Schweizerische Meteorologische Anstalt
Institut suisse de météorologie
Istituto svizzero di meteorologia
Swiss Meteorological Institute

No. 178

AMETIS II - MANTRA

Working Progress Report 1993

Daniel Cattani, Dagmar Engfer, Jacques Ambühl, SMI
Edoardo Amaldi, Frédéric Aviolat, Eddy Mayoraz, Vincent Robert, EPFL

June 1994

Arbeitsberichte der SMA
Rapports de travail de l'ISM
Rapporti di lavoro dell'ISM
Working Reports of the SMI

Schweizerische Meteorologische Anstalt
Krähbühlstrasse 58, Postfach
CH-8044 Zürich

Tel. (01) 256 91 11, Fax (01) 256 92 78, Telex 81 73 73 met ch

Schweizerische Meteorologische Anstalt
Institut suisse de météorologie
Istituto svizzero di meteorologia
Swiss Meteorological Institute

No. 178

AMETIS II - MANTRA

Working Progress Report 1993

Daniel Cattani, Dagmar Engfer, Jacques Ambühl, SMI
Edoardo Amaldi, Frédéric Aviolat, Eddy Mayoraz, Vincent Robert, EPFL

June 1994

Arbeitsberichte der SMA
Rapports de travail de l'ISM
Rapporti di lavoro dell'ISM
Working Reports of the SMI

Schweizerische Meteorologische Anstalt
Krähbühlstrasse 58, Postfach
CH-8044 Zürich

Tel. (01) 256 91 11, Fax (01) 256 92 78, Telex 81 73 73 met ch

AMETIS II - MANTRA
WORKING PROGRESS REPORT 1993

by

Daniel Cattani, Dagmar Engfer and Jacques Ambühl (SMI)
&
Edoardo Amaldi, Frédéric Aviolat, Eddy Mayoraz, Vincent Robert (EPFL)

June 1994

Résumé

Ce rapport intermédiaire du projet AMETIS II présente les travaux effectués en 1993 ainsi que les premiers résultats obtenus en collaboration avec l'équipe MANTRA de l'Ecole Polytechnique Fédérale de Lausanne (EPFL).

Après une introduction rappelant les objectifs de ce projet, sont décrits les nouveaux instruments installés sur le site de Kloten pour compléter le réseau de mesures. La démarche et les techniques utilisées dans le cadre de ce projet par l'EPFL sont ensuite discutées avec l'appui des premiers résultats.

Zusammenfassung

Dieser Zwischenbericht zum Projekt AMETIS II stellt die bisherigen Arbeiten des Jahres 1993 sowie erste Resultate vor, die in Zusammenarbeit mit der Eidgenössischen Technischen Hochschule Lausanne erhalten wurden.

In der Einführung werden kurz die Hauptziele des Projektes vorgestellt. Als nächstes sind die neuen Instrumente beschrieben, die in Kloten zur Vervollständigung des Messnetzes installiert wurden. Schließlich werden das Vorgehen und die in diesem Projekt von der ETH Lausanne verwendeten Methoden mit Hilfe der ersten Resultate erläutert.

Riassunto

Si presenta in questo rapporto intermediario del progetto AMETIS II, i lavori effettuati durante l'anno 1993, ed i primi risultati della collaborazione con il gruppo MANTRA del EPFL (l'Ecole Polytechnique Fédérale de Lausanne).

Nel introduzione si descrivono gli obiettivi del progetto, per poi presentare i nuovi apparecchi introdotti a Kloten per completare il campo di misure. L'approccio scientifico e le differenti tecniche utilizzate sono discusse ed illustrati da qualche risultati.

Summary

This intermediary report of the AMETIS II project, presents the work done during 1993 and the preliminary results obtained in collaboration with the MANTRA team at the EPFL (l'Ecole Polytechnique Fédérale de Lausanne).

First a short introduction defines the aims of the project, the new instruments installed at Kloten in order to complete the measurement network are then described. The scientific approach and the technics used are also discussed and illustrated by the first results.

AMETIS II - MANTRA
WORKING PROGRESS REPORT 1993

Table of contents

Résumé / Zusammenfassung / Riassunto / Summary

Table of contents

1 Introduction

Human observers perform aeronautical meteorological observations at the Swiss airports in Geneva, Zurich, Altenrhein, Grenchen, Bern, Lugano and Sion. These observations are made every thirty minutes round the clock at Geneva and Zurich. They are broadcasted to the International Aeronautical Telecommunication Network as METAR Telegrams, elaborated following the International Civil Aviation Organisation (ICAO) standards.

The aim of the AMETIS II project is to design, realize and put in operation a system providing in the night hours, at Zurich and then Geneva and regional airports, fully automatically generated METARs.

The coding rules of the METAR have been amended in a session of the ICAO held in 1990, in order to allow such an automated generation. Member States ratified the amendment and the new coding rules have been put in operation on the first of July 1993.

The structure of such a new METAR is given here below:

LSGG 0640Z 23012G25kt 1200S 6000NE Br Sct015 Bkn050 05/M00 Q1012

Valid for Geneva airport at 6 h 40 GMT. Wind is blowing from 230° at 12 knots with gusts up to 25 knots. Minimum visibility of 1200 m. occurs in the south direction, maximum visibility of 6000 m. in the north-east. Mist is present. A first scattered layer of clouds in present at 1500 ft, a second broken layer is present at 5000 ft. Temperature is 5 degrees Celsius, dew point is just below freezing, and the see level reduced pressure is of 1012 hectoPascals.

New technologies play a crucial role in the project. Physical sensors providing global radiative balance of the sky, as well as devices automatically generating present weather reports, will be introduced. Neural networks and expert systems will assume the interpretation and management of the large amount of sophisticated information emanating from the sensors.

2 Organisation

The success of the first phase of the AMETIS II project may be considered as a result of a good collaboration between three departments of the Swiss Meteorological Institute (SMI) and the MANTRA team at the Ecole Polytechnique Fédérale de Lausanne (EPFL).

2.1 Swiss Meteorological Institute

The three following departments of the Swiss Meteorological Institute are involved. The *FlugWetterZentrale* (FWZ), at Zurich airport. Project main management is assumed by Mr. Hack, head of the FWZ. Miss Engfer, meteorologist, works as a first assistant. The main set of trial instruments, including, pyrgeometers, ceilometers, forward scattometers, present weather sensors, is installed at Kloten (Zurich airport).

The *Station Aérologique de Payerne* (SAP), manages all the problems related to the meteorological observation: atmospheric and instrumental physics, raw data computer handling, purchases, construction, calibration, operational duties. At the SAP, Dr. B. Högger and Mr. G. Converso are associated to AMETIS II project.

The *Centre Météorologique de Genève* (CMG), assumes the coordination between the Mantra group at EPFL and the Swiss Meteorological Institute (SMI). The CMG is in charge of elaborating the learning tasks presented to the neural networks, including the design of the pre-processing algorithms involved. This work was attributed to Dr. D.Cattani and J.Ambühl in 1993.

