EUROPEAN
CLIMATE
SUPPORT
NETWORK

DNMI Klima
05/2000

EUMETNET - ECSN

EUMETNET - ECSN

Project: Climate Databases

PROCEEDINGS FROM THE OSLO WORKSHOP 11. – 12. OCTOBER 1999

Margareth Moe
Editor

EUMETNET - ECSN

DNMI - REPORT

NORWEGIAN METEOROLOGICAL INSTITUTE

BOX 43 BLINDERN N-0313 OSLO

PHONE: +47 22 96 30 00

ISBN 0805-9918

RAPPORT NR.

05/2000 KLIMA

DATE

27.01.2000

TITLE

Proceedings from the Oslo workshop 11. - 12. October 1999

EDITOR

Margareth Moe

SUMMARY

This volume of the Workshop Proceedings comprises the full set of abstracts and presentations received before and after the workshop.

In the proceedings, the Climate Database of each country is described. In addition, a presentation of the quality control system at DNMI, and presentations of WEB-interfaces to climate data and climate / station-information are included.

Many of the countries are using Relational Database Management Systems (RDBMS), while some are using other commercial database / file-systems. The Climate Databases holds large amounts of data (several Gigabytes) and problems related to storing and retrieving the data are recognized among the participants.

KEYWORDS

1. Climate database
2. Workshop
3. EUMETNET - ECSN
4. International collaboration

SIGNATURE

Margareth Moe	Bjørn Aune
Margareth Moe	Bjørn Aune
Project leader	Head of Climatology Division

CONTENTS

FOREWORD

For some years now commercial databases (RDBMS) are used for climate data at many NMS. Standarized database tools and application tools together with common need for data and applications are all arguments for cooperation and exchange of information and experience between the different NMS. To accomplish collaboration in this field, a workshop on Climate Databases in Europe was held in Norrköping in 1996. This workshop agreed there was a clear need for more contact and exchanges of experience and information.

In september 1998 the EUMETNET Council decided Climate Databases to be one of the projects in the ECSN-programme.

To share knowledge and experience in database products, tools and implementations, acchieve easier exchange of data between countries and establish an European network among the climate database experts, a workshop concerning these topics was planned in Oslo 11th to 12th of October 1999.

This volume of the Workshop Proceedings comprises the full set of abstracts and presentations received before and after the workshop.

I want to thank the participants to the Workshop for their delivered contributions. A special thank to Bjørn Aune who have given this Workshop his full support and backing in the arrangement work. I also want to thank ORACLE NORGE and SGI for their contributions to the Workshop.

Oslo 24 November 1999

SUMMARY

PRESENTATIONS

All the presentations at the workshop gave a very good overview of the different climate databases in the participating countries. Problems using commercial database products, opportunities given by using RDBMS, design and optimization questions, storing techniques and retrieval possibilities were addressed at the workshop. Some very interesting presentations were given on climate applications, retrieving data directly from the database using a web-browser as the end-user interface.

APPLICATION DEVELOPMENT TOOLS

The use of web-tools in presenting information in climate databases was discussed. All agreed that this gives great possibilities for use of climate data everywear in the world. To achieve acceptable access to the data, attention must be given to the design and organization of the database and the data. Application development tools were also discussed. Especially, some of the countries had experienced limits in the 4.th generation development-tools, when developing database applications. All agreed that programming languages such as FORTRAN and C will be the most important programming languages for still some years, though report generators, JAVA and new programming-tools are used in some of the countries.

FURTHER COOPERATION

A closing discussion at the meeting was how to develop the cooperation further and how to achieve better databases and access to data. Some fields where cooperation would be very useful is undoubtfully the RDBMS systems and design of climate databases, to gather experiences and optimal solutions. Both quality control procedures / algorithm and GIS-based applications were recognized as fields where cooperation would be useful (As for GIS, an own ECSN project is dealing with this). Another area is ease of access to the data and a standardized interface to the data across borders. This would ease access to climate data both between NMS and could also be used by people in general and sure save much parallel work at the different NMS. Such a system should be WEB-based and the databases should be situated in the respective countries (distributed), though accessible through WEB-based user interface. To develop such a system, recommendations and specifications both for the WEB-application and the database (datawarehouse) should be worked out and developed.

WORKSHOP OVERVIEW

Bjørn Aune

The European meteorological services store the same types of data in their databases and use the data stored in them for more or less similar services and products. One would therefore assume that there exists a close co-operation between the services concerning climate databases. But that is not the case. Most climate databases in Europe has been developed with very little or no co-operation with other services. It is the same situation for the operation of the databases and few or none of the services have easy communication between their climate databases.

This lack of collaboration was the reason why ECSN organised the Workshop on Climate Databases in Norrköping, Sweden in 1996. That was the first European workshop for they who work directly with climate databases, and the main objective was to get an overview of the European databases, get they who work with them together and build at foundation for future collaboration. The further plan was to make such a workshop an annual event and develop an organised collaboration between national services. But since the reorganising of ECSN took longer time than expected, more than 3 years went before we were able to arrange this second workshop. And I sincerely hope that we now have laid most of the difficulties behind us and can start a new era in the field of climate databases and associated work. We have seen in other fields that international collaboration can work and that everyone gains by working together.

The general technical development since Norrköping has been great, and some of the future possibilities discussed there have already become realities. The commercial database systems have become better and communication is much easier. A decentralised European climate database is much closer now than then, both technical and psychological. The Internet has proved to be a new interesting system for communicating with and between databases. Web based systems are already in operation in some services.

Collaboration and communication demand standardisation. The more we collaborate and communicate, the more standardisation do we need. And if we extend the collaboration to include other than climate services also, standardisation becomes even more important. To standardise can often seem logical in theory and when planned, but can be painful in practise if we have to change yearlong traditions. But the process is necessary if we want to survive and prosper in a steadily more competitive world.

The databases and their environment have been better but also more complex. You can do more than before with new systems, but developing these systems demands more expertise than before. More and more climate services may soon find that they have to go outside their own ranks if the want to keep abreast with the advancements. This may at first look like a setback, but it can also be an advantage. And if we at that time have managed to create a good European collaboration system, we stand much stronger than if we still operated on our own.

The time between the workshop in Norrköping and this one shows that the start of a process that shall lead to a closer collaboration between meteorological services, is difficult. It needs time, patience and a lot of stubbornness. But the success is dependent of you who take part in this workshop and your colleagues. You will have to make the collaboration a part of your daily life and organise it so that it works on a permanent basis. You must become one team.

THE KMI CLIMATOLOGICAL DATABASE

Hans Van Hauteghem
Royal Meteorological Institute
Ringlaan 3, 1180 Ukkel, Belgium

phone: +32 2 373 06 74
fax: +32 2 374 75 62
e-mail: hans.vanhauteghem@oma.be

ABSTRACT

The climatological database at the RMI is currently a mixture of binary and text files, accessed by Fortran programs. Since this is a typical legacy system, the RMI started a project in association with the University of Antwerp to convert the totally unstructured file system to a new object-relational database management system. After some study of different database systems, the RMI decided to buy Oracle8i with JDeveloper and Time Series Cartridge.

1.0 INTRODUCTION

The meteorological data of the RMI is stored in hundreds of binary and text files, all with a different structure and different programs to access the information. The RMI was divided until the beginning of 1999 in five distinct departments. Each department consists of two or three sections. Every section as its own research staff and programmers who all write their own programs. Apart from the climatological information, the RMI also has data from the four other departments (Aerometric studies, Aerology, Applied Meteorology and Geophysics). Only the climatological data structure will be described in this paper.

2.0 THE CLIMATOLOGICAL DATABASE

WHICH INFORMATION IS STORED?

The database of the climatology department consists of six different files:

- Two-hourly information of Uccle
- Daily information of Uccle
- "Metagri" (Meteorology and Agriculture) information
- Hourly synoptic information
- Rain and temperature
- "Bulletins"
- Metadata

All this data is stored in binary files that can be accessed by Fortran programs. The metadata of the stations is stored in a MS Access database consisting of one table.

The Climatology department has three different station networks. The Metagri (Agricultural Meteorology) network has 50 stations measuring important information for the agricultural sector. There are more than 150 stations that measure the rain and temperature. The network that is responsible for the bulletin-files consists of 19 stations of the RMI, the Meteowing (Military) and Belgocontrol (Flight Guidance).

Two-hourly information of Uccle

The two-hourly information of Uccle is measured since 1901, but the data is digitally stored since 1931. These files contain the following information, mostly with a frequency of 2 hours:

- Temperature with extreme values
- Humidity with extreme values
- Vapor pressure with extreme values
- Air pressure with extreme values
- Wind speed and direction
- Sunshine duration
- Precipitation quantity and duration
- Temperature in the ground
- Pluviometer, Cloudiness and other information at 8am

Daily information of Uccle

The daily information is calculated from the hourly information and contains 35 parameters:

- Maximum, minimum and average temperature
- Precipitation quantity, duration and intensity
- Evaporation
- Average and maximum wind speed and direction
- Average, maximum and minimum humidity
- Average, maximum and minimum air pressure

All the information concerning temperature is for both open and closed thermometer huts.

Metagri information

The Metagri network consists of a combination of synoptic and thermo-pluviometric stations. Depending on which station transfers the information, the following information is stored:

- Wind speed and direction at 6am
- Maximum, minimum and average temperature from 8am to 8am
- Maximum and minimum temperature from 6am to 6pm
- State of the ground
- Snow cover code
- Cloudiness

Hourly synoptic information

This is a combination of the hourly and three-hourly synoptic information of the belgian stations who send their data on the GTS-network.

Rain and Temperature

The daily rain, the minimum, maximum and average temperature of more than 150 stations is stored in these files. We have 277 stations who measure the rain, where 165 of them also measure the temperature.

Bulletin

This information is used to create a monthly bulletin with 35 daily parameters of 19 stations of the RMI, Meteowing and Belgocontrol. The following data is stored:

- Average air pressure
- Cloudiness at 12pm, 6am, 12am, 6pm
- Sunshine duration
- Maximum, minimum and average temperature
- Minimum temperature in the ground at 2, 5, 10, 20, 50 and 100 cm depth
- Minimum and average relative humidity
- Fog duration with visibility <1000m, <500m and <200m
- Precipitation quantity and duration
- Thickness of the snowcover at 6pm and 6am
- Days with snow, hail and storm
- Average, minimum and maximum wind speed
- Direction of the wind at 12pm, 6am, 12am, 6pm

Metadata

The metadata, which is managed by a small MS Access program, contains only information about the stations of the "Rain and Temperature"-files:

- Station name and codes
- Geographical, Lambert coordinates and height
- Observator name, address and begin and end date of observation
- The responsible department and section of the RMI
- A description of the surroundings in all directions
- Instrumentation
- The network where it belongs to
- ...

DATA MANAGEMENT

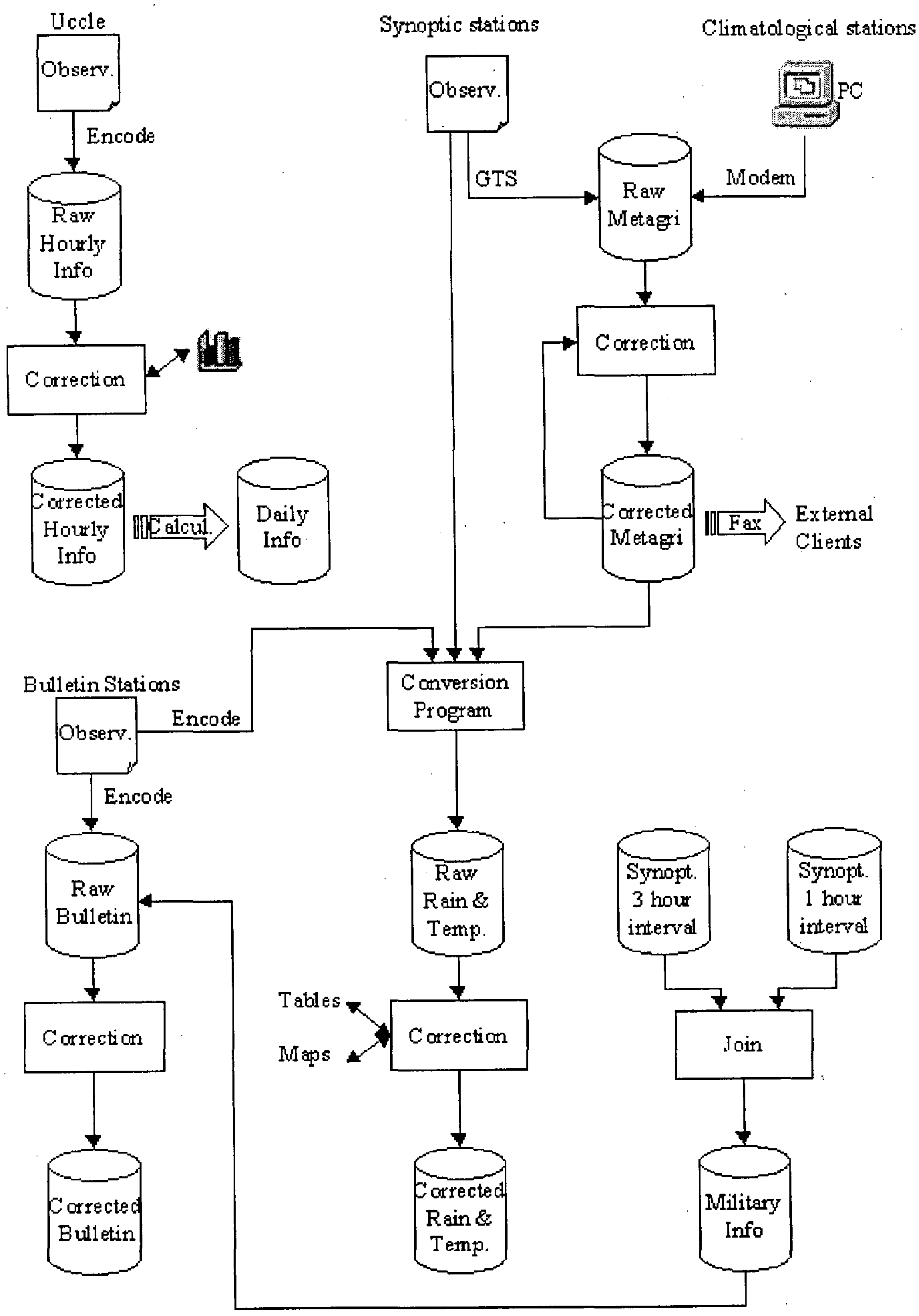


Figure 1. Schematic overview of the datastreams

Information Retrieval

All the observations at the RMI are done manually. The observator writes the measurements on paper, which are then encoded to digital format by one of the assistants at the computer science department.
The data from the synoptic stations is put directly on the GTS and is captured at the institute by the department of Synoptic Meteorology, who then forward the files to the climatology department. The climatological stations have observators who send the data by mail (post) at the end of the month or by phone the day itself to the institute. All this data is also encoded by someone at the RMI and stored in the appropriate files.

Error Correction

There are no automated correction methods at the RMI. All the corrections are done manually by looking at the tables or by creating maps and graphs to find irregularities. If the error can not be detected, the observator is contacted by phone to ask for his opinion. In the future, we will start developing methods to do this process as automated as possible.

Archiving

After the correction the data is stored in different files than the original or raw data. This means that all the information is stored twice, to be sure that the original data still exist after the corrections. A backup of the complete database is performed every week and an incremental backup every evening. At the end of each month a complete backup is done, which is kept for a longer period than the weekly backups.

3.0 PROJECT: THE DESIGN AND IMPLEMENTATION OF AN OBJECT-RELATIONAL DATABASE FOR THE ROYAL METEOROLOGICAL INSTITUTE OF BELGIUM

INTRODUCTION

Currently various sets of data (historical, observational, numerical model-output, satellite, radar, etc.) are stored in a hybrid way. Each application requires a time consuming development and testing of a large number of specific programs. The introduction of a simple and user- friendly way of consultation and exploitation of the data can only be achieved by the implementation of an object-relational database. The implementation of such a concept will provide a faster and a standardized access to the data by using a high-level query language. More attention could be drawn to the exploitation of the data and to the development of new applications. In the future the handling and the storage of huge volumes of datasets produced by new observation techniques could be simplified.
The project to design this new database is in cooperation with the ADReM (Advanced Database Research and Management) group at the University of Antwerp and is scheduled for a period of two years. If the database is not finished after these two years, the project will be extended for another two years.

CHOICE OF DATABASE SOFTWARE

We have decided after a thorough investigation of different software products, that we will buy the Oracle8i database with the Time Series Cartridge and the JDeveloper Suite.
The main reasons for this decision were the large support division of Oracle in Belgium, the use of Oracle at our neighboring institutions, the big market share of Oracle and the scalability and development tools of the database.

DATABASE SCHEME

A preliminary database scheme is already developed and is based on the old file system. Almost all the tables have the same structure with a station_id, a date/time stamp and all the parameters of the old file. Only the metadata is completely reorganized because of the lack of normalization in the old MS Access-table.
The satellite and radar data will also be integrated in the database by storing the metadata of the images in the database and keeping the files on the storage devices.

THE FUTURE

In the near future, the database software will be installed and the old files will be converted to the new database tables. We will try to convert almost all the Fortran programs to new Java applications with increased functionality. The programs that are to difficult to convert will be integrated with the database using embedded SQL. If the database is functional, we will start developing new quality control routines and improved customer related applications, like a new web server and automated delivery of the climatological products.

THE AUSTRIAN CLIMATE DATA BASE SYSTEM

Wolfgang Lipa
ZAMG
Central Institute for Meteorology and Geodynamics
Hohe Warte 38
A – 1190 Vienna

Phone: +43-1-36026-2207
e-mail: lipa@zamg.ac.at

ABSTRACT

ZAMG's climatological data base system is implemented on Sun Servers with the data base management system SYBASE. At present the second table design of relational data base system is used. Its structure, quality control and application will be described in the following paper.

1.0 INTRODUCTION AND REVIEW

ZAMG, the Central Institute for Meteorology and Geodynamics of Vienna, looks back on a long period of registering meteorological data. Since 1851, the climate was observed and stored on hand written journals, therefore the institution is unique in the world.

Austria's climate data base includes long term (more than 100 years) daily observation series of the towns Vienna, Innsbruck, Kremsmünster, Graz and Salzburg, further on of the world's first meteorological mountain observatory Sonnblick.

The first digitised data were monthly climate values registered on punch cards, and the change to magnetic tapes took place in 1979.
It was a big step in 1984 to start a climate computer network which included data of whole Austria checked automatically by 45 rules. The file system was based on a hand written data management system.

ZAMG changed to UNIX-systems and a commercial relational data base management system SYBASE in 1990. It took 3 years to store all digital data in the data base server used since 1992 that registered one table per time period. This solution offered great advantage in data access but also subjected the problem to differ between original and corrected values.

To test new structures ZAMG received a new department data base server in 1996 for calculation and storage of online hourly climate data sets. The new structure succeeded, so a new data base server with full duplex system with 2 GB memory, 4 engines and 500 GB storage got installed.

2.0 PRINCIPLE COMPONENTS OF THE NEW DATA BASE STRUCTURE

The following list explicates the design of the

1. New table structure:

Four tables for each period of time

- The table with the suffix **-orig** (original) contains the original values
- The table with the suffix **-wert** (values) contains the final values
- The table with the suffix **-flag** (flagged) contains the description of the table values
- The table with the suffix **-ersatz** (substitute) contains proposal values for the computer and is only temporal

2. The flag table

consists of only 6 different values because a more explicit classification would complicate the action:

0 = unknown	blue
1 = original and ok	green
2 = wrong	red
3 = suspect	light red
4 = not original but ok	light green
5 = not original and not checked	brown

In the beginning a meta table required the station information, today ZAMG uses a system with 25 related tables that include station history, location, observer and instrumentation.

For users there are programmes based on Fortran and C as well as SQL nested in c-shell and perl script, Excel and Access connection with ODBC. The highest applications are GIS-Systems based on ARC View and Arc Info that combine these data sets with lightening and satellite images and the Austrian topographic elevation model.

3. Quality control:

Hourly values can easily be calculated by incomplete 10 minutes messages. Those values are signed as incomplete and brown (not original and not checked) in the **flag-table**. The original values stay unchanged in the **orig-table**. Each change of the **orig-table** is registered in the **flag-table**.

The last check of data is made with the software GEKIS (geographical climatological information system) developed together with the company Datamed.
Based on their geographical position the values are shown in the map with the colour depending on the **flag-table**.

3.0 MAINTENANCE OF THE DATA BASE

The image shows the tasks of the two groups that both are responsible for the data base:

	Computer Division DB-group	Climate Division Data-group
Hardware	***	
Server Software	***	
Administration	***	
Security	*	*
Disk Space	*	***
Table Definitions	*	***
Contents		***
Interfaces	**	*
STAFF	**3**	**12**

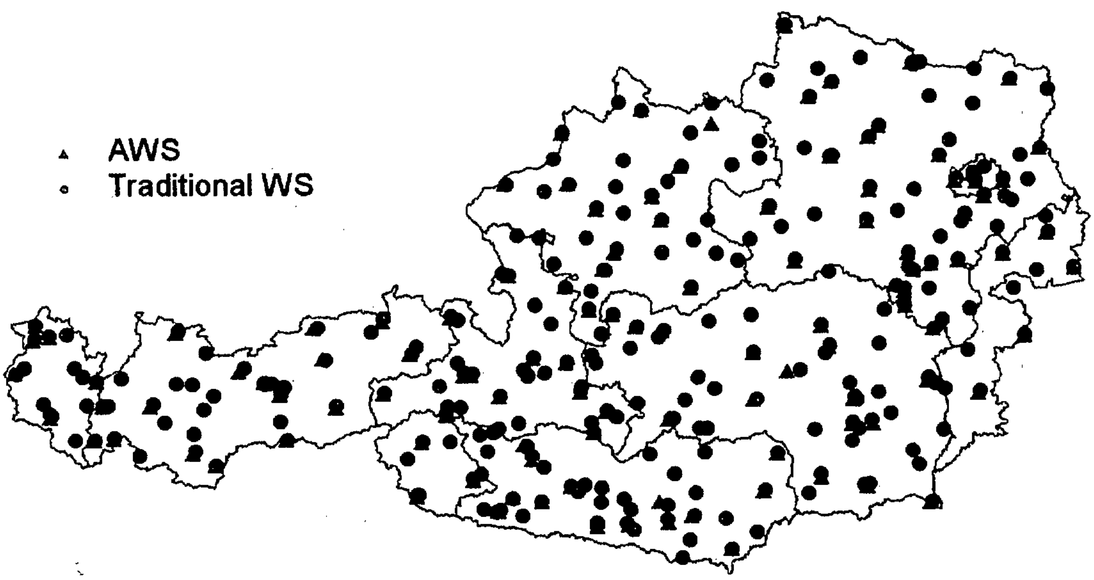
4.0 OUTLOOK

For the future there are two main points to gain:

- To bring all old data into the new structure
- To develop computer based checks to correct the station values with help of radar and satellite images (at the moment still check with an image)

ENCLOSED:

- This map shows the actual situation of Austria´s data station network and the stations´ density (250 – 300 stations since 1960). About more than 100 daily data of traditional WS and AWS online in universal time (radar and satellite images) are available.



DESCRIPTION OF DATABASE AT HYDROMETEOROLOGICAL INSTITUTE OF SLOVENIA

Mojca Dolinar
Hydrometeorological Institute of Slovenia
Vojkova 1/b, 1000 Ljubljana, Slovenia

phone: +386 61 327 461
fax: +386 61 133 13 96
e-mail: mojca.dolinar@rzs-hm.si

ABSTRACT

Presently on our institute we use the VMS database system on VAX computer machines. It is based on the different types of station networks. For applications we use mainly clasical data, sometimes also automatic weather station data. Our future plans are to improve our metadata base, to establish an archive of spacial data and to replace the existing database with ORACLE database.

1.0 INTRODUCTION

Presently on our institute we use the database system in VMS operational system on VAX computer machines. Our plans for the future are to replace the existing system with a new commercial database system (ORACLE). The present database system is based on the different station networks. For archiving data we use 3 VAX computer machines. In the article our network of meteorological stations, the structure of the archive and the tools we use in our work with data will be presented.

2.0 CLASSICAL CLIMATOLOGICAL DATA

The network of classical climatological stations presently consists of 45 stations. In the past we had more classical climatological stations, but now their number is reducing. The measurements on these stations are carried out three times per day (at 7h, 14h and 21h in solar time). The observations for non-measured parameters are performed through the whole day. 12 climatological stations are sinoptic stations, where meteorological parameters are measured every hour. Meteorological parameters that are measured or observed on the climatological stations are:

air pressure
air temperature (2 m)
maximum daily air temperature (2 m)
minimum daily air temperature (2 m)
air humidity (2 m)
wind on 2 m

precipitation
sunshine duration
visibility
cloudiness
other observed parameters (snow cower, heavy winds, fog, dew, frost, showers, storms, hail, ...)

The data is collected once per month, for 22 main stations collection is made daily. After monthly collection control tests are made on completeness, plausibility and one station control test with combination of many parameters. Missing values are interpolated manually by operator (spatial interpolation) and flagged. After the control procedure data is archived on VAX. There is one file for every station per year. Every file has separate record for each day. Monthly derived statistics of climatological data is archived in separate directory on VAX. In one file there is data for all stations for one year. Climatological data series has been archived as described since 1961. For some particular stations even longer series are archived (for example for Ljubljana we have data from 1866 on). For searching data we have two Fortran codes: one for retrieving daily data and the other for retrieving monthly sums and averages.

Among climatological stations there are 38 stations with registration instruments for temperature and 44 with registration instruments for humidity. Hourly averages and terminal values are archived in special archive on VAX.

3.0 CLASSICAL PRECIPITATION DATA

The network of classical precipitation stations presently consists of 196 stations. In the past we had more than 300 classical precipitation stations, but like the number of climatological station also the number of precipitation stations is reducing now. Precipitation is measured once per day (at 7h, solar time). Beside the precipitation also the height of new-fallen snow and snow cover are measured. The observation of parameters related with precipitation is made through the whole day and are as follows:

shower, storm, hail ... (also the time is noted)
dew, frost, rime ...

The procedure of data collection, testing and archiving is similar as on the climatological stations. The data is collected once per month. After the collection control tests are made on completeness, plausibility and one station control test with combination of many parameters. Missing values are interpolated manually by operator (spatial interpolation) and flagged. After the control procedure daily data and monthly sum of precipitation is archived on VAX computer machines. For daily data there is one file for every station per year. Every file has separate record for each day. Monthly derived statistics of precipitation are archived in separate archive on VAX. In one file there are data for all stations for one year. Precipitation data series have been archived as described since 1961. For some stations longer series are archived. For external application we use the same two Fortran codes as for other climatological data.

