes

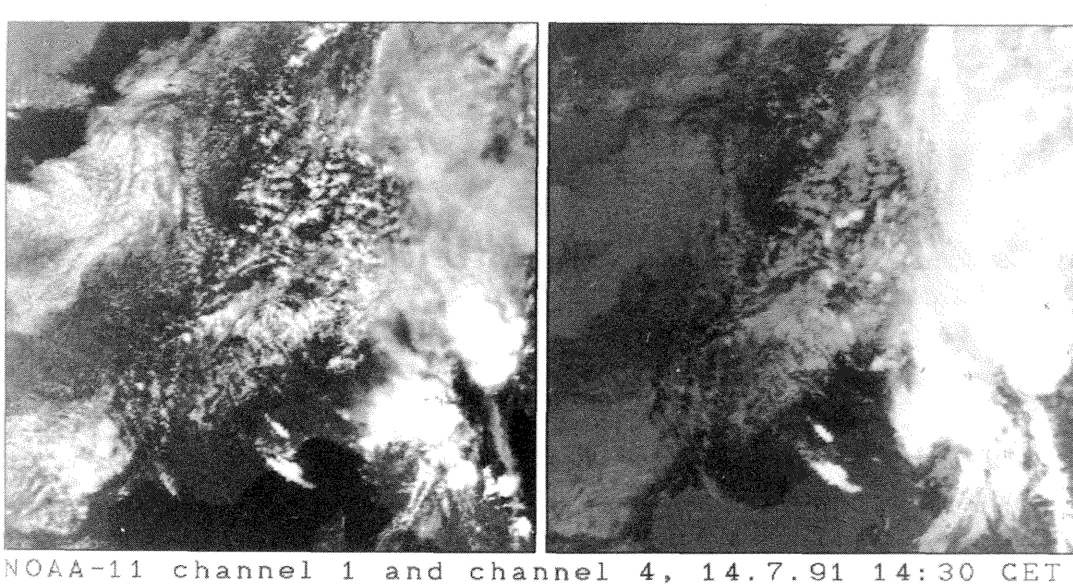

Figure 2: NOAA-11 14.7.91 channel 1 and channel 4

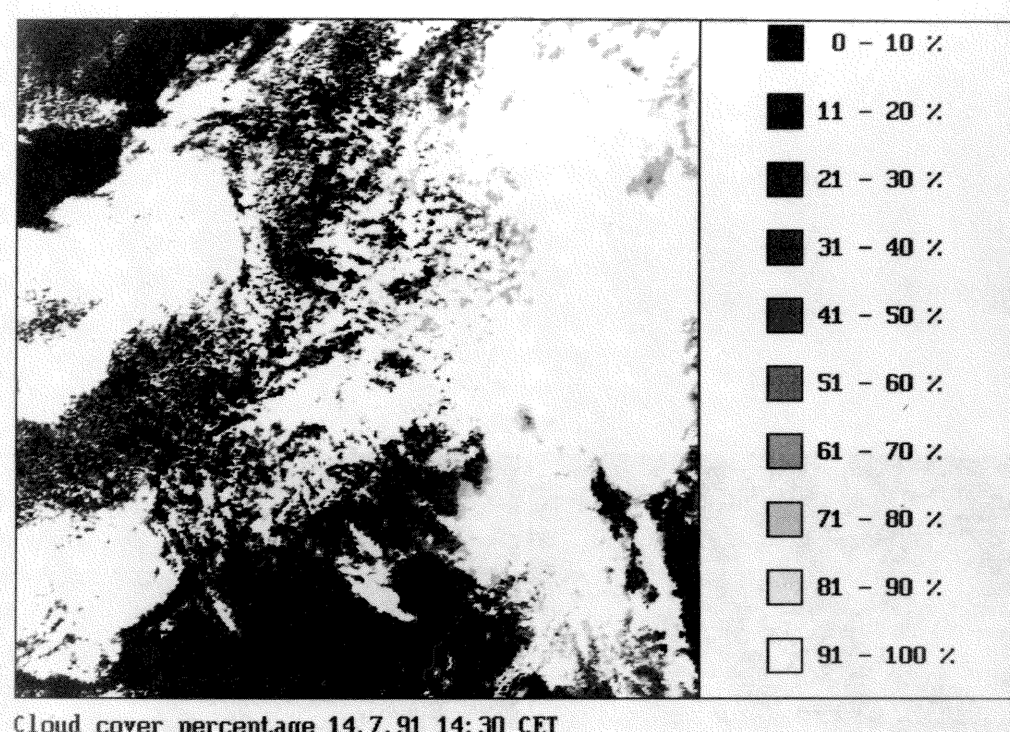

Figure 3: Fuzzy logic classification: cloud cover

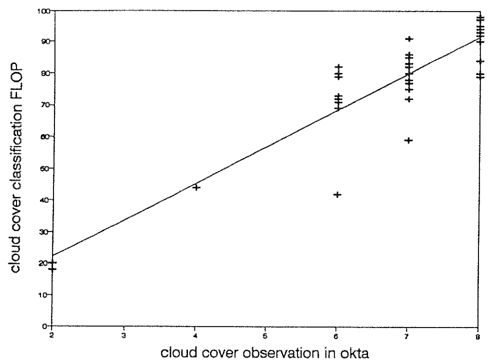

Figure 4: Regression analysis ground observation - FLOP

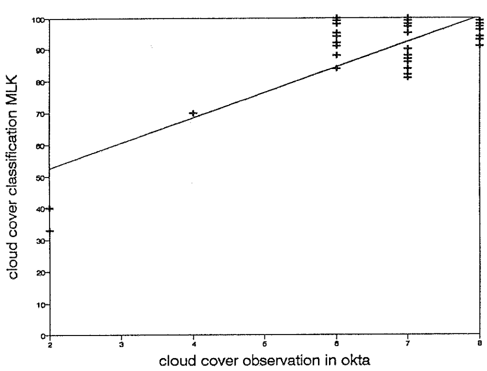

Figure 5: Regression analysis ground observation - MLK

according to it's distance from the center and cloud height. This operation generates an image of cloud cover as observed from a ground station. The template size was set to 17 pixels (about 56 km).

In figure 4 and figure 5 classification results are plotted against data from 44 DWD (Deutscher Wetterdienst) ground stations. Date of observation was 14:30 CET on 14.7.91. The ground stations are located in southern Germany, where convective clouds covered parts of the sky during observation.

This type of cloud is difficult to detect using the MLK, because cloud size is small compared to pixel resolution. Cloud cover is overestimated because a great number of mixed pixels are classified as cloud. The regression coefficient is 0.62 for the regression ground data - MLK. For the fuzzy logic algorithm the correlation coefficient is 0.81. This is explained by a better classification of inhomogenous parts of the cloud classes (like cloud edges).

Considering that these results are only a first step in the development of an appropriate classification method for time series of satellite data, they show that fuzzy logic seems the better means to describe the characteristics of clouds for classification.

REFERENCES

Arking A., Childs J.D., 1985. Retrieval of Cloud Cover Parameters from Multispectral Satellite Images. Journal of Climate and Applied Meteorology, 24:322-333.

Bachmann M., Bendix J., 1991. Fog Studies in the Alpine Region with NOAA/AVHRR. In: Proceedings of the IEEE. Helsinki, 1713-1716.

