



# The MeteoSwiss Data Warehouse System

In its role as national weather service, MeteoSwiss has the mandatory task of collecting, storing and managing a vast amount of meteorological and climatological data.

In order to guarantee the long-term and efficient execution of this task MeteoSwiss operates a data warehouse system. The system comprises applications and a central databank which has been optimised for analytical and information purposes, the so-called Data Warehouse (DWH). It is a high-performance and expandable infrastructure which provides tools for the preparation and processing of meteorological and climatological data (aggregation, quality control, correction).

The DWH system offers in-house and external users customised data access. Both current data and historical measurement series can be obtained. In addition, the data from the DWH form the basis for creating the majority of meteorological and climatological products.

To date, several billion datasets have been archived in the DWH and every day measurements are added from a multitude of sources and in many different formats.

## **The architecture of the MeteoSwiss data warehouse system**

In the DWH system elements of classical relational databanks are combined with a corporate information factory: the data warehouse technology. The system contains four levels: the source, staging, storage and analysis & access levels (see figure 1).

A metadata repository can be accessed from all levels, containing a context databank which stores and administers climatological metadata (e.g. measuring station information). This ensures that the same context data (metadata) are used at all levels.

A brief overview of the individual levels:

### **Source level**

The data flowing into the DWH originate from a number of different sources:

- observation systems (surface level measuring network, camera network, bio-meteorological networks, radio soundings)
- numerical weather-forecasting models
- the WMO Global Telecommunications System (GTS)
- weather radars and satellites
- data archives with historical data series
- data from observation systems run by partner networks

### Staging level

The staging area contains the working database and all methods and techniques necessary for the collection, processing and loading, as well as the quality control and validation of the meteorological data.

### Storage level

The actual MeteoSwiss DWH is located at the data storage level. It consists of an analytical databank which puts an integrated dataset at the disposal of meteorologists and climatologists in order to assist them in weather forecasting and climate research. The so-called data marts are specific sub-areas of the data warehouse which provide data pools customised to meet certain functional requirements and the needs of specific users or fields.

### Data analysis & access level

The data analysis and access area contains the applications for carrying out queries in the data warehouse and in the data marts. With the application CLIMAP, for example, measured data can be displayed quickly and clearly in the form of tables, charts or process diagrams.

Both MeteoSwiss internal users and interested institutions and individuals have access to these data, using flexible software which can be individually configured such as CLIMAP-net or IDAWEB. In addition, the data contained in the DWH are used for creating a number of MeteoSwiss products that can be found in the shop or on the current weather pages.

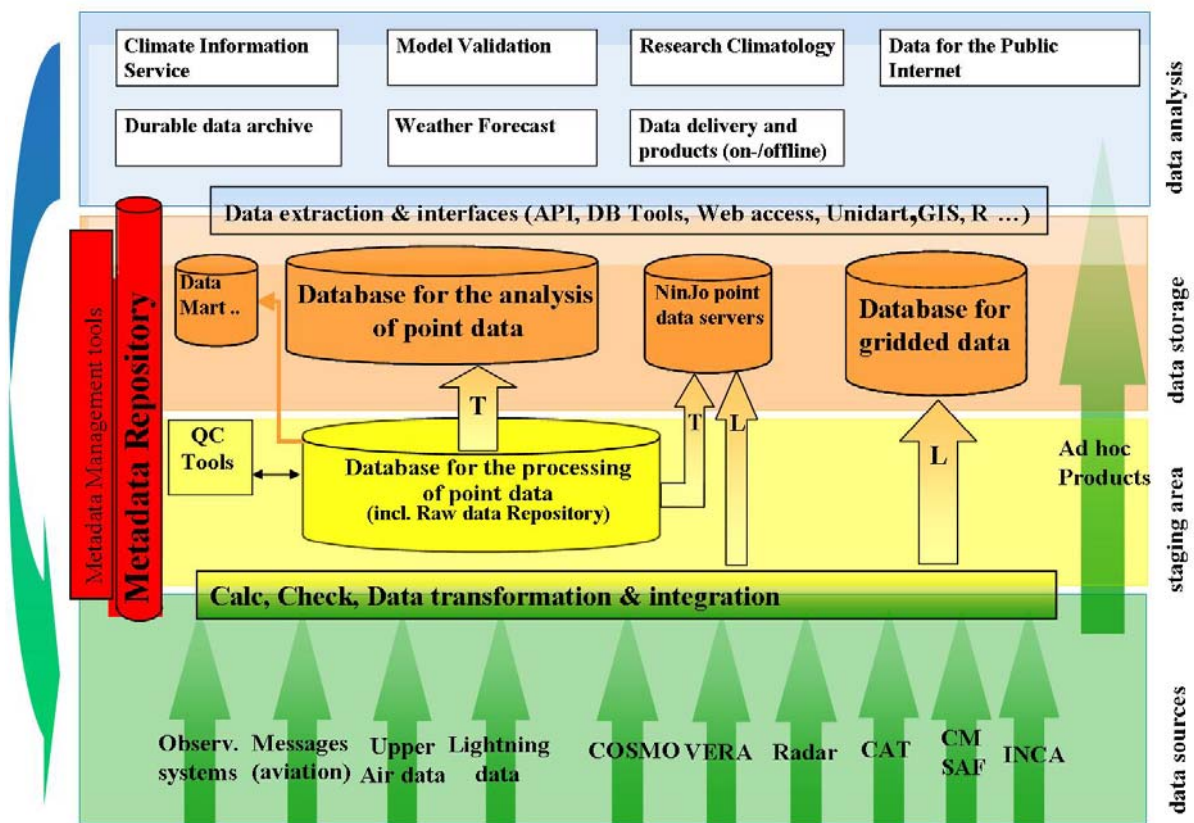


Figure 1 – Four-level architecture of the MeteoSwiss DWH

## History of the DWH

The DWH system was developed in a project that officially started in the year 2000. The resulting reorganisation of the database landscape was based on an expert report by Professors Nievergelt and Zehnder (Department of Computer Science, ETH Zurich) as well as on an expert opinion by Professor Dittrich (Department of Informatics, University of Zurich).

In the course of the development process the new infrastructure became operational by stages, enabling the de-activation of several previous systems. Meeting all the demands resulting from the MeteoSwiss fields of activity required 7 releases; towards the end of 2009 the implementation was complete.

A subsequent project called „RENAISSANCE“ aims at modernising the production services grafted onto the DWH and at expanding data supply processes. It is here that the existing architecture is consolidated and adjusted to new requirements which are the result of current issues.