4.0 CLASSICAL RAINGAUGE REGISTRATION DATA

The network of classical raingauge registration stations consists of 36 stations. There is historical data from raingauge registration instruments for some more stations. Two months ago 25 new raingauges were set up in the Soča watershed in the framework of MAP (Mezoscale Alpine Programme).

The sheets from registration instruments are collected once per month. The data is digitised in time resolution of 5 minutes. After the digitalisation plausibility test is made with the sum of 24h precipitation measurements. The missing data is not interpolated. If suspected values occurred, they are changed into missing values. The raingauge data is archived in one file per station per year. Every record in a file is reserved for one hour of data. Raingauge data is archived from 1970 on, for two stations even from 1948. For searching data we use two Fortran codes: one for retrieving raw data and the other for getting different time maximums of precipitation.

5.0 DATA FROM SPECIAL WIND MEASUREMENTS

On some climatological stations and also on other locations (depending on the customer's requests) special wind measurements were carried out. The lengths of data series are different – from 3 months to 10 years. The archive of this data is structured as regard to the type of the wind-measuring instrument (electronic anemometer, mechanical anemometer, observations...). Hourly velocity averages, hourly velocity maximums and prevailing hourly wind directions are stored in the file with extension, that indicates the type of wind measuring instrument. In every file there is the data for one year and one station. There are special Pascal codes on VMS operational system for retrieving this data from archive and also making all the necessary statistics.

6.0 AUTOMATIC WEATHER STATION DATA

The network of automatic weather stations presently consists of 22 stations. The first station began to work operationally in 1993. Meteorological parameters measured on these stations are:

air pressure
humidity (2 m)
humidity of leaf
air temperature (2 m)
air temperature (0.5 m)
air temperature (0.05 m)
temperature at 5 cm below ground
temperature at 10 cm below ground
temperature at 30 cm below ground
global sunshine radiation
diffuse sunshine radiation
UVB radiation
wind on 2 m
precipitation
radioactivity

Not all parameters are measured on all stations, but on all of them air temperature, humidity, wind and precipitation are measured. This is on line data and is transferred through telecommunication channels to our centre every half an hour.

The archive of automatic weather station data is separated in two parts. One part is like classical data on VMS operational system. During data collection some logical controls are done. Missing data are not interpolated. 30 minutes averages, terminal values and for some parameters minimum and maximum values are stored in a file. There is one file for every station per day. The data for last two or three months is on VAX, older data are archived on tapes and are put on VAX if needed. This data we use in our operational work, but there are problems with large gaps of missing data and quality of data. Fortran code is made for searching this data from the archive.

Automatic weather station data is archived also in ORACLE database system on VAX. After the collection data is loaded with SQL loader into ORACLE tables. One table is reserved for one parameter. There are loose logical constrains on the tables. If the data exceeds the prescribed limits or the data is not described into control files then it is written into BAD or DSC files. For searching data in ORACLE database there is a DELPHI application for control purposes and test use.

7.0 SOLAR RADIATION ENERGY DATA

Radiometers that are measuring solar radiation energy are mounted on automatic weather stations. The energy of global solar radiation is measured on 19 stations, diffuse energy is measured on 8 and UVB energy on 2 stations. The data series are of different length, the longest is a series for global solar radiation energy in Ljubljana (35 years). The data is collected every day, when limits test is done. At the end of the month hourly missing values are interpolated with an objective procedure and flagged. After the control procedure the data is archived separately from other automatic weather station data. The archive is structured as regard to the type of solar energy. Hourly averages are archived in a file. In every file there are the data for one year and one station. File extension indicates the type of solar radiation energy. Pascal codes are used for searching the data and making simple data statistics.

8.0 CLASSICAL SUNSHINE DURATION DATA

The network of classical sunshine stations has presently 23 stations. There is historical data from solarigraph for some more stations. The sheets from solarigraph are collected once per month. The data is digitised in time resolution of 1 hour. After the digitalisation graphical test by eye is made. Fixed obstructs are interpolated manually and an objective procedure is used to correct data because of obstacles. The missing data is interpolated manually by the operator (spatial interpolation) and flagged, most of it only on clear days. After the interpolation procedure data is archived in a file on VAX computer machine. In every file there is the data for one year and one station. Pascal codes are used for searching the data and making simple data statistics. 7 stations have 39 years long series of data, data series on other stations are shorter.

9.0 METADATA

The data that describe working and canceled stations is collected in a file on VAX computer machine. For every station each change in geographical coordinates, height above sea level and parameters that are measured is written in a record of file. At the time we are preparing a new format of metadata. It will contain more information (description of microlocation, description of instruments..) and will be more user-friendly (operating in ORACLE). Presently it is in phase of collecting supplement information.

10.0 APPLICATION TOOLS

As described above, for retrieving data from archive there are several Fortran and Pascal codes in VMS on VAX computer machines. There are also some special codes in VMS written for routine work that must be updated every year (calculating of returning periods..). For final work we use PC tools like Statistica, Surfer, Idrisi, CorelDRAW, MS Excel, MS Word.

11.0 FUTURE PLANS

As mentioned above, presently we are working on many tasks concerning data and databases. First, and at the moment the most important task is to complete and use operationally the new metadata. The second task, which is also implementing now, is to establish an archive of nonclasical data (satellite images, radar images, soundings, analysis of pressure and temperature fields and discharges). Satellite images, radar images and soundings have already been archived. We need to organise the archiving of other data and to develop application tools for searching all this data from the archive. One of the momentary running tasks is also development of a system for automated spatial quality control.

Our most important task is to replace the existing database system with a new ORACLE database system. As described above, real–time data has already been archived in ORACLE. Application of this archive is in test phase and it is used now (from September 7th till November 15th) for MAP intensive observation period. The PRO C code is developing now for loading and controlling the data in ORACLE database. The most difficult task now is how to combine real-time data with classical data and how to make this database as efficient as possible.

THE CLIMATE DATABASE AT DMI

Claus Kern-Hansen
Danish Meteorological Institute (DMI)
Lyngbyvej 100 - DK 2100 Copenhagen, Denmark

Phone +45 39 15 75 00, Fax. +45 39 27 06 84
http: www.dmi.dk - e-mail: ckh@dmi.dk

ABSTRACT

The Danish Meteorological Institute (DMI) is responsible for all Meteorological services in Denmark, the Faeroe Islands and Greenland, as well as the surrounding waters and airspace.

As part of its activities DMI operates a central database (located at the head-office in Copenhagen) where all observation data (GTS as well as non-GTS) from the geographical area covered by DMI's responsibility is collected.

The current Climate database is a Relation Database Management System (RDBMS) (Open INGRES), organised in a classic table structure designed mid 1980'ies.
The content comprises all surface observations (SYNOP, automatic/manual Climate-stations, daily precipitation stations, sunshine recorders etc.) since 1960, upper-air soundings (TEMP) since 1993, as well as older data (since 1873) from some stations and data from special-purpose station-networks (the Danish lightening recorders, the Danish slippery road network and others).
At the moment the database stores more than 20 Giga-bytes, divided into several thematic databases, including a station directory.
Most retrieve, storage and data-management applications are FORTRAN programmes, and in addition to these a tailor-made time-series and wind-rose plotting programme as well as standard software (MS Excel, MS Access and SPSS) are used for data-manipulation and plotting.
During the last 10 years new user requirements, new data-flow formats and the increased number of automatic stations has called for new dedicated tables including daily, monthly and grid tables.

On the doorstep to the new century, the challenges to be met are related to: Quality-control/labelling, a more flexible (table) design, and an up-to date user-interface.

HANDOUTS FROM THE GIVEN PRESENTATION ARE PUBLISHED IN APPENDIX A.

CLIMATE DATABASE AT FMI

Pauli Rissanen and Antti Samuli
The Finnish Meteorological Institute
BOX 503, Vuorikatu 24, 00101 Helsinki, Finland

Phone: +358 - 9 - 19 29 1
Fax: +358 - 9 - 19 29 57 79
e-mail: pauli.rissanen@fmi.fi
antti.samuli@fmi.fi

ABSTRACT

At this moment Oracle version 7.3.4 is run on servers Digital Unix 4.0d. In 2000 Oracle 8 will be taken into use. Real time database (RTDB / NEONS, 24 G bytes) consists of different tables that include weather observations, grid data and point forecast data (Figure). Data is stored for some days or weeks mainly. In this environment different real time products for customers are produced.

1.0 CLIMATE DATABASE

The basic description of Climate and AWS tables (12 G bytes) are given on publication "Climate Databases in Europe" (1996). The Climate database and AWS database are described as "climate database".

After that some changes have been done:

- Table **ST_HISTORY** (all tables) compensates the table SLOG
- New table **AUTOSYNOP** has been created for automatic airport observations
- Some new parameters have been inserted to different tables
- New **METADATA** tables have been created

2.0 METADATA TABLES

The need of metadata information has increased a lot. Ordinary users need background data of stations and different classifications of stations. Also new quality control programs need control tools to handle varying parameter fields in time and space.

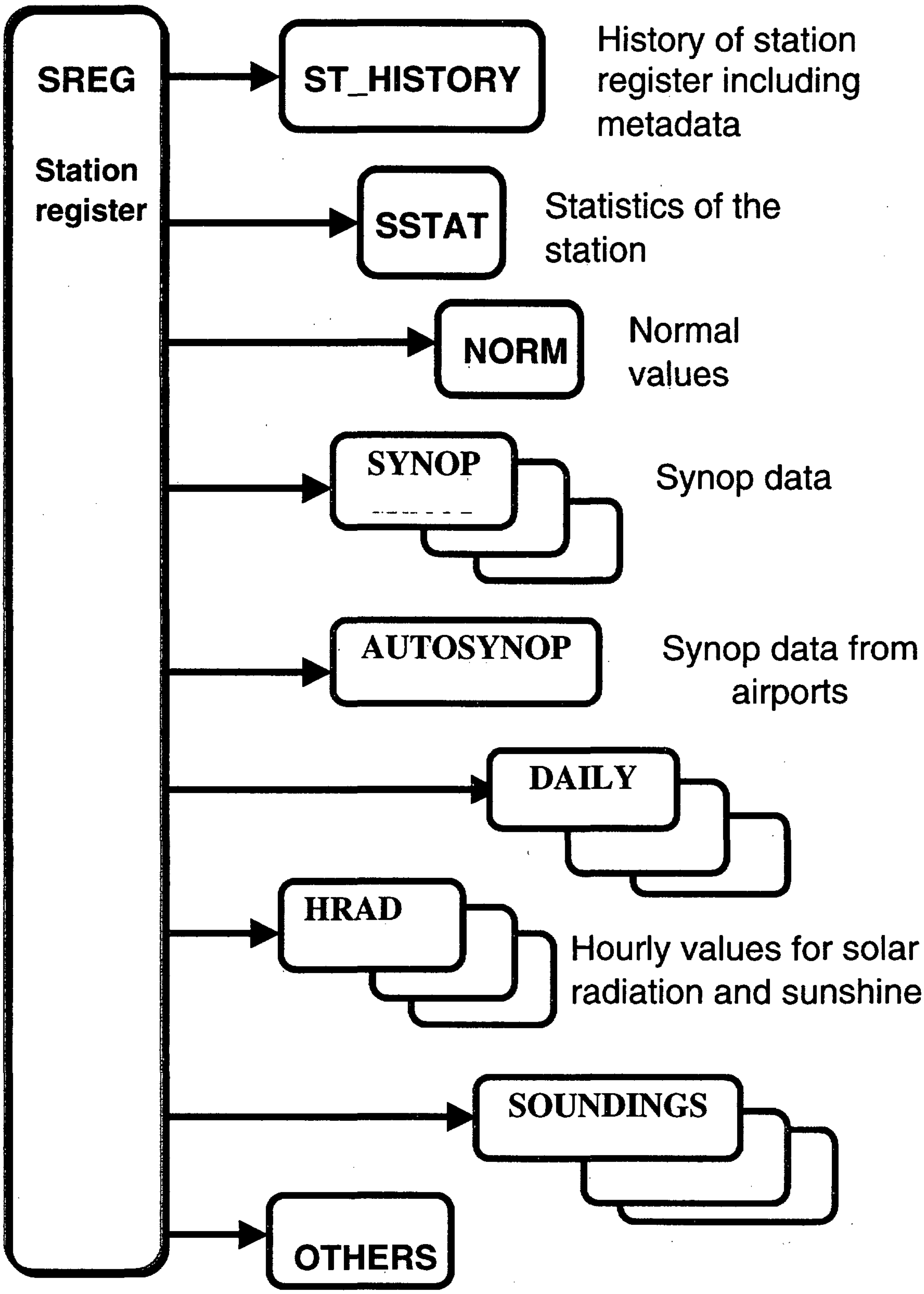
Metadata tables at this moment:

- **AREA_OF_WATERS** Percentage of water areas around the site
  - **ST_CLIMATIC** Climate area represented by the site
  - **CLIMATIC_CLASS** Main class of climate area
  - **CLIMATIC_SUBCLASS** Sub class of climate area
  - **CLIMATIC_SPECIAL_AREA** Special climate area of the site
- **COUNTIES** List of counties
- **REGIONAL_COUNCILS** Regional councils
- **PROVINCES** Provinces
- **CENTRE_OF_COUNTY** Center of the county
- **COUNTY_STATION** The most representative sites of the county
- **DATA_COMMUNICATIONS** Data communication information of the site
  - **MICRO_TYPES** List of different PC's
  - **MODEM_TYPES** List of different modems
  - **TEL_AREA** Area codes of telephone numbers
  - **TEL_TYPES** Type of the telephone (line, NMT, GSM)
  - **ELSTAT_CODES** List of methods used in defining site elevation
- **EXPOSURE** Exposure descriptions
  - **EXPOSURE_DATA** Exposure to different sectors
  - **EXPOSURE_METHOD** Method used in defining exposure
  - **EXPOSURE_OBSTACLE** List of obstacles
  - **TIME_SCALE** List of periods (year, month, …) defining prevailing periods of the exposure
- **NEIGHBOURS** 10 closest sites of the site
- **ADDRESS** Address information
  - **ADDRESS_TYPE** Type of the address (post adddress, visiting address, home address)
  - **PERSON** Personal information
  - **PERSON_GROUP** To which groups person belongs to (rights)
  - **PERSON_GROUP_LIST** List of above groups
  - **RESPONSIBLE_PERSON** Responsibility (area) of the person
  - **RESPONSIBLE_AREA** List of above areas
  - **TELEPHONE** Telephones
  - **TELEPHONE_TYPE** Type of the telephone number (work, home, mobile, …)
- **ST_DESCRIPTION** Descriptions of the site
- **ST_DESC_REPORT** Inspection reports attached to descriptions
- **ST_DESC_TYPES** List of description types
- **ST_TYPES** Types that site is representing
- **ST_TYPE_CAT** List of description types (new version)
- **SREG_STATION_TYPES** List of different site types (old types 1-9)
- **VALOKUVAT** Photographs

Following metadata tables will be used in near future:

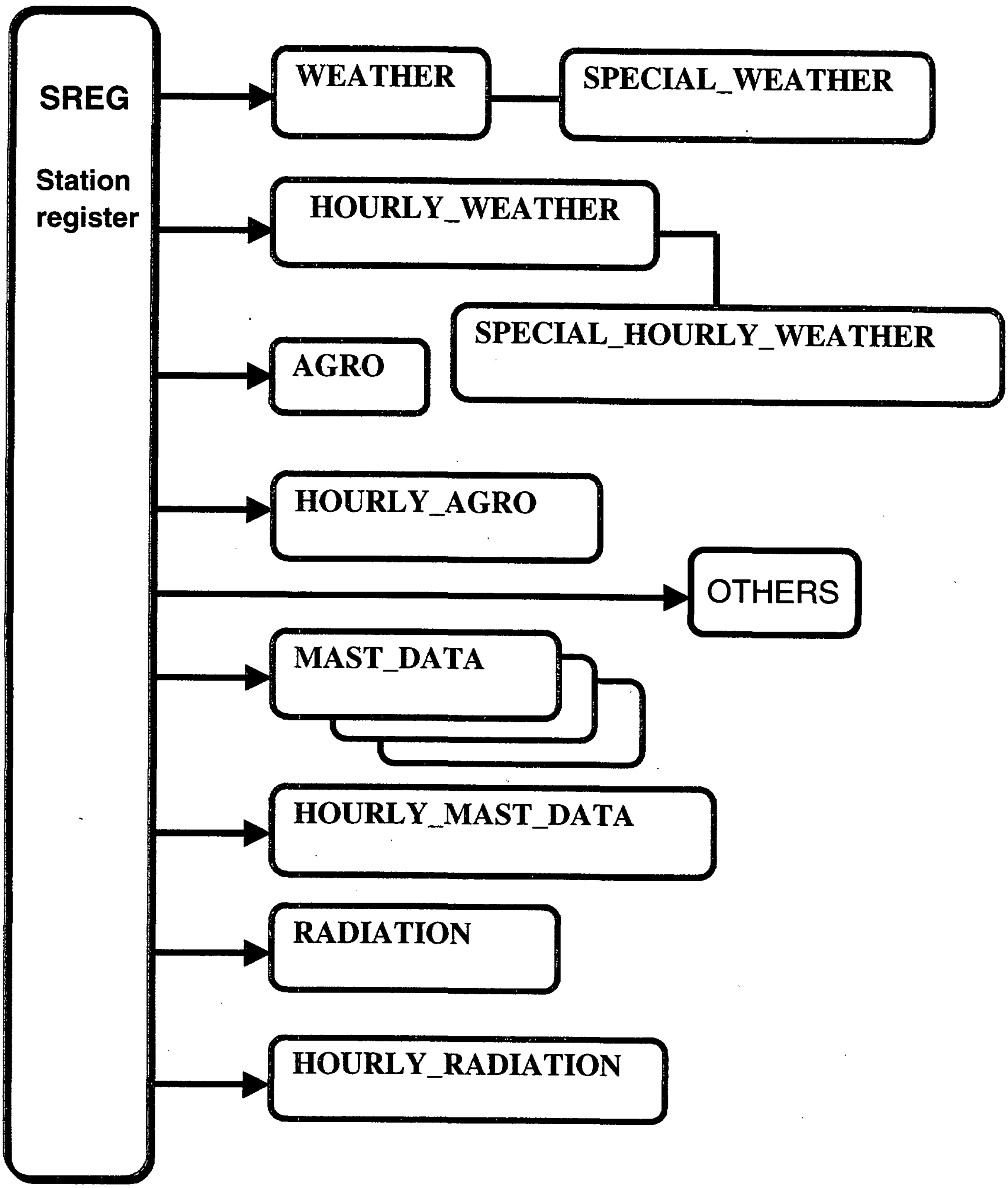
- **CONTACT** Contact information
- **CONTACT_GROUP** Contact group ()
- **CONTACT_GROUP_LIST** List of contact groups
- **DEVICE** Instrument
- **DEVICE_LOCATION** Place where instrument is located
- **LOCATION** List of location places
- **MEASURE** List of observations
- **DEVICE_MEASURE** Links between observations and instrument types
- **DEVICE_TYPE** Instrument types
- **OBSERVATION** Observation / measurements at the site
- **EQUIPMENT** Instruments at the site
- **EQUIP_OBS** Links between instruments and observation at the site

CLIMATE DATABASE AT FMI



10.09.1999 /
AA / PR

AWS - DATABASE AT FMI



10.09.1999 /
AA / PR

KLAR - CLIMATOLOGICAL DATABASE AND RETRIEVAL SYSTEM AT SMHI

Nils-Åke Andersson
SMHI

e-mail: nanders@smhi.se

ABSTRACT

The SMHI database system KLAR represents all climatological databases and retrieval systems at SMHI and some of the climatological presentation and analysis systems.

Most of the climatological databases and retrieval systems at SMHI are running on Alpha/OpenVMS computers. We are using MIMER as database handler in the Alpha/OpenVMS-environment. MIMER is a Swedish developed relational database management system and the SYSDECO MIMER AB are selling the product all over Europe.

The largest climatological database is BAAK (about 45 GB) containing climatological observations as

- *SYNOP data* *(digitised from 1951)*
- *precipitation and snow data* *(digitised from 1961)*
- *temperature data* *(digitised from 1961)*
- *radiation data* *(digitised from 1983)*
- *METAR data*
- *aerological data*
- *hemispheric SYNOP data*
- *catalogue with station identities and station history*
- *data catalogue*

The BAAK-system has been running since 1987.

Another database contains monthly data for

- *the new climatological reference period 1961-90*
- *precipitation and mean temperature starting from 1722.*

The third database REKORD contains extreme values and it is running on PC and is using Microsoft Access. Version number 3 was released in June 1999 containing temperature, precipitation, snow depth and sunshine. The development continues and in future versions more parameters will be included and there will also be a system for extreme weather situations described in words.

To access the climatological databases we have a rather userfriendly system, STATGEN, which can extract the users data or combine data for statistical analysis.

1.0 WHAT HAS HAPPENED SINCE THE FIRST DATABASE WORKSHOP IN 1996?

The main task during the last 3 years has been to millenium-secure (y2k) the systems and databases and to migrate them from the VAX/VMS platform to the Alpha/OpenVMS platform.

We have continued to improve the climate databases, especially the database of extremevalues REKORD.

We are using GIS (Arc/Info and MapInfo) more frequently for analysis and presentations of the historical climatological data. We have even used GIS (Arc/Info AML) for automatic plotting of monthly (seasonly, yearly) data.

A new database with gridded climatological data has been developped using MESAN (analyse system on mesoscale). MESAN combines observation station data together with radar and satellite data. The database contains the last 2 years with monthly temperature, precipitation and cloud cover in a 11 km grid and other variables in a 22 km grid. We are planning to do re-analyses to get a longer climate series with gridded data.

2.0 FUTURE PLANS?

In the next years we will concentrate on:

- A new production system for KLAR with an intranet approach
- GIS-applications
- Gridded data
- Perhaps we will start looking for a new RDBMS and a new database model for climatological data
- Homogeneity testing

THE CLIMATOLOGICAL DATABASE AT MET EIREANN, IRELAND

Liam Keegan
Met Eireann,
Glasnevin Hill, Dublin 9 Ireland

Phone: +353 1 8064243
Fax: +353 1 8064216
e-mail: liam.keegan@met.ie

ABSTRACT

Met Eireann uses Computer Associates Openingres 1.2 RDBMS running on a Sun server. Most data are held in the RDBMS but satellite and radar data are held as files. We are reasonably happy with the current setup but have some concerns about future software support.

1.0 CURRENT POSITION.

Openingres 1.2 on a Sun machine.
A 5 year contract for maintenance and support on Openingres is near its end.
The database is very reliable, we are only aware of 1 error.
The front end tools are old and very limited.
Their PC windows 98 client is very fragile.
We are concerned about long-term third-party support for Openingres.
We are concerned about the long-term commitment of Computer Associates to the database.

2.0 APPLICATIONS.

More than 101 FORTRAN programmes, some c programmes, shell scripts.
Computer Associates Openroad, which is fast enough on modern machines.
Microsoft Excel via ODBC.
Web based tools. These are very useful because no deployment or licensing issues arise and they can deal with file based storage as easily as with the database.

3.0 DATABASE STRUCTURE.

Raw data tables contain data as reported, even if obviously wrong.
Final tables contain quality controlled data. Flags indicate if data is estimated or not standard.
Hourly, daily and monthly data back to 1939 are stored in the database, as are data from ships and buoys and from roadside stations. Radar and satellite data are stored in files.

4.0 PROCESSING.

Most data are available to users soon after receipt. These data are only checked for gross errors. Quality control takes from 1 week to 1 month for synoptic stations, and longer for climatological and rainfall stations. The quality control is slipping further back because of errors and missing data from automatic stations. We are now improving the quality control to better handle missing data.

5.0 METADATA.

At this point, little or no metadata are digitised.

6.0 NEW SINCE THE LAST MEETING.

We have added satellite, radar and lightning data.
We have added roadside data without quality control.
We have installed a new database server and upgraded the database software.
We have started to change from Unix to PCs on the client side. The PCs are not as reliable.
We are doing quality control on older climatological data.
We have been coping with problems caused by automatic weather stations.
We have been developing new data flow and quality control systems.
We have introduced an experimental Web interface.
We have assessed optical character recognition. We believe that this is not yet capable of fully replacing human operators for data entry.

7.0 FUTURE PLANS.

Metadata.
Further development of quality control.
Digitise older daily data.
Introduce G I S ??
An automatic climatological weather station project where we have been able to set the specification.
Watch developments with Computer Associates software.
Watch developments with optical character recognition.

SPANISH METEOROLOGICAL INSTITUTE (I. N. M.) CLIMATOLOGICAL DATABASES REPORT

Andres Chazarra
INM

e-mail: clp@inm.es

ABSTRACT

At the moment the climatological information in INM is distributed in 15 Regional Meteorological Centres (CMT's), each one with a Regional Database containing the data coming from the stations within its area of responsibility.

Besides, at the INM Headquarters in Madrid there is a National Database containing all the data from the Regional Databases.

On a daily basis the data from the Regional Databases are transferred automatically to the National Database.

The Database Management (both for the National and the Regional Databases) is with ADABAS(C) on SUN SPARC servers with Operating System Solaris 2.5.1 Sun OS 5.5.

The development language is NATURAL.

The occupied space at the National Database is 5 Gbytes with UNIX compression.

The communications go through dedicated lines under TCP/IP Protocol.

Currently a copy of the National Database is being made with SGBD ORACLE on a HP-K460 server, with Operating System S HP-UX. The development language is standard C with embedded SQL sentences.

An automatic process does the transfer of data from the ADABAS to the ORACLE versions of the National Database so as to have direct access to the data from a Geographic Information System Arcview with Spatial Analysis and other GIS tools.