2.2 Mantra team

Mantra is an acronym for Machine Neuronale en TRAnches de Réseaux Autonomes. Mantra is funded by the Swiss Research Program SPP-IF 1992-1995 (Schwerpunktprogram Informatik Forschung), oriented towards massively parallel computing systems. Five laboratories of the EPFL are involved in the Departments of Electricity, Computing Sciences and Mathematics. Ametis II project is included as an application sub-project in the main Mantra stream. Led by Prof. Hasler, four graduate engineers worked for Ametis II in 1993, Mr. E.Amaldi, F.Aviolat, E.Mayoraz and V.Robert.

3 New Instruments

The measurement equipment at Kloten airport was completed by three new instruments, installed during summer 1993. A present weather sensor (LEDWI), a forward scattometer (FD12P) and pyrgeometers (EPPLEY).

3.1 The present weather LEDWI (ScTI, Model OWI-240)

Ledwi means Light Emitting Diode Weather Identifier and is the name of an instrument from ScTI for measuring the precipitation type, amount and intensity. The LEDWI is a kind of a Present Weather Identifier.

Functional Description:

This instrument is able to identify the present weather with the parameters precipitation state (yes/no) and precipitation type (rain/snow).
The LEDWI is equipped by a transmitter optical assembly with an infrared emitting diode and a receiver optical assembly that receives the scintillated light (figure 1).

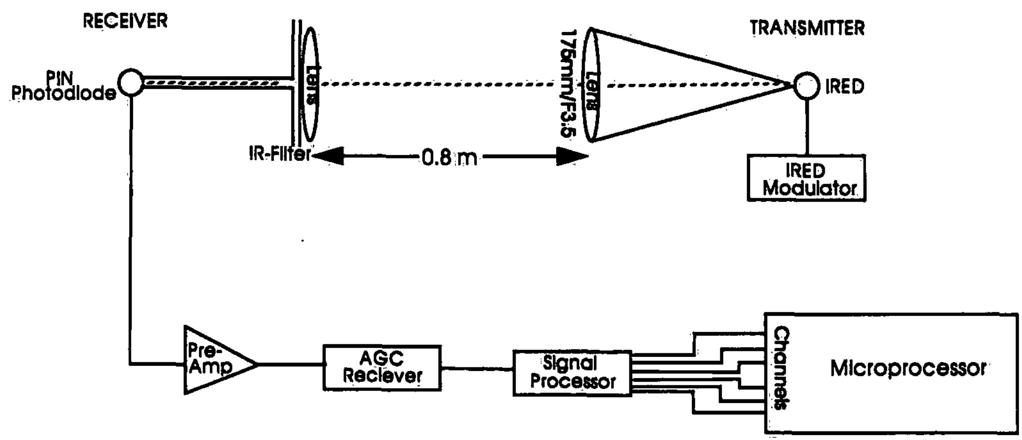

Figure 1 Functional diagram of the LEDWI.

The functionality of the LEDWI is based on light scintillation induced by weather particles. The detected temporal frequency spectrum varies according to the size and the velocity of falling precipitant (rain induced scintillation gives frequencies above 1KHz; snow induced scintillation gives frequencies below some hundreds of Hertz). To exclude the background light contamination (i.e. sunshine, runway lights) the IRED (=InfraRed Emitting Diode) is driven by a modulator.

The transmitter optic collects the IR-light from the IRED to a partially coherent beam of about 50 mm diameter. This light beam is sent to the receiver optics installed about 0,8 m away. Particles falling through this beam induce a light scintillation that can be measured from the receiver optic. The receiver optic detects the scintillated light with a horizontal line aperture to measure only the vertical precipitant with a PIN photodiode. The detected light is sent from the PIN photodiode to the pre-amplifier and to the AGC (=Automatic gain controlled) receiver. The signals from the different channels get from the signal processor to the microprocessor from where they can be transformed to a readable telegram (figure 2).

To distinguish rain/snow and the intensity the LEDWI uses two different frequency bands and a particle counting channel:

high band: 1 - 4 KHz
low band: 25 - 250 HZ
particle channel

With these three channels, it is possible to detect the precipitation state and with the ratio from high to low band the precipitation type can be assessed.

The algorithms used for the detection can be resumed as follows:

1. The one minute average for the 3 channels is taken as raw data value
2. When the particle channel reaches 50 counts the precipitation is set on TRUE
3. Precipitation is only set on TRUE if 50 counts are measured for the two 30 sec segments in the 1min average
4. If it is set on TRUE the signals from the high and low channels and the ratio of the to channels are measured.
5. The signal strength of the high channel is taken for intensity and accumulation of rain
6. The signal strength of the low channel is taken for intensity of snow
7. The ratio from high to low channel is taken for the water content of rain and snow.

LEDWI: WR-P1.17S000000X437100L020017K308127H172075
LEDWI: W?3P0.00S0000EAX436100L003012K189127H099075
LEDWI: WS-P0.02S0000FAX441001L026000K351104H07507

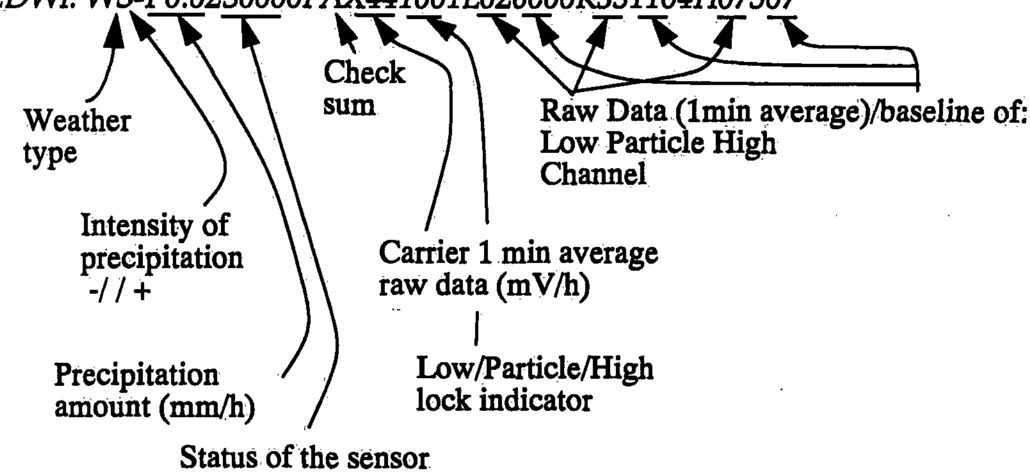

Figure 2 Examples of telegrams sent from the sensor LEDWI.

3.2 The Forward scattometer FD12P (Vaisala):

The FD12P is a forward scatter optical sensor. It measures visibility, precipitation type and intensity; precipitation type is related to the weather code used in WMO standard message SYNOP and, as additional information, is as well given in the NWS (US National Weather service) abbreviation code.

