# Data Assimilation from Research to Operational Application

Roland Potthast

**COSMO General Meeting**

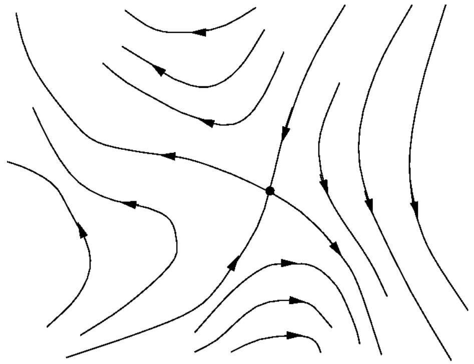
Moskow, Sept 6-9, 2010

Germany DWD Deutscher Wetterdienst, Switzerland MCH MeteoSchweiz, Italy USAM Ufficio Generale Spazio Aereo e Meteorologia, Greece HNMS Hellenic National Meteorological Service, Poland IMGW Institute of Meteorology and Water Management, Romania NMA National Meteorological Administration Russia RHM Federal Service for Hydrometeorology and Environmental Monitoring

# Contents/Goals

1. Reflect the *general framework* of Data Assimilation (DA) for Numerical Weather Prediction (NWP)
2. Sketch upcoming *Challenges* for DA in the COSMO Framework
3. Provide a survey about the *scientific communities* involved in the development of DA
4. Describe a *Data Assimilation Development Architecture* (DADA)

# 1.1 Data Assimilation Tasks and General View I



**Simulation:**

- ➔ Dynamical System M: $(t,x_0) \to x(t)$ in a state space X
- ➔ Initial state $x_0 = x(0)$

Here: M = **NWP** System Simulation Components, COSMO-EU/DE … etc

**Assimilation:**

- ➔ Measurements y in a measurement space Y
- ➔ Measurement Operator H: $X \to Y$ (state to measurements)

**Determine $x_0$ given measurements y in Y at particular times t!**

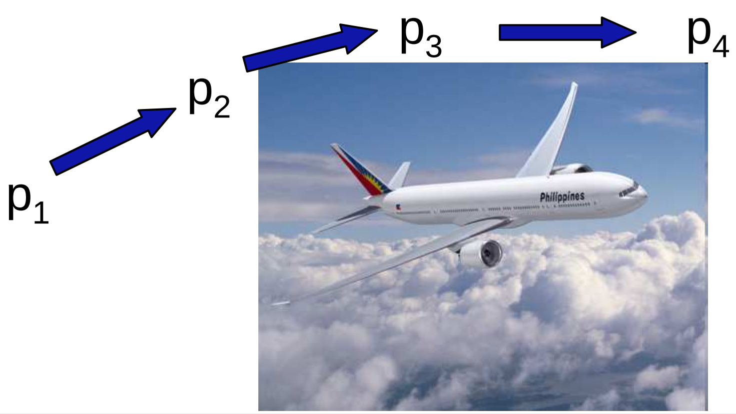
# 1.1 Data Assimilation Tasks and General View II

- ➔ Measurements    $y$    in a measurement space $Y$
- ➔ Measurement Operator    $H: X \rightarrow Y$    (state to measurements)

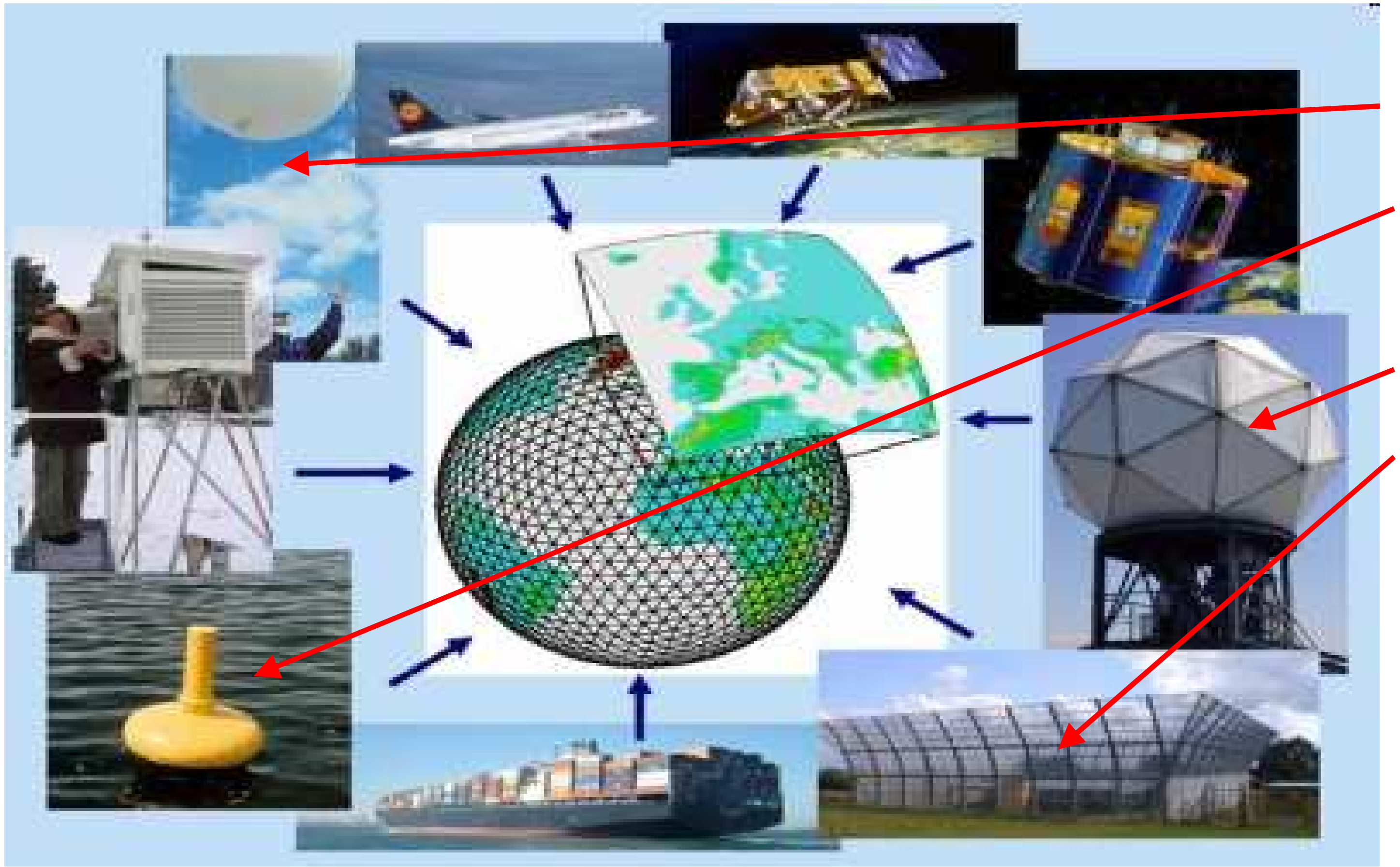
Usually    $y = (y_1, y_2, y_3, y_4, \dots, y_n)^T$    with $y_j$ in $Y_j$    for $j=1,\dots,n$