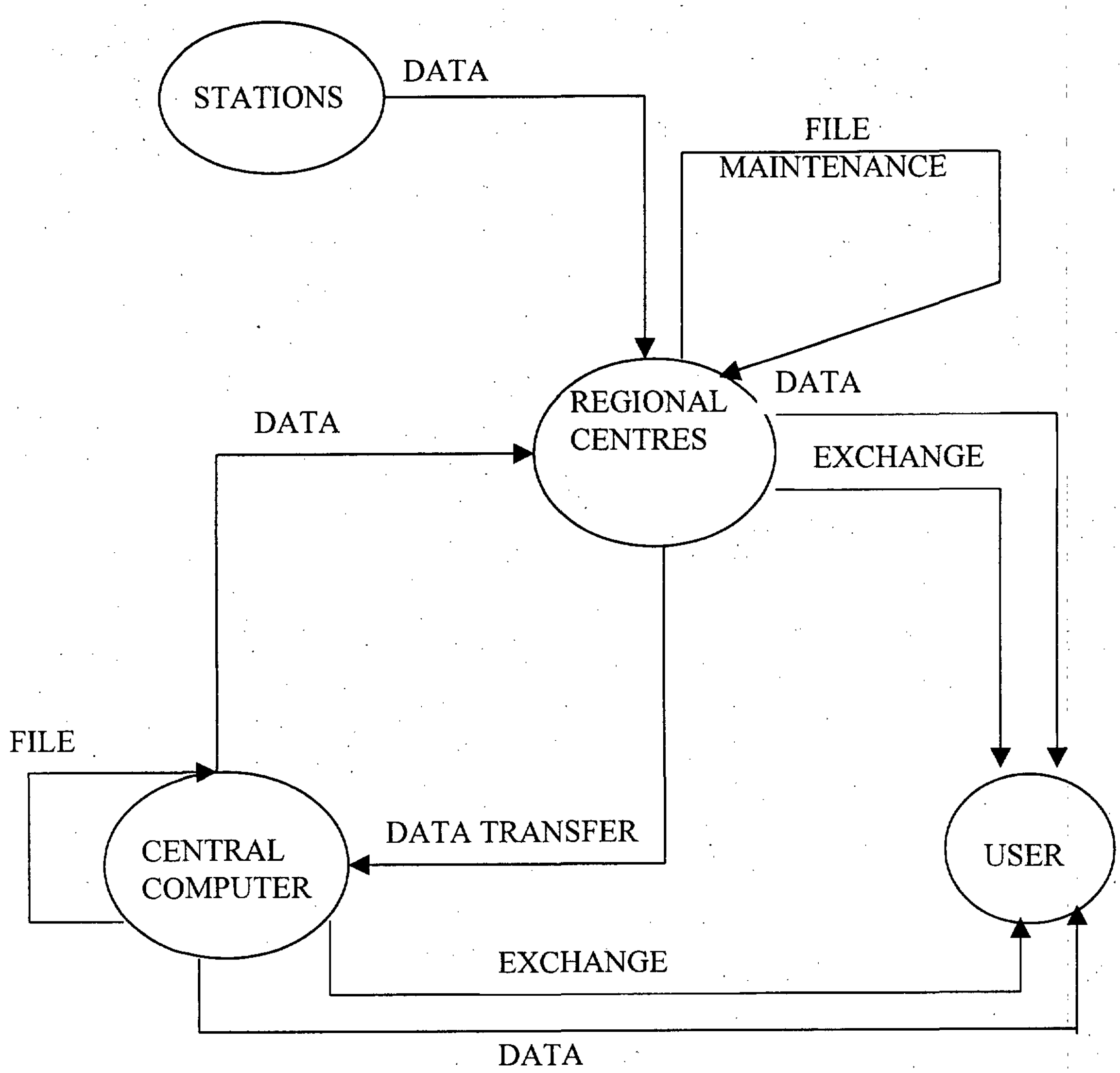
Climatological Database Files (ADABAS)

CLBFATEP (91)	DATOS (TEP)-	Contains the daily data for pluviometric stations. Each record has data from all the month.
CLBFAEFE (92)	EFEMERIDES	Historic extreme values for each variable and station. There are monthly and annual extreme values, with a periodic group for the monthly values
CLBFARES (92)	RESUMENES	For stations with only monthly data (TT and RR)
CLBFAEST (95)	ESTACIONES	Data about the stations (metadata)
CLBFACOM (98)	DATOS (COM)	Hourly or at fixed hours data from principal stations

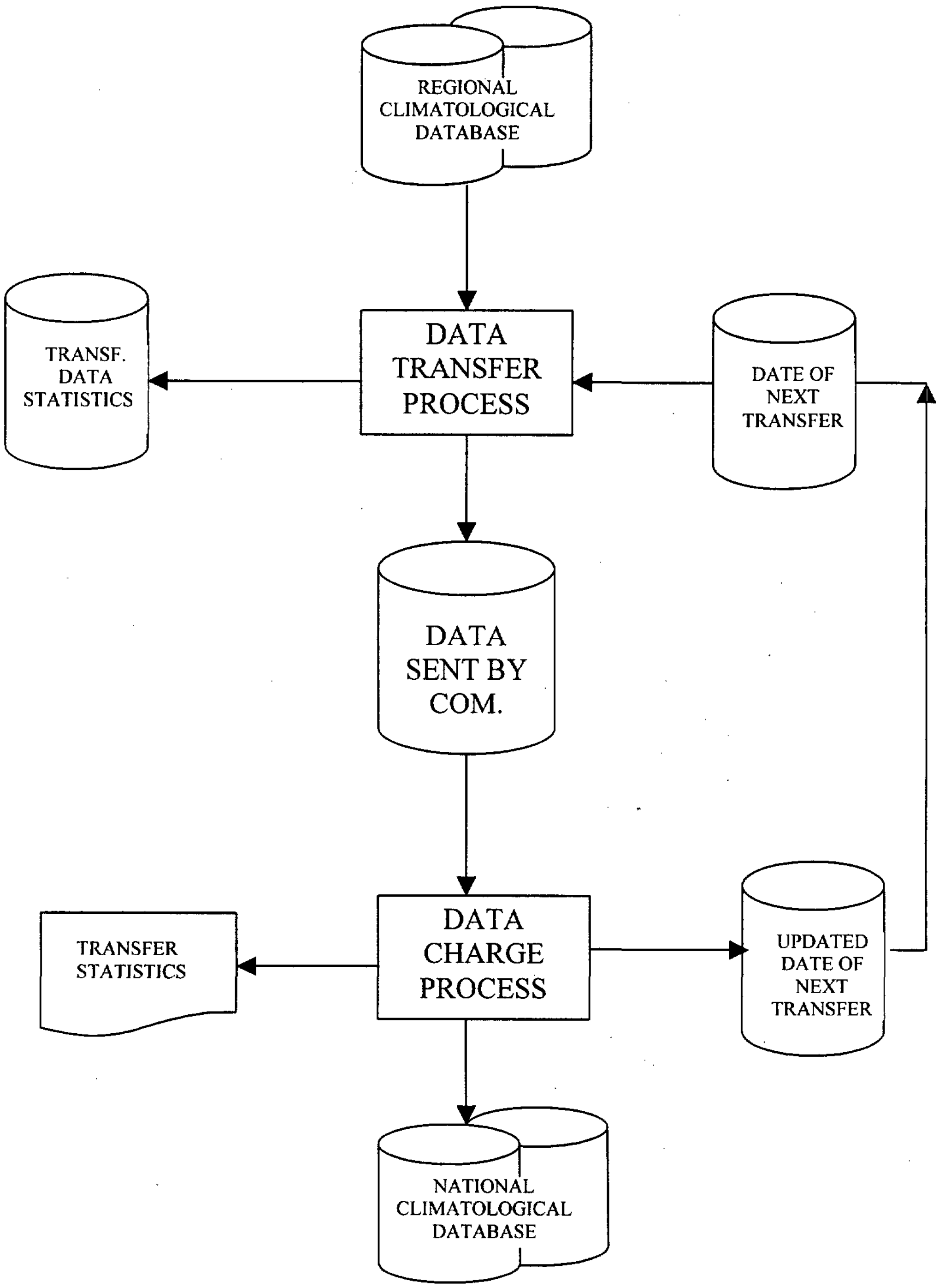
Table List

VARIABLE	Meteorological variables codes
TIPOS DE ESTACIONES	Station type
CUENCAS	Hydrological basins
CUENCAS PARCIALES	Partial basins
NUBES	Cloudiness codes
METEOROS	Codes for meteorological phenomena
GESTORA	Codes for ownership of the station
CODIGO POSTAL	Postal codes of Spain
MUNICIPIO	Municipality codes for Spain
PROVINCIA	Province codes
CENTROS	Regional centre codes
AUTONOMIA	Spanish autonomies (political regional units)
INSTRUMENTOS	Codes for meteorological instruments
EMPLAZAMIENTOS	Type of location codes
PERSONAL	Codes for personnel
METODOS	Types of statistical methods for filling missing values
USUARIOS	Kinds of access to the database

GENERAL STRUCTURE OF THE DATA FLOW



DATA TRANSFER PROCESS FROM THE REGIONAL TO THE CENTRAL DATABASE



MIRAKEL,
THE CLIMATOLOGICAL DATABASE SYSTEM AT DWD

Johannes Behrendt
Deutscher Wetterdienst (DWD)
Postbox 100465, D-63067 Offenbach, Deutschland

Phone: +49 69 8062 2955
Fax: +49 69 8062 3182
e-mail: johannes.behrendt@dwd.de

ABSTRACT

The climatological database system MIRAKEL implemented at DWD is based on a Silicon Graphics (SGI) data server holding an ORACLE 8 database. The data model for observational data as well as for derived data is described shortly. The volume of the German climatological database (including historical data and metadata) sums up to about 60 GB (including indices, with a yearly increase of 6 GB) in 112 tables (including 19 tables with more than 1 Mill. rows). The application system for importing and retrieving large data volumes is based on SQL and 3-GL-programs (C, FORTRAN). The ORACLE-tools DESIGNER, DEVELOPER and ORACLE BROWSER are used for interactive applications and ad-hoc-queries. Some aspects of quality control description and international data exchange are discussed in detail.

1.0 INTRODUCTION

The Deutscher Wetterdienst (DWD) as German National Meteorological Service has the responsibility for the collection of national climatological data. All relevant climatological data, including - as far as possible - historical data dating back to the last century, have been stored on computer readable archives organised as time series. During the last years new data types like data from automatic stations, numerical models and remote sensing systems with increasing data volumes are becoming more and more important for climatological applications. Also, the requirements for the availability of data are increasing rapidly. To meet these new challenges the former file orientated database system was changed towards a modern data management system using a relational database management system (RDBMS), the related tools and new communication technologies like internet/intranet.

In 1995 DWD has established a project to implement a RDBMS for all meteorological data, called MIRAKEL (" Meteorologische Informationen in einem RDBMS ablegen und komfortabel und effektiv lesen"). The project has been realised in close co-operation of the departments for Technical Infrastructure (TI) and Research and Development (FE). At the end of this project, we are able now to handle the real-time data used for operational weather forecast as well as data used for climatological applications with help of the same database management system. The implementation for the national climatological data and the related metadata reached its operational state in 1998. After installation and population of the database, the important question is been discussed, how to meet the requirements of retrieving and maintaining the database with suitable applications. Of special interest are the interfaces between database system, web–applications and Graphical Information Systems (GIS). Of special interest are the implementation of suitable quality assurance procedures and of procedures for an effective international data exchange.

2.0 THE NATIONAL DATABASE KLIDABA

The national climate database (KLIma DAtenBAnk, KLIDABA) contains the climate data of Germany collected from the station networks of the DWD. The KLIDABA contains data from about 4500 stations with different observation programs. The data are generated from different station types: fully automatic stations, automatic stations with observer and stations without any automatic data registration. The number of meteorological parameters being measured or observed varies from up to 5 at specialised stations (for e.g., precipitation or wind) to about 20 at ordinary SYNOP-stations. The station networks are characterised by the number of stations, the time interval between the observations and the meteorological parameters being observed:

- about 10 upper air stations, 4 obs/day,
  values of the WMO codes TEMP and PILOT
- about 120 SYNOP stations, 24 obs/day,
  values of the WMO code SYNOP
- about 600 climate stations, 3 obs/day,
  values similar to SYNOP
- about 4500 precipitation stations, 1 obs/day, precipitation (incl. snow)
- about 120 soil temperature stations, 3 obs/day
- about 100 - 200 stations in special networks, 24 obs/day, values: wind, temperature, relative humidity, sunshine duration
- about 100 automatic stations, 1 obs/10 minutes,
  subset of SYNOP values

In addition to these observational data, derived data are stored, e.g., monthly values (means, totals, extremes) and monthly means for 10 or 30 years (normals).

The migration of the KLIDABA has been accomplished for all data sets (except 10-minute-values, aerological data and SYNOP-data before 1990). The present data are available in MIRAKEL since January 1998. The integration of additional data (e.g., maritime data, phenological data) is on the way

3.0 THE IMPLEMENTATION OF MIRAKEL

The entity relationship model (ER model) for the climatological data in MIRAKEL can be divided into two major parts: **data** and **metadata.** The model for the data (Fig. 1) consists of entities for observations and for derived data. The names of tables and attributes are originally in German but are here for better understanding in English notation. The climatological data are grouped into single-level-observations like surface data and multi-level-observations like upper air data and soil temperatures. Entities for derived data like daily data, monthly data and normals complete the scheme. The data are identified by primary keys containing:

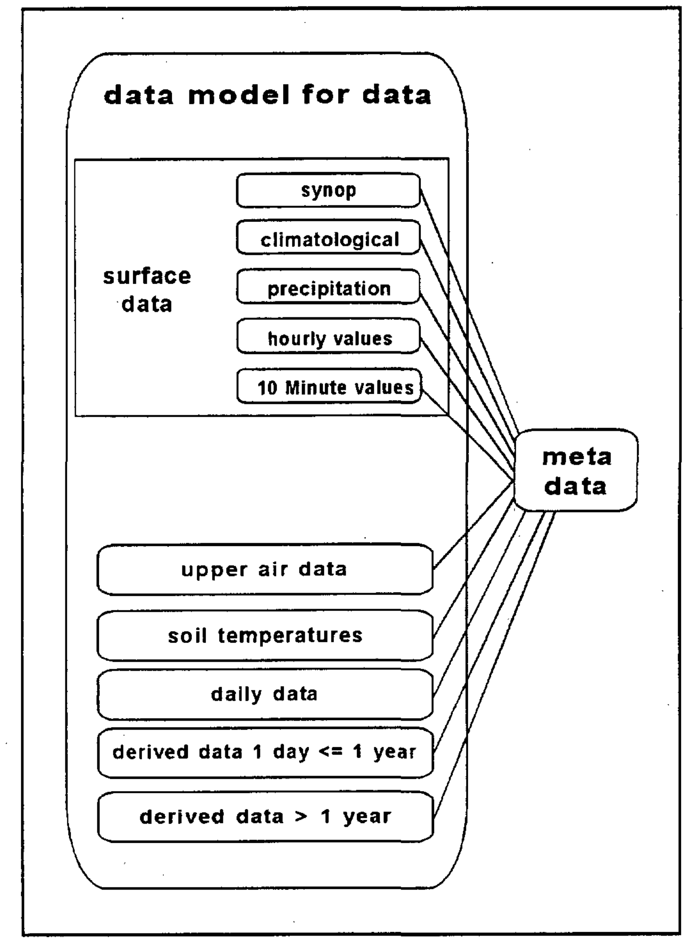


Fig.1: Data model for data

For all entities

- station_id
- date/time
- data_set_id
- coding_id
- quality_level_id

The **station_id** is an unique identification for any observation station. The **data_set_id** identifies the data source. The **coding_id** identifies differences in a data set due to changes in the coding or in the rules for the observation procedures. With the introduction of a **quality_level_id** it is possible to keep data of different quality levels in the database, e.g., original and corrected data. The key entities for the metadata are 'station' and 'data set'. Information like location, station name, station height, instrumentation or data source can only be derived from these metadata tables using the foreign keys **data_set_id** and **station_id**.. For some entities additional attributes for the primary keys are needed:

multi level data:

- level (depth or pressure)

derived data with integration > 1 day and <= 1 year (e.g., monthly data):

- type of index (e.g., mean, extreme, sum, number of days with, ...)
- period (e.g., week, decade, month, season, year)

derived data with integration > 1 year (e.g., normals):

additional to the primary key for derived data <= 1 year:

- type of integration over the years (e.g., mean, extreme, ...)
- integration period (e.g., 10 years, 30 years)

All other attributes are defined by a list of meteorological elements (e.g., temperature, wind speed) and the related quality bytes (e.g., temperature_qb, wind_speed_qb).

DWD has decided to use the ORACLE RDBMS and the related ORACLE tools. The MIRAKEL system has been implemented on a client/server environment with SGI servers and clients using X11, WINDOWS NT and WINDOWS 3.11. All here described data are stored in central databases. A remote access from the regional offices to the central databases is possible.

The volume of the German climatological database (including historical data and metadata) sums up to about 60 GB (including indices of about 30 GB) with a yearly increase of about 6 GB. The data are stored in 112 tables including 19 very large tables of more than 1 Mill.rows. The largest table has 90 million rows with totally 4 GB

4.0 THE APPLICATION SYSTEM OF MIRAKEL

DATABASE ACCESS

One of the reasons for the choice of ORACLE was the system of tools which ORACLE offers for the development of applications. These tools, range from ones for the database design (DESIGNER) to tools for the development of interactive applications (DEVELOPER including FORMS, REPORTS, GRAPHICS and PROCEDURE BUILDER) and ad-hoc-queries (ORACLE-BROWSER). They allow a fast and comfortable software development but are highly proprietary to ORACLE standards.

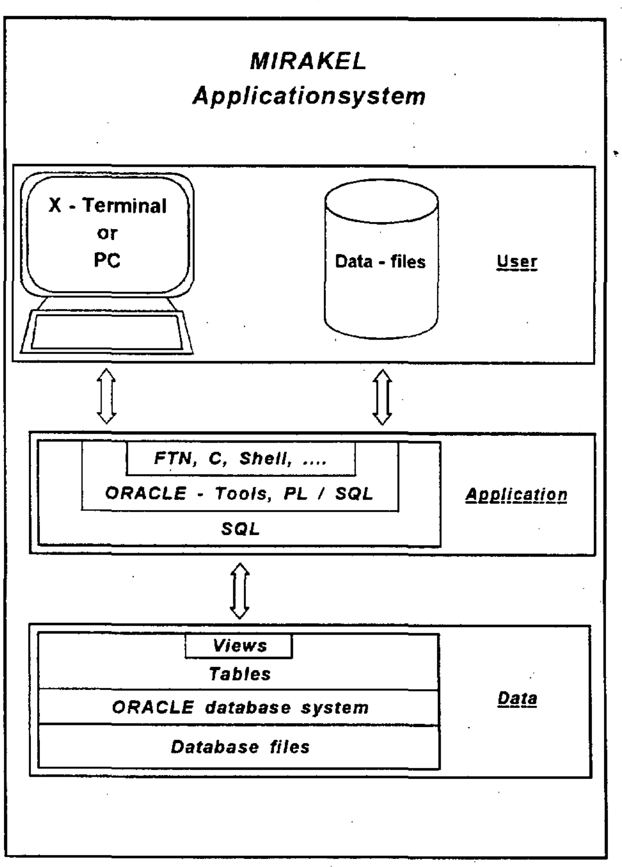


Fig. 2: Applicationsystem

The interactive access to the MIRAKEL database is realised by applications developed with DEVELOPER 2000, ORACLE BROWSER and/or SQLPLUS. At present the operational web-applications are realised by using PL/SQL and appropriate packages.

A typical application in the operational business is the processing of large data volumes. Since we are still using the existing old software systems for quality control and climatological evaluations with an interface to sequential files, we have to satisfy these file interfaces also for data retrieved from the MIRAKEL database. We retrieve large data sets with optimised SQL-scripts and do the time consuming sorting, joining and formatting of the result with 3-

GL-programs and shell-scripts. The performance for the retrieval of large data files (of more than 100 000 rows) from the MIRAKEL database is now acceptable with more than 20,000 rows per minute.

The population of the database has been accomplished by an application system using shell-scripts and C-programs with embedded SQL. All data, historical data coming from the old time series archive as well as actual data collected from the operational station networks are processed and put into the MIRAKEL database with the same software system.

Graphical applications like maps or presentation-graphic may be done using the DEVELOPER-tool GRAPHICS, but we decided to use the standard graphic software at DWD named Interactive Data Language (IDL) to meet the requirements on graphical presentation of climatological data. As GIS-software we will use ArcView (with probably some add-ins). The SDE (Spatial Data Engine) will serve as interface to the ORACLE-database system. The development of a powerful graphic interface to MIRAKEL is still in the conceptual phase.

After having implemented an application system which is capable to handle the operational needs we can step forward now to new applications which benefit more and more from a modern database system. Some of the major requirements to be met are:

- Support of quality assurance with FORMS-applications providing information not only of observations but also of model output fields and remote sensing data.
- Graphical data presentation with IDL and ArcView
- Support of a common web-based exchange facility for climatological indices

Last but not least, a lot of optimising and tuning is necessary to meet better the requirements in the performance of loading and retrieving large data volumes.

QUALITY CONTROL

The quality description of data in MIRAKEL is done by the **quality level** related to each row and by the **quality bytes** related to each attribute in the rows. The **quality level** describes which quality control procedure has been applied to the data. The quality level changes by applying a quality control procedure, irrespective of finding any errors. The **quality bytes** describe whether an individual value has been found suspicious, has been corrected or has been checked without finding any problem. The data are corrected by FORMS-applications. The quality bytes are set by comparing the corrected data with the uncorrected data.

DATA EXCHANGE

The exchange of climatological data is normally performed by sending files with extracted data and corresponding file descriptions or by using the Global Telecommunication System (GTS). e.g., for data in the WMO-formats SYNOP, TEMP, CLIMAT or CLIMAT/TEMP .Up to now at DWD a direct access to the climate database by web-applications is supported only for metadata like data set descriptions, station descriptions and 'what is available'.

A direct access to the databases needs some preparations if you do neither want to learn the specific properties of the different database systems nor to open the operational database systems for an external access. Since the national database systems are tailored for the specific needs (language, structure of historic and actual data) it will not be possible to achieve one common system for all climatological data. But, it may be possible to define a common data warehouse with a defined data structure and a data dictionary describing the data. This data warehouse receives input from the national databases and could be physically distributed. The user interface should use web-technology.

5.0 SUMMARY

The climatological database system MIRAKEL is operational and contains (besides other data) almost all German observational data. The database structure, the quality control procedures and the application system is tailored for the national needs and need still to be improved. The data exchange will be supported by GTS.A distributed data warehouse using web-technology, a common data structure and a common data dictionary is proposed.

CLIMATOLOGICAL INFORMATION AT THE PORTUGUESE METEOROLOGICAL INSTITUTE

Luís Cardoso and Maria de Fátima Espírito Santo
Instituto de Meteorologia
Rua C - Aeroporto de Lisboa
1700 Lisboa
Portugal
Phone (351)1 8483961
Fax (351)1 8402370
e-mail:luis.cardoso@meteo.pt
fatima.coelho@meteo.pt

ABSTRACT

In 1996 the Portuguese Meteorological Institute (IM) has chosen the RDBMS INFORMIX to manage some of its climatological data. The application "BDMN" (National Meteorological Database) covers, so far, the daily and the hourly surface information, total ozone and sulphur-dioxide (Brewer) since the beginning of the measurements took place (1923 and 1972, respectively). The main characteristics of the IM stations are also available. The data is stored in a database called "synopdb". "Synopdb" is located on a dedicated Unix server (Alpha Server 8200 DG/UX). It is accessible, up until now, only to IM users through SQL and an application developped using INFORMIX-Dynamic4GL. This application allows users to access data in character mode, GUI mode and Web browser mode using the same executable program. Data is loaded into the database through pre-defined procedures (daily and hourly surface information) and whenever available (the remaining information). The next steps should be to improve quality control of available data and to store data from automatic stations.

1.0 INTRODUCTION

The Portuguese Meteorological Institute (IM) is responsible for the storage of current and historic climatological information. Aiming to improve the availability of that large amount of data, making it easier and faster to access on-line, IM has decided to purchase a hardware and software solution, easier to administrate, operate and to gradually replace the existing data archives.

2.0 THE CLIMATOLOGICAL DATABASE

The portuguese climatological database system, **"BDMN"**, includes a database that we called **"SYNOPDB"**, several tables, programs, procedures (either Unix scripts and stored SQL procedures) and an application that is the interface between the users and the database server.

INFORMATION STORAGE

"**BDMN**" has its historical information source at the portuguese meteorological archives, stored under *traditional sequential indexed files* and accessed via third generation languages - mainly FORTRAN routines.
At the first stage the following information was agreed to be loaded into the database:

- Brief information about the main characteristics of the portuguese meteorological stations;
- Surface hourly information (available since 1923);
- Surface daily information (available since 1923);

Surface hourly and daily information is stored in two ways:

- Real time, for the information that refers to the synop. Hourly previously defined procedures load automatically data into the related tables.
- Deferred time, at the end of each month, for the remaining information keyed from paper forms.

FIELDS ADDED

Beyond the existing fields in the original archives, additional fields were defined in the tables to store information about quality control, the information source (real time or deferred) and, the type of network it is coming from (classic meteorological stations or automatic stations).
The information quality control is performed at two levels:

- On real time information (synop) establishing acceptable boundaries for each parameter, according to WMO's "*Guide on the Global Data Processing System*"
- On information, which was gathered on paper forms, last entered at the end of each month, through the systematical application of programs of quality control and just before it is loaded up to the database.

OTHER DATA

Information about total ozone, sulphur-dioxide (Brewer) and raw data from automatic stations was also loaded.
We have information about total ozone and sulphur-dioxide (Brewer) available to users (years 1972 and 1975 through 1998).
Raw data from automatic stations is not yet available to the users (years 1995 through 1998 are stored).

ACCESSING THE "BDMN"

The information mentioned above is available, so far, on-line, only to IM users. To other people, interested in that information, it is made available by request. So, IM users can access the information either:

- Via **character-based terminals** - Unix environment, through SQL commands (interactive mode or through stored procedures), programs developped using

INFORMIX programming tools (such as I-4GL, I-D4GL, I-ESQL/C, I-ESQL/COBOL);

- **GUI interface** - client/server system using TCP/IP, running an application, without modification, to the source code developped using INFORMIX-Dynamic4GL, across client platforms including Windows 3.11, Windows9x/NT and the Web (HTML and, later this year, Java) clients.

DISK SPACE

Disk space used by 76 years of daily and hourly information is about 500MB and 1.1GB and the monthly increase of this space is about 0.6MB and 1.5MB, respectively. Information about total ozone and sulphur-dioxide (Brewer) uses about 15MB of disk space and raw data of automatic stations about 650MB. This last information is not yet handled.

3.0 HARDWARE AND SOFTWARE

The database server is a DEC AlphaServer 8200 with 2 CPU, 512MB RAM and 30GB of hardisk.
The operating system is DIGITAL/UNIX V 4.0E.
The RDBMS is an INFORMIX DYNAMIC SERVER V7.22 and the following tools - I-SQL, I-ESQL/C, I-ESQL/COBOL, I-4GL, I-Dynamic4GL. For pre-processing purposes, we have installed MicroFocus COBOL.