Functional Description:

The FD12P uses the effect of light scattering caused by precipitation particles to measure the precipitation state, type and intensity. The light scattering from a precipitation particle is proportional to its volume.

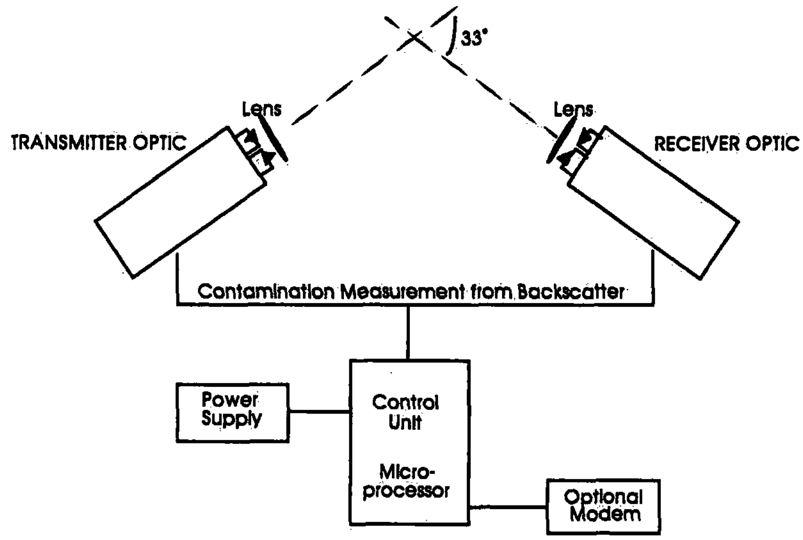

Figure 3 Functional diagram of the FD12P.

The instrument FD12P is equipped with a transmitter and a receiver card (optical assemblies), a control board and an interconnecting cable. The transmitter card consists of an infrared light emitting diode (LED) and an optical lens (with a diameter of 71mm). The transmitter generates infrared light pulses (figure 3).
The light pulses are sent at a frequency of 2,3kHz. The wavelength of the infrared light is at about 875 nm. The scattered light is detected by a PIN photodiode in the receiver. The receiver is equipped with a phase sensitive lock-in amplifier. The light signal that is generated by the PIN photodiode in the receiver is converted to a frequency and analysed by the microprocessor in the control board. This board controls also the instrument operations.

The transmitter and receiver optics are aligned at an angle of 33° and as well tilted 20° downwards.

A backscattering system controls the optical system. It measures the contamination and aging of the optical assembly. For the backscattering the FD12P consists of a second PIN photodiode in the transmitter optics, to monitor the transmitted light intensity, and a second LED in the receiver optics.

As an additional element the FD12P consists of a DRD12 rain detector to obtain better results in the measurement of precipitation intensity (mm/h) and the water content of the precipitation. There is also integrated a DSR13 ambient temperature sensor that is needed for the precipitation type.

The calculation of the values from the light signals is, in short words, as follows:
The FD12P operates with a small sample volume of about 0,1 l. to obtain the information related to precipitation, in order to detect each independent particle.

1. The converted signals are sorted by frequencies, and frequencies changes distribution is determined in order to obtain a signal profile.
2. When the signal passes the precipitation limit within 4 minutes, the precipitation state is set on TRUE.
3. The precipitation is calculated with an algorithm able to define which part of the signal profile is decisive for the calculation of the average signal.
4. The intensity is calculated with a rain intensity scale.
5. The precipitation type is calculated with the difference in the optical intensity and in the DRD12 intensity as key factor. The temperature is also considered in this calculation for the precipitation type, so that at a temperature above 8°C the precipitation type is rain, below -5°C snow and between -5° and 8°C unknown. For each other specific precipitation type there are different and more detailed criteria.
6. For the visibility the converted signals can be classified as follows:
   signal of 0.1Hz corresponds to a visibility of 20 km
   signal of 1 Hz corresponds to a visibility of 4500 m
   signal of 10 Hz corresponds to a visibility of 800 m
   signal of 100 Hz corresponds to a visibility of 150 m

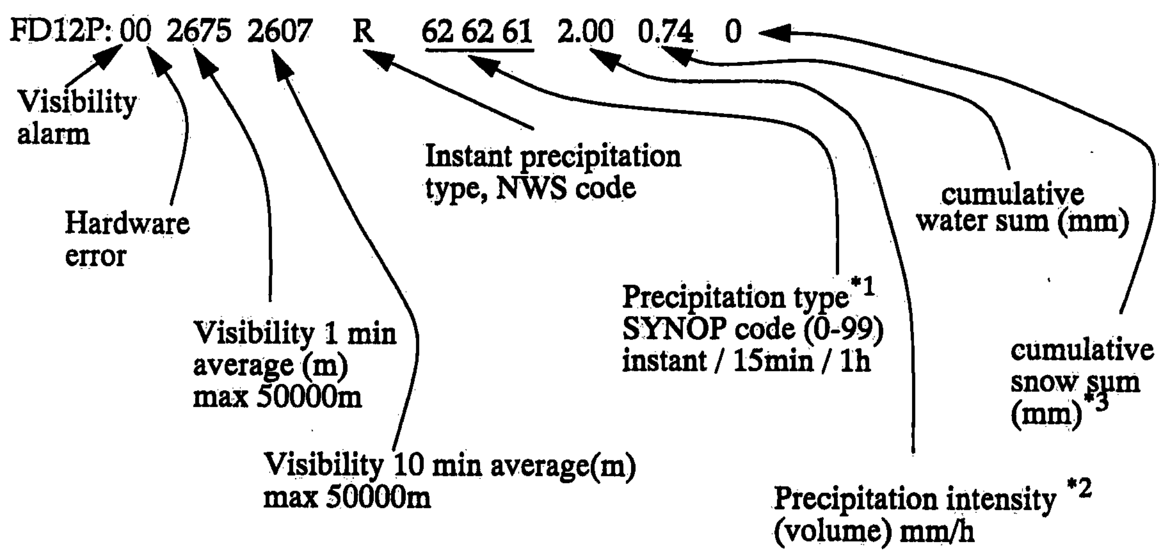

Figure 4 Example for a telegram sent from the FD12P.

*1 precipitation type:
instant = the precipitation type, which was measured most frequently during the last 15 sec.
15min = the precipitation type, which was measured most frequently during the last 15 min.
1h = the precipitation type, which was measured most frequently during the last 1h.
*2 precipitation intensity = highest intensity during the last 15 sec.
*3 cumulative water sum = cumulative sum until the value 999 is reached, then it is reset to 0.

3.3 The pyrgeometer (EPPLEY: PYR)

The pyrgeometer from EPPLEY (model PYR) is introduced to help the estimation of the cloudiness during the night time (see [Ambrosetti 91]).

