Primary author: **Haeberli, Christian** (Swiss Federal of Office of Meteorology and Climatology MeteoSwiss, Meteorological Data Coordination), christian.haeberli@meteoswiss.ch

Co-authors: Estelle Grüter, Martin Brändli, Marc Musa, Monika Oswald, Catherine Stocker, Marlen Kube (All: Swiss Federal of Office of Meteorology and Climatology MeteoSwiss)

Abstract ID: 1O4

Data Integration and long-term Planning of the Observing System as a cross-cutting Process in a NHMS

More than five years ago MeteoSwiss brought the new Data Warehouse System into operation. In the meanwhile this system has become the "enterprise-wide" data integration platform. Besides the centralized metadata repository the Data Warehouse System is the anchor for most data procurement and data delivery processes. This includes large parts of meteorological and climatological products and services. A successful data warehouse comprises organizational issues as well as technical solutions. MeteoSwiss decided to assign the operational tasks of the data warehouse as a cross-cutting activity to the cross-divisional unit "meteorological data coordination". The presentation will give a review of the existing system including some details about the organization. It will give an outline of the links between the planning and operation of the observing system, the data procurement and the data delivery processes. Finally some of the present challenges will be discussed.

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Bundesamt für Meteorologie und Klimatologie MeteoSchweiz

Data Integration and long-term planning of the Observing Systems as a cross-cutting process in a NMS

ECSN Data Management Workshop Nov 4th 2009

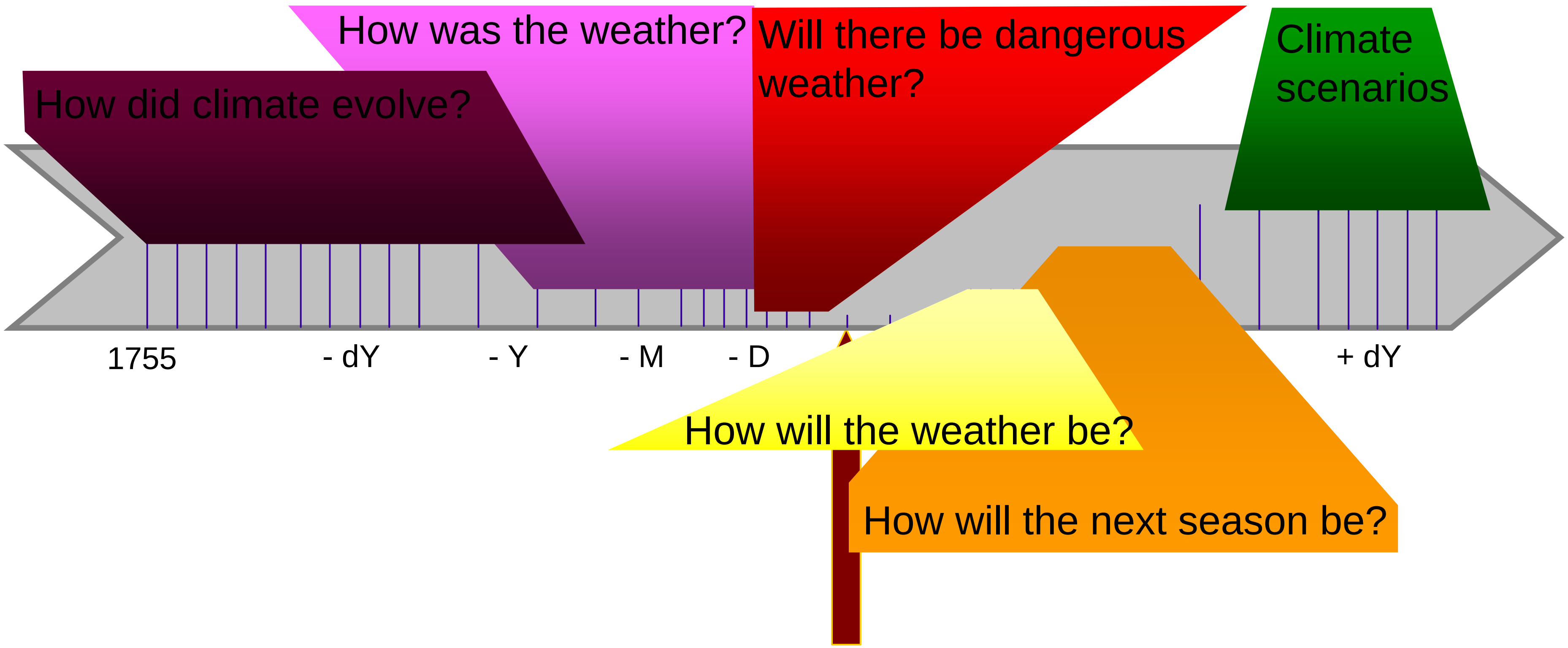
Ch. Häberli
Deputy Head Climate Division/Head Meteorological Data Coordination

Contributions from: Ch. Appenzeller, S. Schärer, Ch. Frei, R. Stöckli, R. Erdin, B. Dürr, U. Germann

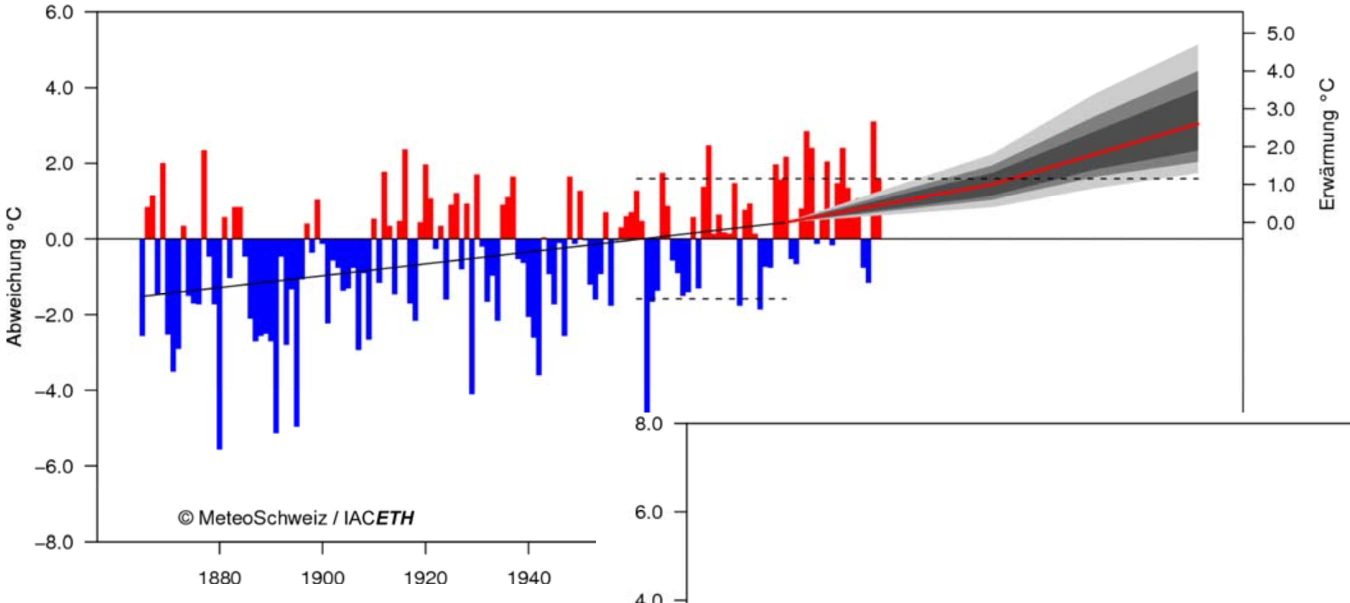
Agenda

- Challanges
- Basic Strategies, Concepts and Blueprints
- An Example: Improving the precipitation picture
- Consequences
- Summary

The Seamless Time Dimension

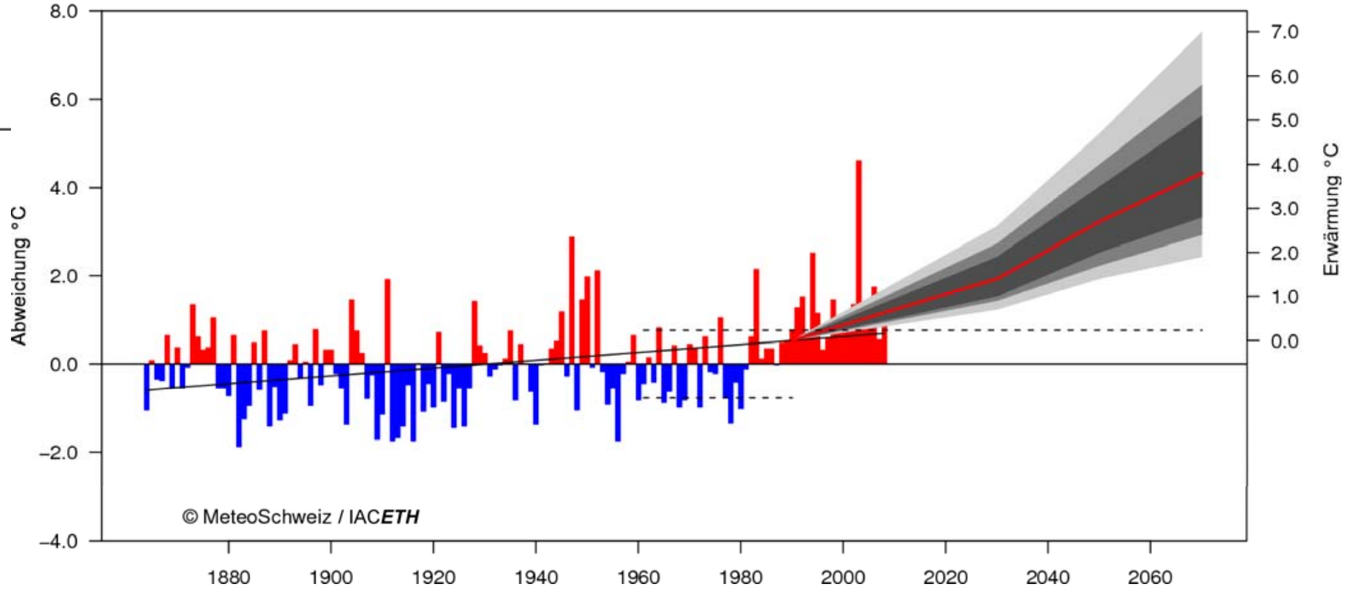


The Seamless Time Dimension

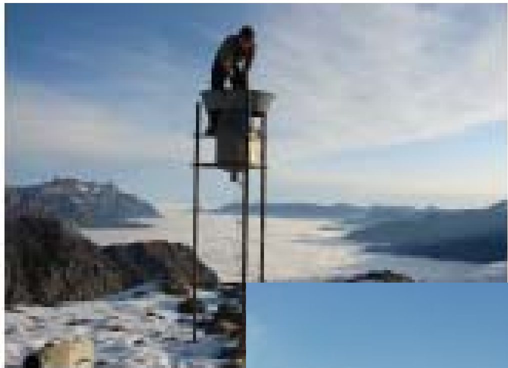


winter

summer



Evolving Technologies



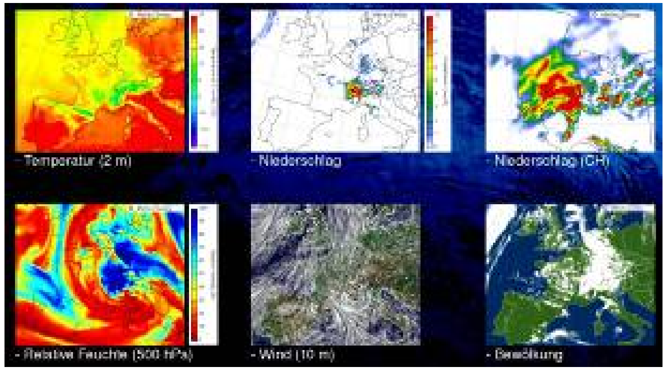
Integration of various Data Sources

Observations (remote sensing / in situ)