4.0 THE NEAR FUTURE

We think that the next steps are the following:

- Hiring more qualified staff to work on this project
- Improving data quality control
- Making available automatic stations information
- Installing the most recent versions of INFORMIX software

CENTRALIZING THE RETRIEVAL OF ARCHIVED DATA AT KNMI

Jan Jans
The Royal Netherlands Meteorological Institute (KNMI)
PO Box 201, 3730 AE de Bilt, The Netherlands

phone: +31302206644
fax: +31302210407
e-mail: jans@knmi.nl

ABSTRACT

At KNMI various data is collected and stored for real-time retrieval by meteorologists and for the purpose of applications. KNMI has started projects to centralise this data in a few main databases and make it accessible by one userinterface using metadata for every authorised user. For radar- and satellite-images we are realising a new archiving system now. The archiving of the hourly synoptic data from the national observational stations is well organised by the Climatological Information System (KIS). This is a modern Oracle database with good validation and retrieval facilities. Other climatological data is, due to separated developments in the past, not archived in an organised manner or archived in isolated systems. Only dedicated employees can retrieve this data. Maintenance and system-management gives here also cause for concern.
Recently the Climatological Services have started a project to improve the accessibility of these data. At the moment we are exploring the ways we can realise this aim.

1.0 INTRODUCTION

The tasks of the KNMI are dramatically changed at the beginning of 1999.
All commercial tasks (radio, TV, newspapers, farmers-, traffic information, etc) are take over by commercial meteo services. The KNMI will be a data and research centre.
Important tasks are now:

- collecting data
- storing and archiving data
- providing and distribution data and basic products, according to the ECOMET rules
- research and development
- some public tasks (for national security)
- climate advise

Besides KNMI has to reduce the costs of personal by automation as much as possible.
For instance by automation of the manually observations and the production of forecasts.
Because these developments, the role of research and development will be more important.

Consequences are:

- other requirements for retrieval,
- more interest in the content of the data,
- the exchange of data witch other research departments will be more important.

KNMI will anticipate for these changes by realising a new design for the entire KNMI data infrastructure. For this raison KNMI has started the HOPWA-project.
In the next pages we will explain more details about :

- this new concept for the KNMI infrastructure
- a current project, the realisation of a image archive
- a the existing system, the Climate Information System (KIS)
- and a project that had just started :
  The open question how to make accessible the other climate file-systems and archives at KNMI

2.0 DESIGN CONCEPT FOR THE KNMI INFRASTRUCTURE

KNMI has started a project (known HOPWA) to redesign the entire KNMI data infrastructure by centralise this data in a few main databases :

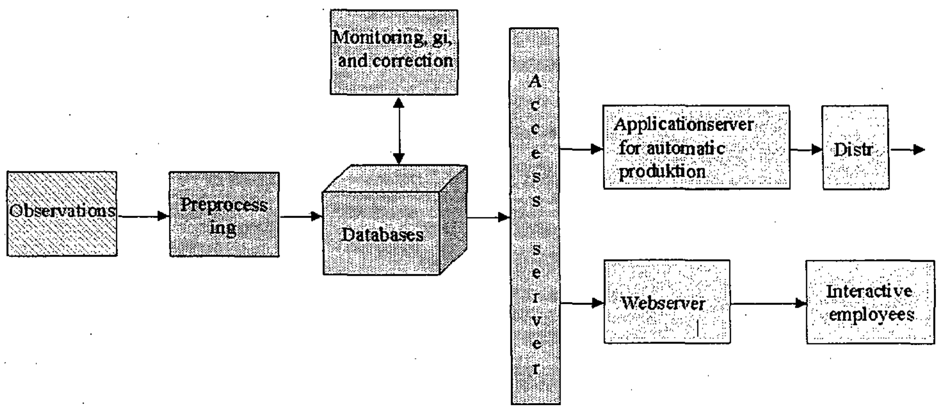
- real-time observations (flat file)
- time series and model data (oracle db)
- images (radar, satellite, oracle db)
- forecasts
- aviation (oracle db)
- the climatological archive KIS (oracle db)

The data of all this main databases must be available for all authorised users at KNMI
The access of these data will be organised by one interface that arranged the querying, authorisation, the priority and the accounting. The role of the metadata is very important in this process. Where possible we will generate the metadata automatically.
The metadata will be stored in a structured way.
All new databases at KNMI will be based on RDBMS technology like ORACLE.
New interactive user-interfaces will be realised witch Web-technology.
For data exchange international standards are used:

- for metadata the ISO-TC211 standard or the European CEN TC 287 standard (CEN = comite europeen de normalisation)
- for interfaces SQL, JDBC, ODBC, OLE-db.
- for data formats HDF, GEOTIFF, GRIB and other WMO-formats.
- for development JAVA, HTML, XML, ORACLE Webdb, PL/SQL, Visual basic (webclass), Perl, GIS-components like Mapobjects, SDE, ARC-explorer.

KNMI will separate the functionality in different layers for databases, access, application and user-interface.
In the HOPWA project we will realise an automatic production process step by step.
We started by organising the databases.
The databases for climate data, aviation data and model data are already well accessible ORACLE databases.
As a prototype for the new concept we are now realising an image archive with ORACLE-8 .

The real-time observation databases is not yet a RDBMS. We must replace this database in future. A product or forecast database does not exist at the moment. We have to build it in future.



Concept for the KNMI data infrastructure
(The HOPWA-project)

3.0 THE RADAR- AND SATELLITE IMAGE ARCHIVE PROJECT

As an example for the new developments we describe a current project, the realisation of an image archive. The role of metadata, access by webbrowsers and the image format HDF (= Hierarchical Data Format of the national centre for supercomputing applications, US) are essential in this system. People can browse in the indexes, the metadata and see the quick looks with the web browser.

There are three storage levels for the images :

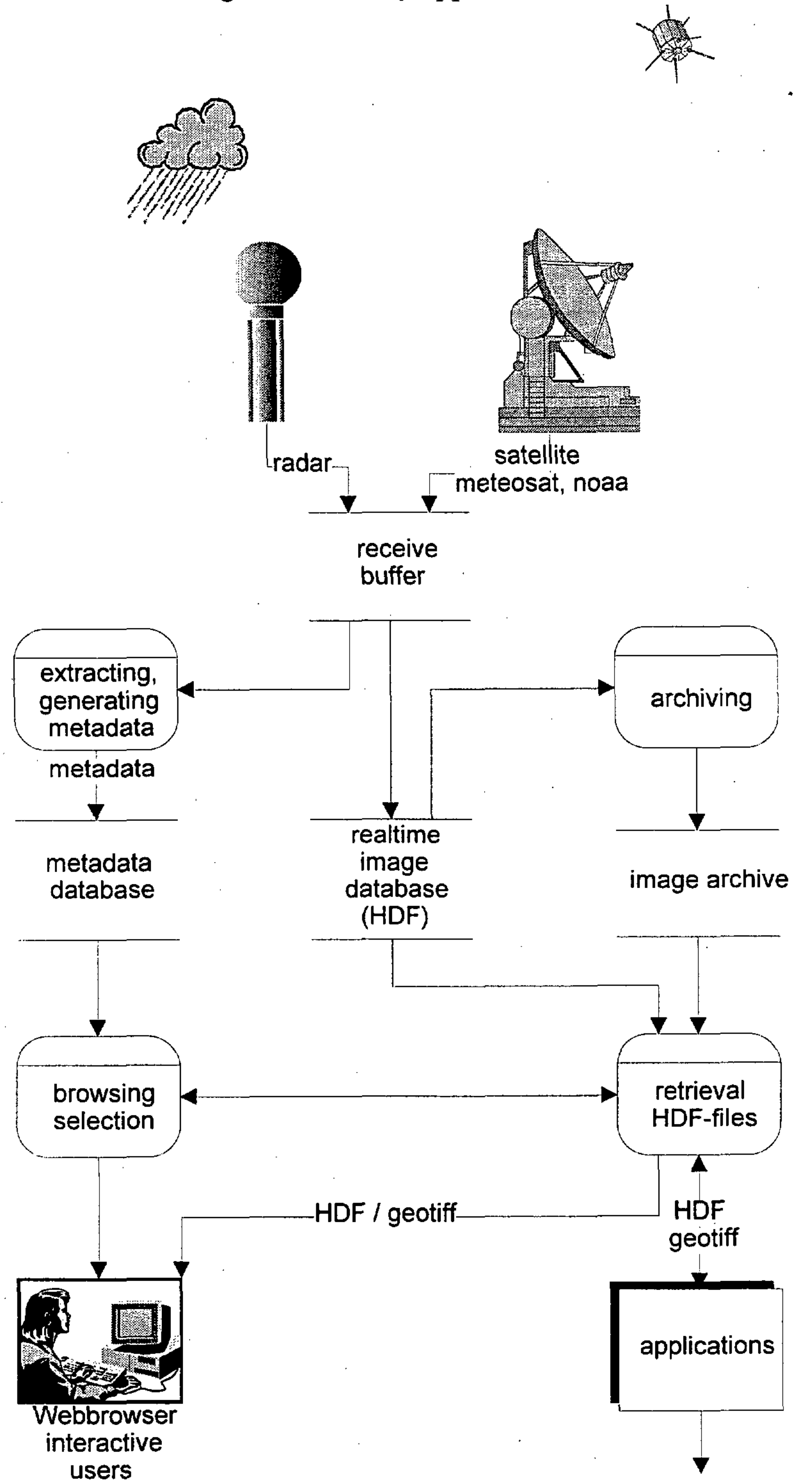
1. A circulating real-time database (14 days)
2. A short term archive database (3 months)
3. A long term archive

We are using international standards as ORACLE 8 and SQL for the databases, ISO-TC211 (CEN) for storage of the metadata, HDF for storage of the images and GEOTIFF for import in GIS applications.

Functionality's are:

- Receiving the data from the radar- and satellite-image process systems.
- Possible converting the data-format to HDF
- Extracting and generating of the metadata
- Storage of the metadata in an ORACLE database
- Storage of the image-data file in the database
- Periodical archiving of the image data to tape robot
- Users are allowed to browsing indexes, quick looks and other metadata by web browsers.

- After selection they can retrieve of the image data in HDF or GEOTIFF
- Retrieval of the image data files by applications in HDF or GEOTIFF



The radar- and satellite-image archive project

4.0 THE CLIMATE INFORMATION SYSTEM KIS

An example of an existing modern database at KNMI is the Climate information system KIS. This is a client sever application originally build with the INGRES RDBMS. KIS was operational in 1993. Because there where problems which the INGRES organisation, KNMI was forced to look for a other RDBMS. KNMI selected ORACLE. KIS has been converted to ORACLE-7 in 1997.
The KIS database is primary built for use by advisers of the climatological service division KIS distinguish validated and not validated data
KIS Contains:

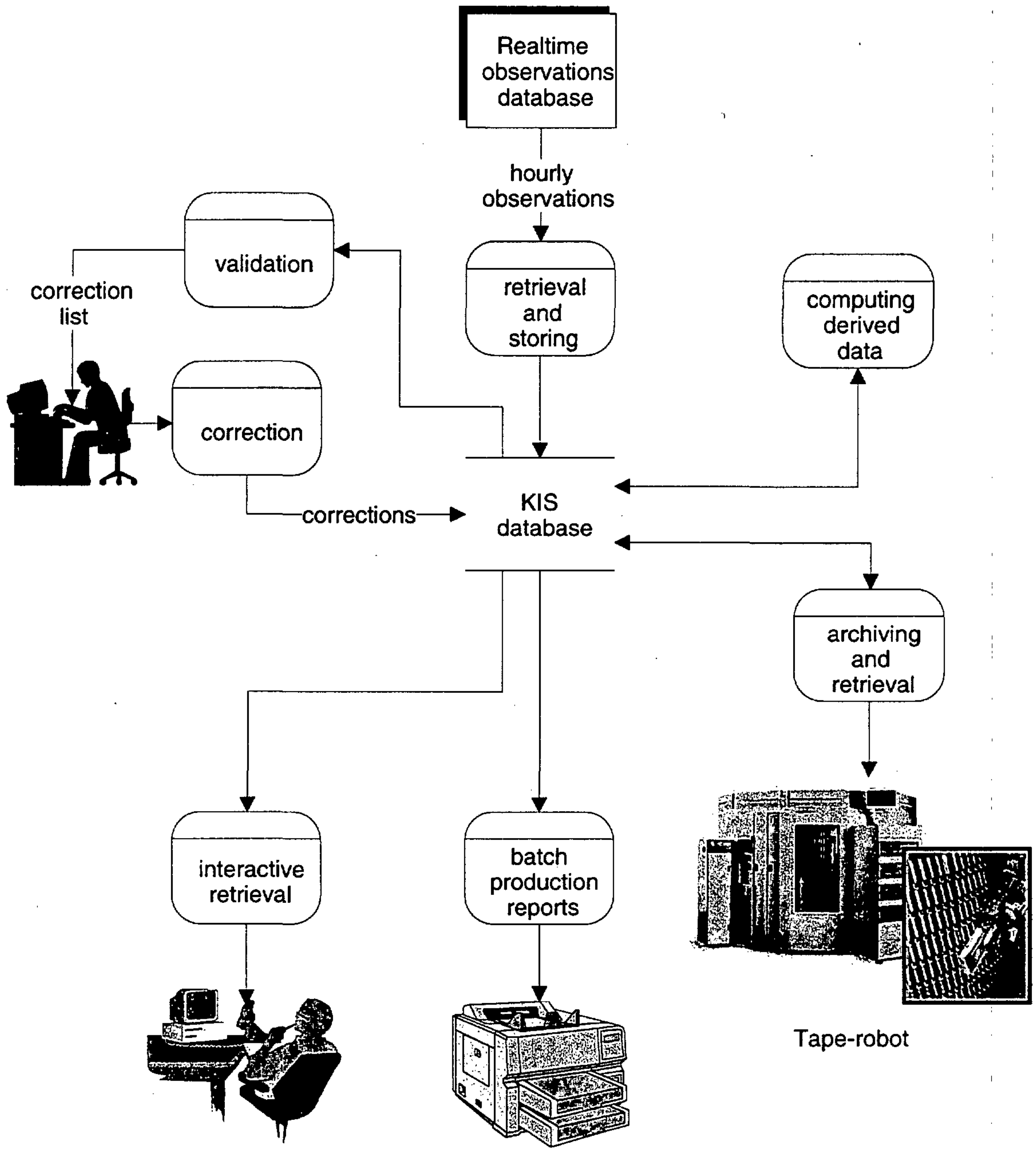
- Metadata concerning stations and observations
- Hourly synoptic data
  - 5 year online
  - older data is stored in the tape-robot, but accessible with KIS
- Derived data for day-, decade-, month- and year values are, depending of the station, online available from 1850
- The normal values

KIS has facilities for:

- Daily automatic retrieval of hourly observations from the real-time observations database
- Storing this data in the database as unvalidated data
- A validation process that produces correction lists
- manual correction of the hourly data in the database. And when this is done the data is stored as validated data.
- Computing of derived data at several levels (day, decade, month, year, extreme, average, highest, lowest, etc.)
- Archiving to and de-archiving from the tape-robot of hourly data older then 5 years.
- Generating of reports
- Daily batch production of several reports
- Retrieval by user interfaces (day-, week-, month-view) and SQL

KIS is a concept from 1990. So we have to update the KIS system for the new requirements, for example

- Upgrade to ORACLE-8 or 8i
- Extend the interfaces which more graphical retrieval possibilities (GUIs) using GIS and Web technology.
- Possibilities for easier access for long term timeseries
- redesign the datastructure and make the metadata more visible by separate this from the data and store it in a standard structure
- and we will make KIS suitable for storing and retrieval of other sorts of data by extending the database



The Climate Information System (KIS)

5.0 OTHER CLIMATE FILE SYSTEMS AND ARCHIVES AT KNMI

Until now we described the new concept at KNMI, running projects and well organised databases at KNMI. But there are also numberless archives as isolated file systems on servers or PC at KNMI. Most employees can't get data out of don't even know the existence of these information systems.
At the climatological services we have several other climate file-systems and archives, as:

- The Precipitation radar-image archive
  These are radar-image files in a specific, not regular, format.
  These files are stored on a server and accessible by an old application on a PC.
- Archive of lightning (ASCII files on a server)
- Historic forecasts (files on a PC)
- 10-minute sample observations (files stored in on the tape-robot)
- Maritime data set (file-system)
- Unstructured data (as old tapes, old land- en maritime observations on paper)

These data is difficult accessible and the problems here are :

- there are several sorts and formats of data
- it are isolated systems
- there is less or no information about the content of the data
- only a few dedicated employees can retrieve the data
- only a few dedicated employees can maintain these systems
- no indexes or metadata are available

The climatological service has the intention to make this data accessible for all employees using one interface, like KIS. We will to organise these data in one or more open databases and store these data in a structured way, using metadata and other standards.
Recently we start a project to realise this. Questions have to solve in this project are:

- How to make this data accessible and available for all KNMI employees, at least at the climatological service?
- How to integrate these isolated systems in the data-infrastructure, like KIS ?
- Which data will we archiving (archiving strategy) ?
- How will we organise the system management ?

QUALITY ASSURANCE OF DATA FLOW AND SYSTEM MANAGEMENT AT DNMI

Petter Øgland
The Norwegian Meteorological Institute
Postboks 43, Blindern, N-0313 Oslo, Norway

phone: +47 22 96 30 00
fax: +47 22 96 30 50
e-mail: petter.ogland@dnmi.no

ABSTRACT

As computers are gradually taking over more and more of what has been traditionally handled by manual means in the area of data collection, database management and production of climate overviews and statistics, the consequences when something goes wrong may be significantly more severe in the digital and automatic environment than it was when everything was controlled by experts with hands-on experience of the various stages in the processes. In this paper we present some of the main structures of quality assurance, such as statistical process control (SPC), that have been applied at DNMI in order to put the present processes under control and help improve the processes as we gradually implement methods of artificial intelligence (AI) in order to accommodate the demands of the future.

1.0 INTRODUCTION

From our perspective, a possible future scenario at DNMI might look something like this:

- The weather stations networks consist exclusively of automatic weather stations (AWS)
- Standard data processing, such as quality control, correction and interpolation, is completely automated.
- Standard climate overviews and statistics are automatically prepared and are available for free on the Internet.

The human intervention in the observations and climate services, in this context, should then be reduced to system monitoring in the aid of system quality control and improvement, [1] and [2], and by making computer programs that simulate the human way of performing the routine tasks by use of techniques of artificial intelligence [3].

In this paper we give some examples on how a general management method based on these principles, [4], has been applied to the area of quality assurance of data flow and systems management of meteorological observations.

2.0 A SYSTEMS VIEW ON DATA PROCESSING

One of the basic principles of modern statistical quality control is that the process should be described from a systems point of view, [1], and one of the most important steps in this approach is to identify the input/output channels for the system, and to measure quality at these places which, in our case, have been identified as three main interfaces.

- DATA COLLECTION
- CLIMATE PRODUCTS
- PROCESS CONTROL AND IMPROVEMENT

Our type of data collection consists of real-time data, climate data and meta-data. The quality of observations and the quality of the routines for each type of data collection may have direct impact on how well we are able to perform quality assurance during the data processing stage.

The main output channel from the data processing stage is the climate products, which may also include types of statistics used within meteorological forecast products and applications for internal or external research projects. The quality of each product may be strongly related to the quality of performance during the data processing stage.

The third type of interface is the continuous monitoring of the data processing by use of methods of process control and process improvement in order to check that the quality of performance during the data processing stage is stable and find ways in order to make it improve.

3.0 COLLECTING REAL-TIME DATA

At present there are four real-time data systems monitored by methods of statistical quality control at DNMI. All the systems are operated by the crontab scheduler on the database mainframe computer.

- SYNO_INN
- META_INN
- PIO_INN
- AUTO_INN

The SYNO_INN system is responsible for collecting observations from 216 SYNOP stations, 166 national stations and the remains from neighbouring countries. The SYNO_INN program is executed every 10 minutes all day.

In the META_INN data collection system we collect observations from 58 METAR stations. The program reads binary files, similarly to the SYNO_INN program, and stores the observations in ORACLE data tables. The program is run every 3 hours.

The PIO_INN routine consists of reading observations from semi-automatic weather stations (SAWS) and stations that are electronically connected to the DNMI computer system by Personal Computers (PIO stations). This population consists of 16 PIO and 15 SAWS.

In case of the AUTO_INN system there are 64 automatic weather stations (AWS) collecting hourly observations into the KLIBAS database system at the moment. The program is run once every evening.

4.0 ON-LINE QUALITY CONTROL DURING DATA COLLECTION

For all of the four routines described above, some kind of elementary quality control, such as performing format checks, the simplest checks described in the WMO quality control procedures [5], are done. Problems are written to log files and statistical analysis is performed automatically.

In the statistical quality control (SQC), we have followed two main principles.

- Plotting Shewhart diagrams in order to monitor quality over time.
- Plotting Pareto diagrams in order to focus on the most relevant problems.

The total number of problems recorded for a program during each run or each day is handled as a stochastic process $X(t)$. Statistics are then plotted on a daily scale for the past 12 months in order to see whether the system is improving, getting worse or being under control. Empirically given upper and lower control limits are also plotted and used for deciding whether the process is under control or not.

A simplified Pareto analysis is performed by producing frequency lists, relating causes and consequences for each problem. In the case of programs like SYNO_INN and PIO_INN, the analysis consists of listing stations according to the number of problems registered by the program as related to each of the particular stations.

5.0 ON-LINE QUALITY CONTROL FOR CLIMATE PRODUCTS

What does it mean that climate products are of high quality? The answer to this question, from our point of view, is to satisfy internal and external users by developing computer programs designed to check the climate products in the same manner as we usually check products by manual inspection.

In order to develop this type of quality assurance, we have focused on two main climate products.

- KA_H_STAT (monthly climate overview)
- STATUT (weather statistics for the past 30 days)

As with the PIO_INN data collection format control, the CHECK_H_STAT program is run on a daily basis, collecting defects on the KA_H_STAT output according to specifications, sorting and logging the defects on file, and producing Shewhart diagrams and Pareto diagrams in order to focus on relevant problems and improve the process.

The same principle is used with the CHECK_STATUT program, automatically making quality control checks and producing sorted lists of problem stations in order to give directions to the general quality control data processing.

6.0 QUALITY CONTROL SPECIFICATIONS

Since 1992, development of quality control and quality assurance has been driven by participation in national and international projects.

- KLIBAS 7.2 (1992-93)
- FREYR (1996-97)
- NORDKLIM 1.2 (1999-)

In the KLIBAS project, run internally by DNMI, specifications for all types of meteorological data were collected, although with a special emphasis on raingauge data. The project was completed in May1993, [6].

The Nordic FREYR project continued focus on raingauge observations, based in part on previous experience in the KLIBAS project and similar projects in the other Nordic countries. The project was completed in February 1997, [7].

Using experience from FREYR and other projects, the more ambitious NORDKLIM 1.2 project is designed to look at quality control for all types of meteorological observations, although the main focus for 1999 has been on real-time data, [8-10].

Even though the daily work on improving the quality control systems is based on prototyping management, [3], and not the traditional waterfall model of software development, [11], the projects, nevertheless, contribute to thinking long term even though the action is short term, which seems to be a perfect blend in terms of our experience so far.

7.0 DATA PROCESSING MODULES

As a consequence of the FREYR project, the data processing at DNMI has been conceptually structured into three main components.

- ERROR DETECTION
- ERROR CORRECTION
- USER INTERFACE

Error detection in this context means identifying suspicious values or values in error, and then having these values filtered on to log files where various logical and statistical tests may be performed.

The error detection methods involve a great number of different types of tests, including gross validity checks, range checks, missing value checks, format/code checks, time-space consistency checks for a single parameter and inter-parameter consistency checks.

New tests are implemented all the time, in order to accommodate needs in the various routines, and the tests are mostly specifically written in order to suite the observations programme at each type of stations, adapted to whether, say, the observations are being made ever hour or only once a day or less.

As there has been an emphasis, in the NORDKLIM project, on the differences between automatic quality control (AQC) and human quality control (HQC), the user interface is treated as a singular module in the general data processing system.

8.0 CORRECTING REAL-TIME DATA

Automatic correction of real-time data is presently done only for the SYNOP/TELE database, and in the case of this routine we have found in convenient to separate between three types of error correction.

- STATISTICAL INTERPOLATION
- INTERPOLATION BY USE OF FORECAST DATA
- THE EXPERT SYSTEMS APPROACH TO DATA CORRECTION

In the case of statistical interpolation, we have based our approach on continuous refinement of a methods developed at DNMI in the 1970s, and which has been analysed and improved ever since, [12-13]. The method is used for interpolating missing values of air temperature, relative humidity, air pressure, cloud cover and precipitation.

In most cases, interpolation by this statistical method is superior to using estimates generated by the HIRLAM forecast model, but as the statistical method depends on the density of the station network and how much data being available at the time of check, in many cases the forecast data are invaluable.

The expert systems approach to data correction consists of coding functions that copy how simple error situations are handled manually by routine without subjective involvement.

9.0 THE CLIMATE DATABASE SYSTEM

As a consequence of the gradual automating of processes for handling real-time data, automated process on handling climate data must follow. At the moment there are four climate data routines being monitored by methods of SPC and AI.

- Manual weather stations data processing routine (ALV routine)
- Manual raingauge stations data processing routine (ALN routine)
- Automatic climate stations data processing routine (AANDERAA routine)
- Digital wind observations data processing routine (AERO/XVIND routine)

For the manual weather stations and semi-automatic weather stations (ALV routine), the quality control system of the 1970s continue to be maintained and updated, while daily statistics on quality flags are being collected. There are also plans for changing and modernising the routine, a main consideration for the internal DATRUT project [14].

In case of the manual raingauge stations (ALN routine), work related to creating a full automatic routine (ALN2) in parallel with the manual routine has started. Output from the manual and the automatic routines are being systematically compared.

For the automatic climate stations (AANDERAA routine), elements of statistical process control are presently being implemented as a part of the quality assurance of the data collection.

In the case of digital wind observations (AERO/XVIND routine), some elements of statistical process control have been implemented.

10.0 CONCLUSIONS AND PROSPECTS

The general positive experience of applying methods of total quality control (TQC) for data flow and systems management prospers mostly in the benefit of being able to make software development priorities on an objective basis.

By having the system logging all types of internal errors and warnings, and by performing a statistical analysis on the diagnostics data, we have introduced metrics that give daily numerical values on the general quality of the system, and thereby having easy access to visualise whether the system is improving or declining, and, most importantly, by performing a Pareto analysis, we can automatically and objectively focus on programs and modules that are fouling up the system.

Our development has been motivated by researching two areas.

- STATISTICAL PROCESS CONTROL (SPC)
- ARTIFICIAL INTELLIGENCE (AI)

The system is being modelled and evaluated on a continuous basis by SPC. All relevant problems are being logged, and objective analysis decides further maturity and development.

Problems identified by the process control are solved by first analysing manual solutions, and then the solutions are gradually formalised until they may be more or less described as algorithms to be run on the computer, indistinguishable from the human approach and thus satisfying the Turing criterion of AI.