The radiation on the earth may be divided in two parts. The *short-range solar* radiation which belongs to the wavelength range of 0.3 to 3.0 µm. The *long-wave terrestrial* radiation which comprised of the incoming atmospheric component (i.e. downward emission by gases of the atmosphere; water vapour and carbon dioxide especially) and the outgoing terrestrial component (i.e. upward emission and reflection by natural surfaces and atmospheric gases). The wavelength range of these components is 3 to 40 µm. The terrestrial radiation which reaches the ground from clear sky above, is mainly due to emission by atmospheric gases. Within these conditions around 50% of the radiation comes from the first 30m of the atmosphere above the ground. The presence of relatively dense liquid water clouds in the atmosphere act as blackbodies, and thus strongly modify the radiation balance [Coulson 75].

The *pyrgeometers* are long-wave radiometers, they inherently measure the exchange of radiation between a horizontal blackened surface (the detector) and the target viewed (sky, ground, ...).

The EPPLEY PIR pyrgeometer is designed to have a detector automatically compensated, allowing isolation of the radiation from the target impinging on the detector [Eppley]. A thermistor-battery-resistance is incorporated to precisely compensate the detector temperature.

The detector (a thermopile) is isolated from solar short-wave flux by a hemisphere of silicon on which an interference filter was deposit, as a result the detection wavelength window is of the range of 3 to 50 µm. A crucial system of ventilation plus heating is associated to the detector when used outside (figure 5).

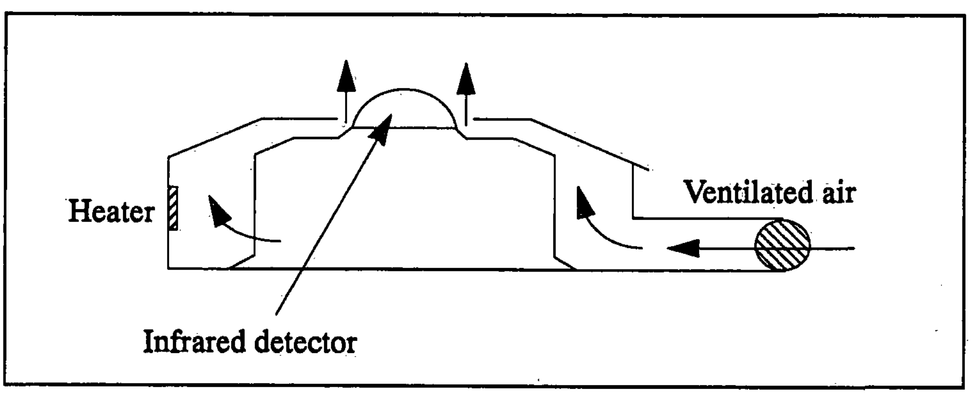

Figure 5 Schematic view of a pyrgeometer.

Since august 1993 two pyrgeometers are installed in Kloten, the distance between the two instruments is about 2 meters.

The message emitted every minute from the pyrgeometers is presented on figure 6.

```
93-12-21 00:02:28
PYRGEOMETER: -.32601,-.31756,-.33343,.00328,160.98, 161.24,-.32424,
-.31576,-.33015, .00324,161.11, 161.83,258.97, .16819, 164.82, 184.59,
15.195, 13.806, 0
```

Figure 6 Typical message sent by the 2 pyrgeometers, containing the thermo-tension values, their standard deviations, and temperatures of the dome and the case for each pyrgeometer, together with data concerning the ventilator.

According to several references reviewed in [Ambrosetti 91], it is possible to estimate of the amount of the clouds or the type of the clouds present in the sky by coupling the infrared emission values with a set of atmospheric parameters. Empirical relations are established in order to reduce the influence of the first layers of the atmosphere on the measured radiation, and in this way enhance the impact of the clouds on the radiation balance.

But the results are not valuable enough, and further problems may occur due to seasonal or diurnal variations of weather conditions.

4 Data acquisition

A personal computer is used to collect all the messages emitted by the different instruments. They are all interconnected using modem lines. A sequence of measurement is recorded every minute, it contains: the present weather sensor, the scattometer, the 2 pyrgeomenters and data obtained by a ceilometer (heights of the base of clouds). Typical sequence is presented in figure 7.

```
93-12-21 00:02:28
PYRGEOMETER: -.32601,-.31756,-.33343,.00328,160.98, 161.24,-.32424, -
.31576,-.33015, .00324,161.11, 161.83,258.97, .16819, 164.82, 184.59,
15.195, 13.806, 0

LEDWI: W P0000S0000BAX431111L-06-07K111107H073071

FD12P: 00 50000 50000 C 0 0 0 0.00 48.37 5

CEILOMETER:    X0CCF    ////    ////    ////    2EE0////A9FF18/////
001R=============================================================
=================================>>=======//////////A7
```

Figure 7 Sequence of telegrams recorded every minute.

The files containing the data of the new instruments are combined together with data from the main database of the SMI in order to build *tasks* required by the data treatment. Each task is on a matrix format (example on figure 8), it is based on the comparison on past or recent weather measurements - and manual weather observations in METAR or SYNOP form.

```
{0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0
 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0
 0 1 0 1 1 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0
 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0
 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0        ceilometer data
 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0
 0 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0
 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0
 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0        2 pyrgeometers
 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0}
{20 357.4099 357.1933 357.2037 357.5137 357.5779 357.2597 357.9704 358.5811 358.4460 358.8014 }
{30 359.0022 359.1564 359.8435 360.3273 360.1499 360.2625 360.3328 359.6680 360.0726 360.4842 }
{40 359.5869 359.2238 359.5030 359.5645 359.9812 360.1114 359.4604 359.2638 359.1554 359.4770 }
{20 352.5200 352.7571 352.4254 352.6860 352.5537 352.4351 352.7304 353.1609 353.1956 353.9079 }
{30 354.3209 354.5692 354.9234 355.1956 355.1781 355.0865 355.0621 354.6900 354.7075 355.5894 }
{40 355.1242 354.2098 354.2860 353.9603 354.7489 354.7907 354.0245 354.1503 354.2272 354.1118 }
{20 118 115 947 9616 0 131 110}
{30 119 113 901 9616 1 125 103}        automatic observations
{40 117 113 904 9616 1 124 102}
{{ 1993 8 27 20 40 4200 58 2 400 5 2000 8 3300 }}        synoptic observation
```

Figure 8 Example of the data matrix corresponding to the 30 minutes preceding the synoptic observation. Four parts are present; the ceilometer data, the 2 pyrgeometer measurements, the automatic observations and the synoptic observation including: date, time of the observation, visibility, type of weather and reported clouds.

The past weather data are recovered from KLAZY (SMI database), either from the one-hour database or the 10 minutes database (figure 9). The SYNOP observations are generally used for learning purposes: their description of cloud amounts (9 classes) is more precise than the one of the METAR observations (4 classes).