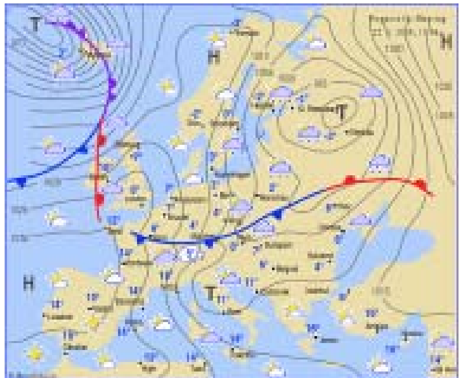
ECMWF

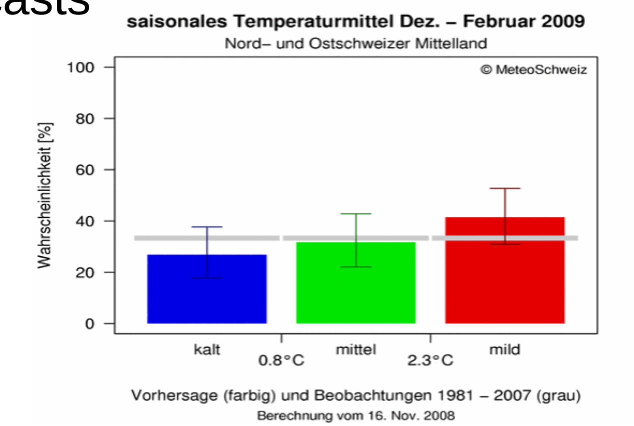
COSMO-2

Output from NWP



Weather and Climate forecasts





Basic Strategy

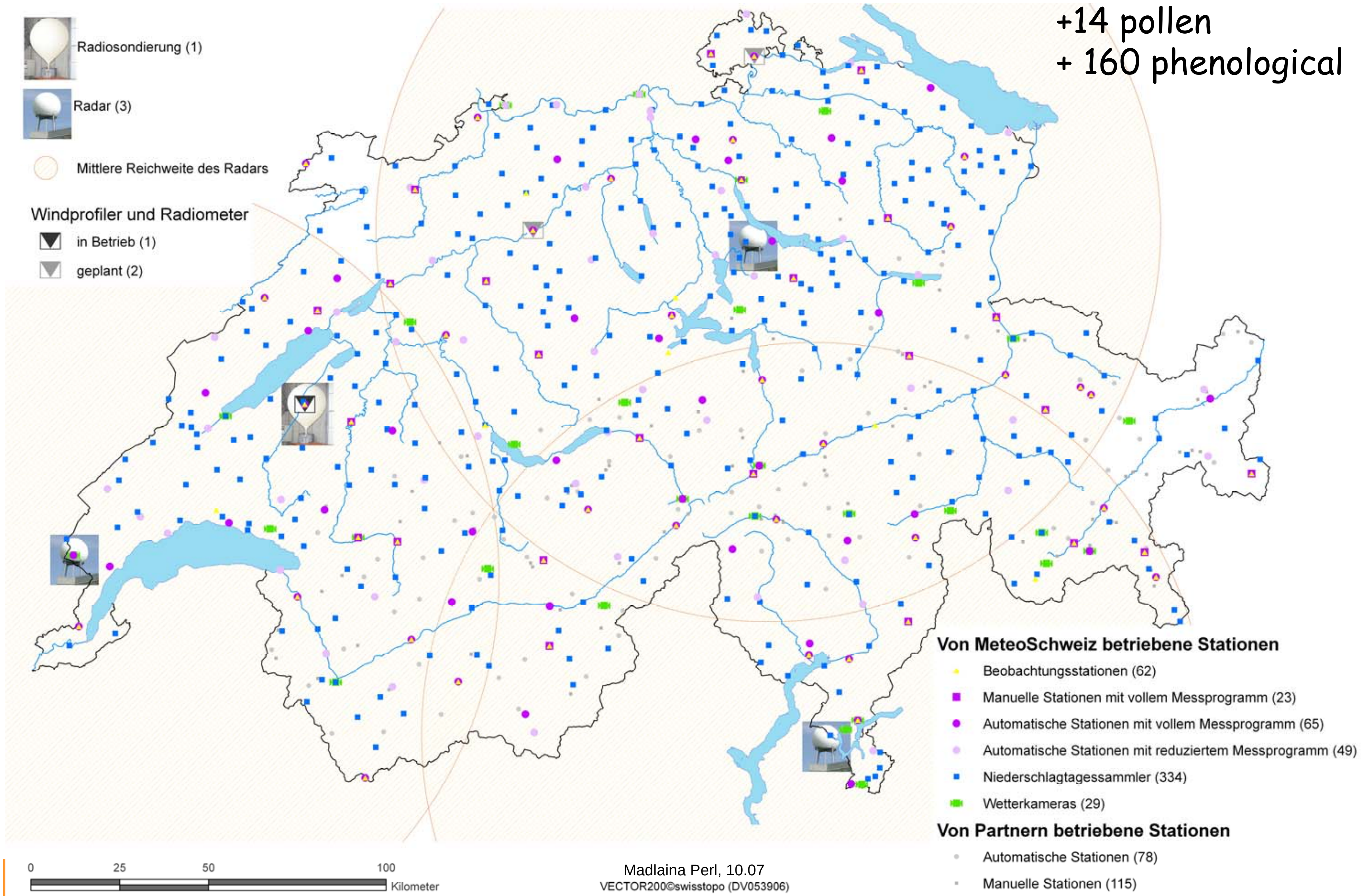
- MeteoSwiss operates the backbone of the monitoring systems in Switzerland and the nationwide data integration platform for measurements in the atmosphere
- The data procurement and data processing strategy of MeteoSwiss is an <u>integration strategy</u>.
- Aspects of integration
  - various observing technologies and platforms
  - observing systems outside MeteoSwiss
  - datea und meta data

→ **<u>comprises all aspects of „Enterprise Data Integration“</u>**

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Bundesamt für Meteorologie und Klimatologie MeteoSchweiz