REFERENCES

[1] Mitra, A., 1993: Fundamentals of Quality Control and Improvement, MacMillan Publishing Company, New York, 1993.
[2] Shearman, R.J., 1992: Quality Assurance in the observations area of the Meteorological Office, Meteorological Magazine, **121**, 1992, 212-216.
[3] Chaudet, H., Pellegrin, L., 1998: Intelligence Artificielle et psychologie cognitive, Dunod, Paris, 1998.
[4] Juran, J.M., 1979: Managerial Breakthrough, revised edition, McGrawhill, New York.
[5] Abbot, P.F., 1986: Guidelines on the Quality Control of Surface Climatological Data, WMO, WCP-85, TD-No. 111, Geneva, May 1986.

[6] Kjensli, P.O., Lystad, S.L., Øgland, P., 1993: Kontroll av data–spesifikasjonsrapport, KLIBAS-report no. 01/93, DNMI, Oslo, May 1993.

[7] Vejen, F., Rissanen, P., Øgland, P., 1997: FREYR. Specification report on a system for quality control of meteorological data, draft version 1.3, DMI/FMI/SMHI/DNMI, February 1997.

[8] Øgland, P., 1999: Quality Control of Meteorological Real-Time Data: Results from the S-T-F system, KLIMA-report no. 11/99, DNMI, Oslo, March 1999.

[9] Øgland, P., 1999: Quality Control of Climate Data: An Automatic Approach, KLIMA-report no. 12/99, DNMI, Oslo, March 1999.

[10] Øgland, P., Hellsten, E., Jacobsson, C., Madsen, H., Rissanen, P., Vejen, F., 1999: Human Quality Control (HQC) in the Nordic Countries, KLIMA-report no. 32/99, DNMI, Oslo, October 1999.

[11] Yourdon, 1993: Decline and Fall of the American Programmer, Prentice Hall, 1993.

[12] Øgland, P., 1994: Some Recent Developments Regarding Quality Control of Precipitation Data in Norway, The 14th Nordic Meteorology Meeting, 6-10 June 1994, Kristiansand, Norway.

[13] Øgland, P., 1997: Evaluation of a double exponential correlation weighted interpolation method by statistical experiments, KLIMA-report no. 25/97, DNMI, Oslo, November 1997.

[14] Andresen, L., Kjensli, P.O., Toporowski, Z., Vidal, Å.M., Øgland, P., 1999: DATRUT-prosjektet: Skisse av dataflyt, datakontroll og databasestruktur i en fremtidig helhetlig behandlinsrutine ved klimaavdelingen, KLIMA-report no. 25/99, DNMI, Oslo, September 1999.

WEB- INTERFACE IN MONITORING OBSERVATIONS AND METADATA AT FMI

Pauli Rissanen and Ari Aaltonen
The Finnish Meteorological Institute
BOX 503, Vuorikatu 24, 00101 Helsinki, Finland

Phone: +358 - 9 - 19 29 1
Fax: +358 - 9 - 19 29 57 79
e-mail: pauli.rissanen@fmi.fi
ari.aaltonen@fmi.fi

ABSTRACT

*Internal web- interface at FMI provides common tools for users to find information concerning different observations and summaries, error handling, metadata, station network, database structures, weather cameras and other information. This service is meant for FMI's stuff, firewall is protecting illegal attempts to connect into the system. In some cases customers receive special service via www. Oracle Application Server, PL*SQL, JAVA and ArcView elements are used in procedure.*

1.0 INTRODUCTION

One problem for users is "how to know what's going on" in station network, observation data, measuring techniques, surroundings of the site, operation and so on. Nowadays the number of different database tables is huge. Few basic tables include normal (synoptic) observation values and some statistics, but automatic station data, metadata and in general maintaining of the data have increased the number of database tables to several tens.

Automation of sites effects in many ways on data. It is not just pure synoptic observations in 3 hours period that we are receiving – it is much more. Automatic station reads data from sensors within few seconds intervals. After that normal temperature readings and other weather parameters for our use are calculated. Also different average, standard deviation, maximum and minimum values are defined. The automatic station itself controls digital readings and makes error reports.

We easily get 1 minute, 10 minutes, hourly etc. data. How much we store data depends on the capacity of databases and needs of users. AWS data is routed to different database tables.

Metadata includes tens of tables for different purposes. Maintaining and visualisation of these data requires special methods. Photos are a good tool to get idea of the site or even get information of the weather on the site.

2.0 WEB- USER INTERFACE

For ordinary users and for quality control purposes above described data flow is impossible to handle without good user interfaces attached to data flows and databases. At FMI the answer is web. It provides an easy link for all users to get information and data.

Following data are available for internal users:

- Station network (lists and maps)
- Station descriptions (text, photos, exposure, etc.)
- Observations (AWS, synop, map and text presentations of precipitation and temperature, temperature and wind curves from AWS, ceilometer data, tower data)
- Weather cameras
- What synop and AWS messages we got?
- Missing synop and AWS messages
- MILOS 500 errors
- Battery charges at stations with own energy input
- Problems at the site?
- Database descriptions
- Telephones
- Links to other sources (Vaisala, road stations, schools, companies) observing weather
- Cameras (Broadcasting company, National Road Authority, others)
- Other weather services

These features will be demonstrated at the workshop.

Some intranet pages are following;

- Front page, which helps users find information concerning different items.
- Some examples of WEB products.

Observations Branch

Finnish Meteorological Institute
Observations Branch
Sahaajankatu 20 E
00810 Helsinki
tel +358-9-19291, fax +358-9-1929 5779

IN FINNISH

New

- Changes in station network
- Havina News
- Other News of the Branch

Contact Information

- Telephones
- Group email addresses
- Contact information to FMI

Station network

- Station lists and maps
  - stations in counties
    - stations in municipalities
  - reporting weather stations
  - automatic weather stations
  - manual weather stations
  - climate stations
  - daily reporting prec. stations
  - monthly reporting prec. stations
  - FEI stations
  - weather radars
    - weather radars and products
  - mast stations
  - sunshine duration stations
  - solar radiation stations
  - air quality stations
  - non-operating stations
- NEW!

  representative stations in municipality

Observations

- Weather cameras
- AWS observations
- SYNOP observations
- Airport observations
- Daily temperature data
  - maps
- Monthly precipitation sums
  - maps
- Weather data from Milos 500 stations (graph)
- Mast observations (nuclear powerplants)
  - Kivenlahti mast 1
  - Kivenlahti mast 2
- ceilometer observations (graph)
  - Herttoniemi ceilometer
- weather in Wire

Others operations

- ASMO - update program of station register
- Description of AWS routine
  - loading data into database
  - data in database / day
  - missing data in database / 15 days
- Missing SYNOP observations
- Error reports from Milos 500 stations
- System for error reports
  - error reports
  - make a new error report
- Battery charges of AWS

- Last changes in station register
- List of HaSy stations
- Wind sensors at stations

- In addition every stations has its own www page:
    - locations and fotos
    - sensor, observation and technical information
    - surrounding and representativeness descriptions
    - history and neighbouring stations
    - contact information (password protected)
    - exposure descriptions
- for example 1304 05184 Hattula Lepaa

- Battery charges of AWS

Databases

- Climate database description
- AWS database description
    - WaWa codes
- Some tables of metadata database

Others

- Collection of links
- Feedback page

Public pages

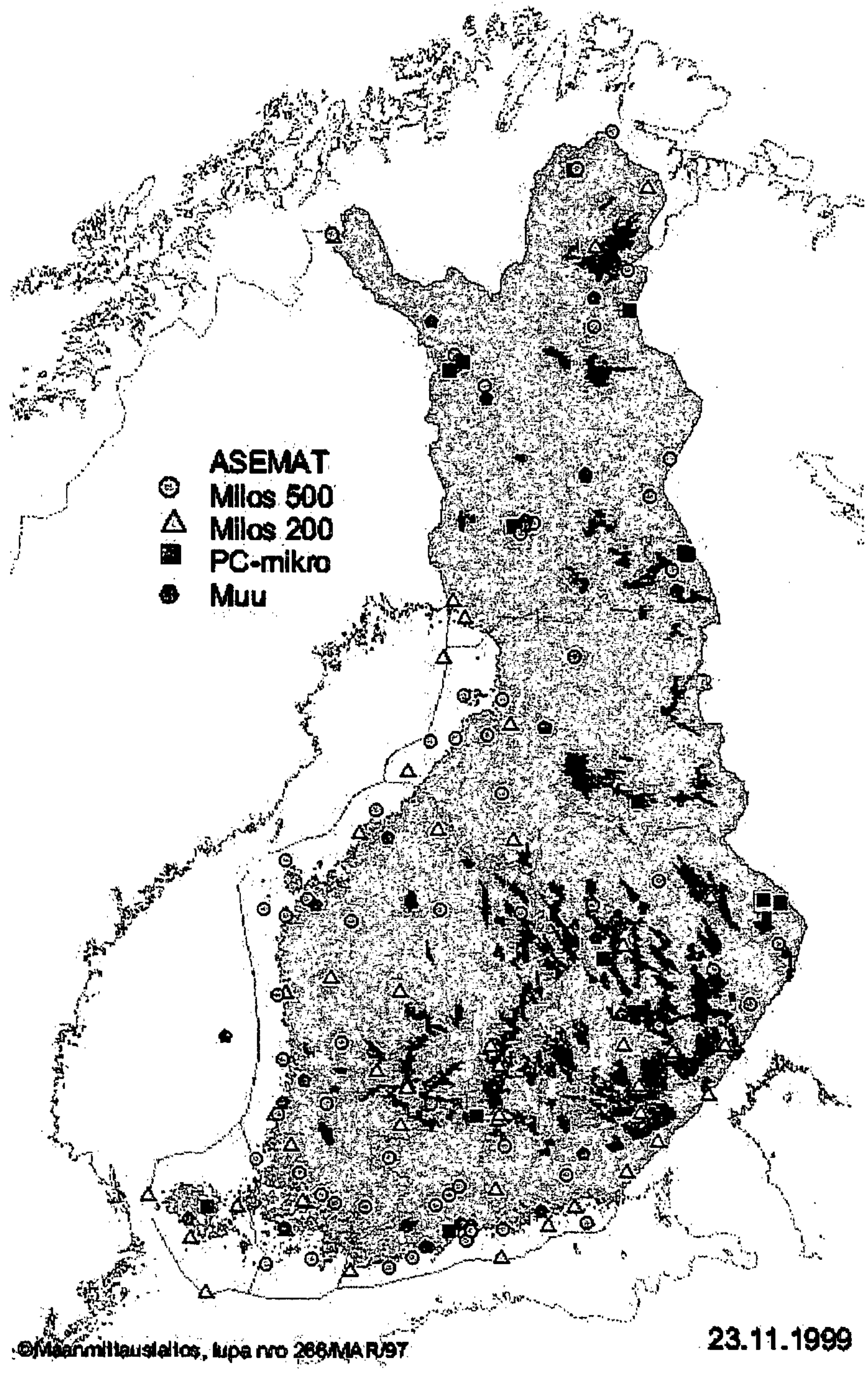
- Observations Branch on public pages

Wire

- Observations Branch in Wire

Editorial staff:
Harto.Ahonen@fmi.fi
Ari.Aaltonen@fmi.fi
Väinö.Lappalainen@fmi.fi

Automaattiset asemat



Kartta: Heikki Uhrman@fmi.fi

Asemaverkko | Yhteystiedot | use password | Avoimuusluvut

Automatic Weather Station
Piikkiö Yltöinen
lpnn: 0103 (national)
wmo: 05583
aws_id: 05583
(address)
MTT/PUUTARHATUOTANNON
TUTKIMUSLAITOS
TOIVONLINNANTIE 518
21500 PIIKKIÖ

In use: from 01.06.1927

In database: from 01.01.1959

Station screens towards the north (Veikko Helminen 30.05.1983)

Station type data:

Solar duration measurements 01.01.1927 - 31.12.1983

Solar radiation measurements 01.01.1968 - 31.01.1978

Localisation data:

The station is located about 5 km south-east of Piikkiö and about 15 km East of Turku.

Coordinates (WGS-84):		Latitude N: 60° 23` Longitude E: 22° 33`
National grid:		x: 3254682 y: 6705227
Altitude		6.0 m above mean sea level
Drainage basin:	no: 82	south-west coast, Aland
Adminstrational area	Province:	Prov. of West Finland
	Municipality:	Piikkiö
Climatological region:	Main region:	South-Boreal, Southern Taiga
	Subregion:	South-West Finland
Percentage water surface:		17 %. Percentage surface covered by lakes or sea within 10 km radius of the station

Elevations of the sensors above ground level (m)

Wind sensors: 11.0

Temperature sensors: 2.0

Precipitation gauge: 1.5

Observations:
HaAWe (Milos500 with automatic data collection)
Synoptic observations: 8 per day (all automatic)
Synoptic observations: 0 minutes before full hour
observation data collected 8 times per day
1. collection begins at 00:10 AM and it is repeated at intervals of 180 minutes

---

Communication:
Normal subscriber telephone

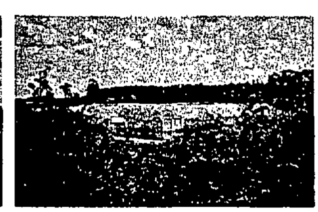
Descriptions:

Exposure of the precipitation measurement, Valid: 02.07.1990 - 28.11.2002

Direction :	N	NE	E	SE	S	SW	W	NW	ka.
Vertical angle (0 -90):	2	2	2	10	10	20	10	10	8.3 (relatively open place)

Temperature: Photographs from the east, south, west, and north (12.6.1995).






Close Environment: from 01.01.1984: Kaukoranta/03.07.1997

An open garden in a low-lying meadow, on the north coast of a sea bay, 100 m from the coastline. Relatively high rocks towards west.

Environment: since 01.01.1984: Kaukoranta/03.07.1997

The station is located on the coast of Paimio bay, which reaches far inland . The surrounding area is field with rocks of various sizes. Some of the rocks are partly covered with trees.

Temperature, representativeness, Since 01.01.1984 : Kaukoranta/03.07.1997

Minimum temperature during calm and bright nights on average 1-3 degrees too low due to the low-lying location.

Wind, representativeness, since 01.01.1995: Kaukoranta/01.02.1998

Open garden on the north coast of the bay, valley. Most exposed from the sector between south-east and south-west. A narrow strip of forest (h. 15-20 m) between the station and the coast, at a distance of about 100 m. A relatively high forest-covered rock in the sector between north and west considerably affects north-westerly winds:

Wind sensors, photographed from the east, south, west, and north, (28.11.1997)

History (changes):

Date	Case (change)	Old	New
16.11.1995	Number of observations/day	4	8
	Local operator	Mttk/Hilma Kinnanen	Mtt
	Station type	4	6
01.01.1981	Number of observations/day	3	4
01.11.1980	Height above mean sea level (m)	4	6
01.01.1972	Height above mean sea level (m)	28	4
01.09.1971	Station activity started	01-Jan-59	01-Sep-71
01.08.1971	End of station activity	-	01-Aug-71

Neighbouring stations, distances (km):

Other than precipitation stations			Precipitation stations	
km	Type	Station	km	Station
16.5	AWS	0117 05792 TURKU KUPITTAA	21.2	1106 LIETO TAMMENTAKA
21.6	Aviation WS	1101 02972 TURKU LENTOASEMA/RUSKO	25.0	1126 TARVASJOKI LIEDONPERÄ
26.0	Climatological S	0113 05392 KEMIÖ LÖVBÖLE	28.4	0107 KEMIÖ MATTKÄRR
29.4	AWS	0116 05295 TURKU RAJAKARI	32.2	0009 RYMÄTTYLÄ KK
30.4	Climatological S	0205 05683 SALO KÄRKKÄ	35.3	1107 NOUSIAINEN KOLJOLA

Corrections by E-mail to: Ari Aaltonen@fmi.fi
Updated 01.09.1999 Input truncated to 9 characters

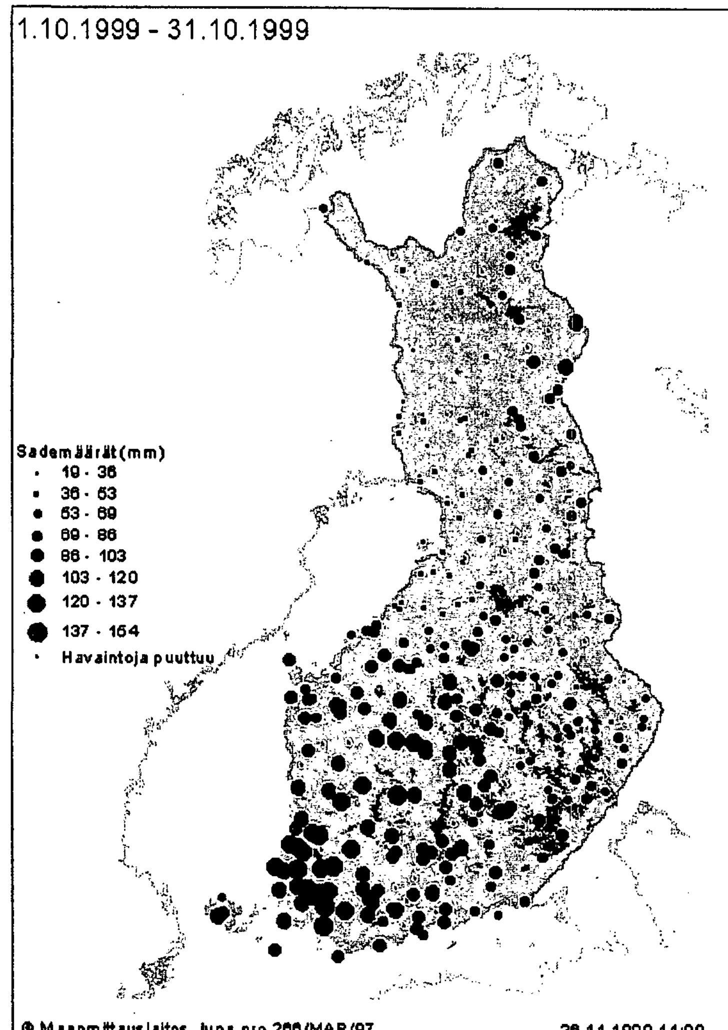
[Havainnot-sivulle]

Eri kuukaudet: [1998] [1] [2] [3] [4] [5] [6] [7] [8] [9] [10] [11] [12] [2000] [Kartta]

Sademäärät lokakuussa 1999

Sivun tiedot perustuvat DAILY, SSTAT ja NORM - *taulujen tietoihin ja tiedot voivat olla tarkastamatta.* Taulukko on päivitetty 26.11.1999 10:58

24 tunnin sadetietoja on tallennettu 198 sadeasemalta (yht 6026 havaintoa) ja 119 muulta asemalta (3681 havaintoa). Havainnoista sadetta (R>=0mm) on havaittu sadeasemilla 69.4 % ja muilla 73.8 % tapauksista. Sadetapauksissa keskimääräinen sademäärä on sadeasemilla 3.7 mm ja muilla 3.6 mm.



Ylläolevan kartan saat omana sivunaan klikkaamalla sitä.

Sum Sadekertymä lokakuussa 01.10.1999 06utc - 01.11.1999 06utc (mm). Jos havaintoja puuttuu, niin kertymä on suluissa.

Rmax Suurin yhden vuorokauden sademäärä lokakuussa 1999 (mm)

L-k-m Sadepäivien lukumäärät: ensimmäinen on niiden päivien lukumäärä, jolloin on satanut >= 0.1 mm, sitten >= 1.0 mm ja kolmas >= 10.0 mm

N Keskimääräinen kuukausisadesumma normaalikaudella 1961 - 1990 (mm)

Evrk max Ennätysrekisteristä löytyvä lokakuun suurin yhden vuorokauden sademäärä asemalla (mm) ja tapahtumispäivämäärä

Ekk min Ennätysrekisteristä löytyvä lokakuun pienin kuukausisadesumma asemalla (mm) ja tapahtumisvuosi

Ekk max Ennätysrekisteristä löytyvä lokakuun suurin kuukausisadesumma asemalla (mm) ja tapahtumisvuosi

Huom: Ennätyksissä ei ole mukana kuluva kuukausi. Jos ennätys on varhaisemmalta ajalta kuin aseman aloituspäivämäärä, niin kyseessä on jonkun aiemman lähistöllä toimineen aseman ennätys !

Taulukon rivit voi järjestää uudelleen valitsemalla jonkin otsikoista LPNN, Aseman nimi, R tai Sum. Sivun ylälaidan linkeillä pääset muiden kuukausien sivuille.

Lpnn	Aseman nimi	Sum	Rmax	L-k-m	N	Evrk max		Ekk min		Ekk max	
0002	Korppoo Utö	94.0	25.6	23-14-4	60	30.5	01/10/1912	11.9	1951	174.4	1974
0005	Korppoo Korpogård	106.2	18.0	19-13-4							
0008	Geta Östergeta	61.7	12.8	19-12-2	59						
0009	Rymättylä Kk	113.2	23.9	17-14-5							
0011	Jomala Jomalaby	88.6	16.3	17-13-3		29.0	30/10/1990	25.9	1976	148.5	1980
0020	Jomala Södersunda	112.9	27.5	21-14-3		20.2	18/10/1998	29.7	1996	117.9	1998
0101	Hanko Russarö	102.7	25.5	20-15-4	58	35.9	20/10/1970	10.8	1953	129.5	1974
0103	Piikkiö Yltöinen	141.5	29.1	22-14-5	69	39.0	19/10/1970	17.0	1979	154.3	1980

CLIMATOLOGICAL DATAWAREHOUSE APPLICATIONS AT DNMI.

Geir Anker
Norwegian Meteorological Institute
P.O. Box 43 Blindern, N-0313 Oslo, Norway
E-mail geir.anker@dnmi.no

ABSTRACT

At the Norwegian Meterological Institute (DNMI) we are working with a GIS related web-tool programmed in JAVA. This system will be available at our intranet and internet.

INTRODUCTION

The system is using a 3-layer structure: Database – server – client. The database (Datawarehouse) contains calculated daily and monthly values. The server produce dynamic HTML documents containing tables and graphs for selected parameters using servlets. The client contains the user interface, including an interactive map for selection of stations. The map functionality is programmed in JAVA1.2 and is running in the Java plugin environment.

WEB CLIENTS.

We want to use internet technology in our Climatological Datawarehouse applications because it will increase the accessibility for our data and it's platform independent. Web technology makes the system easy to maintain and there is no installations needed on the clients. Internet standards offer tools for handling security problems including monitoring of traffic (customers).

3-LAYER STRUCTURE: DATABASE – SERVER – CLIENT.

This configuration makes thin clients. Only the user interface is loaded over to the client. The database logic and heavy operations are placed on the server. This configuration also makes it easy to design an object-oriented system. We can separate the "business logic", the user interface and the presentation of data in separate modules.

SERVLETS

Servlets is serverside Java programs for generating dynamic HTML-documents or other data-objects. This technology replaces the traditional CGI-scripts, and it's fast. We decided to use servlets because is flexible and we can use already developed Java modules. The servlets have basically two main functions. First it produces the station map showing stations that have the selected parameters for a given period. Second it produces a HTML document showing the data for selected stations.

THE USER INTERFACE.

The user interface consists of three parts. First the dialog for selection of report type, parameters and period, second the map interface for graphically selection of stations and third a station table. When the user select a report type, a dialog for required inputs is popping up.
When the user has finished the required inputs, the map shows actual stations. The next step is to select one or more stations graphically in the map or in the station table. When the user has finished the selection, the result will appear on a new frame or window.

The user interface is programmed in Java1.2. Using this technology we can make build advanced user interfaces running in a browser. The functionality allows the user to zoom in and out in the map, paning, zoom into selected stations and zooming to the extent of all stations.

CONCLUSION.

Using internet technology allows us to increase the accessibility of Climatological data. It has been significant technical challenges to overcome. But our focus on building a general platform will be paid back in further development of the system.

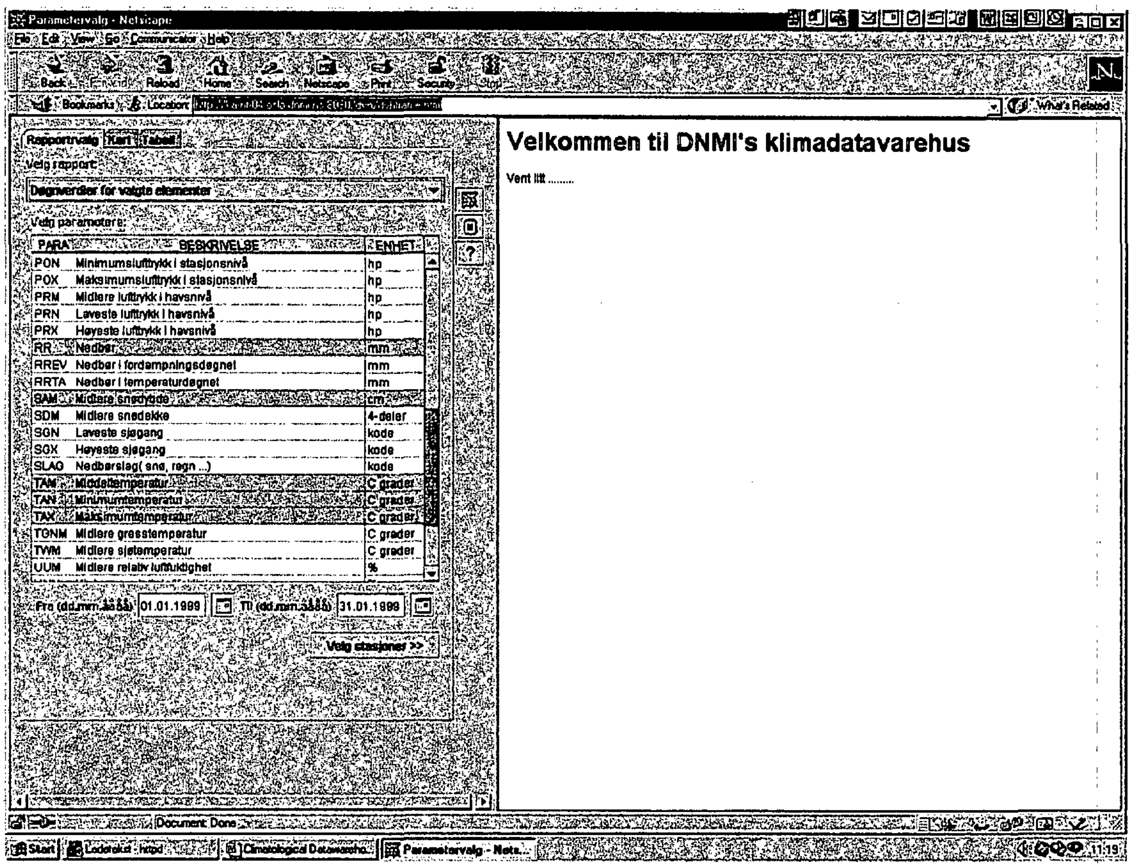


Fig 1. Selecting a report.