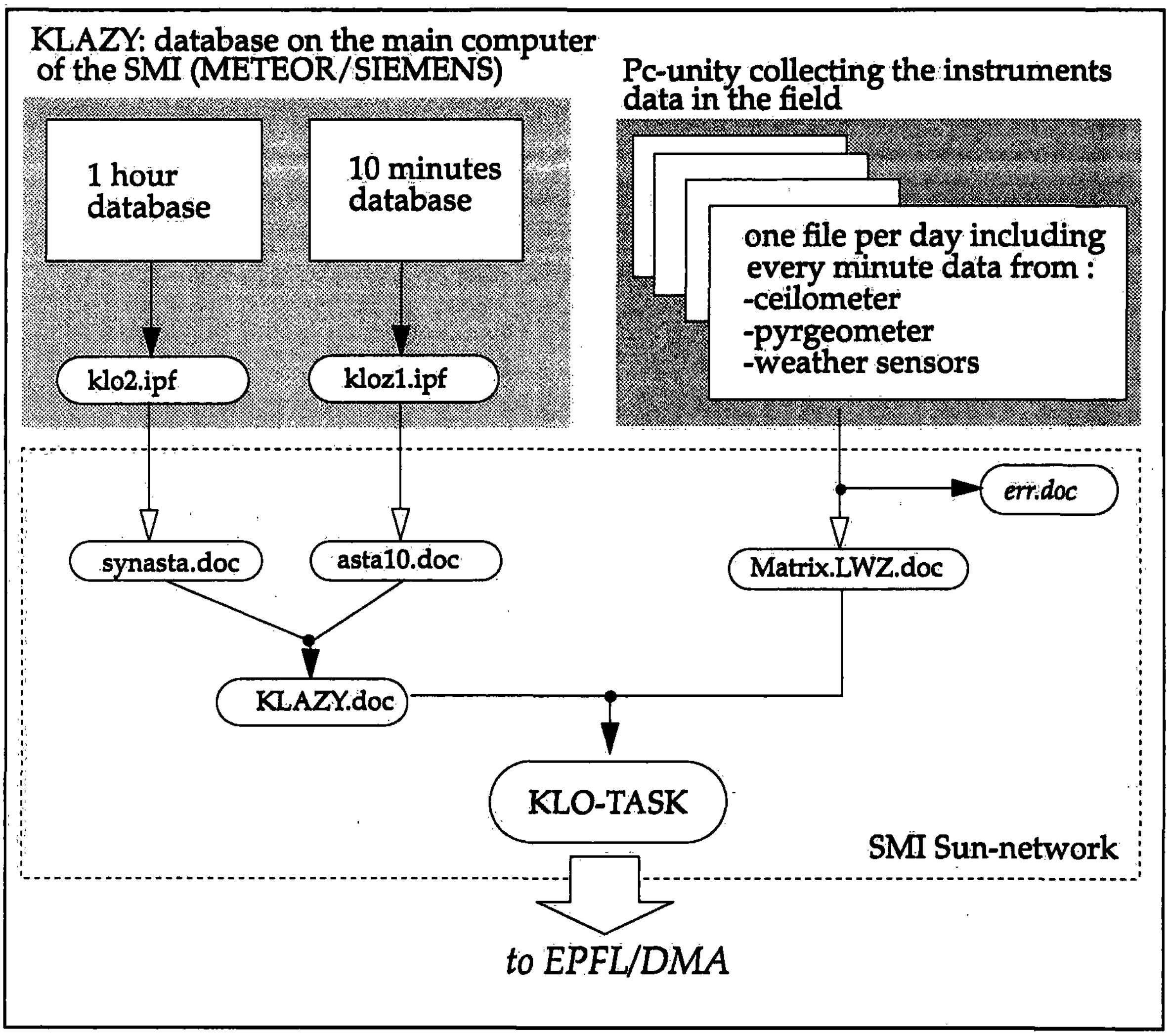

Figure 9 Software diagram route used to construct a complete task.

The amount of a cloud layer, when reported in 9 classes, refers to 0 (clear sky) up to 8 eighths of sky covered (sky completely covered), - in 4 classes (new METAR code) SKC sky clear, SCT scattered (1 to 4 eighths of sky covered), BKN broken (5 to 7 eighths of sky covered) and OVC overcast sky completely covered.

5 Artificial Neural Networks for Classification

Data treatment is based on Artificial Neural Networks used in this application for classification. As presented by [Ambühl 93] neural network is an interesting tool to help the analyse of an important flow of meteorological data.
In the project AMETIS II, artificial neural networks are used to establish a "meteorological context" from which the characteristics of the cloud coverage may be deduced. All the inputs of the network are automatically measured parameters such as temperatures, precipitation amount, i.r. radiation, wind, ...
No physical pre-treatment or empirical rules are applied at this stage of the investigations.

5.1 Artificial Neural Networks

A neural network is characterized by its architecture, including its number of units and their pattern of connection, the kind of computation performed by each unit and the learning method used to adjust its parameters. We will concentrate on multilayer perceptrons, as they are best suited to classification problems (figure 10).

A feedforward multilayer neural network is composed of n input units, followed by one or several layers of processing units; the last layer contains m output units. There is no connection between units on the same layer and each unit receives its inputs only from the preceding layers. Input units don't participate in the calculations; they are only present as an interface. Units of the intermediate layers are called "hidden units" as they are neither input nor output units.

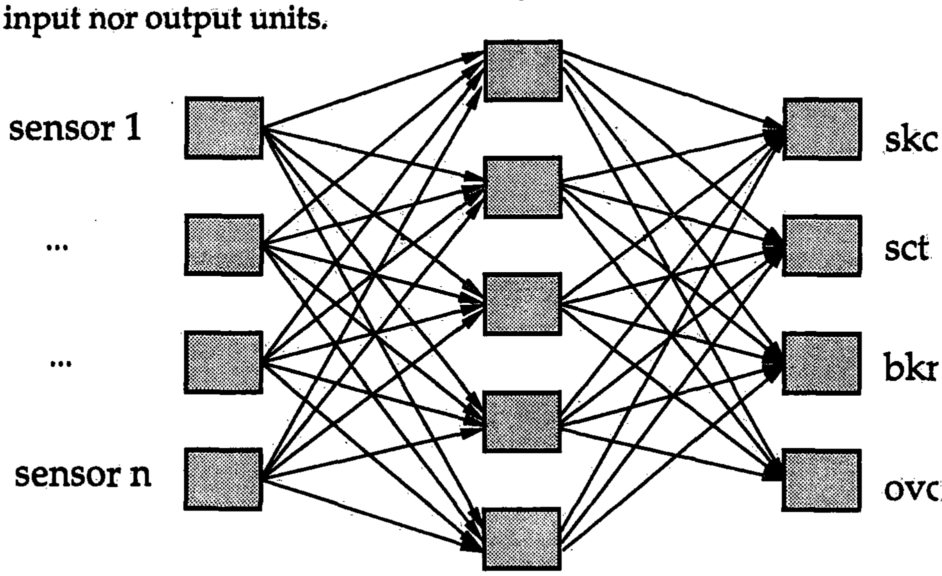

Figure 10 Schematic representation of a multilayer neural network with three layers and four output units.