+14 pollen
+ 160 phenological

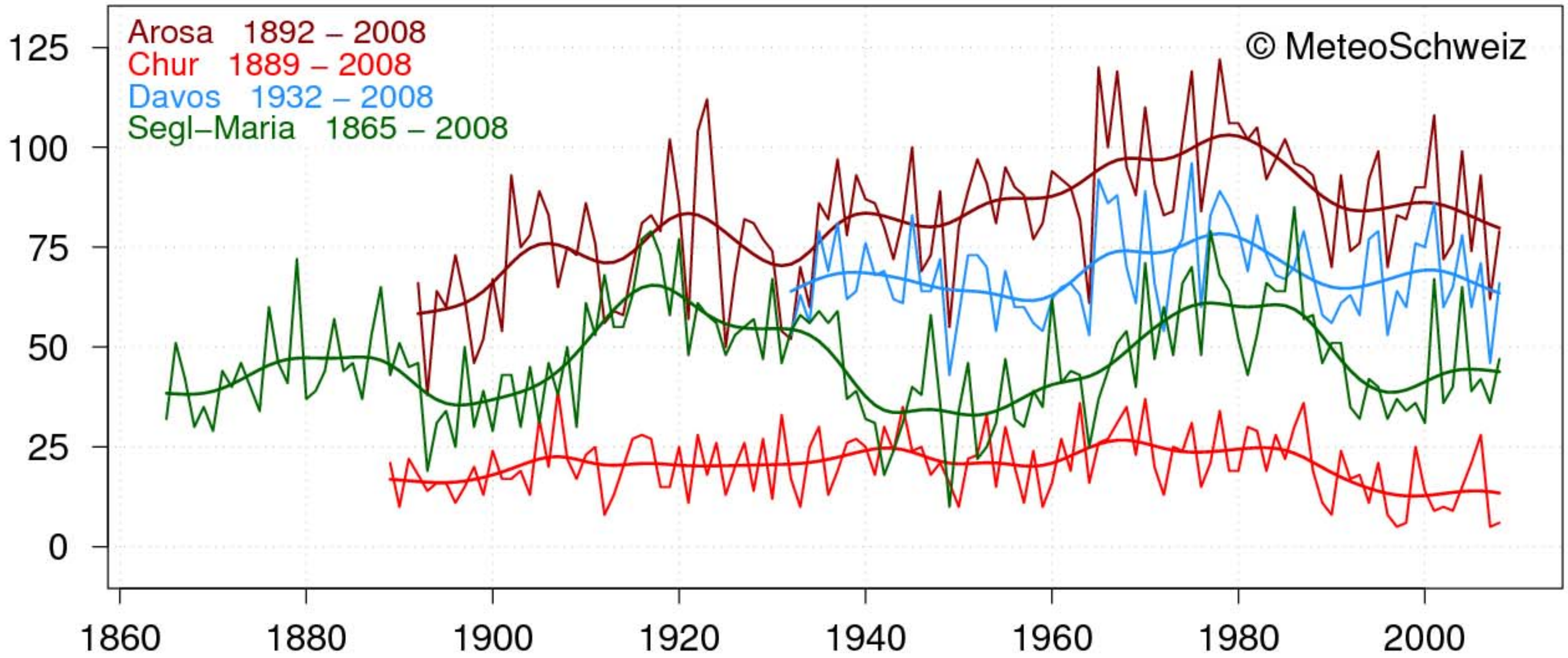


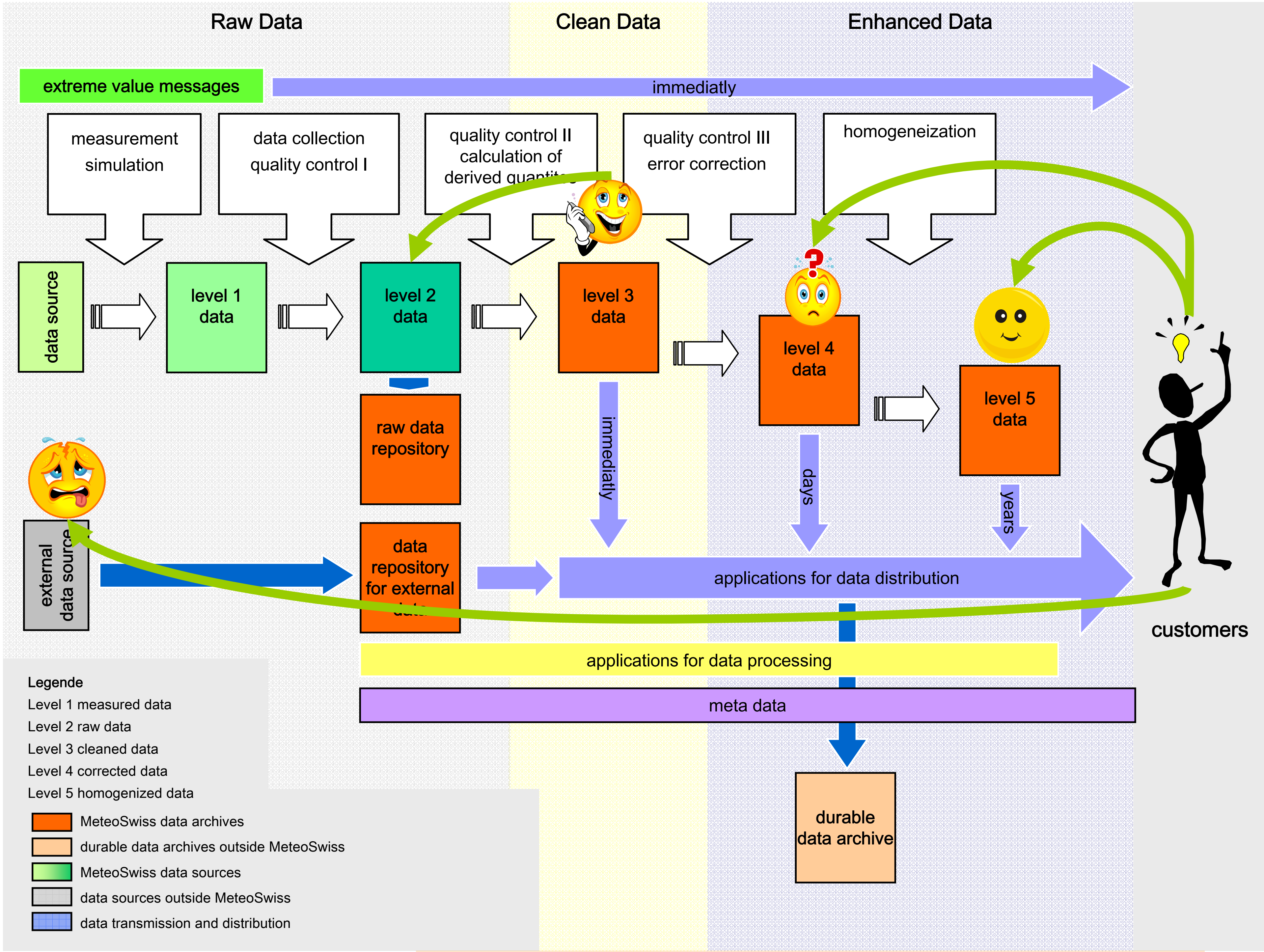
Madlaina Perl, 10.07

The backbone of the atmospheric monitoring systems

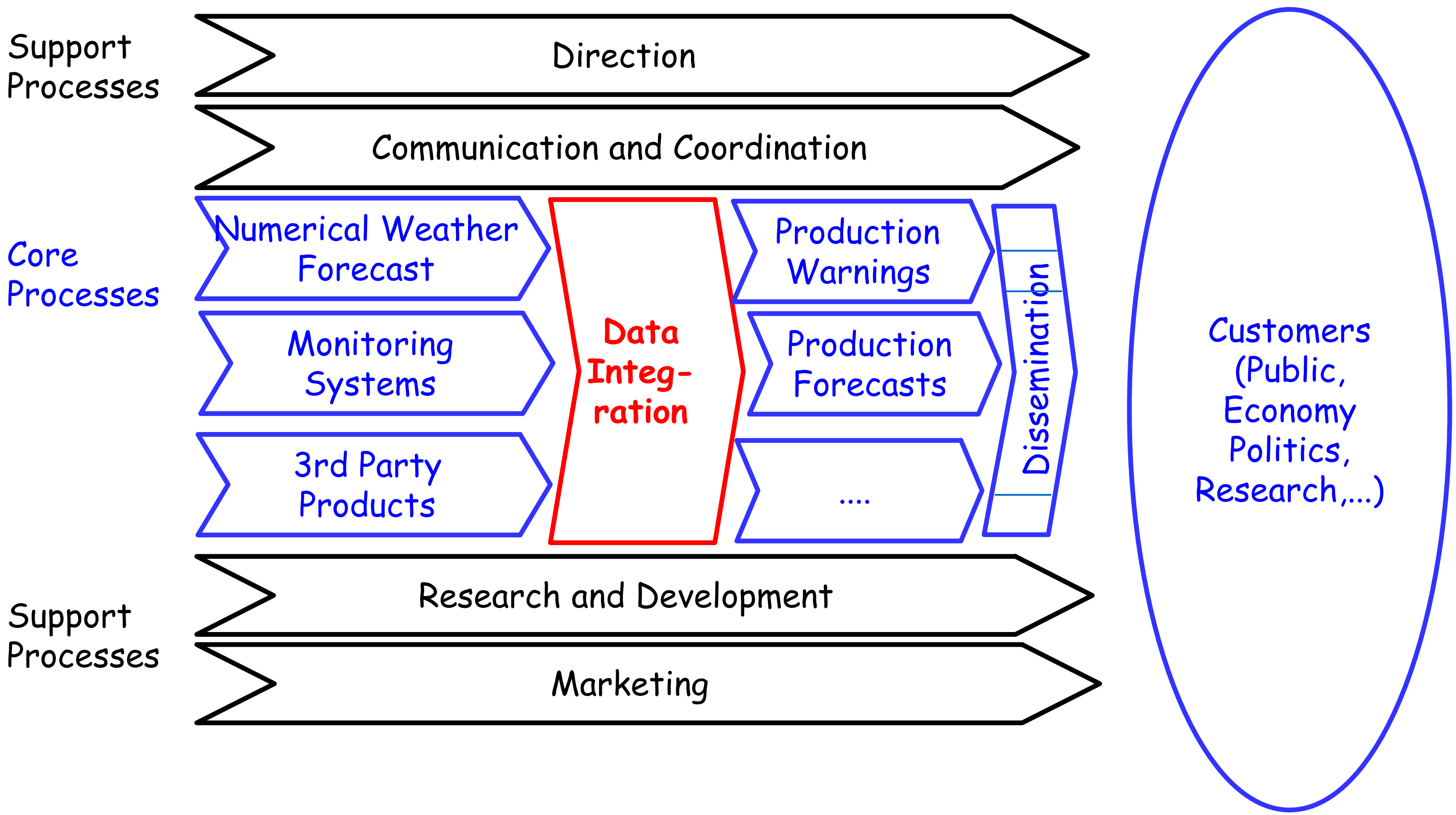
Very long Swiss snow series (1864-)

long-term trends of days with snowfall ≥ 1 cm

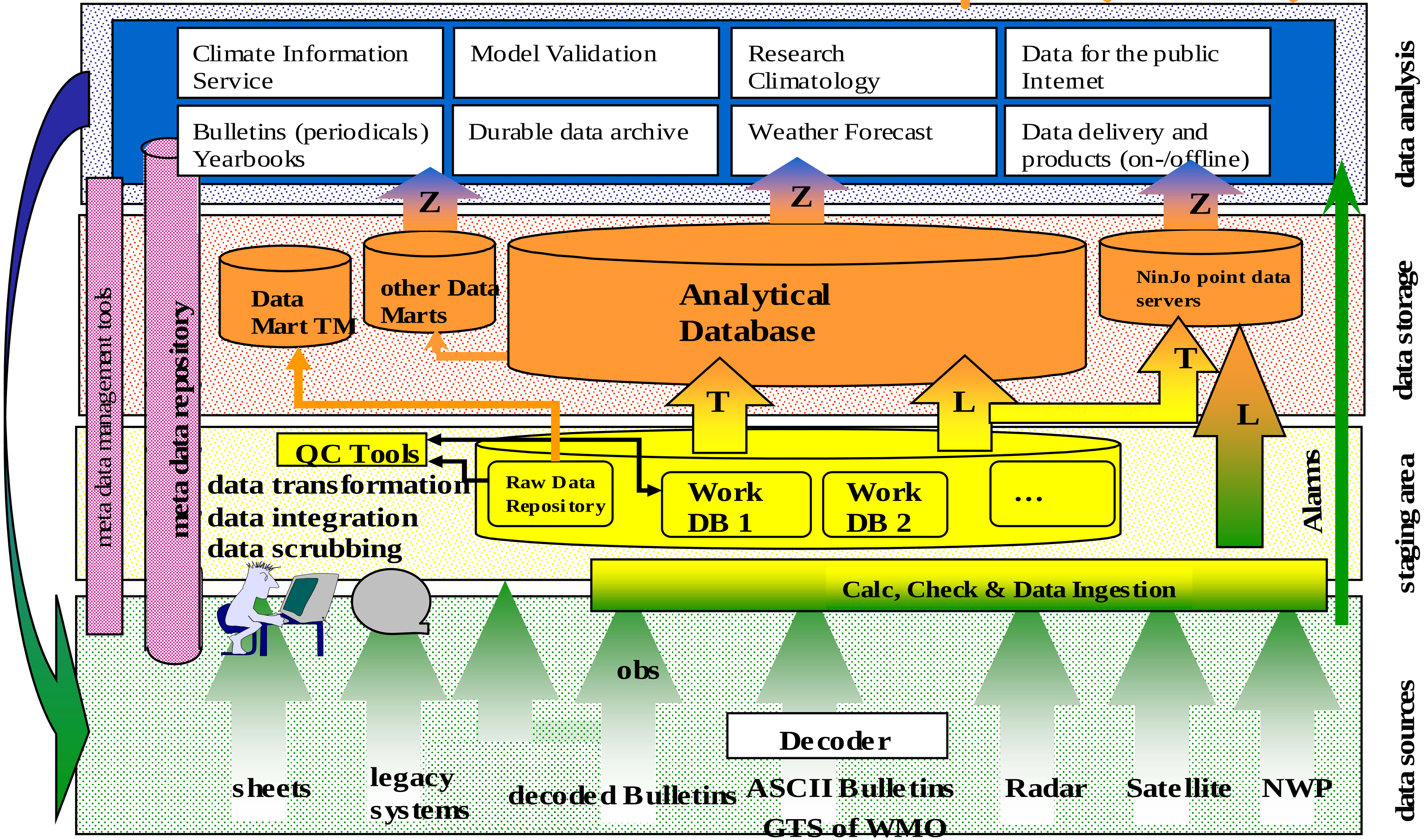



Raw Data
Clean Data
Enhanced Data
extreme value messages
immediatly
measurement simulation
data collection quality control I
quality control II calculation of derived quantit
quality control III error correction
homogeneization
data source
level 1 data
level 2 data
level 3 data
level 4 data
level 5 data
raw data repository
immediatly
days
years
external data source
data repository for external data
applications for data distribution
customers
applications for data processing
meta data
durable data archive
Legende
Level 1 measured data
Level 2 raw data
Level 3 cleaned data
Level 4 corrected data
Level 5 homogenized data
MeteoSwiss data archives
durable data archives outside MeteoSwiss
MeteoSwiss data sources
data sources outside MeteoSwiss
data transmission and distribution

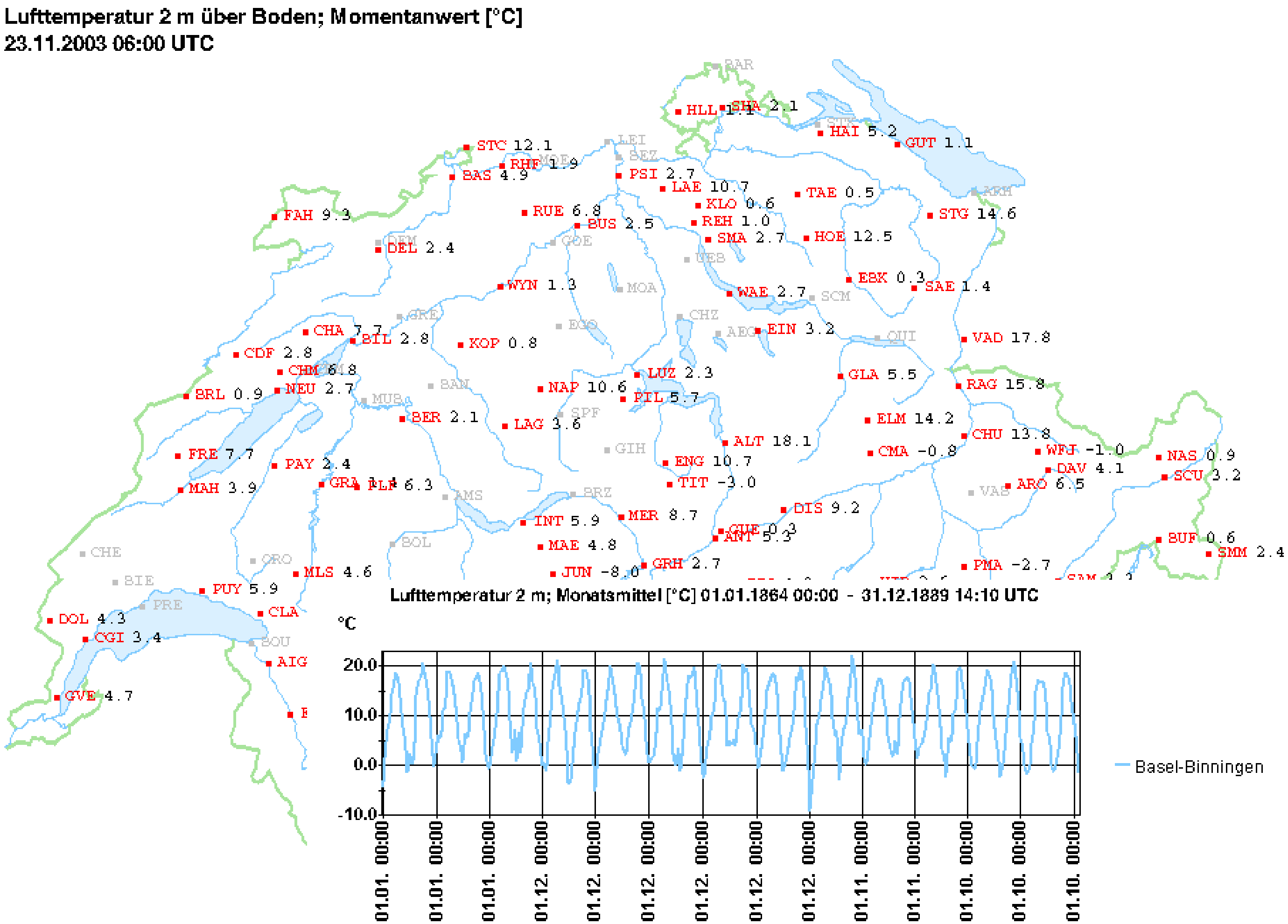
Organisation Concept



Architectural concept (CIF)