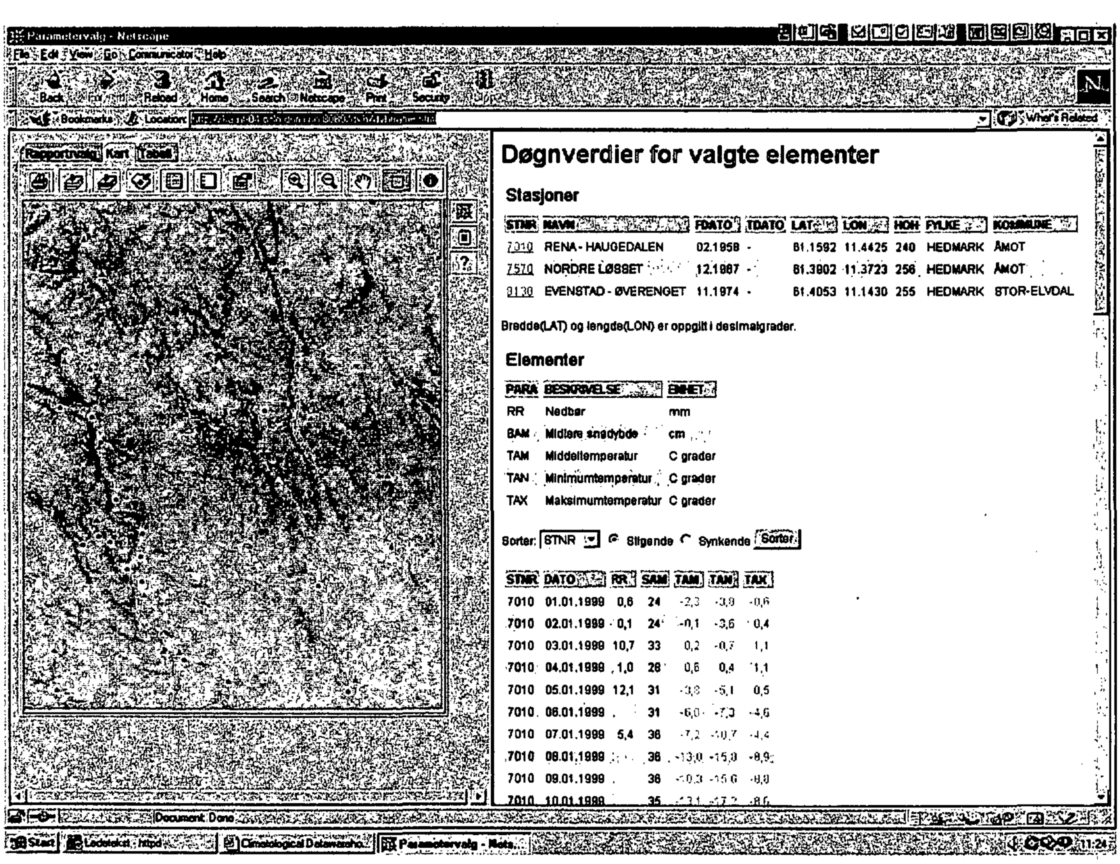


Fig 2. Select stations in the map and show the data.

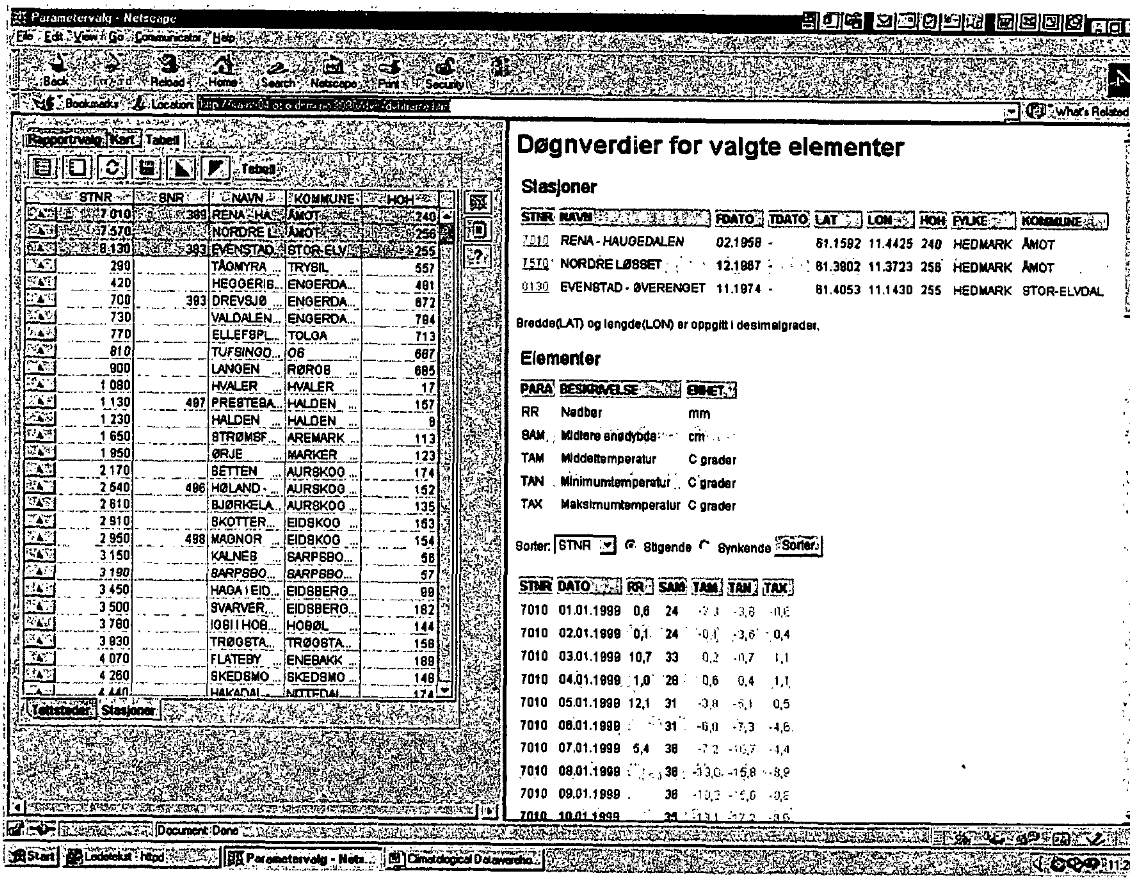


Fig 3. Use the station table to select stations.

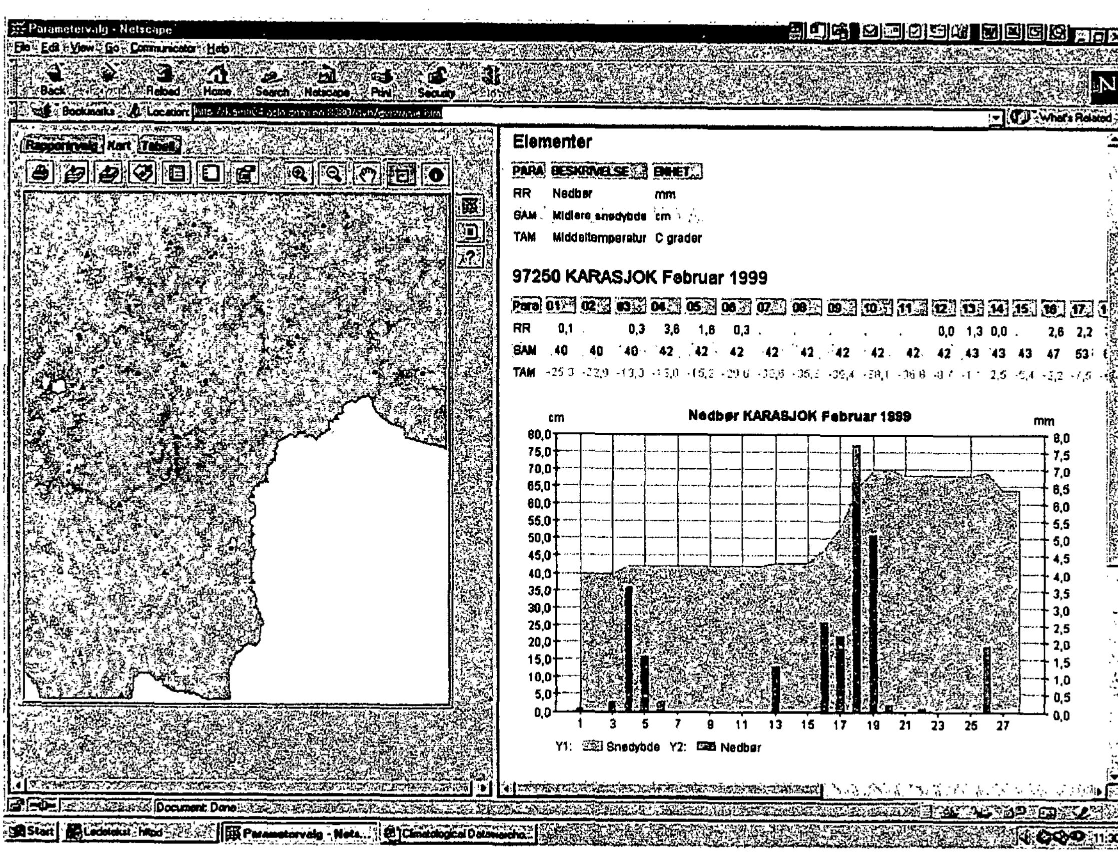


Fig 4. A report type containing a graph.

APPLICATION OF ORACLE REPORT TO CLIMATOLOGICAL DATABASE

Hanna Szewczyk-Bartnicka and Per Ove Kjensli
Norwegian Meteorological Institute
P.O. Box 43 Blindern, N-0313 Oslo, Norway
phone: +47 22 96 30 00
fax: +47 22 96 30 50
E-mail hanna.bartnicka@dnmi.no, per.ove.kjensli@dnmi.no

ABSTRACT

In the paper we discuss applications made with the Oracle Report 3.0, which is a part of the Developer 2000, We focus on the report applications to the DNMI climatological database. We present our experience and discuss the advantages and disadvantages of the Oracle Report as a routine tool for our database.

INTRODUCTION

In 1993 the climatological database at the Norwegian Meteorological Institute was designed. It succeeded to perform basic functions. The loading of new data and correction of data worked well, however some of the dynamic applications used for the presentation of data and calculation did not work as well as expected. The reason for this under performance was a not optimal structure of the database and the absence of a common library for the subroutines used for similar applications.

Presently, the programs developed at 1993 are still used to obtain different types of climatological reports as ASCII files. These programs are executed on the database server using a dynamic user input from PC or Unix terminals.

In the past few years, new advancements in technology (Internet, Intranet) created new possibilities and therefore new solutions. The introduction of *Oracle 8*, a relational, object oriented database structure provided an efficient tool for different applications.

ORACLE REPORT 3.0

Following the previous successes of using *Oracle Forms* in making data transaction applications, the next step was to investigate the data reporting tools in *Oracle Developer 2000*. Initial experience with the *Oracle Developer 2000* indicated that it had the potential for an efficient solution of our previous problems.

With *Oracle Report* it is possible to create traditional client-server applications or WEB-client-server applications, both featuring dynamic user input. These applications support important output file formats like for example, HTML, PDF, and ASCII.

Oracle Report provides a standard framework: a data model to define relations between data, including an *SQL* query to retrieve data from the database, *PL/SQL* procedure to calculate additional values, and a layout model to describe layout properties. In principle, this framework includes all what is needed to create an advanced, dynamic and detailed report, which can be easily implemented on WEB.

An attractive and important tool of *Oracle Report* is the Wizard. This tool is easy in use and fast in creating a simple report in an acceptable, user specified format.

Our experience with *Oracle Report* has shown, however, that along with some advantages come surprising disadvantages. Both, main advantages and disadvantages are listed below.

Advantages

- Supports dynamic applications. The user may choose locations, time periods, meteorological parameters and various types of report.

- The Report applications can be run with the interface to Oracle Forms or Java applications.

- Easy to create a WEB version of the application. With the use of Oracle Application server and Report server it is possible to execute a dynamic report from outside of DNMI (could be used by customers, for example).

- Supports several output file formats (e.g. PDF, HTML, and ASCII). The user has a choice, which format him/her most during the execution of the program.

- Has a very good response time for both the client-server and WEB version.

Disadvantages

- Complicated tool. Oracle Report is relatively easy to use for simple applications. In more sophisticated applications, its use is not straightforward and becomes difficult, sometimes very difficult.

- Some more demanding applications require additional use of advanced SQL and PL/SQL in order to run faster.

- Before proceeding to make a report, a detailed specification of the requested results is needed. A small change in this specification may implicate a restart from scratch.

- Presently, supports only HTML 3.0. Support for HTML 4.0 is necessary.

- Problems with output format. PDF files require a large amount of memory on printers (or other output devices). It is difficult to produce closely packed printouts.

- A run time licence is needed.

SUMMARY OF PRESENT EXPERIENCE AND FUTURE PLANS

Our main impression of *Oracle Report 3.0* is that it requires long time practical training, especially for more sophisticated applications. In addition, for a fast execution of applications the use of advanced *SQL* and *PL/SQL* is necessary.

Presently, we have not yet succeeded to introduce the application on Intranet, due to our inability to implement the *Report* server on an NT machine. It is possible that *Report* functions much better on a *UNIX* machine, but this would indicate farther costs to purchase a Unix version of the program. The decision and how to proceed is not yet conclusive.

As we continue to wait for a *Report* server we use the client version of the program. In the future we will probable use *Oracle Report* to a limited extend, and our main effort will focus on developing of *Java* applications.

EXAMPLES

The next pages show a paper printout examples of climatological reports.

DNMI

KLIMA

90450 TROMSØ

driftsperiode: 1920 -

kommune: 1902 TROMSØ

stasjonshøyde: 100

NEDBØR 1980

	01	02	03	04	05	06	07	08	09	10	11	12	
01	5.1	1.3	8.3	0.0	0.5			0.0	20.4	6.5	0.0	6.5	01
02	1.9		0.1		13.0				5.3	0.7		12.4	02
03	0.8	0.0	0.8		0.1	7.6			0.0	0.0	2.0	6.0	03
04		0.2	0.5	0.0	1.5	14.5		1.0	0.5	1.3	7.0	5.0	04
05		0.7		0.5	3.6	0.3		4.2	17.0	0.1	1.9	5.5	05
06		3.0	2.0	6.1	6.7	5.9		2.2	0.1		3.2	3.7	06
07	0.5	0.3	2.6	13.1	0.8	3.6		5.2	7.9	0.0	6.0	1.4	07
08	0.2	0.3	0.6	6.2	1.8	0.0			1.2	0.0	6.6		08
09	0.0	0.4		1.2		5.4			2.2		5.1		09
10		0.2			0.1	0.2			12.0	1.4	0.6	8.8	10
11					0.1	2.3			0.5	2.8		8.9	11
12	0.0			8.3	6.8	5.7	0.0		2.2			7.5	12
13	7.6	0.0		4.0	0.7	6.2	1.2		0.4	0.0		4.7	13
14	5.2	0.0		3.7	0.0	1.8	0.0			0.4		2.4	14
15	3.8	0.7	0.3	1.3		0.1	0.8		0.5	3.5		3.9	15
16	5.9				0.5				2.5	4.0	1.0	6.2	16
17	8.1	2.9		0.2	0.1				0.3	16.5	0.6	17.7	17
18	15.2	0.7		1.6	10.1				3.3	7.1	5.8	0.1	18
19	3.4	10.5	0.6		1.3				7.5		6.5		19
20	0.5	1.9	0.0			5.2		3.5	12.2		1.9		20
21		0.1		0.5		0.0		1.8	1.5	1.3			21
22			1.4	6.6	0.5				0.1	9.7		0.0	22
23		0.0	0.5	2.0	1.1	16.0	5.3		4.5	10.5			23
24		4.2	0.5	1.2		2.3	0.4		0.5	4.9	2.4	0.2	24
25		0.0	5.2	0.8		0.3	16.1		0.3		8.0	1.6	25
26	6.2	0.2	1.0	0.0		0.0	1.5	1.0	0.0	4.1	1.4	1.3	26
27	1.3	5.3		0.1		0.0		0.2	0.3	12.4	1.8	0.0	27
28	8.5	10.9			0.7	0.5		0.0		2.1	5.2		28
29	12.4	16.8							7.4		0.4	1.3	29
30	9.9			1.2					0.3	4.1	0.3	6.5	30
31	0.1		0.0					0.0		8.8		17.7	31
	96.7	**64.4**	**24.2**	**58.6**	**50.1**	**78.8**	**25.3**	**19.1**	**111.0**	**102.2**	**76.7**	**129.5**	

AVRUNDETE VERDIER

	01	02	03	04	05	06	07	08	09	10	11	12	År
Sum.	97	64	24	59	50	79	25	19	111	102	77	130	837
Nor.	95	87	72	64	48	59	77	82	102	131	108	106	1031
Proc.	102	74	33	92	104	134	32	23	109	78	71	123	81

17-12-1999 12:40

DNMI

KLIMA

18700 OSLO - BLINDERN kommune: 301 OSLO

driftsperiode: 1937 - stasjonshøyde: 94 m o.h.

MÅNEDENS NEDBØR

	01	02	03	04	05	06	07	08	09	10	11	12	ÅRET
1951		89	30	51	9	76	70	303	73	33	119	48	901
1952	32	15	35	50	73	73	87	82	104	75	33	59	718
1953	34	26	6	88	69	71	117	117	63	51	44	42	728
1954	61	64	50	19	53	84	95	95	122	86	92	71	892
1955	31	37	5	37	81	41	16	31	92	51	60	55	537
1956	53	7	19	25	12	104	57	87	88	17	34	28	531
1957	45	60	25	32	48	81	103	145	100	75	50	3	767
1958	39	37	5	38	70	74	93	104	41	109	43	56	709
1959	60	2	58	105	28	13	66	44	15	131	133	144	799
1960	72	28	30	22	19	107	181	144	44	78	131	65	921
1961	51	41	34	18	40	66	63	67	115	163	90	37	785
1962	60	49	33	59	88	43	94	149	74	41	64	57	811
1963	4	5	21	76	116	48	54	142	101	72	100	7	746
1964	15	20	0	34	32	109	76	77	97	158	41	68	727
1965	56	9	21	55	82	110	115	115	229	21	34	70	917
1966	34	87	40	28	67	40	54	113	67	120	107	137	894
1967	58	59	54	36	90	71	55	116	57	139	93	80	908
1968	51	49	40	17	59	89	79	28	71	112	62	20	677
1969	102	23	7	54	58	41	72	41	77	23	64	42	604
1970	25	17	33	61	7	80	166	42	72	101	180	24	808
1971	61	40	72	22	56	50	77	57	43	41	81	35	635
1972	25	36	42	55	54	97	80	112	30	22	11	63	627
1973	33	54	13	52	58	41	97	53	71	12	36	69	589
1974	88	46	27	0	27	54	112	55	181	81	121	46	838
1975	106	18	34	29	37	38	59	71	115	66	84	27	684
1976	35	39	6	16	41	17	33	12	42	181	88	51	561
1977	66	25	53	76	26	78	69	37	56	63	80	49	678
1978	69	43	104	20	17	33	125	89	79	22	28	24	653
1979	12	17	115	80	60	38	81	105	84	75	104	58	829
1980	16	26	31	5	57	126	76	146	80	115	37	52	767
1981	19	23	64	11	68	94	107	12	74	88	115	37	712
1982	27	34	92	39	75	32	35	82	116	70	134	105	841
1983	27	13	35	47	103	27	43	25	165	97	18	45	645
1984	79	20	21	28	50	108	86	95	96	142	51	95	871
1985	40	42	74	75	31	88	114	161	131	35	46	74	911
1986	66	1	53	21	65	33	58	126	28	72	75	112	710
1987	17	24	77	19	48	126	40	82	136	173	99	22	853
1988	139	89	107	41	22	34	175	228	115	80	26	27	1083
1989	11	46	83	85	38	60	59	178	53	59	74	33	779
1990	77	87	20	63	10	89	64	63	60	77	48	87	745
1991	59	24	64	22	3	97	43	33	70	62	115	21	613
1992	14	32	69	59	36	38	89	135	64	35	141	52	754

STATISTIKK

	01	02	03	04	05	06	07	08	09	10	11	12	ÅRET
St.a.	29	23	30	25	27	31	36	59	41	45	39	31	
Nor.	49	36	47	41	53	65	81	89	90	84	39	55	763
Mid.	48	36	43	42	50	67	82	95	85	79	76	55	756

29-10-1999 11:07

HUNGARIAN METEOROLOGICAL DATABASE

Zsuzsa Kover and Szalai Sandor
Hungarian Meteorological Service
P.O.Box 38
Budapest, H-1525
Hungary
E-mail: szalai@met.hu

ABSTRACT

The Hungarian Meteorological datase is a ORACLE7 database running on a HP200 server. Upgrading to Oracle8 is planned in next year. Oracle tools as precompilers for C and FORTRAN and DEVELOPER2000 are used. Data from synoptic, climate, precipitation and aerological stations are stored in the database. Data are controlled and corrected. Different types of summary or extracted values are stored in DAILY, MONTHLY, DECADE and EXTREM tables.

1.0 DATABASE AND DATA ARCHIVES

The HMDB (Hungarian Meteorological Database) consists on two connected parts (Figure 1.). One of the parts is the data archives ordered into determined files (ARCH), and the other is the real database, which is under control of ORACLE7 server (ODAB Oracle DAtaBase).

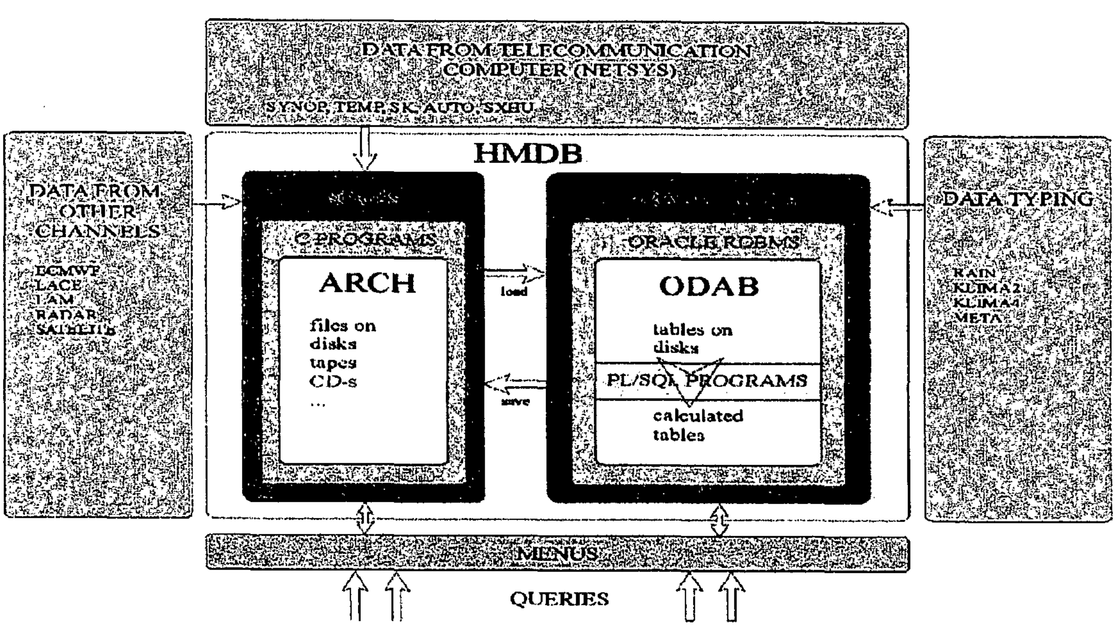


Figure 1

Archives (ARCH)
The ARCH contains the shorter preservation time data, the rare used data, the pictorial information, data which will come into the ODAB later and the security salvages of the ODAB.

Oracle database (ODAB)
The ODAB contains the meta data, the long preservation time data, and the unstinted preservation time data.

Hardware
The ORACLE7 server runs on HP K200, therefore it have be in possession of large memory and hard disc space (768Mb memory and 30Gb disks). The capacity of the direct reaching disc is depend on the on-line reaching data. If the archiving runs also, the secure saving is essential.
The departments can reach the database through PC and Workstations. It's very important to built up the PC-s for this purpose.

Software
We bought an 12 concurrent user's ORACLE7 (7.3) system to the HP K200. Following the demands we will increase the number of users. We also bought the Programmer 2000 precompiler set, which contains the „C" and „FORTRAN" precompiler. The PC Developer 2000 (FORMS 4.5, REPORTS 2.5, GRAPHICS 1.5) system was bought for 4 users. Next year (2000.) we will change ORACLE7 to ORACLE8.

Human resource
Nowadays the developing of the database runs by the Database Group with the lead of the Database Manager. The group consist of 5 persons, included the Database Manager. The professional supervisor of the group is a 4-member committee, who represent the main fields of meteorology. We have got the chance to involve different departments to the process of developing and realisation.

Tasks

- META
- KLIMA
- SYNOP
- AUTO
- SXHU
- AERO
- RADAR
- SAT
- LEV
- SOLAR
- RADIO
- AGRO
- NUM
- KIADV
- LONG
- HOUR

We have too many tasks and we cannot do everything in the same time. Therefore we had to align the jobs and to decide which are the most important.
Our first task was the loading of the META tables. We cannot do anything with the other tables before this, because we refer to the stations through the META tables. These tables also include many conditions of the controlling.
The second job was the processing of the traditional climatic stations' continuously arriving data. This system's name is KLIMA that is handling many kinds of input.
We control and store data of the European and Hungarian synop telegrams making use of the SYNOP system. We process the observed data of the main stations within the compass of this system.
The mentioned three tasks occupy the capacity of the database group therefore the development of the AERO system will be the job of the aerological department.
The improvement of the automatic stations (AUTO) devolves on our new colleague who joined at the beginning of the last year.

2.0 INPUT DATA

The most of meteorological data comes as character information through international channels and from the national stations on phone-lines or mail.

Types of station

(S1) Synoptic station with aws
We make an effort to have automatic weather station in every hourly-observer, international related stations. Today the number of these stations is 30. They make 10 minutes measurements also, what they send to the centre of Hungarian Meteorological Service (HMS) though the connected phone lines.

(S2) Synoptic station with observer
There are 25 synoptic stations, where human observation exists.

(K1) Climatic station with aws
We locate aws with limited measurement program to the climatic stations. Nowadays there are approximately 50 this kind of station.

(K4) Traditional climatic station
They observe 3 times a day in the main terms. The data of the night term are superseded from the writer instruments. The data are sent to the centre 2 times in a month in mail. The number of these stations is decreasing, now 20.

(K2) Traditional small climatic station
They make the measurements 2 times a day: in the morning and evening. They send the data to the centre in every month. Nowadays there are approximately 30 this kind of station.