Each processing unit (*neuron*) realizes a simple computation:

$$v = tgh\left(\sum_{j=1}^{n} w_j \cdot x_j\right)$$

It performs the weighted sum of its inputs x_j and evaluates its output $\sigma(\nu)$ by applying a non-linear function σ. This kind of formal neuron is historically called a *perceptron* (figure 11). In our case, σ will be the hyperbolic tangent function, which is well suited to continuous input and output values.

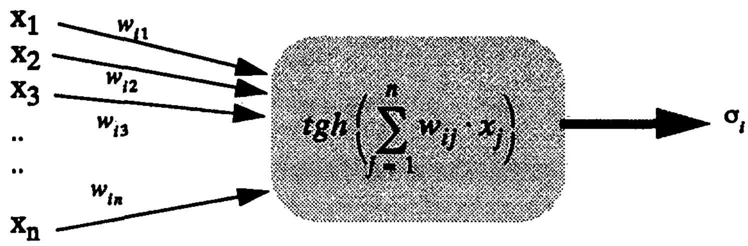

Figure 11 Schematic representation of a formal neuron.

When presenting a input vector to the network, by assigning each component to the corresponding input unit, the processing units evaluate their outputs layer by layer, from the first hidden layer to the output layer. The activations of the output units are the components of the output vector produced by the network. Thus, the network performs an application from R^n to R^m, which depends on the weight parameters w_{ij}.

The goal of *supervised learning* is to memorize associations between *p* input vectors $\underline{a}^k$ and *p* output vectors $\underline{b}^k$. These *p* associations $\{(\underline{a}^1,\underline{b}^1), \ldots ,(\underline{a}^k,\underline{b}^k), \ldots ,(\underline{a}^p,\underline{b}^p)\}$ form the "*task*" that must be realized by the network. Given a task and a network, learning consists in determining values for the parameters w_{ij} of the network, such that it gives the desired output $\underline{b}^k$ for each input vectors $\underline{a}^k$.
Usually, learning is formulated as an optimization problem where the goal is to minimize an error function, measuring the performances of the system for the considered task. After learning is complete, it is interesting to look at the quality of the system's responses for input vectors that don't belong to the task. This amounts to evaluate the "generalization" properties of the system.

5.2 Algorithmic Developments

Training a feedforward multilayer neural network with continuous activations is classically performed by the backpropagation algorithm [Rumelhart 86]. This amounts to minimize the sum of the quadratic errors over the associations of the task using a simple gradient descent technique. Unfortunately this algorithm has many drawbacks: it has a very slow convergence rate and it gets often trapped in poor local minima of the error function. In the extensively used stochastic backpropagation, one doesn't even consider the total gradient of the error function.

In order to improve the training of the parameters of the network, more sophisticated optimization methods can be adopted. A first improvement is to combine several descent directions, as in the conjugate gradient algorithm [Minoux 91]. However, instead of using only the first order derivatives of the function to optimize, its second derivatives are of interest. This kind of second order methods are referred to as Newton methods. Unfortunately, the resulting Hessian matrix is very time consuming to compute.
An approximation of this matrix is used in general, giving rise to the quasi-newton methods. Broyden, Fletcher, Goldfarb and Shanno developed such a method, referred to as BFGS, for which they proved robustness and good performances. In the framework of neural networks there is usually a large number of variables; therefore the matrix approximating the Hessian can be quite large. In order to avoid manipulating such large matrices, a variant of the BFGS algorithm was developed, requiring the use of vectors only. This method is called *Memoryless* BFGS, because it doesn't store any matrix. It has been shown to be as robust as the matrix based BFGS, thus also allowing for inexact unidimensional search, and leading to results that are almost as good [Shanno 78].

6 Tasks and Results

Four different tasks were established during 1993. The purpose was to approach step by step the final complexity of the so-called definitive task at Kloten while all the new instruments were neither fixed nor connected.
The selection of the crucial parameters, the formatting and the normalisation of the data have to be selected. The desired output is based on the manual synoptic observations where clouds are reported. Data was divided into 2 sets of equal size: one used for the learning part and the other one to evaluate the network performances. The goal of the learning phase is to determine a set of weights that provides the best performances on the training set. Evaluation of the results on the testing set was performed by computing the distance between the class given by the network and the desired class, either with the original 9 classes or these reduced to the 4 classes necessary for the METAR messages.

6.1 Task Locarno

As a first goal, we have chosen to evaluate the total cloud amount from the meteorological background given by the automatic station and the pyrgeometer, assuming that the base height of the main cloud layer is known. The data were collected at Locarno meteorological site, where the first pyrgeometer was installed.

6.1.1 Preprocessing of Data

First we had to find a good preprocessing of the data. It is an important step so that the network can learn the task well. It evolved through the several tasks we considered. An appropriate normalization of the input data is performed by centering it on the mean and reducing it by the variance. A more sophisticated treatment has been adopted for some of the inputs. The extensive simulations permitted to select the most relevant set of inputs.

6.1.2 BFGS (Broyden, Fletcher, Goldfarb, Shanno) Algorithm

The first retained architecture has one hidden layer with a few sigmoidal (hyperbolic tangent shaped) units and one sigmoidal output unit. The discrete output values (the classes) were transformed into continuous ranges of equal length. The objective of the learning phase is to minimize the quadratic error over the training set. Learning was performed by a *Memoryless* BFGS algorithm with inexact one-dimensional optimization search. A thorough study was carried out to determine the most appropriate size of the network. Simulations showed that few hidden units suffice to obtain very good results on the testing set.

6.1.3 Trial: Weigend penalty

In order to improve the performances of the networks over the testing set and generalization ability of the network, several improvements were investigated. To avoid an overfitting of the data, a penalty term similar to that proposed by [Weigend 91] was added to the error function. Such a modification did not lead to significant improvement while making the training process more complex.

6.1.4 Added: Validation

In order to devise an optimal stopping criterion for the optimization procedure, we extracted from the training set a subset of data to be used as a "validation set". The learning phase is stopped as soon as the error on the validation set increases sensitively. This way to limit overfitting led to a significant improvement in the generalization performances of the networks. Moreover, results became more robust with respect to the number of hidden units.

6.2 Task Payerne

This is the first step towards serial data. The goal of this task is similar to task Locarno and consists of evaluating the cloud amount of the main cloud layer. The input vector now contains measured fluctuations on the pyrgeometer values. In the data set from the Payerne station, the pyrgeometer values are measured every 10 min., and all the measurements of the current hour are employed.

6.2.1 Importance of Pyrgeometer

In contrast to task Locarno, results were improved here by including a better contextual environment for this important input. Each bar chart of Figure 12 shows the number of correctly classified examples at the bottom in black, then above the number of examples with output at a distance 1 from the desired value, and so on. The first chart was obtained without pyrgeometer measures, the second one with the current pyrgeometer measure and the last one with all 5 measurements of the current hour. The tests made without the pyrgeometer input (left bar on figure 12) were not as good as the two others, showing that this measure is essential in order to determine the level of cloudiness and thus justifying the need for such a instrument.