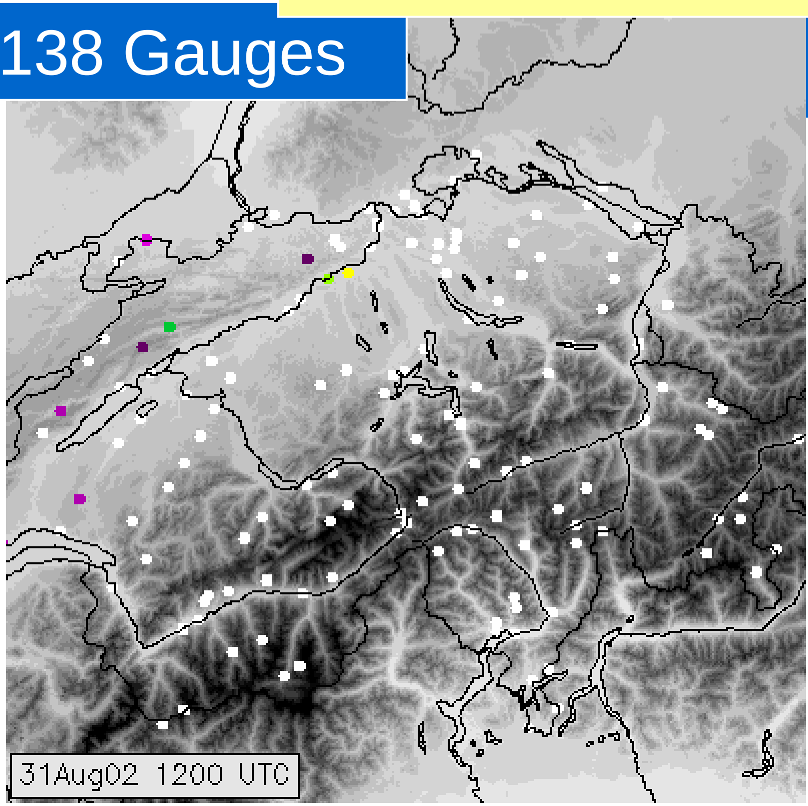


The backbone of the atmospheric monitoring systems



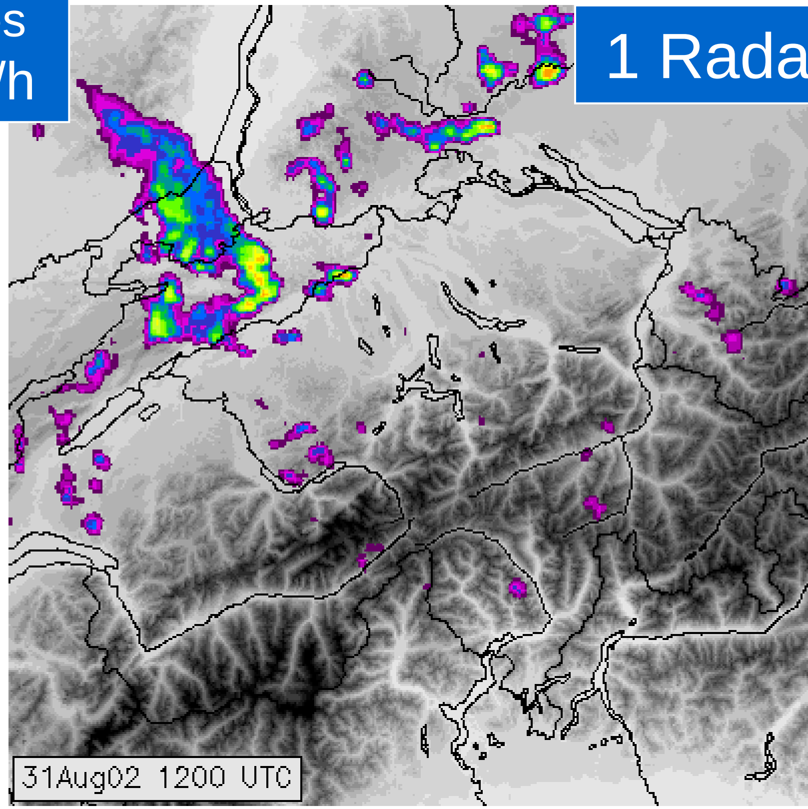
Here, we artificially double the number of gauges by creating for each existing automatic raingauge a hypothetic random daughter-station some km to the east and to the north.

138 Gauges



Red hues >40mm/h

1 Radar



Ex 1: Combination Radar – surface in situ precipitation

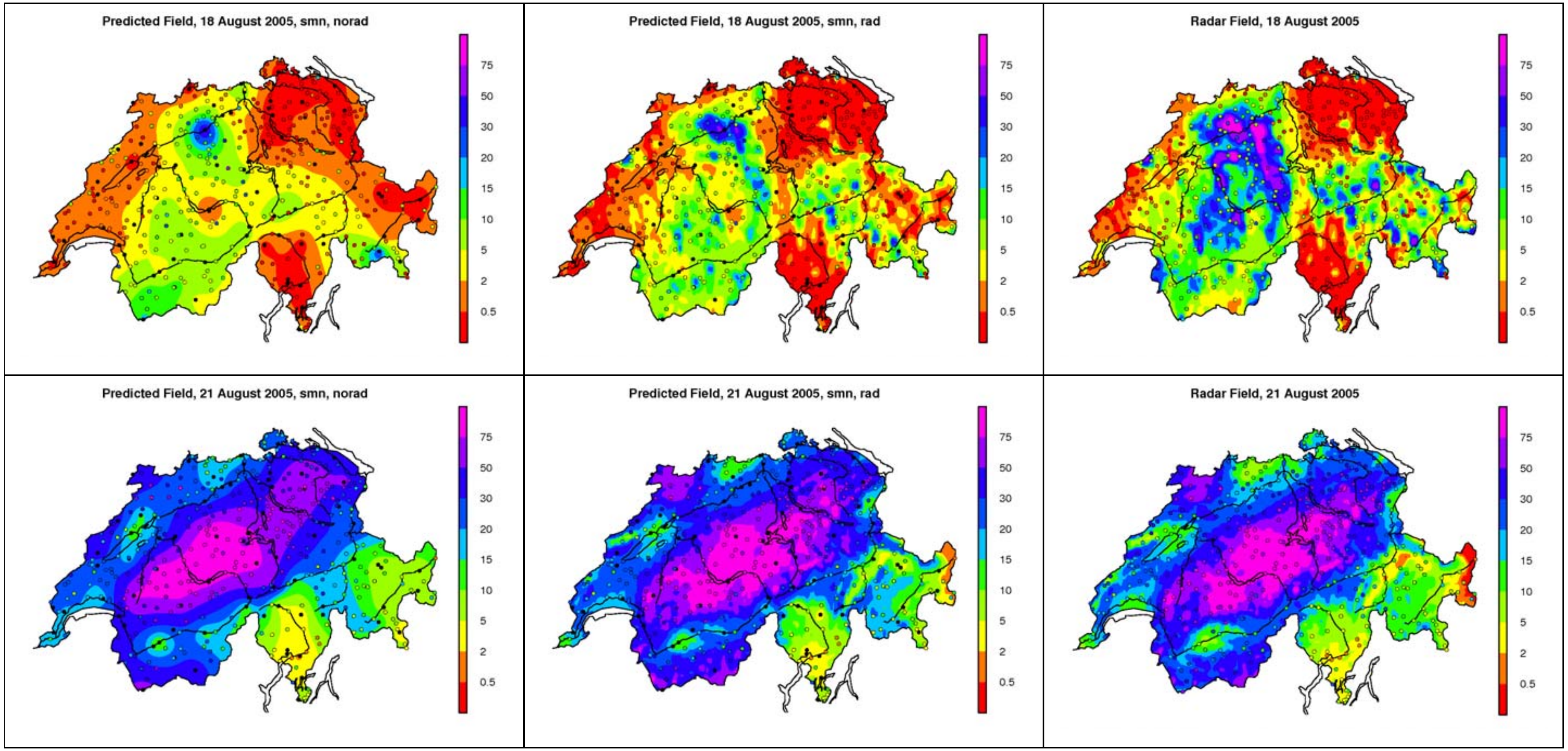
station data

+ accurate values at stations
- coarse resolution (15km, 24h)
- uncertainty between the stations

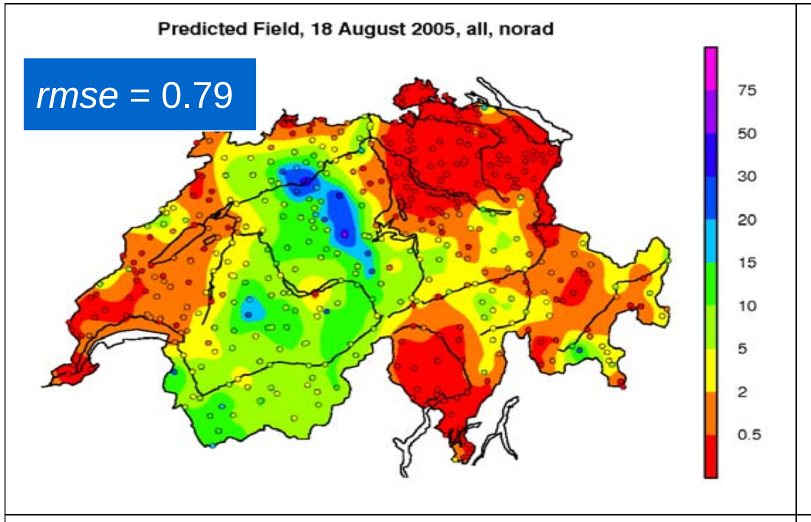
radar data

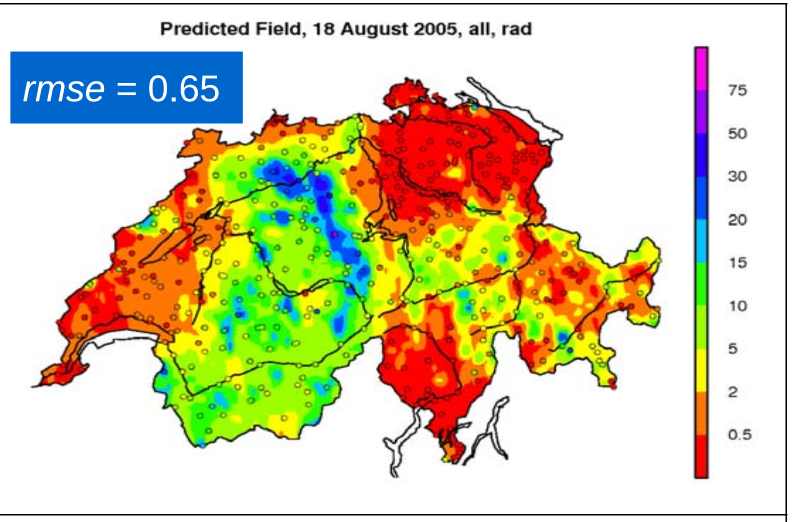
+ high spatial and temporal resolution (1km, 5min)
- uncertainty in absolute values, often bias

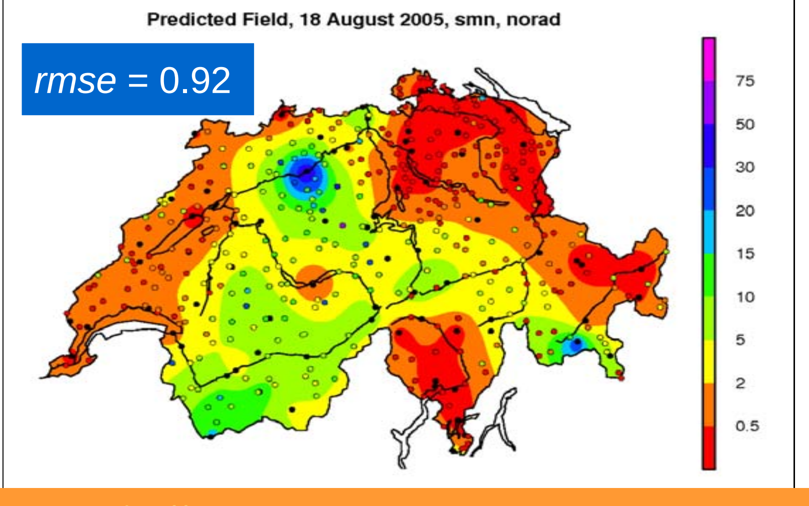
Ex 1: Combination Radar – surface in situ precipitation