(R1) Rainfall measure station
They make precipitation and snow-depth measurements once a day. They observe the sort of the precipitation, the beginning and closing time of precipitation and some special meteorological elements, like fog, dew, clouds, ... They pass the data to the centre once a month. We have got about 550 rainfall stations.

Aerological observations have done in 2 stations: Budapest and Szeged. (Figure 2.)

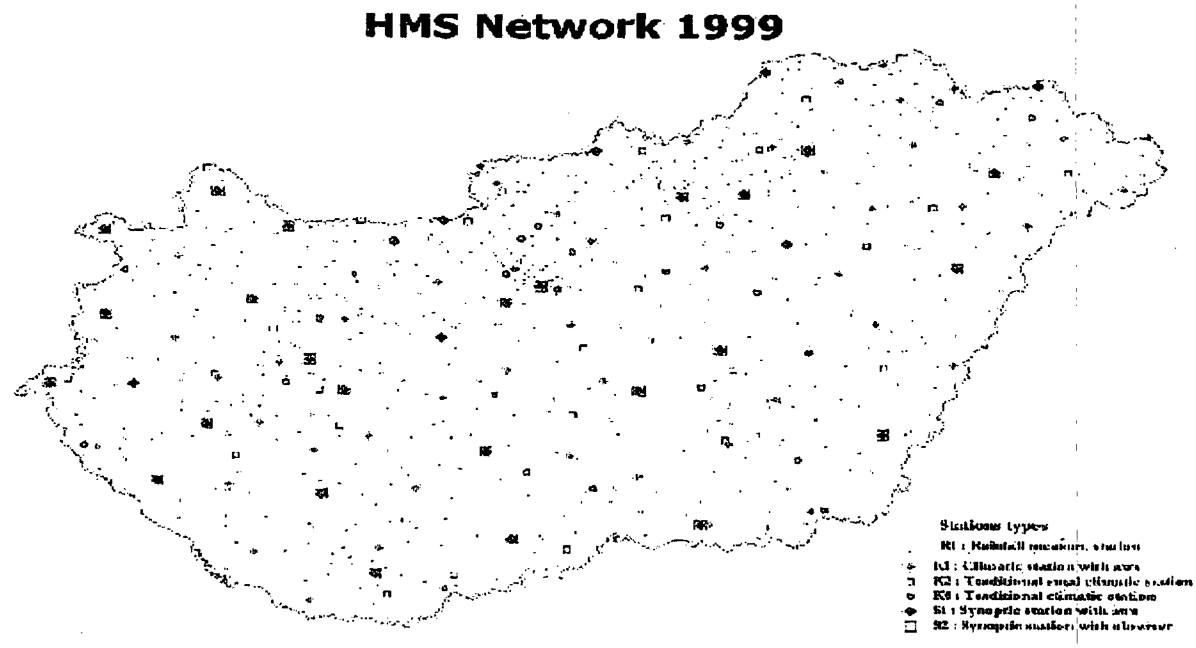


Figure 2

Input types

RAIN (from the stations: K2,K4,R1)
We got a climatlogicaly very valuable homogeneous data, because the stations have sent the same type of cards with the same contain to the centre since the 1920`s. From 1951, the amount and type of precipitation getting from the cards have been recorded completely. The data from the previous years are available only on cards.
The new system (ODAB) can adapt all the information from the card, so we can get more data about the fog, dew, rime, storm, clouds.

SYNOP (from the stations:S1,S2)
We adapt the synop messages of 1 hour measurements of the Hungarian stations, and the 3 hours measurements of the other European countries. After finishing the build up of the whole system we plan to fill it up with the previous years (1975-1995) uncontrolled data, saved on tapes and disks.

AUTO (from the stations:S1,K1)
We store the 10 minutes data in the on-line table of automatic data.

KLIMA2 (from the stations:K2)
The recording of data occurs from the certificates used since 1970`s. We have to record the data of the certificate in character form. We will load to the database the data adapted and saved on tapes, since 1972.

KLIMA4 (from the stations:K4)
The recording of data occurs from the certificates used since 1960`s. We have to record the data of the certificate in character form. We will load to the database the data adapted and saved on tapes, since 1961.

KLIMA8 (from the stations:S2)
The data of the 8 main terms (from the hourly measurements of the synop stations) has written on the so called „8-terms-certificate" since the end of 1960's. The certificates have been sent to the centre to type them. Later the datafiles have been completed going back to 1951, using the same form of cards. Stopping the 8-terms-certificates, the measured and observed data of synop stations are collected in the ODAB, getting from SYNOP, AUTO and SXHU massages.

SK (from the stations:S1,S2,K2,K4)
The importance of the rainfall-telegrams comes from their speed and the fact they cover the whole country, because they are sent from about 100 stations. Spreading of automatons their importance will decrease. The data of the rainfall telegrams become obsolescent, when we can get the data from a more reliable source.

SXHU (from the stations:S1)
The telegrams come from the main stations since December of 1995. They contain the measured data of that traditional instruments, which are already not included on synoptic messages, and they also contain the observed special meteorological phenomenon. Dismantling the traditional instruments the telegrams going to be shorter.

AERO (from Budapest and Szeged stations)
All the Hungarian aerological data, measured and archived from 1962 will constitute the part of ODAB, included the main point and the isobar-level data also. Developing the system we have to take good care make the system able to adapt the TEMP messages of the neighbouring countries.

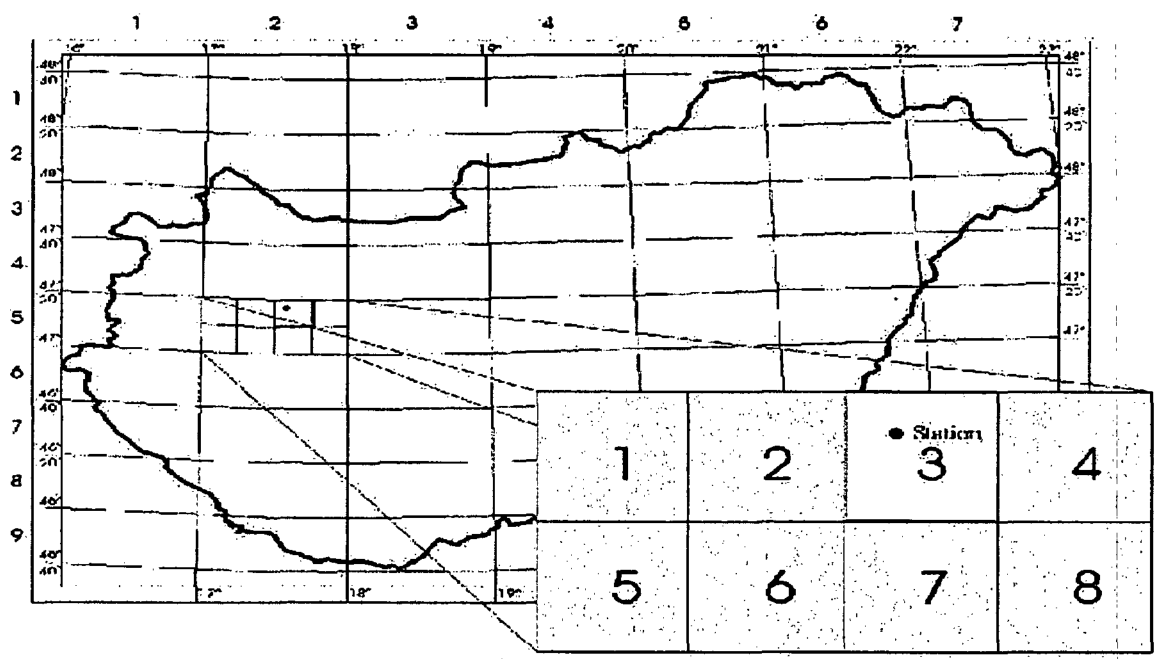
3.0 DATABASE TABLES AND THEIR COMMON COLUMNS

To plan the tables we considered different respects. It is very important to separate the data of automatons from others because of the well arrange, easy survey and comparison.
To save place it is practicable to separate in saving the data measured hourly and 1 or 2 times a day. We make effort to group the data (for example in KLIMA tables) with minimum of empty cells.
It is quite difficult to satisfy this requirement because the different stations record not the same data, and times of measuring are also different. To select the appropriate order of columns we can avoid the useless storing. The numbers of tables have to be not so many and not so little because it could cause difficulties in adaptation.
It is also important to choose good names (not too long, unambiguous, consequent).
The first 5 columns of the data tables are the same: identity number of station (stno), date of measure (dat), date of the last modification of row (upd), source of data in the row (source), quality code of data in the row (qcode).
stno NUMBER(6) station number

Each meteorological station must have an own unique station number. This number will appear in the database in all row of all data-tables. We have to identify the European

synoptic station also because SYNOP telegrams come to our database from whole Europe. In this case the international WMO station number is the suitable.
Every Hungarian station had been identified in five-figure number, the characters of which mean the following:
In Hungary considering the latitudes in 20 minutes and the longitudes in degrees we get a 7x9 net (this is the 1st and the 2nd numeral). This sorting is similar to the available military maps. Inside these areas the maps are divided to 8 smaller parts (3rd numeral). In the part there are 100 possibility to identify the stations (4th and 5th numeral). The stations are lying on the ordered to the smaller co-ordinates. The starting point is nord-west, the upper-left corner of map (Figure 3.).

The new Hungarian identification of stations



Number of station is 0253xx in the database, where xx is a number between 0 and 99.

Figure 3

We have to keep in mind the foreign stations that do not have WMO number, because we have got data from the data exchange program of neighbouring countries. This kind of numbers made from country number and a special number.
To separate the 3 types of station numbers inside the database the id. Number will be completed to 6 figures, the following:
- if *abcde* is the WMO number of the station, then inside the database *1abcde*,
- if *abcde* is a Hungarian station number, then inside the database *0abcde*,
- if *abcde* is foreigner but not WMO number, then inside the database *2abcde* will be the **stno.**

dat DATE the date of measurement

Using the DATE type we got the possibility to apply date functions, we do not have to take care of leap-year, end date of month, because the Oracle arranges it.
It is important to save the Hungarian climatological data in Central European Time (CET) because of the easier counting. The foreigner data will be stored in UTC.
upd DATE the date of last modification

It is a good idea to keep the date of last data change in the rows (like in BDCLIM at METEO-FRANCE).

source VARCHAR2(1) source of data

We refer to the data type of input in the data rows of tables. We made a priority succession between the sources. (For example the rainfall telegrams come more quickly than certificates, but they are less accurate, so the certificates have higher priority.) If higher priority data arrive from the same station, then the lower priority data will be overwritten. (Some possible sources are the following: *r* (RAIN), *t* (SK), *2* (KLIMA2), *4* (KLIMA4), *s* (SYNOP), *x* (SXHU),...)

qcode NUMBER(2) data quality code

This field shows control level of the row. The controls of data occur on different levels. The data get a qcode value depends on its control level.

4.0 ACTUAL TASKS

META

In the beginning of the 1980's we just stored the actual place of meteorological stations, so always the last change got to the system. This static storing system had not let us to follow the history of stations either the identification in the past. In the end of 1980's we have developed a system, which is able to store the date of moving or changing the measure program of the stations. The disadvantage of the system is that it let only 16 changes and has not stored text information.

We divided the meta data into 3 groups: geographical location, measure program, and notes about the history of stations.

In our new identification of stations every move of stations followed the change of the id. Number, even there is not obvious change in the geographical co-ordinates. The geographical co-ordinates of stations will be stored in seconds, and the elevation in 0.1 meter accuracy. In case on long data series we have to present the climatological recommended data series from the lists of several stations. It is important to know the path where can reach the data of the right station to get the correct time. Considering this, we store the id. Number of the previous and next station and the date of change. (Figure 4.)

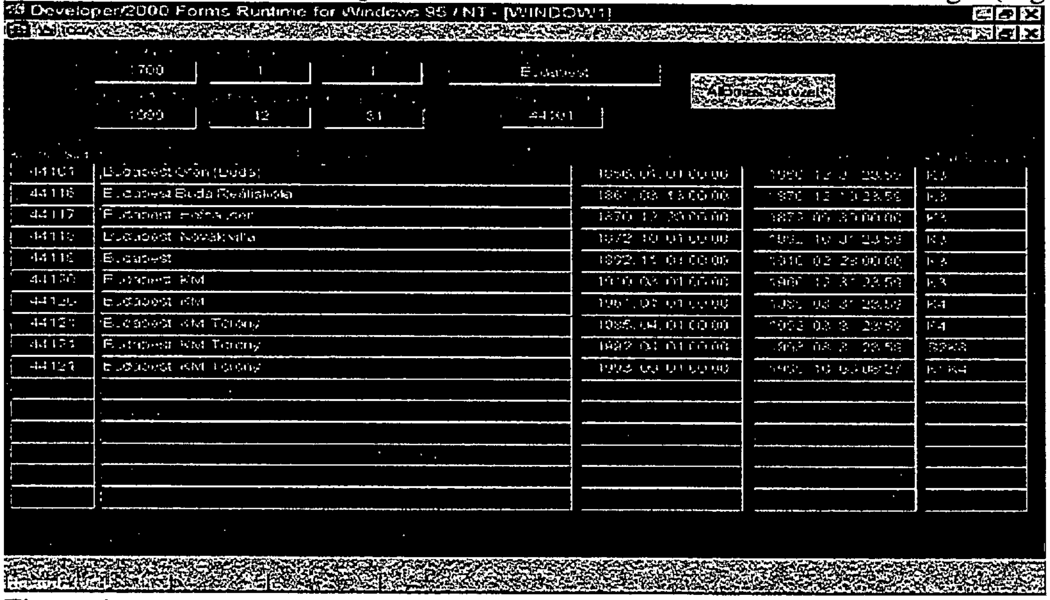


Figure 4

The META tables contain the name and unit of the meteorological elements, the time, accuracy of the measurement, and the information about the instruments, for example height of the location of the instrument.
We provide the possibility to store the hardly codeable and complementary information. The table of the history of stations (HIST) contains the descriptive data in text format (for example: changing of observer, changing the name of station).
The data input, the controlling and query based upon the information of META tables. The data can get into the database only with valid identification number, and if the meteorological element according to META tables is measured on that station. The determination of nearest station during the area control is occurred through the META system. During the query the simplest identification is the number of station, but through the META tables we can sort by the name of station also.

META tables

GEO	This table includes the geographical information about stations. Table's information is used by arcinfo applications and other area checking programs.
STNAME	names and post addresses of stations and observers
STTYPE	types of the stations
TYPES	list of the station types and their descriptions
STNOT	list of area codes which out of border
NATIO	list of national codes and the nation short and full name
REGIO	list of county codes and their short and full name
HYDRO	list of catchment area codes and their descriptions
MPROG	This table contains the measuring program for each station and their measuring equipment at a given time and place.
ELEM	List of all measured and observed element codes that occur at our meteorological service.
INSTR	list of instrument codes and their short descriptions
MTIME	list of time for the measuring program (for missing data checking)
MTINA	list of intervals codes for measured data
HIST	this table is for the station historical note

KLIMA

The tables of the system are CLIMA, CLIMAD, RAIN.
We make effort to minimize the number of empty cells therefore those data are in the identical table that are measured in the same time within a day.
The CLIMA table is a base table, including data that are measured more times a day. There is precipitation sum and form, snow depth, maximum and minimum temperature, clouds, wind, pressure, temperature, visibility, moisture. data in the CLIMA table.
The CLIMAD table is also a base table but it contains data that are measured only once a day for example sunshine duration, maximum wind speed and its direction and time.
The RAIN table is not a base table it is in higher level then the others. It contains daily data of the all stations. There is daily precipitation sum and form, snow depth and some information about the fog, dew, rime, storm, clouds in this table.

SYNOP

The data of the synoptic stations are stored in SYNOP, SYNOPD, PHENO, CLOUD and RIME tables. (Sxhu telegram contains synoptic data, which are missing from synop telegrams. These data are measured or calculated by the observer.)
There are some data derived from both sources. The data of the sxhu telegrams have primary priority.
The **SYNOP** table's data originates from the observer, but these data can measure or calculate the automatic station, too.
The observed phenomena are in PHENO table. This table has no primary key, because it is possible a number of phenomenon in one station and one hour.
The information about cloudiness is in the CLOUD table except the amount of cloudiness.
The rime data appear sporadically the RIME table contains these data.

5.0 CONTROL AND CORRECTION

We aim at having foulness quality of climatic data. Because this is the result of a long process, the data are accessible on every level, and we can label them with quality code. The higher level of control belongs to the higher value of the quality code.

Syntactical and primary control

Syntactical error is not allowed in data. In the primary control we choose the faulty data, which are not suitable for meteorological or physical rules. The chosen data have to be corrected. The syntactical and primary controls are in different ways.

Old data

The old data seem as controlled and incompatible errors free. If nevertheless incorrect data left, during the loading the constraints will shift out them.

Telegrams

In case of SYNOP, TEMP, SK, AUTO and SXHU telegrams the syntactical and primary control occur in „C" language, outside of Oracle. In case of S1 and S2 type stations the incorrect messages are sent back to the station and are corrected by the observer. The errors of other stations are corrected by data controllers in the centre. There is possibility to have accepted the data considered incorrect by the programs.

Typed data

Recording the data can occur only through Oracle FORMS programs. This way recorded data do not contain syntactical and primary errors. In some cases there is possibility to solve the time control by the input FORMS program.

Temporal and area control

After the correction of rough (primary) errors we need a temporal and area control, but we have to take good care in the correction. There had been an area control program in our Service in the 1970's checking the 24h precipitation, which recovered the data automatically (!). Unfortunately it had not labelled the overwritten data, later it has been dismissed. It took much time to restore the original, measured data in data files. (Figure 5.).

Figure 5

The changing in time depends upon the term among the 2 measurements, the measured elements and weather conditions. The used controlling system filters the incorrect data that deviate from the normal change, but the correction will happen just in justified case. If we decide to correct the data, the new data have to comply with all the controlling requirements. During the area correction we got a good method to estimate the missing data. Sometimes it is necessary - specially in reports - to give an approaching data instead of missing data.

6.0 CALCULATIONS

The database tables contain measured and calculated data. We need store the calculated data, because during the fragment query it would take a lot of time to do them in every case. Sometimes we count a column from another column of the same table. For example the dew point and relative humidity are calculated from data of wet and dry temperature (in table CLIMA).

The daily summary table (DAILY) contains the minimum, maximum, mean, sum, etc., of data in basis tables.

In case of stations measuring precipitation several times a day, the values of measurements are stored in the basis tables and the daily sum of them will get to the RAIN table.

The data of MONTHLY and DECADE tables are calculated from monthly and decade values of RAIN and DAILY tables. We need quite often the extreme values of stations for whole measured period. These maximum and minimum values are stored in table EXTREM.

7.0 QUERIES

The users of database interested in the query possibilities what we offered them. They will judge the system by this, so we have to direct our attention to it. The reach has to be simple, but able to receive parameters. (Figure 6).

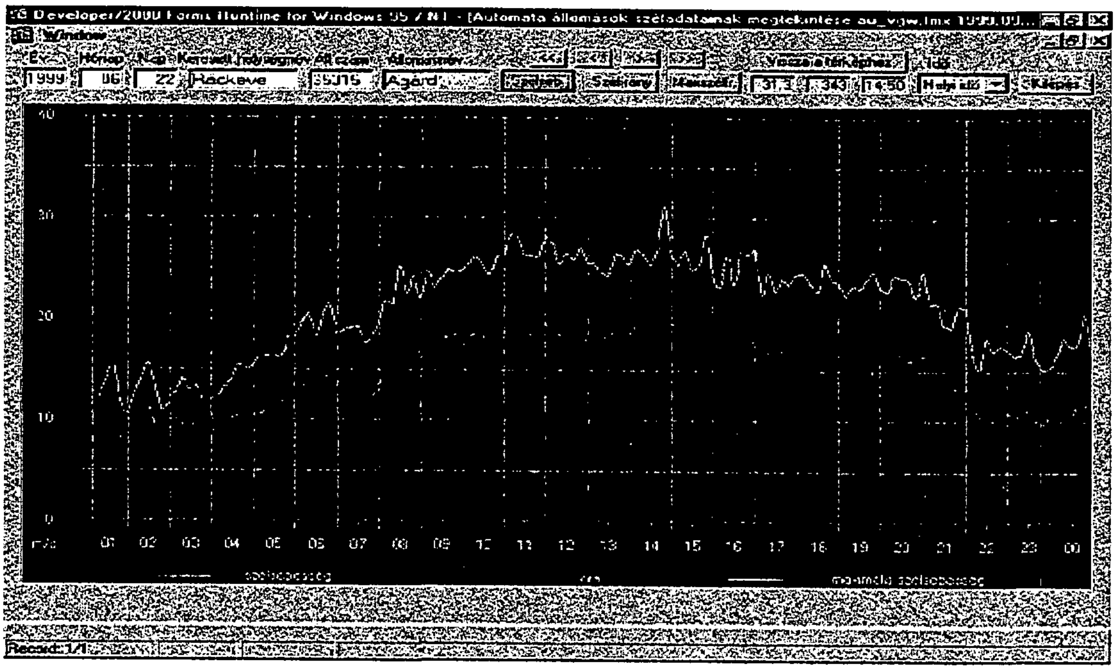


Figure 6

Using the SQL the Oracle provides the possibility to get the basis and counted data systematically, but the knowledge of SQL language is required. In the beginning only a few people have this experience, so we support the data query through the Oracle Tools.

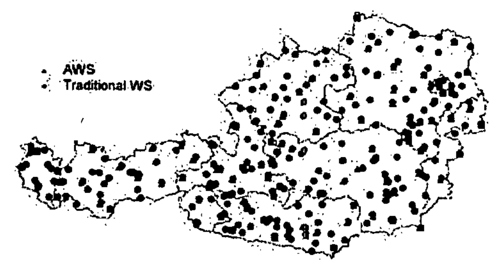
APPENDIX A

Presentations (handouts) from:	Austria (ZAMG)	by Wolfgang Lipa
	Denmark (DMI)	by Claus Kern-Hansen
	France (METEO-FRANCE)	by Pierre Bessemoulin
	Norway (DNMI)	by Åse Moen Vidal

The II ESCN Workshop on Climate Data bases

Oslo
11th to 12th
October 1999

Station Map



Folio 2

Austrian Climate Data Base System

Situated at the Institute for Meteorology and Geodynamics

Review:

1851 Foundation of the institute
Observations on hand written journals
1969 First digitising of monthly climate values based on punch cards
1979 Building of the first climate data base
(Daily and monthly values) on magnetic tapes
1982 Building of the first hourly climate data base on 200 magnetic tapes
1984 Start of a climate computer network for data entry and
Check for daily data values
1987 Building of the first data base for hourly, daily and monthly values
Based on a self written data management system
1990 Change to unix computer systems and building up the first data Sets based on a relational data base management system sybase

Folio 3

Austrian Climate Data Base System

Situated at the Institute for Meteorology and Geodynamics

1992 Start of a data base server with 40 GB disk-storage-system
1995 All climate data sets are stored in sybase
1996 A second data base server was installed for calculation and to store the online hourly climate data sets
1999 Change the old data base server to a new full duplex system
With 2 GB memory, 4 engines, 500 GB storage

Folio 4

Data Base Structure

Old Version

- One table for station information
- One table for each period of time (minutes, hours, days, months)

New Version

Information tables

- One main table for station information station_id, station name
- Additional information tables like which stations are situated in the same climate region

Data tables

Four tables for each period of times

- The table with suffix _orig (original) contains the original values
- The table with suffix _wert (values) contains the final values
- The table with suffix _flag (flagged) contains the data description of the table _wert
- The table with suffix _ersatz (substitute) contains proposal values for the "computer" and is only temporal

Folio 5

Maintenance of the Data Base

	Computer Division DB-group	Climate Division Data-group
Hardware	***	
Server Software	***	
Administration	***	
Security	*	*
Disk Space	*	***
Table Definitions		***
Contents		***
Interfaces	**	*
Staff	3	12

Folio 8

Interfaces

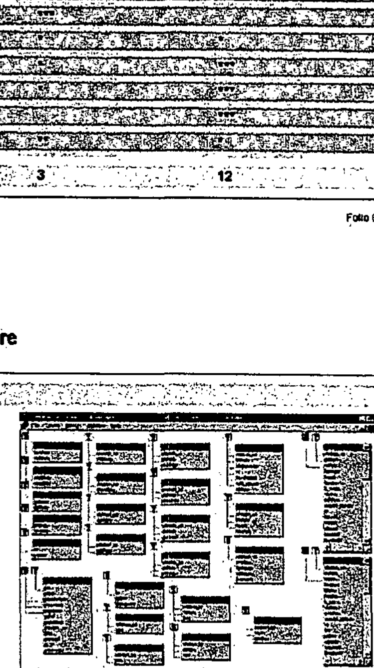
- Own library written in C
easy to link to User programs (Fortran, C)
- SQL nested in c-shell
- perl script
- EXCEL/ ACCESS ODBC
- ARC/ VIEW

Folio 7

Table Structure

Meta Information

- Station history
- Location
- Observer
- Instrumentation



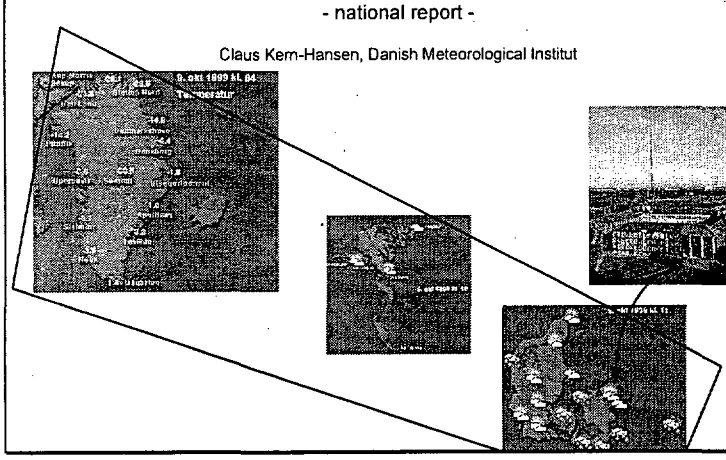
Folio 6

The II. ECSN Workshop on Climate Databases, Oslo 11th-12th October 1999

The Climate database(s) at DMI

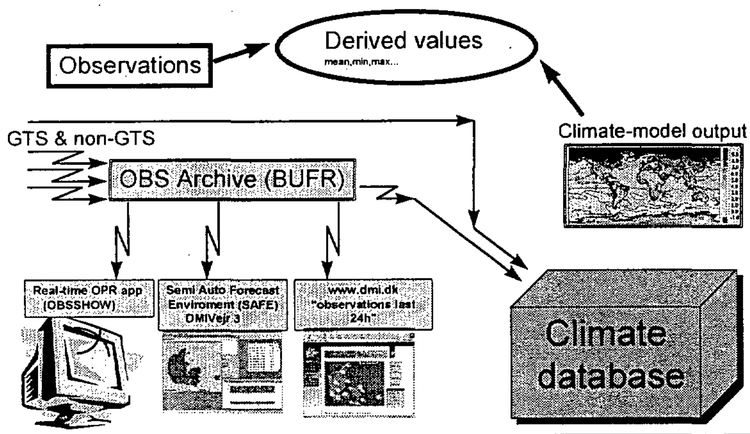
- national report -

Claus Kern-Hansen, Danish Meteorological Institut



Climate database(s)

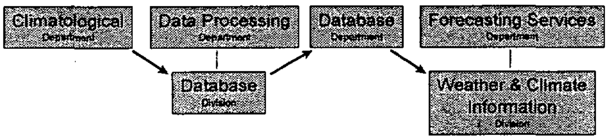
by definition ?