Blonda P.N., Pasquariello G. et al., 1991. An experiment for the interpretation of multitemporal remotely sensed images based on a fuzzy logic approach. International Journal of Remote Sensing, 12:1431-1445.

Cess, R. D., Briegleb, B. P., Lian, M.S., 1982: Low-latitude cloudiness and climate feedback: Comparative estimates from satellite data. Jou. Athmos. Sci. 39:53-59.

Ebert, E.E., 1989. Analysis of Polar Clouds from Satellite Imagery Using Pattern Recognition and a Statistical Cloud Analysis Scheme. Journal of Applied Meteorology, 28(5):382-399.

Franklin, S.E., Peddle D.R., 1989. Spectral Texture for improved class discrimination. International Journal of Remote Sensing, 10:1437-1444.

Heeb, M., 1985. Bewoelkungskartierung ueber der Schweiz mit Wettersatellitenbildern und Bodenbeobachtungen. Geographica Helvetica, 40:133-141.

Hughes, N. A., 1984. Global Cloud Climatologies: A Historical Review. Journal of Climate and Applied Meteorology, 23:724-751.

Hunt, G.E., 1973. Radiative Properties of terrestrial clouds at visible and infrared thermal window wavelengths. Quarterly Journal of the Royal Meteorological Society, 99:346-369.

Hwang, P.H., Stowe, L.L., Yeh, H.Y.H., Kyle, H.L., 1988: The Nimbus-7 Global Cloud Climatology. Bulletin of the American Meteorological Society, 69(7):743-752.

Khazenie, N., Richardson, K.A., 1991. Classification of Cloud Types based on spatial textural Measures using NOAA-AVHRR Data. In: Proceedings of the IEEE. Helsinki, 1701-1704.

Raustein, E., 1989. Use of a Clustering Method for Objective Cloud Classification and Determination of Cloud Parameters from Satellite Data. Meteorological Report Series University of Bergen 2-1989.

Rossow, W.B., 1991: ISCCP Cloud Data Products. Bulletin of the American Meteorological Society, 72(1):2-20.