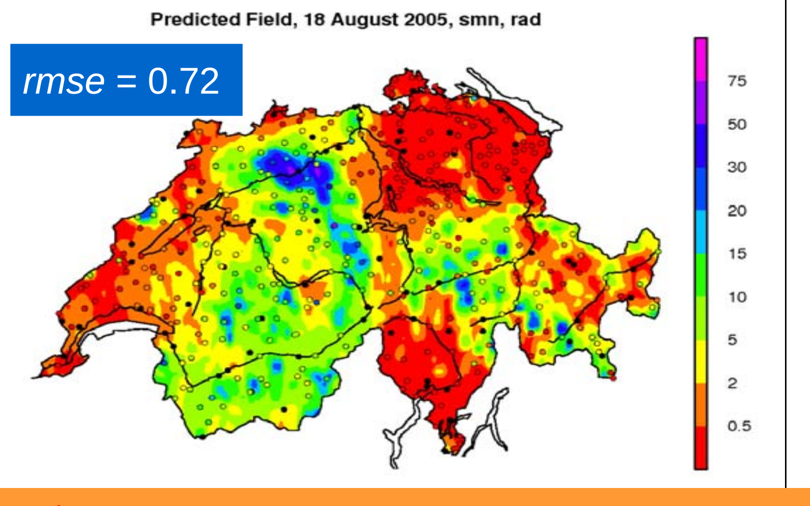


Ex 1: Combination Radar – surface in situ precipitation

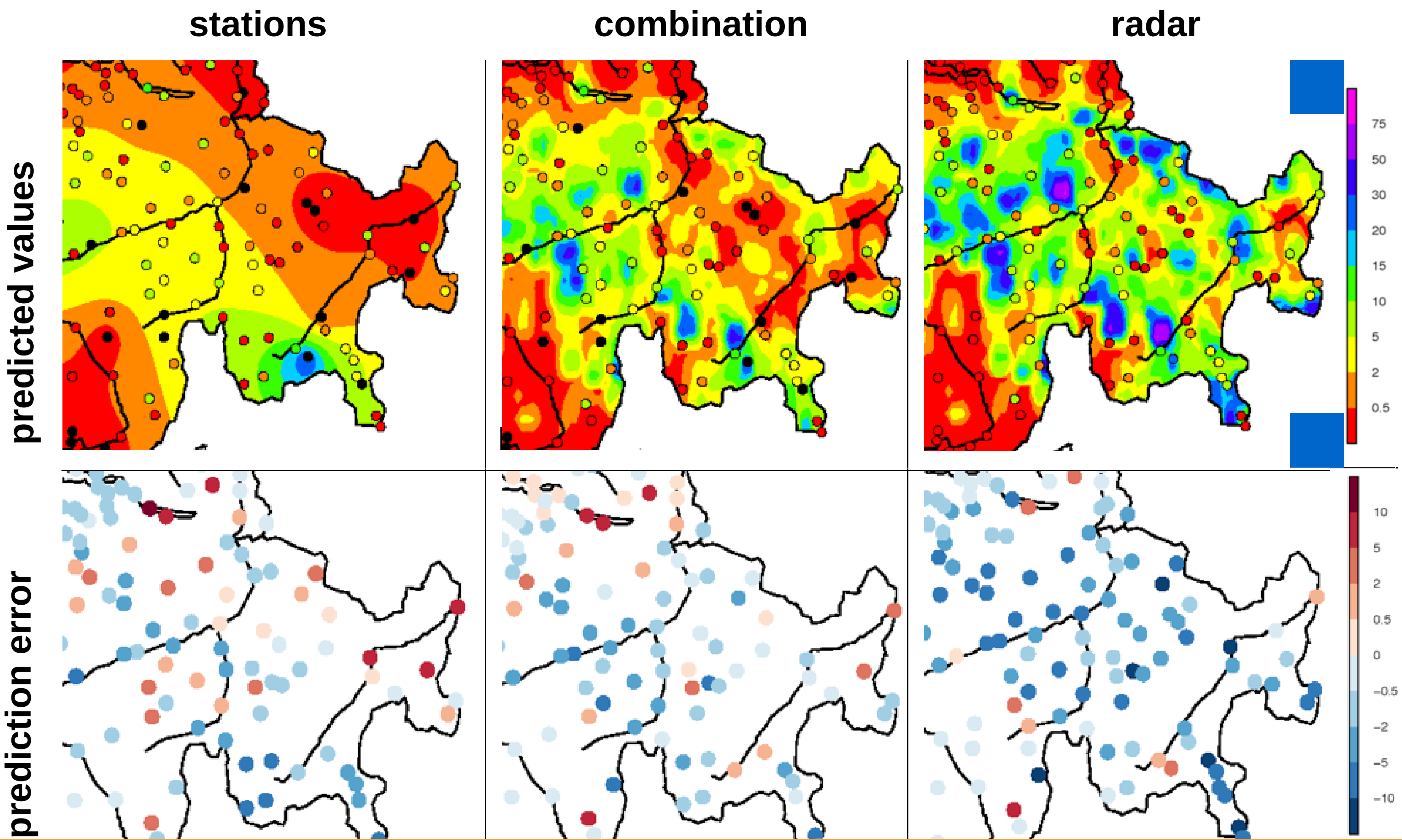






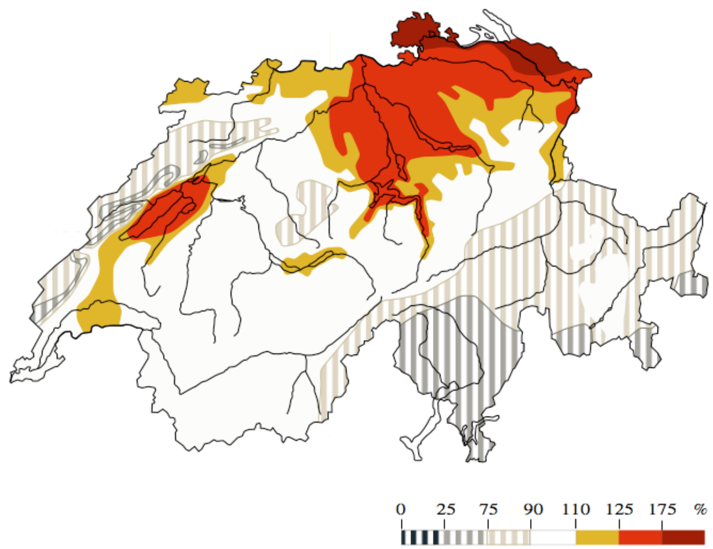


Ex 1: Combination Radar – surface in situ precipitation



Ex 2: Combination satellite – surface in situ sunshine

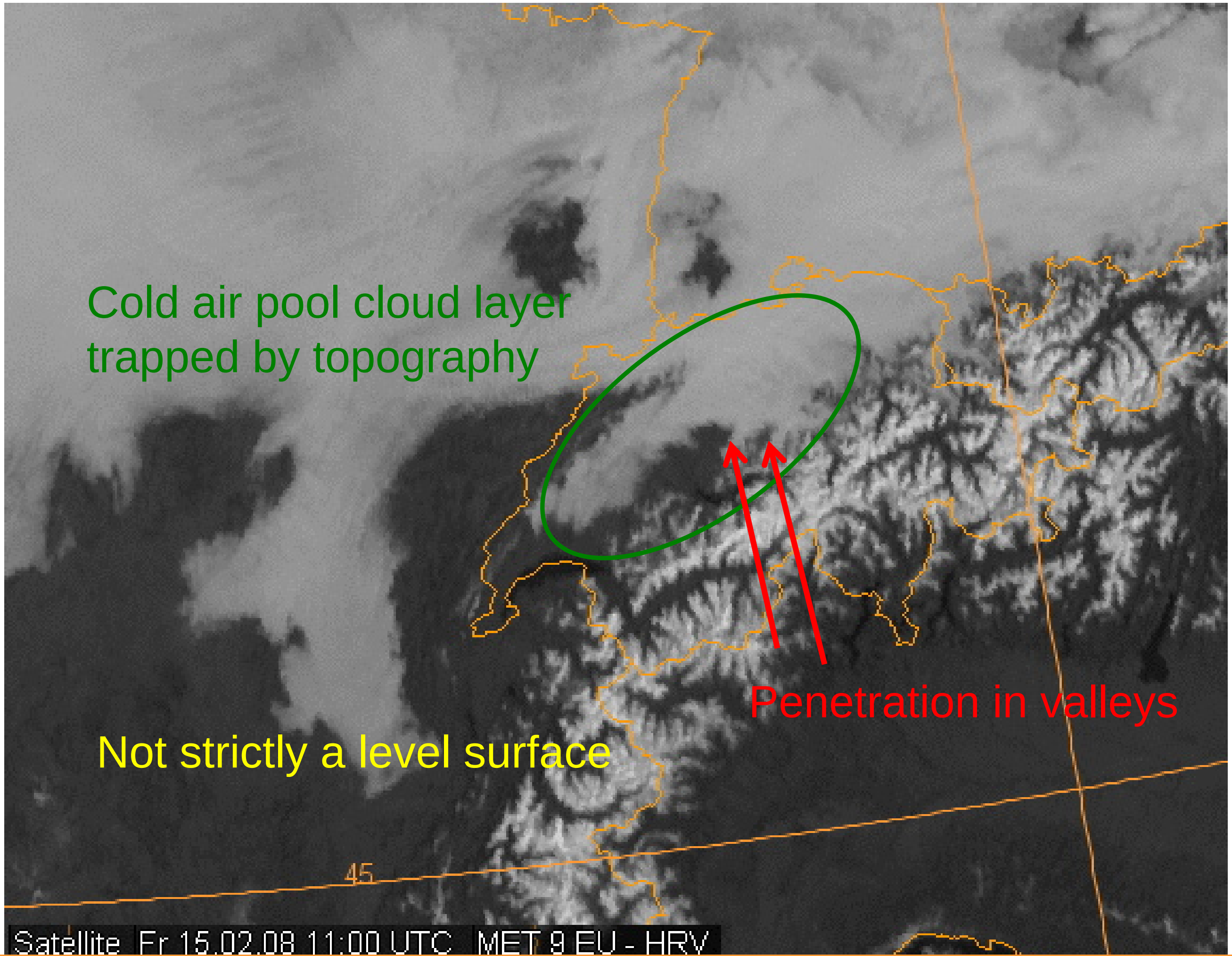
- **Motivation:**
  - SW radiation
    - → forces water cycle
    - → input variable for modelling the effect of climate change
  - relative sunshine duration:
    - → key parameter for long-term monitoring radiation and cloudiness



Sunshine anomaly January 2001 (ratio of 1961-1990 mean).
Hand drawn contours.
Station data + intuition