History & Basics:

DBA Software: **MIMER -> INGRES -> Open Ingres 1.2**
1987 1993 1997

Org. history climatedata:



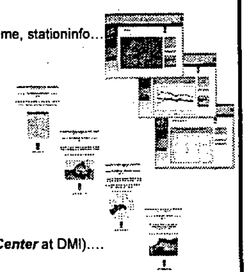
1872... <-..........................~1987-----------------1990'ies-------1996/97

Duties/Purpose of the Climate database:

Storage / Retrieve — Observations / Derived values — The national (Gold) Mine

DMI

Public services:
- telephone
- internet/www -weatherarchive, normals, extreme, stationinfo...
- analysis & publications
  - dataseries, overviews, statistics....
  - weekly, monthly, yearly.....



Commercial services:
- data
- analysis : upon request/by contract
- Tech./Sci. support

Related activities:
- support of Climatemodel research (***Danish Climate Center*** at DMI)....
- digitization of manual measurements/records
- (Q check)
- DBS management; software development.

Keyword: **Utilization**

- Relation Database (Management System) RDMS

DMI

Content: Basically all surface related observations:
DMI GTS; non GTS data
- not image-data (Sat/Radar)

Surface obs: 1987->........................

Upper Air (TEMP): 1993->.......

"General rule": daily/obs: 1961->......................................

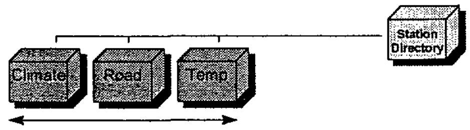
Older data - from projects:
1873->........ yearly, monthly, daily - as made available in digital form

---

Basic layout:

Relation Database

DMI



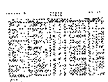
Tables

Sta.type - sta.nr.- date/time - value1 - value2........

Synop
TEMP
Manual Climate
Precipitation
...
...

Basic resolution: Obs (Obs.-time)

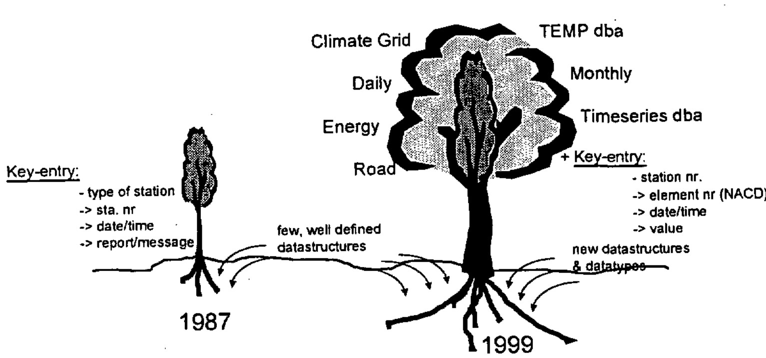
User interface
Applications



- Fortran programmes (>100) -extraction -printouts -calculations->derived values
- Plastic software Excel, Access, SPSS
- Standalone tailormade T-serie plot; windroses

---

a lot of Budding



Last 3 years (since Norrköping):

- Monthly dba
- Clean-up 1961-90-97/98
- Time-series dba (+history)
- Climate Grid (10,20,40 km^2) (Ptn,T,Rs,Epot..)
- -> Open Ingres + new server cluster + new PC-network
- Focus on Long Timeseries (to support Research)
- started Daily dba

Major Challenges:

- Q-Check GI GO *Old sta. network/New sta. network*
- Flexible design New station/dataflow formats; more than one version of a dataset
- User interface

A new organization of Climatological activities at Meteo-France

- **Evolution of the tools available at CDMs** (departmental centers) allowing to assume:

⇨ **the Exploitation in Climatology**

(acquisition, data checks, transfer to central data base)

⇨ **the Production in Climatology**

(statistical data processing, ...)

✌ ***It is based on :***

⇨ **a new telecom and computer environment, including the local climatological database**

⇨ **a new circulation of data between CDMs and the Centr Service**

⇨ **a permanent match between national and local databas which are automatically replicated from the national Data Base**

A NEW TELECOM AND COMPUTER ENVIRONMENT IN CDMs

⇨ **Ethernet LAN**

⇨ **client/server architecture**

- **a climatological database (« BDClim ») on the server**
- **a Climatology application running at the client level**

With ***:***

- **Windows NT (Server and Workstation)**
- **Oracle RDBMS (pl/Sql, Sql*net)**
- **TCP/IP Network**
- **Internet Messaging**

And using :

- **Object technology (OOA Case Tool, C++ ...)**
- **Visual C++, Visual Basic and Word as development tools**

THE NEW ENVIRONMENT FOR CLIMATOLOGY

Expected dates for availability of major components

- **a new National BDClim : end of 10/99**

- **a rescue of the national BDClim (second server) :07/**

- **an operational new circulation of data : 04/2000**

- **new tools to collect, control and manage data :**

  - prototypes : 04/1999
  - operational within the new Production : 09/2000

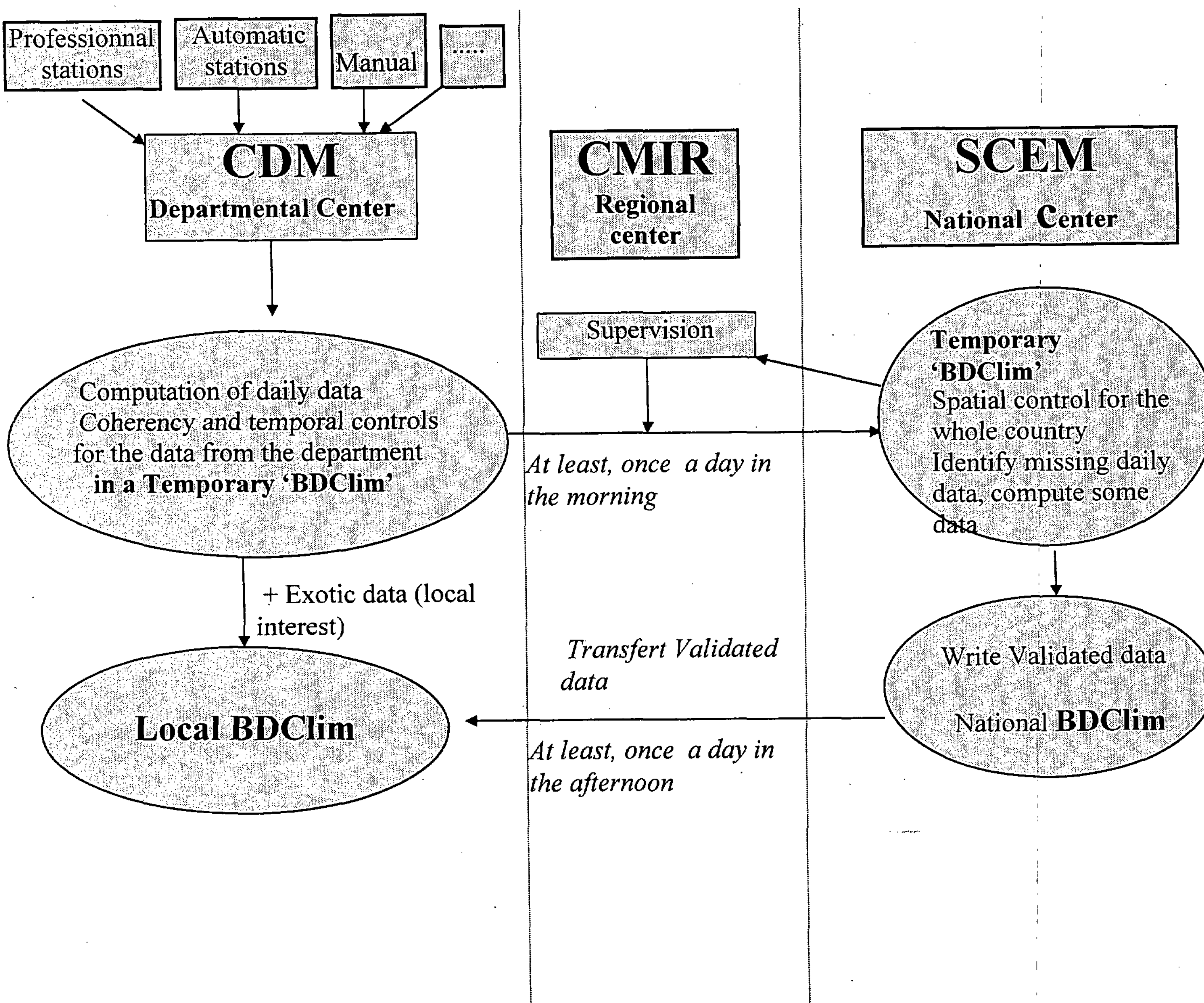
- **production Tools :**

  - prototypes : end 1999
  - operational : 09/2000

THE NEW CIRCULATION OF DATA

Technically : 2 levels involved (CDM and SCEM)

Functionnaly : 3 levels (CDM, CMIR and SCEM)



OUTLINE OF FUNCTIONNALITIES WHICH HAD TO BE DEFINED BY THE « PIC » PROJECT

- **The local BDClim (content, structure)**

  - ***manage different types of stations (professional, automatic, manual)***

  - ***time resolution of the archive (hourly data, 6' data, ...)***
  - ***archive depth and geographical coverage***

  - ***computed data (monthly, decadal, statistics ...)***

- **Data acquisition**

  - ***for each types of stations***

  - ***frequency of acquisition***

- **Security and rescue methods**

- **Rules and formula to calculate data**

OUTLINE OF FUNCTIONNALITIES WHICH HAD TO BE DEFINED BY THE « PIC » PROJECT (cntd)

- **Data Checks**

  - ***types of checks (rules and tools)***
  - ***quality codes***
  - ***means and tools to validate data***

- **Coherency of data**

- **Access to Data** ***(tools for visualisation)***

- **Validation** ***(tools to update original raw or calculated data)***
- **Data dissemination (towards SCEM)**
- **DB initialisation** ***(import historical data)***

- **Management of original data** ***(Hourly, Daily)***
- **Management of « exotic » data** ***(non standard data)***

and some connected functions :

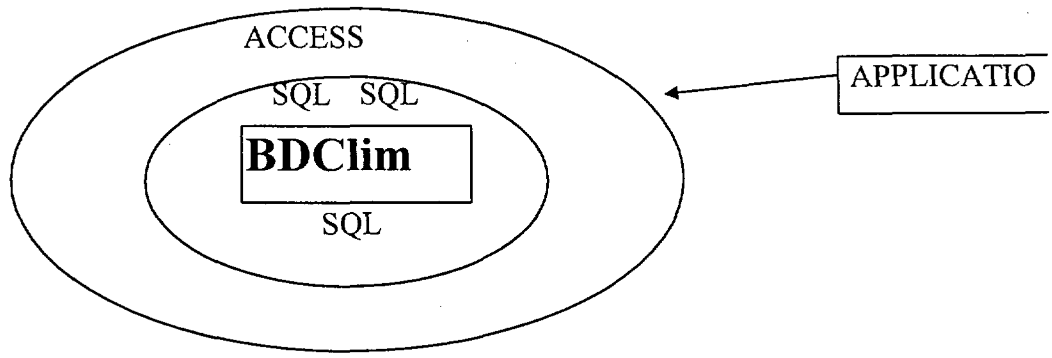
- **Writing of documentations**
- **On line Help**
- **Training of users**

and a specific function : an access Interface

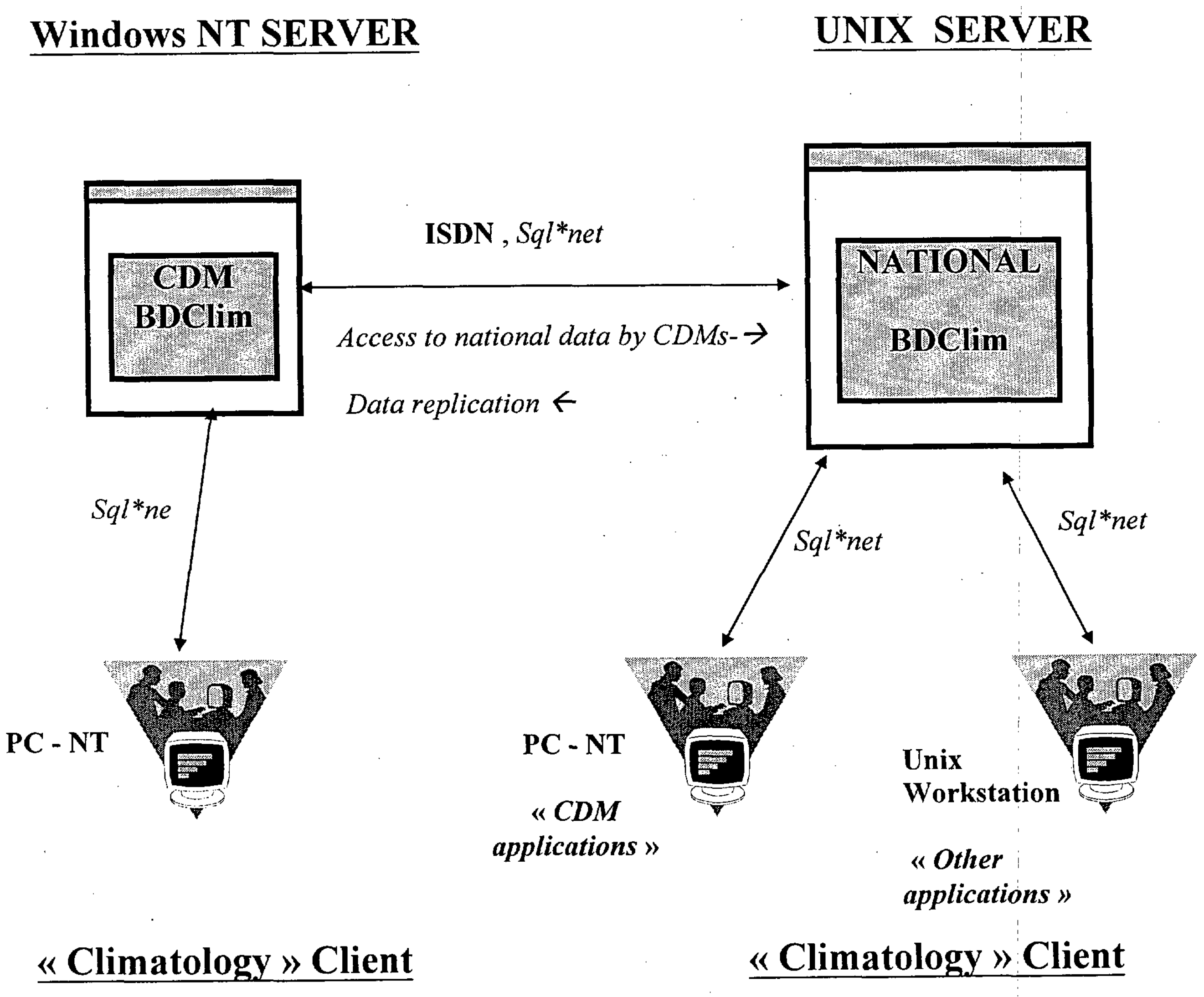
= a software layer over the SQL language (for the developments)

⌧ to make applications independant of the updates of the structure of the BDClim

⌧ to guarantee acceptable performances when accessin BDClim data



OVERVIEW OF THE CLIMATOLOGICAL SYSTEM



AVAILABLE SPECIFICATIONS

- **General Specifications**
- **Structure of the BDClim**
- **Data controls**
- **Quality codes**
- **Management of metadata**
- **The agroclimatological data**
- **Functional analysis of the climatological production**
- **List of Computations and Statistics to be developped**

THE CLIMATOLOGICAL PRODUCTION

- **DEFINITION PHASE OF THE PRODUCT**

  ⇨ ***a tool to access the catalog of data***

  ⇨ ***a tool to select the station***

  ⇨ ***a tool to access metadata***

- **ELABORATION OF THE PRODUCT**

  ⇨ ***Graphical User Interface to specify the product***

  ⇨ ***access to raw data (BDClim ....)***

  ⇨ ***process data (statistics, computations .....)***

  ⇨ ***product derivation (tables, chart, graphs ...)***

  ⇨ ***elaboration of « standard » documents (including texts, charts, ...)***

- **CHECK OF THE PRODUCTION**

  ⇨ ***a general quality code (based on data quality or statistics, ...)***

THE CLIMATOLOGICAL PRODUCTION (cntd)

- **DISTRIBUTION OF THE PRODUCTS**

  ⇨ ***mail , fax, Internet , ftp, …***

- **MANAGEMENT OF THE PRODUCTION**

  ⇨ ***register customers and requested products***

  ⇨ ***automate some tasks (subscriptions, cost estimates and bills,...)***

  ⇨ ***manage the subscriptions***

  ⇨ ***plan the production***

- **Two main functions => two kinds of software**

  ⇨ **DATA EXTRACTION (with some computations)**

  ⇨ **PRODUCT DERIVATION (tables, graphs, charts ..)**

- *Many sub-functions are common with other subjects :*

→ development of a general Application for product derivation dedicated to Climatology and Forecasting

Combination of different items is possible .

- **Developments in Java**

- **A document design standard**

- **Models for the following documents :**
  certificate for bad weather
  summary of weather of the month at the station
  wind rose
  weather survey for building slips
  monthly tables for rainfall
  and some others ...

(requests for such products addressed to different CDMs should obtain th same result)

- **Final product = standard products (appearing in a catalogue of Météo-France products)**

  +

  self-developped products

Planning of the developments for the production

Extraction software: sep 98 - ...
Production Software : Jun 99 - sep 2000

CONCLUSION

- **The French system is not yet fully developped**

- **The first version is close to the frenc organisation**

- **The project is included in a larger project i the CDM with a strong support team**

In the course of 2000,

- **a first version will be operational in MF**

- **it will be then modified to satisfy the needs of th french oversea territories : the oversea system wil be an 'independant package' . It could be a opportunity to include the 'clicom' needs if i seems opportune to some countries . MF coul participate to that .**

- **the general documentation from Meteo-France can be used without any restriction for a met. Usage (no commercial use).**

THE II ECSN WORKSHOP ON CLIMATE DATABASES IN OSLO

DNMI's ORACLE8 Databases for Climatological Purposes

by

Ase Moen Vidal

Precipitation Values

- non measureable (0.0 millimeter)
- No precipitation (-1.0)
- Missing observation (NULL)
- Values > 0.0

-1.0 is also used in the same way for variable wind direction and state of ground

Meta Tables in klibas

- General station information (1)
- Station types (1..n)
- Station parameters (1..n)
- Station instruments (1..n)
- Geographic station information (1..n)

In DWH these tables are summerized in 1 single table

Temporary Tables

- Store temporary data while controlling,updating,interpolating
- One table for each station type
- Data from working tables are transferred to main tables when the data are corrected

Why Data Warehouse

- Making data easily accessible
- Duplication of data in different statistics
- Simplifies programming
- Simplifies ad-hoc queries
- Less demands for machine resources
- End-user data base
- Exclude data of bad quality

DWH data

- Copies of observations
- Derived data
- Empty records for missing data(whole periode)
- Produced by Fortran program
- Average, max, min daily and monthly values
- Summerized meta table data
- Retrived by Java applications / Oracle tools
- Updated by real time data and corrections

Data sets

The periode of data can be from 1 month to 150 years

- Homogeneous data
- Normals
- Long series of monthly values
- NACD data

HISTORY

- Computer from 1957
- Data from before 1957 does exists for some stations
- First commecial RDBMS in 1991 Oracle v.6
- 1500 data tables (1 station pr.table)
- Migrated to v.7.3.3
- Running i test v.8.0.5 by october 1999
- 5 data tables (by station type)
- Migrating to v.8.0.5 before the end of 1999

Underlying Computer Systems

- Silicon Graphics Origin 200
- Operating System IRIX64 Release 6.5
- CPU's 2*225MHz MIPS R10000
- Main Memory 768 MB
- RAID5 storage system
- HA (high availability) / fail safe system (with an Origin 2000)
- Veritas archive system

Database Security

- 2 separated RDBMS
  - Raw Data (Klibas)
  - Data Ware House (DWH)
- On-line archiving
- Full backup every night
- Regular export of tables

Stored Data

- Manual weather data
- Automatic weather data
- Precipitation data
- Precipitation intensity data
- Maritime data
- Metar data

- In total data for 1500 stations

Data tables for station types

- Klibas
  - 5 tables
- DWH
  - 3 tables
  - Daily values
  - Monthly values
  - Datacover

Records (key + data)

Klibas
Station ID + periode + observation
DWH
Station ID + periode + derived values

Data type number

Indexes

- Klibas
  - station number, year, month, day, hour
- DWH
  - station number, date

Only 1 concatenated index

Partitioned Big Tables

- All data tables in klibas and DWH are partitioned
- The partitioning key is station number
- All tables are partitioned equally

3 data partitions
3 index partions

APPENDIX B

THE II ECSN WORKSHOP ON CLIMATE DATABASES
OSLO 11th TO 12th OCTOBER 1999

PROGRAM

Monday the 11th of October 1999

9.00 - 10.00	Registration, light refreshments (DNMI headquarter)
10.15 - 10.30	Welcome to the Workshop by Bjørn Aune, chairman EAC
10.30 - 10.45	Information by Margareth Moe

Review of the development since the Norrköping workshop -1996 with national reports

10.45 - 12.30	Short presentations of delegations and the different NMIs climatological databases (Belgium, Austria, Slovenia, Denmark, Finland, Sweden, Iceland, approximately 10 - 15 minutes each, including comments).
12.30 - 13.30	Lunch
13.30 - 14.00	Short presentations of delegations and the different NMIs climatological databases continue (France, Ireland, approximately 10 - 15 minutes each, including comments).
14.00 - 14.25	MIRAKEL, The Climatological Database System at DWD by Johannes Behrendt, DWD
14.25 - 14.50	Climatological Information at Portuguese Meteorological Institute. by Luis Filipe Soares do Rosàrios Cardoso, INMG.
14.50 - 15.15	Centralising the retrieval of archived data at KNMI by J. D. Jans, KNMI
15.15 - 15.30	Coffee
15.30 - 16.00	How DNMI has organised the Oracle8 databases for Climatological purposes; 1) Database containing raw data 2) Database containing Data warehouse by Åse Moen Vidal, DNMI

Database information

16.00 - 17.00	Presentation of ORACLE (ORACLE) Presentation of SGI (SGI)

Dinner

19.00 -	Dinner at "Hos Thea" Gabelsgt. 11 **SPONSORED BY ORACLE & SGI** (about 20 min. walk from the hotel or 5 min. by taxi).

Tuesday the 12th of October 1999

9.00	-	9.15	Coffee

Quality management of data/databases

9.15	-	10.15	Quality assurance of data flow and system management by Petter Øgland, DNMI

Web interface to the database / datawarehouse

10.15	-	10.45	WEB-interface in monitoring observations and metadata at FMI. by Pauli Rissanen, FMI
10.45	-	11.00	Coffee
11.00	-	11.45	Climatological Datawarehouse applications at DNMI and demonstration of a GIS related web tool programmed in JAVA. by Geir Anker, DNMI
11.45	-	12.30	Application of Oracle Report to Climatological Database. by Hanna Szewczyk-Bartnicka and Per Ove Kjensli, DNMI
12.30	-	13.30	Lunch
13.30	-	14.45	Discussion: Development tools - Use of WEB interface, limits / possibilities
14.45	-	15.00	Light refreshments

Further development and co-operation

15.00	-	15.45	Discussion
15.45	-	16.00	Summing up. Close of the II ECSN workshop on Climate databases

APPENDIX C

Participants

NAME	NMI / Country	E-mail adr.
Pierre Bessemoulin	Meteo-France, France	pierre.bessemoulin@meteo.fr
Hans Van Hauteghem	RMI, Belgium	hans@oma.be
Mr. Lipa	ZAMG, Austria	
Mr.Potzman	ZAMG, Austria	
Hartwig Dobesch	ZAMG, Austria	dobesch@zamg.ac.at
Mojca Dolinar	Hydromet. Instit. of Slovenia, Slovenia	mojca.dolinar@rzs-hm.si
Claus Kern-Hansen	DMI, Denmark	ckh@dmi.dk
Mikael Scharling	DMI, Denmark	mis@dmi.dk
Antti Samuli	FMI, Finland	antti.samuli@fmi.fi
Pauli Rissanen	FMI, Finland	pauli.rissanen@fmi.fi
Nils-Åke Andersson	SMHI, Sweden	nandersson@smhi.se
Luis Filipe Soares do Rosário Cardoso	INMG, Portugal	luis.cardoso@meteo.pt
Maria de Fatima Espirito Santo Coelho	INMG, Portugal	fatima.coelho@meteo.pt
Albert M. G. Klein Tank	KNMI, Netherlands	kleintan@knmi.nl
J. D. Jans	KNMI, Netherlands	jans@knmi.nl
J. W. Nellestijn	KNMI, Netherlands	
Tóranna Pálsdóttir	Vedurstofa, Iceland	tota@vedur.is
Liam Keegan	Met. Eireann, Ireland	liam.keegan@met.ie
Johannes Behrendt	DWD, Deutschland	johannes.behrendt@dwd.de
Andres Chazarra	Spanish Meteorological Institute, Spain	clp@irum.es
Bjørn Aune	DNMI, Norway	bjorn.aune@dnmi.no
Per Ove Kjensli	DNMI, Norway	per.ove.kjensli@dnmi.no
Hanna Szewczyk-Bartnicka	DNMI, Norway	hanna.bartnicka@dnmi.no
Åse Moen Vidal	DNMI, Norway	ase.moen.vidal@dnmi.no
Petter Øgland	DNMI, Norway	petter.ogland@dnmi.no
Geir Anker	DNMI, Norway	geir.anker@dnmi.no
Ole Einar Tveito	DNMI, Norway	ole.einar.tveito@dnmi.no
Margareth Moe	DNMI, Norway	margareth.moe@dnmi.no