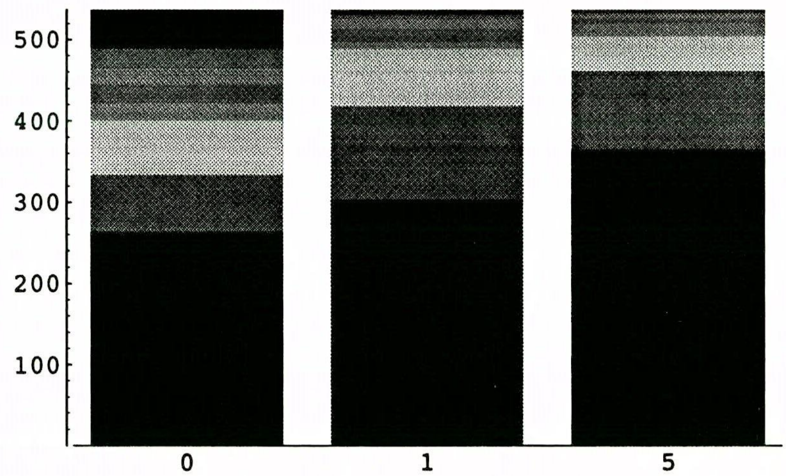

Figure 12 Importance of the pyrgeometer values, (0) no values, (1) actual measures and (5) last hour values. Vertically the number of associations, each color groups the associations with the same distance to the observation. The lowest bar indicates the number of correct associations.

6.2.2 Variation: Learning with 4 classes

Experiments were carried out to have the network learn the task with the classes reduced to 4, as needed for the generation of METAR messages. The quality of the results slightly diminished, but not significantly.

6.2.3 Trial: Binary Outputs

Several additional trials where performed. One of them was to use 9 (or 4) binary output units, one for each cloudiness class. The results obtained with this representation of the classes were clearly not as good as the first ones.

6.2.4 Trial: Radial Basis Functions Networks

Another variation was to use a completely different kind of neural networks, composed of radial basis functions, instead of sigmoidal units. Training was performed by an enhanced version of an algorithm due to [Platt 91]. Although very good results were obtained in other time series applications, this different type of networks did not perform better than the classical ones on the task considered so far. Nevertheless we are planning to experiment them on the new task, which will contain more temporal data.

6.2.5 "Cleaning" of Task

The data used for this trial was thoroughly studied. In particular, all very badly classified cases were identified and all easily recognizable cases were removed. The tests described below (6.2.6, 6.2.7, 6.2.8) were performed on this restricted task.

6.2.6 Balance of Classes

One of the important problems is that the 9 classes are not balanced. For instance, sky clear (0/8) and of completely clouded sky (8/8) examples are more frequent than the others. In order to improve the learning of the small classes, training was performed with extended sets of examples. The first idea was simply to replicate examples of the small classes. The second method generated new examples of the small classes, from the available ones, by slightly perturbing the data. This improved the results, but not very significantly.

6.2.7 One network per class

As mentioned before, we also considered networks with binary output units. Here instead of using one network with 9 (or 4) outputs, we used 9 (or 4) independent networks with one output unit. Each network was trained to discriminate one class against all others. The output of the system is the number of the network giving the highest response. This led to substantial improvement. The obtained results were significantly better than those provided by single networks. As before, small networks were sufficient to perform the task correctly.

6.2.8 Added input data: height of the clouds

Preparing for the definitive task (Kloten) with ceilometer measures, we added the logarithm of the height of clouds as an input. Experiments were carried out both with a single network and with several networks. The results improved significantly. Figure 13 shows a comparison between the use of 1 and 4 networks to determine the cloudiness of the main level, using the height of this level. All networks have 8 hidden neurons and were trained on 4 classes.

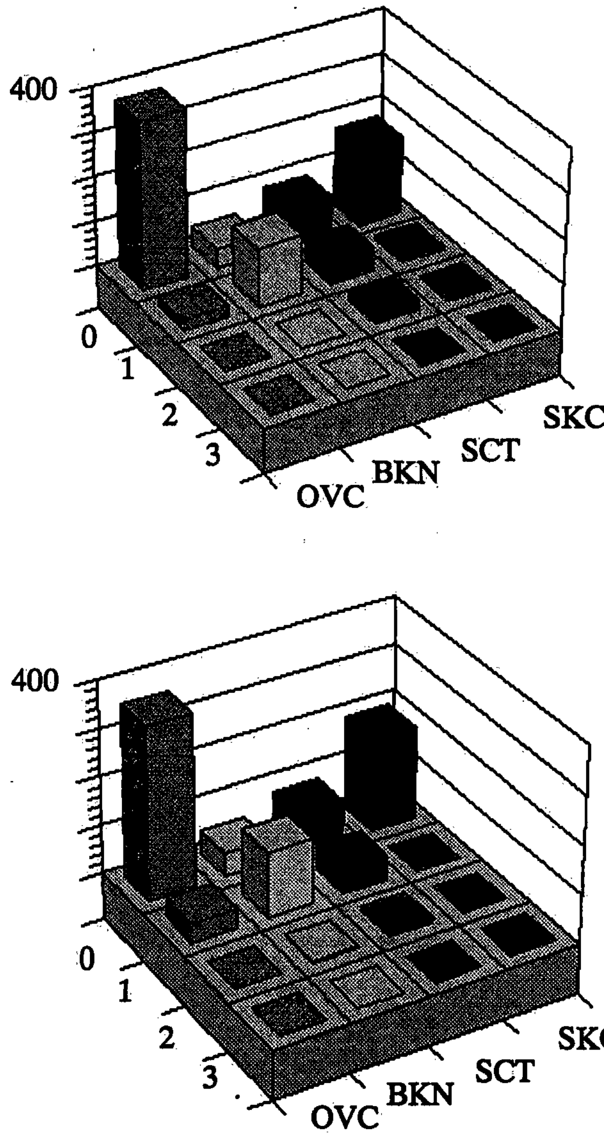

Figure 13 Results of the application of 1 (top) and 4 (bottom) networks. The 3D charts present the numbers of associations with a distance of 0 to 3 from the correct cloud amount of clouds classified in 4 classes. The two distributions look very similar: Improvement appears with respect of the distribution presented on figure 14.

6.2.9 Linear Programming

A different type of algorithm for designing multilayer networks of threshold units is currently under investigation. This constructive technique, proposed by [Mangasarian 93] and based on linear programming, is well suited to classification problems. An interesting feature of this approach is that the network architecture must not be selected *a priori* and it can be easily extended to multiple classes problems. The preliminary tests are promising.

6.2.10 Empirical Formula

Figures 14 presents the application of an empirical treatment in order to deduce the cloud amount from pyrgeometer measurements. The height of the base of the main clouds is assumed to be known.
Both neural networks show a clear improvement with respect to the application of an empirical formula in solving this specific task. Especially in the distinction of the first classes of cloudiness SCT and SKC.