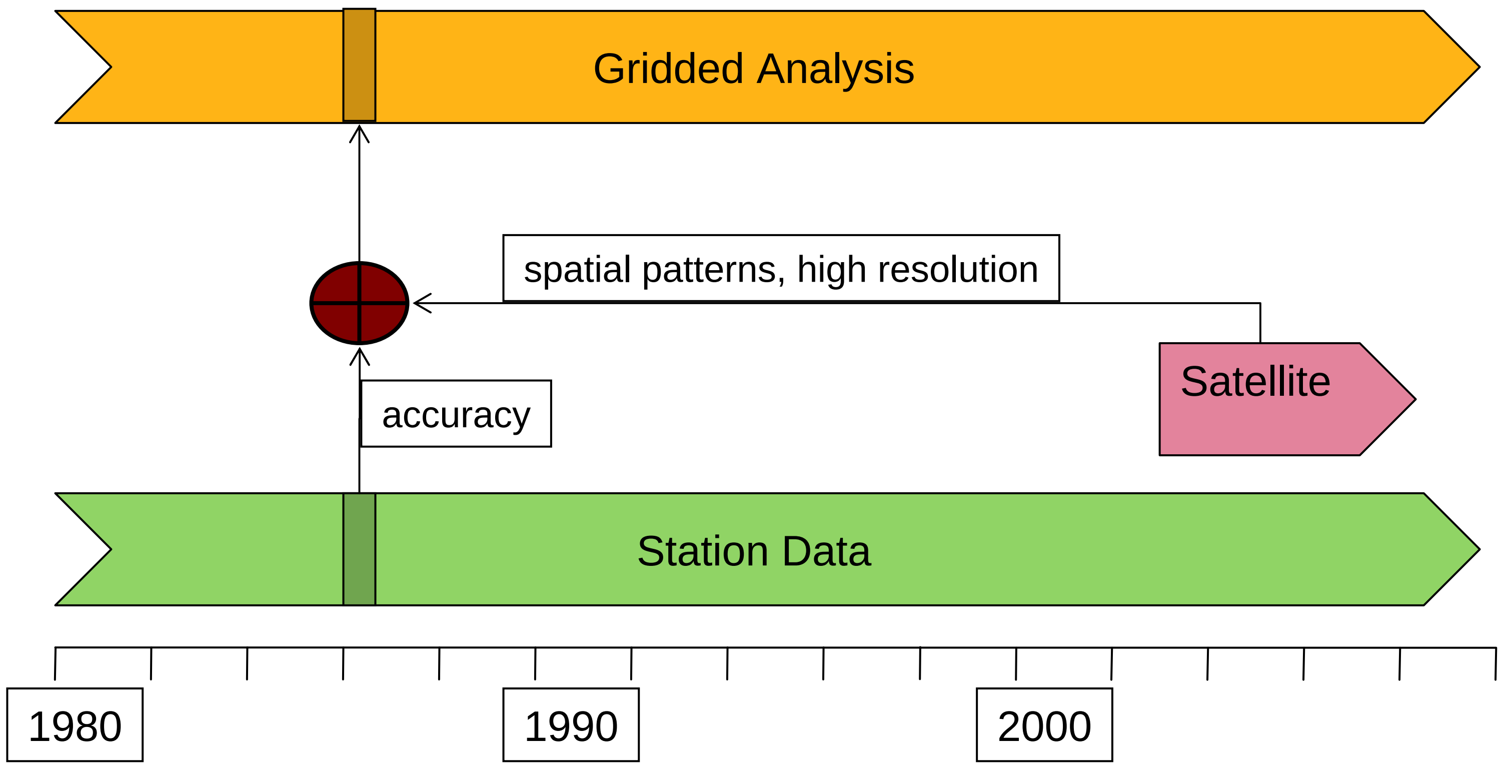
But we have more information....

Ex 2: Combination satellite - surface in situ sunshine



Ex 2: Combination satellite - surface in situ sunshine

- **Non-contemporous information merging**



- **Similar outset like in climate reconstruction tasks**

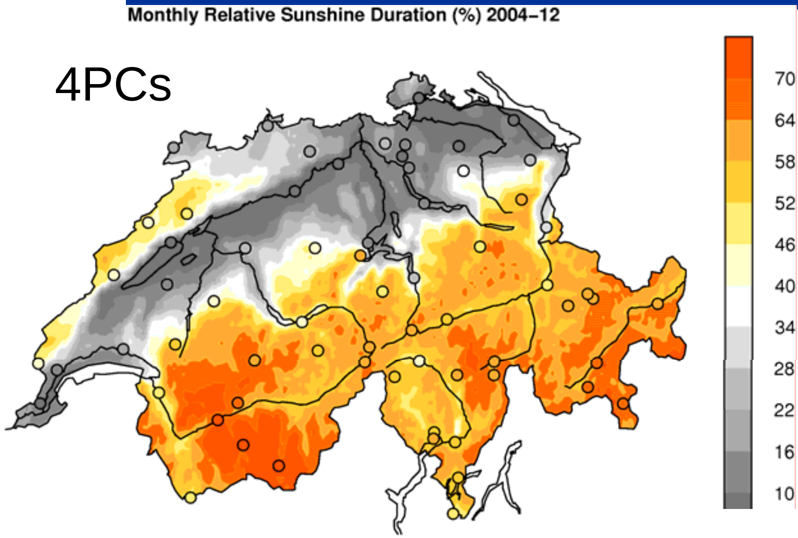
Ex 2: Combination satellite – surface in situ sunshine

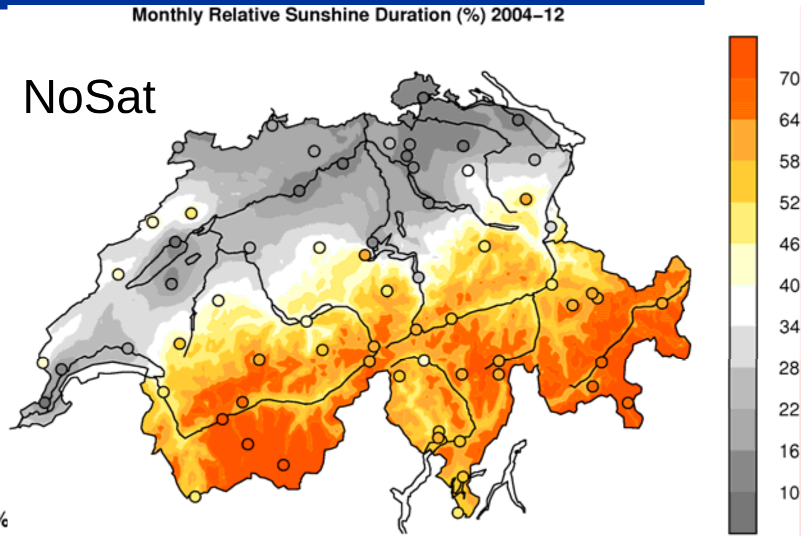
- stations:
  - relative sunshine duration from ~70 stations
- satellite:
  - Heliosat Clear Sky Index (Clearness Index CIN, Cano et al. 1986)
  - 5% (overcast) – 120% (clear)
  - Implemented for the Alps from MSG, special consideration of snow. (Dürr & Zelenka 2009)
  - Part of CM-SAF project
  - Monthly, 2004 – 2008; Resolution ~ 2 km
  - No station data used.
  - Not real-time so far

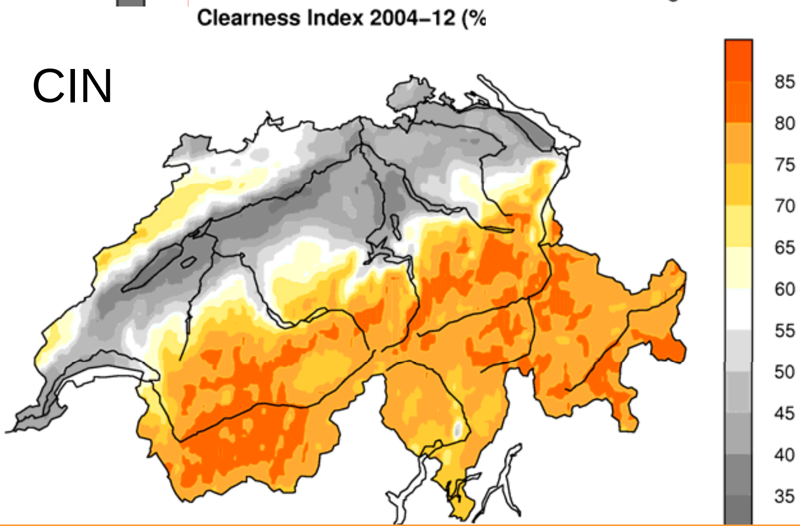
Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Bundesamt für Meteorologie und Klimatologie MeteoSchweiz

Ex 2: Combination satellite - surface in situ sunshine







Ex 2: Combination satellite – surface in situ sunshine

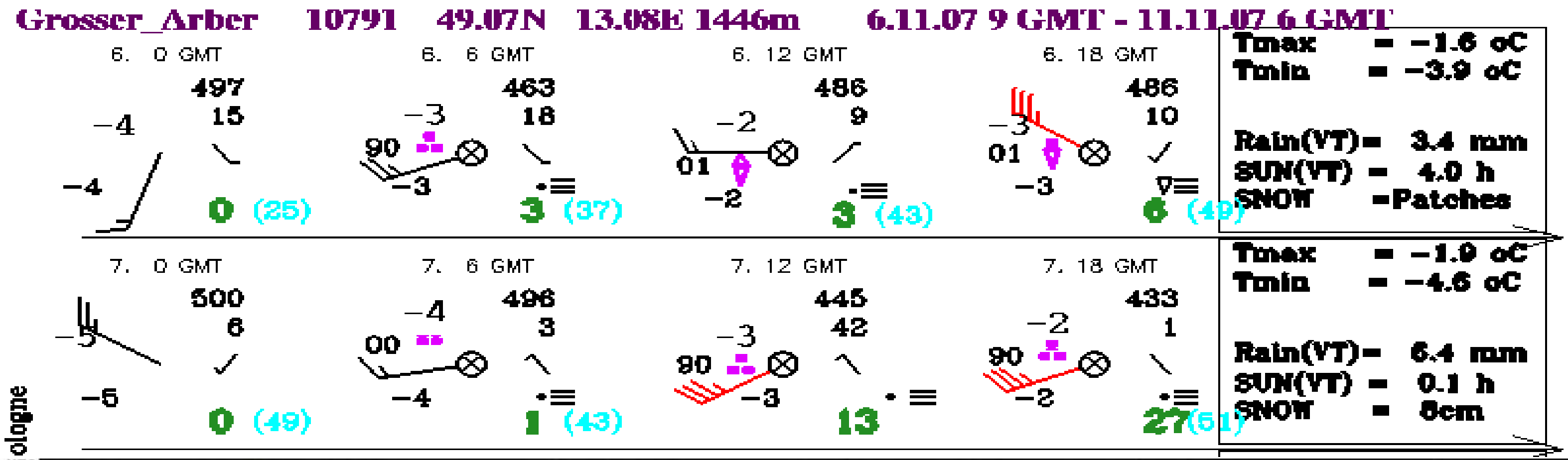
- Combination of satellite and surface in situ sunshine data → added value für gridded data products – also in the pre satellite era
- Satellite data
  - ... introduce plausbile details
  - ... improve gridding, almost always measurable
  - ... of special value in special cases
  - ... can not be replaced by geo-topographical predictors
- Method explains 75-95% of the spatial variance during winter (oct...mar) and 40-80% in summer. Mean absolute errors smaller in summer than in winter.

Consequence → monitorig systems

- seamless time dimension:
  - lasting operation
  - keep the system adaptive (to new technologies, monitoring needs...)
  - multifunctional systems
- automation and modernization
- various platforms (in situ and remote sensing) → syntactic and semantic interoperability

Automation and Modernization

__Challenge__: Increase information content at reduced cost : replacment of observers by cameras (~ - 10 kCHF/Station-year)