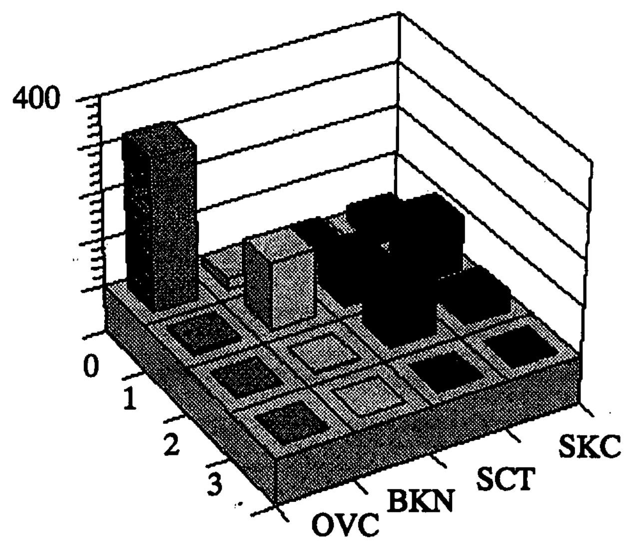

Figure 14 Results of the application of an empirical treatment. The 3D chart presents the numbers of associations with a distance of 0 to 3 from the correct cloud amount of clouds classified in 4 classes.

6.2.11 Summary of sections 6.1 and 6.2

The main results obtained so far are:

- well suited preprocessing of the data
- exclusion of irrelevant parameters such as wind intensity and direction,
- systematic use of temperature differences as input parameters, for example 2m temperature minus the 5cm temperature,
- cleaning of the tasks by removal from the learning task of "trivial" cases, such as the foggy conditions, based on a very poor visibility criterion.

6.3 Task Geneva

This task concerns only the data of a ceilometer, recorded every minutes day and night. In the learning process they are associated with the corresponding METAR messages established by the human observer.

The goal is to prepare an appropriate matrix format and to determine the best preprocessing in order to detect the main cloud layer and the lower cloud layer. Data were collected at Cointrin airport on a ceilometer connected to a PC.

The selected format for the ceilometer data is a matrix filled by 0 or 1. The rows contain the one minute measurement, the columns the levels of the height of cloud basis and the height or their extension. "1" is put when the detected range or extension of a cloud corresponds to a given level (see figure 8).
The selected levels are: [0, 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 6000, 7000, 8000, 9000, 10000] (feet).

A first simple heuristic algorithm was worked out to extract the heights of the clouds from the big amount of ceilometer data. Results seem satisfactory but more data is needed to evaluate this preprocessing.

6.4 Task Kloten

This task gathers the data from the new instruments: pyrgeometers, the ceilometer (every minutes), and the 10 minutes values measured by the automatic station at Kloten. The learning is based on 5 synoptic observations every night. Software programs have been elaborated to construct the matrix used by the learning process. Several months of data are needed to begin the work.

6.5 Conclusion

The learning algorithms used here to minimize the error function lead to great improvements with respect to the standard backpropagation algorithm. They are based on the BFGS optimization procedure, which is a very powerful second order descent method, and require the computation of the first derivatives of the error function. Moreover, the use of a validation set, which is common in backpropagation, turns out to be an efficient stopping criterion.

Note that, as for all optimization methods, convergence always occurs towards a local minimum and no method is guaranteed to give the optimal set of weights. However, the results obtained so far demonstrate that good solutions are reached in a reasonable amount of time, and that the obtained networks have good generalization performances.

7 Concluding Remarks, Outlooks

Teamwork and collaboration between MANTRA-team at EPFL, SMI/CMG and SMI/FWZ is excellent. Recent workshop on SPP Informatik underlines that point. During 1993 personal invest at CMG was 110 days (D.Cattani 60 and J. Ambühl 50 working days), and respectively 80 days at the EPFL by E.Mayoraz, E.Amaldi, F.Aviolat and V.Robert.

Schedule of the 1993 year was exactly respected (see Appendix).

First version of the Kloten main task is ready on the way.

CMG, EPFL and FWZ will continue on this work following a well defined research project for 1994:

1. Elaboration of a neural network as a tool to determine the main cloud layer (height of its basis and cloud amount) using a weight matrix determined by the Kloten learning task.
2. "Learning by hints" methodology will be introduced, using the reports produced by the present weather sensors as hints delivered to the neural networks.
3. The final structure of the Mantra machine, either Hard, or Soft as a neural simu lator, will be fixed in the second half of 1994.

A lot a work is ahead on the way. Personal investment will be as heavy in 94 as it has been in 93.

8 Bibliography

[Ambühl 93] J. Ambühl 1993, *Réseau de neurones en météorologie: Une perspective d'application*, Working Reports No 174, SMI Publication.

[Ambrosetti 91] P. Ambrosetti 1991, *Einsatz eines Pyrgeometers zur Bestimmung der Bewoelkungsverhaeltnisse*, Working Reports No 168, SMI Publication.

[Coulson 75] Coulson K.L. 1975, *Solar and terrestrial radiation*, Academic Press, New York.

[Eppley] Eppley Laboratory, technical notice, *Instrumentation for the measurement of the components of solar and terrestrial radiation*.

[Kunz 91] M. Kunz 1991, *Optimisation d'apprentissage pour reseaux de neurones multicouches*, Travail de semestre, DMA-EPFL.

[Mangasarian 93] Mangasarian 1993, *Mathematical programming in neural networks*, ORSA Journal on Computing, to appear.

[Minoux 91] M. Minoux 1991, *Programmation Mathematique*, Tome 1, Dunod.

[Platt 91] J. Platt 1991, *A resource-allocating network for function interpolation*, Neural Computation 3.

[Rumelhart 86] D. Rumelhart, G. Hinton, R. Williams, 1986, *Learning Representations by Backpropagating Errors*, Nature 323.

[Shanno 78] D. F. Shanno, *Conjugate gradient methods with inexact searches*, Mathematics of Operations Research, Vol. 3, No 3, August 1978.

[Weigend 91] A. Weigend, D. Rumelhart, B. Huberman, 1991, *Generalization by Weight-Elimination with Application to Forcasting*, Advances in Neural Information Processing 3.

Author addresses

D. Cattani
CMG
Case postale 312
CH-1215 GENEVE 15

J. Ambühl
Schweizerische Meteorologische Ansalt
Postfach
CH-8044 ZURICH

D. Engfer
FWZ
CH-8058 KLOTEN

E. Amaldi, F. Aviolat,E. Mayoraz, V. Robert
MANTRA Center for Neuro-Mimetic Systems
INJ-EPFL
CH-1015 LAUSANNE

Appendix

Schedule of AMETIS II project

Year	Task
1993	Project specification (Pflichtenheft) Installation of the instruments on the testing site at Kloten Elaboration of the learning tasks to be presented to the neural network at the Mantra team.
1994	Further studies and developments Material purchases Design and construction of the basic data handling devices (EPFL,ITR)
1995	... Tests
1996	Final material purchases Operational tests (including ICAO compatibility) AMETIS II in operation