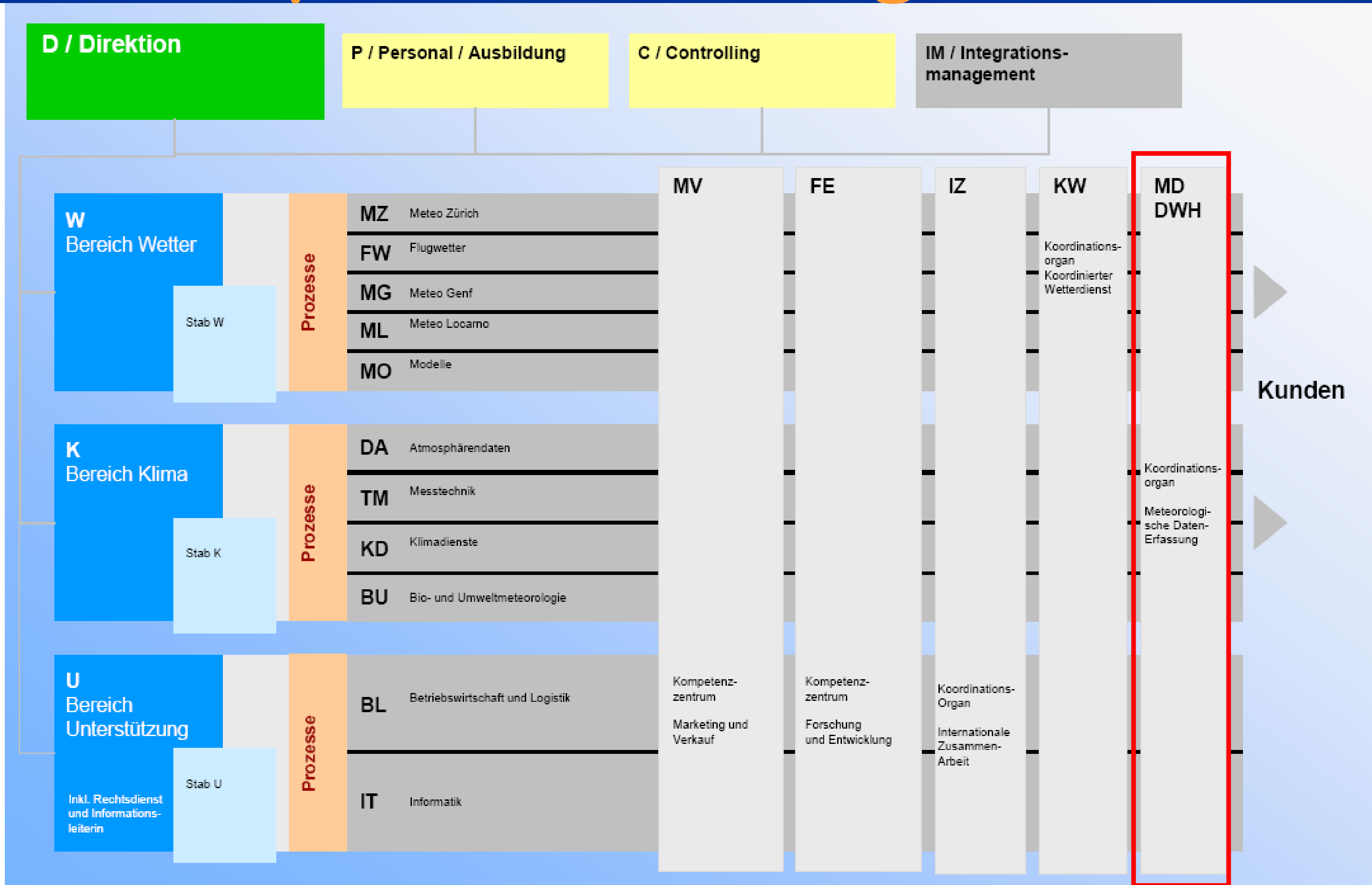
Consequence → data integration

- keep the integration platform adaptive (allow analyses with data from 1864 till 2009)
- allow editing for about 50% of the data set
- high importance of meta data
- data Owner ≠ product owner
- world wide data exchange → data stewardship? data quality?
- permanent increase of data volume
- real-time, 24h/365d, max downtime: 15'
- 1 data source for monitoring the past climate, warnings, forecasts
- continuous new requirements, technologies, regulations (WMO, ICAO) but no classical requirement life cycle → automation and modernization

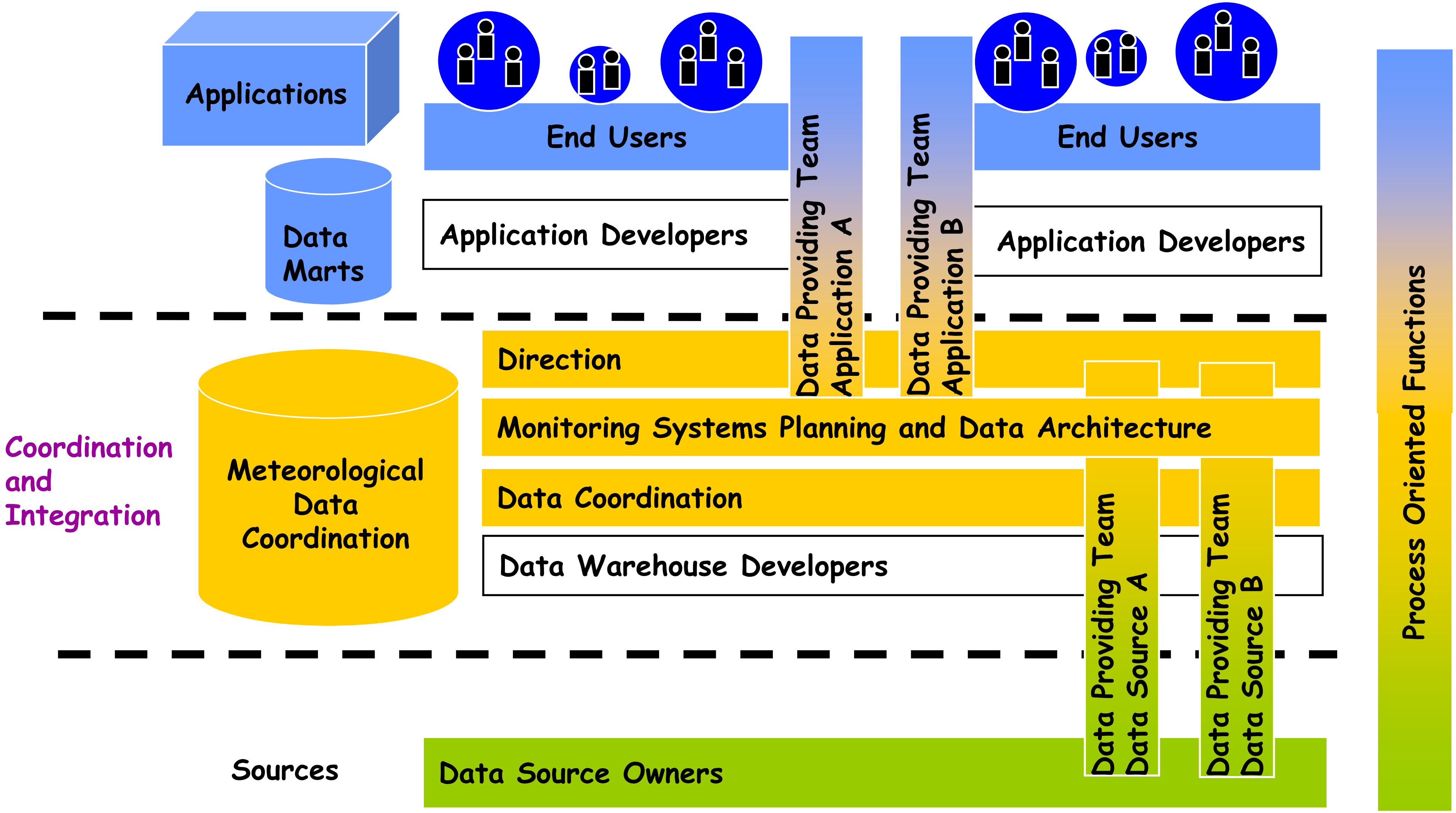
Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Bundesamt für Meteorologie und Klimatologie MeteoSchweiz

Consequence → Organisation



Consequence → Organisation



Consequence → Organisation

Main tasks of the cross cutting unit

- Long-term (strategic) planning of the data procurement chains → architecture blueprints and standards for projects
- Support for projects: fitting the projects into the architecure; change management for the architecture
- Support in daily operations: operations coordination of the data warehouse system

Summary (1)

- Maintaining a monitoring system for the atmosphere is a permanent process and challenging cross-cutting process with a strong Integration aspect.
- The main challenge is to establish a rigorous evolution process by:
  1. regular analysis of the requirements (comprehensive!)
  2. regular review of observation capabilities
  3. gap identification and budget review
  4. implementation and operation

for all parts of any „data chain"

Summary (2)

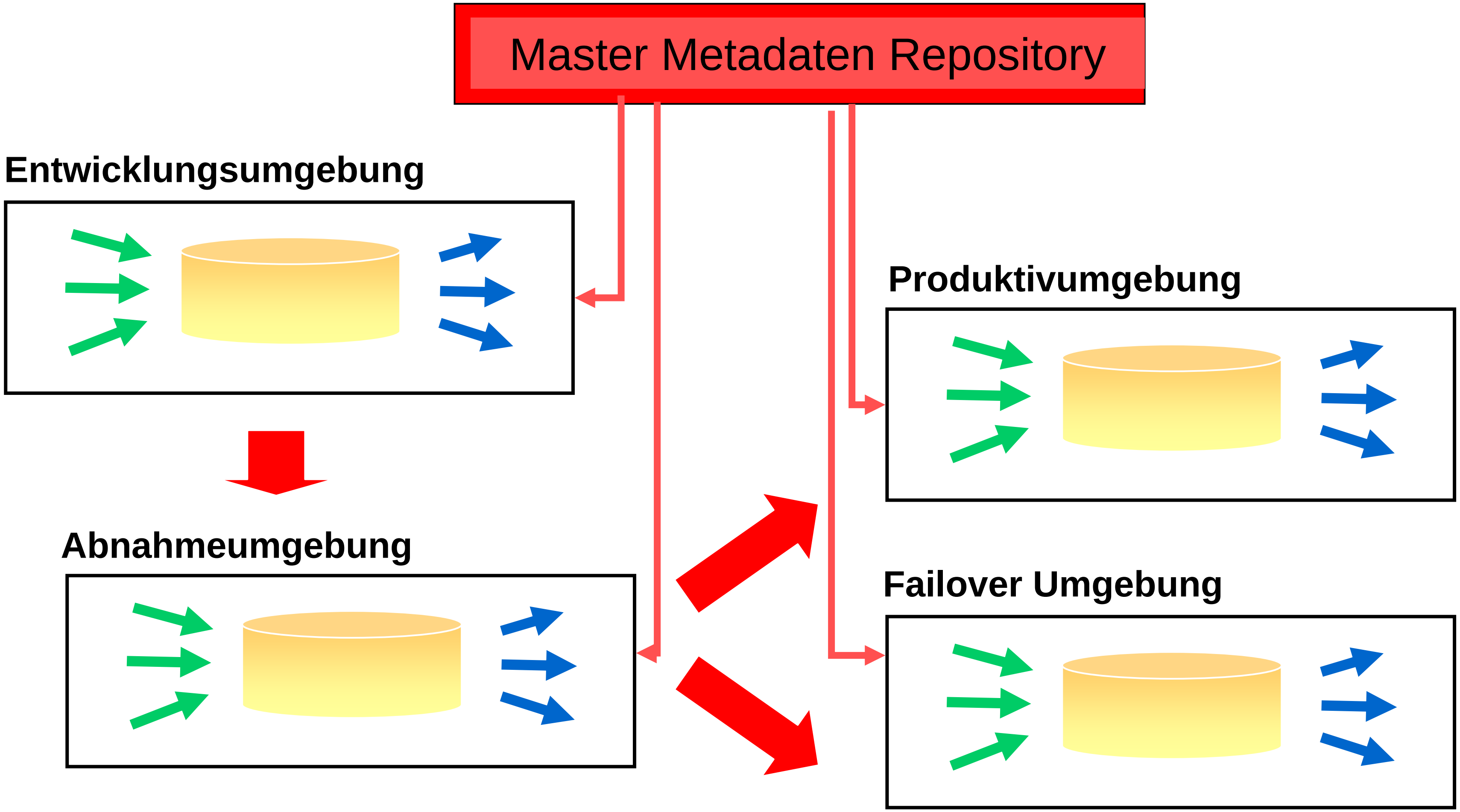
- **The strengths of MeteoSwiss' observing system are:**
  - the mix of observing technologies and platforms (including the access to data from external sources)
  - the balance between advanced, state-of-the-art and conventional technologies
  - a highly „integrated" data management system (which includes data starting 1755)

Thank you for your attention

Kundennutzen

- **Voll konfigurierbare (d.h. flexible) Aufbereitung von meteorologischen und klimatologischen (und ähnlichen) Daten;**
- **Intern:**
  - Eine einzige Datenhaltung für alle meteorologischen und klimatologischen Daten sowie die zugehörigen Kontextdaten
  - Integration von „externen" Datenquellen (Bsp Temperatur: ca 300 Messpunkte verfügbar)
- **Extern (v.a. Partner aus Bundesverwaltung oder Kantone):**
  - Datenhaltung und Archivierung durch MeteoSchweiz
  - Integration von verschiedenen Datenquellen
  - Datenzugriff über Extranet-Applikationen oder Internet

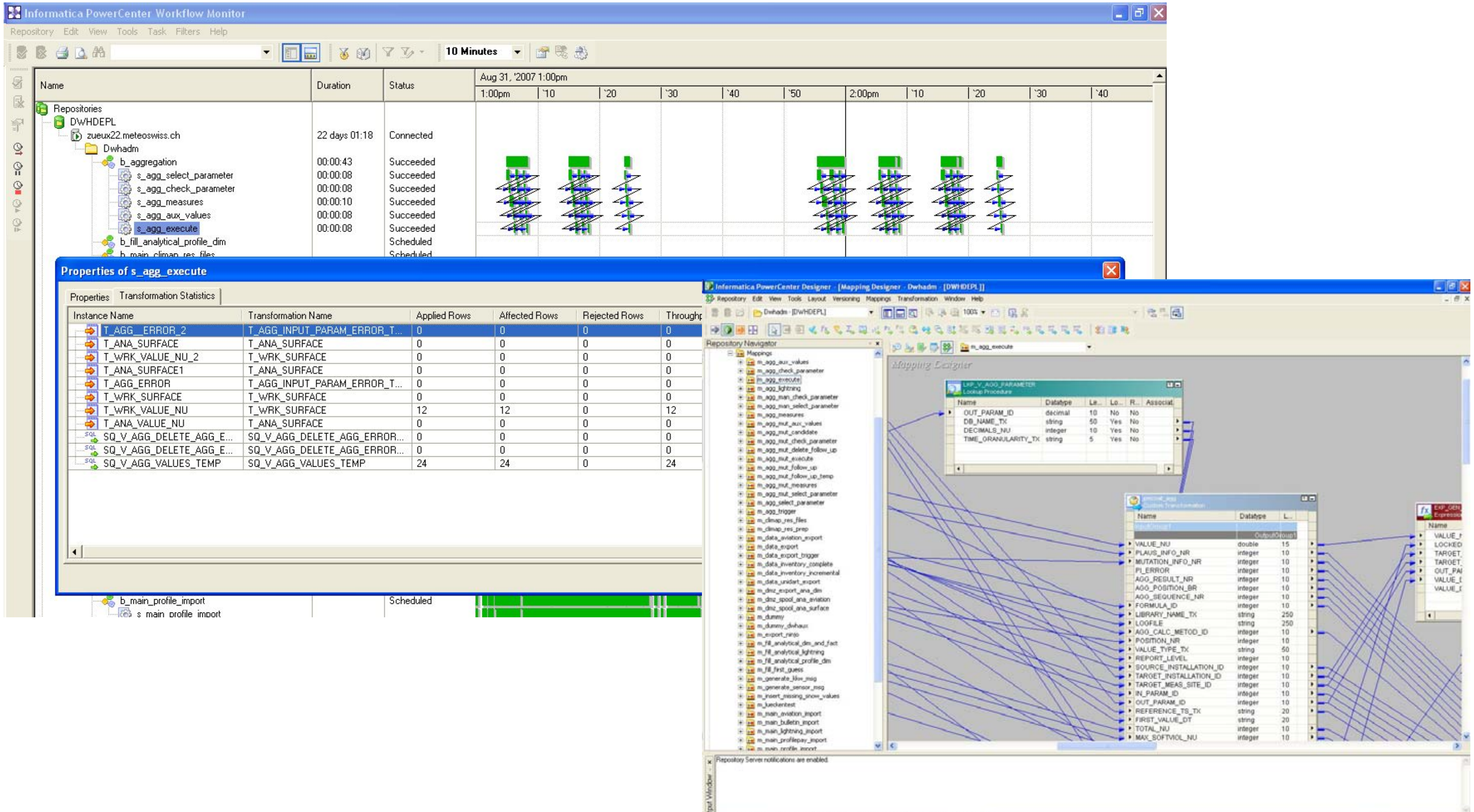
Entwicklungszyklus & Metadatenmanagement



Implementierung

	Production	**Failover Production (hot standby)**
HW	Sun Fire 6800	Sun Fire 6800
OS	Solaris10	Solaris10
RDBMS	Oracle10g	Oracle10g
ETL	PowerCenter 7.1.4	PowerCenter 7.1.4
Datenbestand	1755 bis aktuell	Letzte 60 Tage
Max downtime	Max 15' pro Unterbruch auf Production (mit Failover Production Betrieb während ca 8 Tagen möglich)	

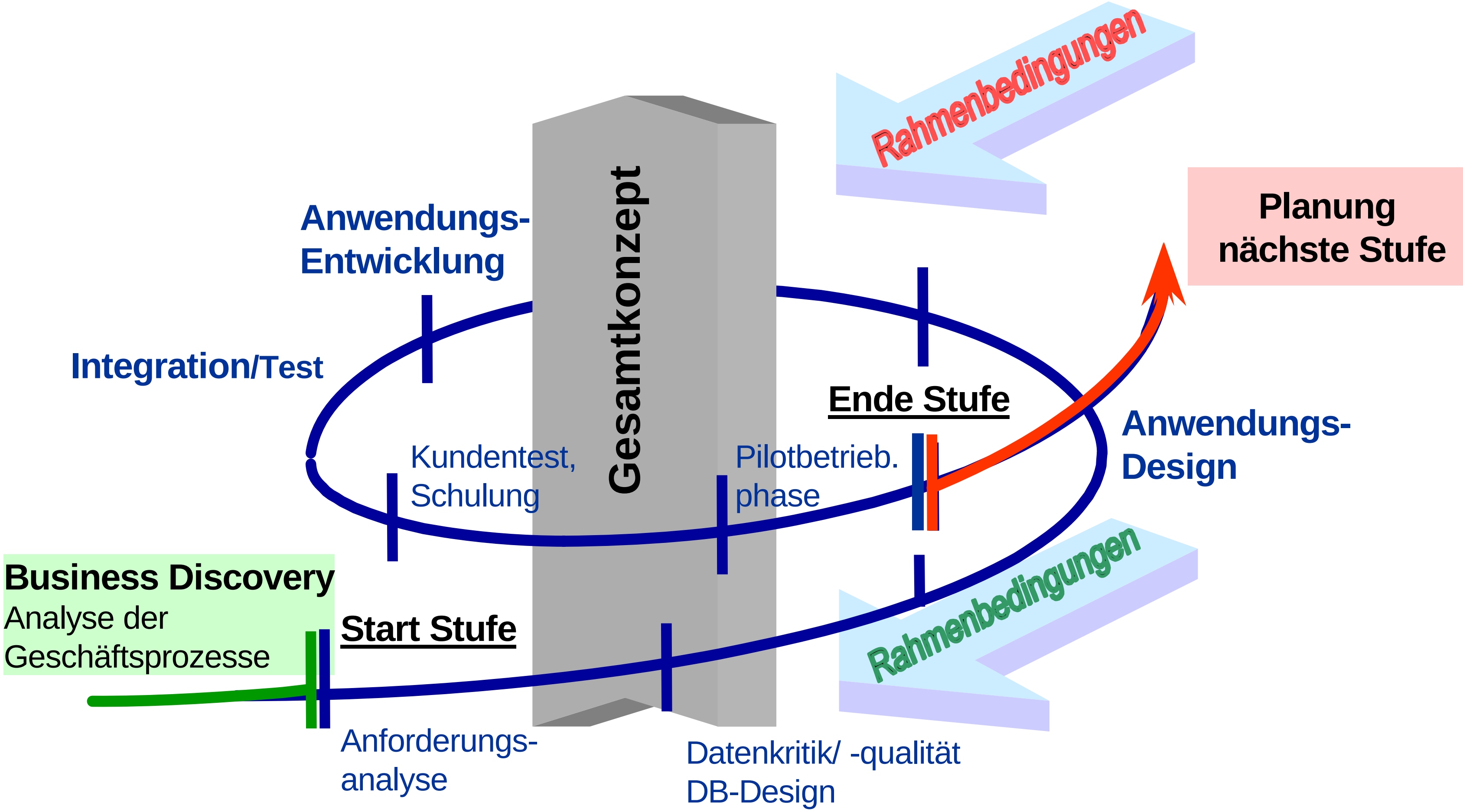
Einsatz von PowerCenter



Projektchronologie

1999	Jahr-2000-Problem: droht der Daten-GAU?
April 2000:	Expertise zur historisch gewachsenen Datenbankstruktur → DWH Architektur
2000:	Anforderungsanalyse/Konzept
März 2001:	Pflichtenheft Pilot WTO Ausschreibung
Dez 2001:	Audit Proof of Concept (Prototyp = Release 1)
Dez 2002:	Beginn produktiver Einsatz
2007:	alle „legacy systeme" ausser Betrieb
2009:	Ausbau für Gitterpunktsdaten/GIS

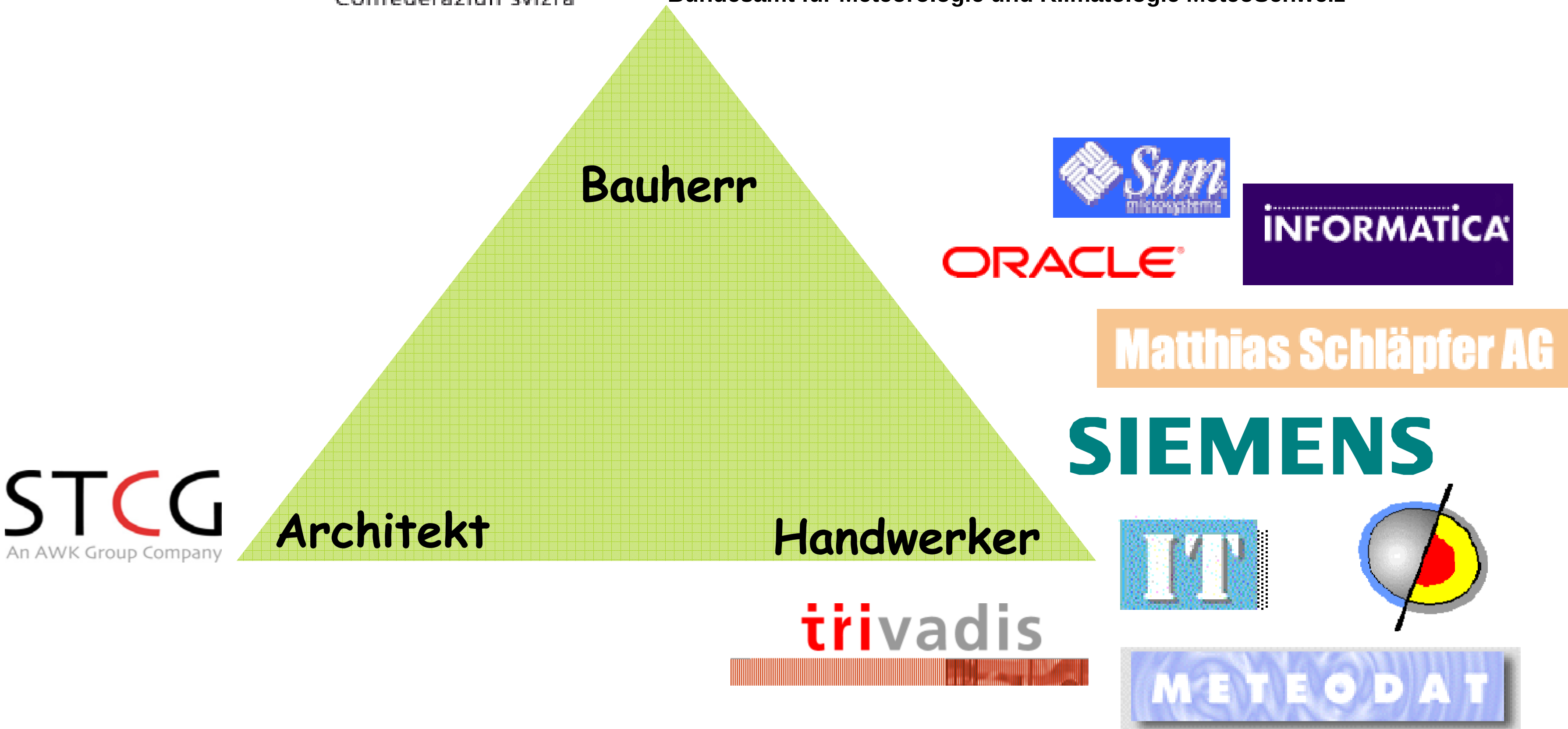
Spiralförmiges Vorgehen



Erfahrungen (1)

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Bundesamt für Meteorologie und Klimatologie MeteoSchweiz



Erfahrungen (2)

Tu es	**Lass es**
Auf die Daten horchen; breit abgestützte Reviews durchführen	Zu sehr auf einzelne Datenanwender horchen
Spiralförmig vorgehen & rasch Betriebserfahrung sammeln; Infrastruktur mitwachsen lassen	Alles auf einmal machen (wollen), Wasserfallvorgehen
Die goldenen Regeln des DB Designs befolgen	Rasch auf Technologie festlegen
Geschäftsabläufe verstehen	Von Beratern Lösungen aufschwatzen lassen
Ab Beginn 1 attraktive Applikation zur Datenanalyse bereitstellen	Segregierte Architekturen