- ➔ Different **types** of Measurements!
- ➔ Measurements are **dynamic**, changing *location*, changing *number*, …

$H = (H_1, H_2, H_3, H_4, \dots H_n)^T$



# 1.2 Data Assimilation Atmospheric Data I



Synop, TEMP,

Radiosondes

Buoys

Airplanes

Radar

Wind Profiler

Scatterometer

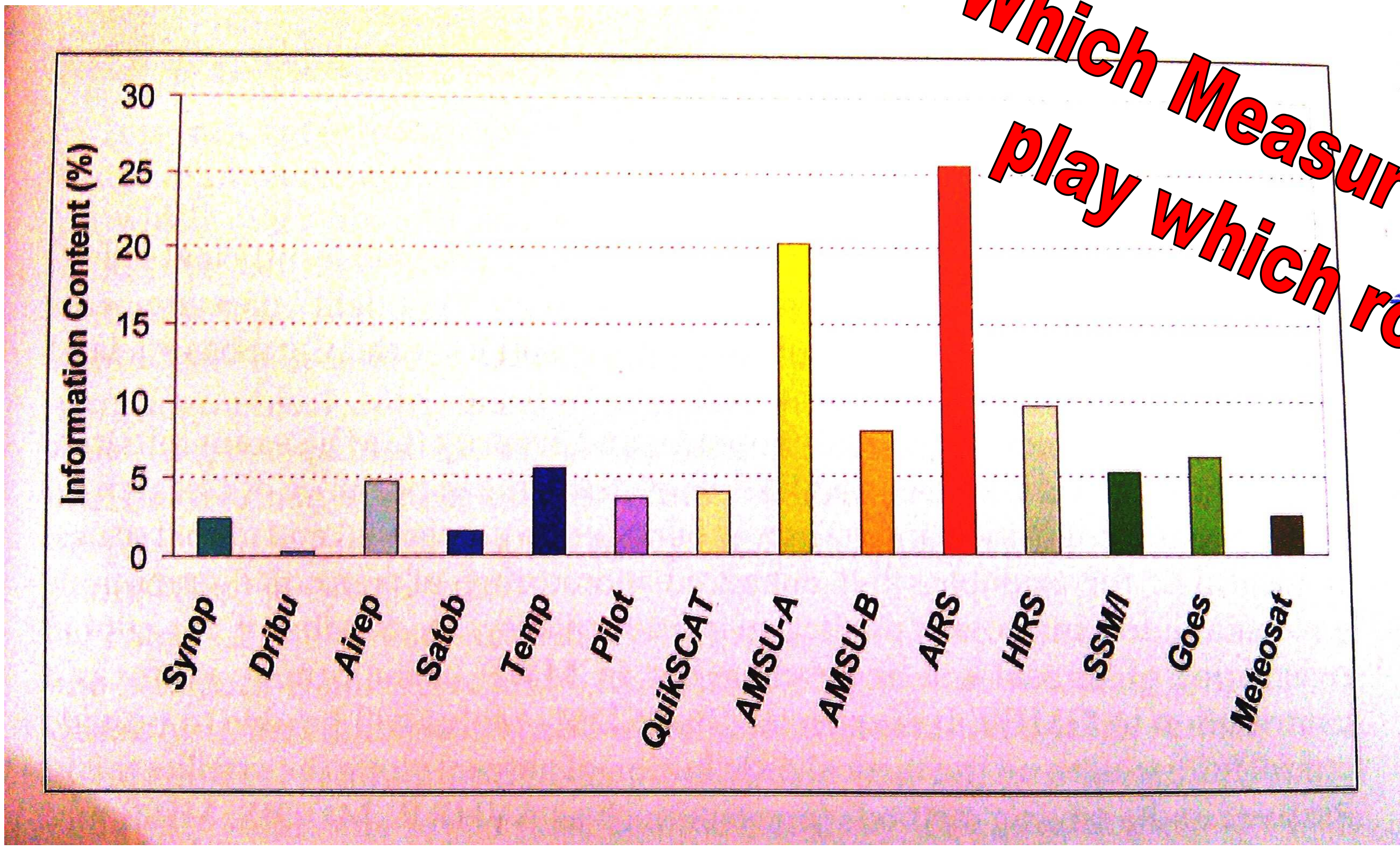
Radiances

GPS/GNSS

Lightning

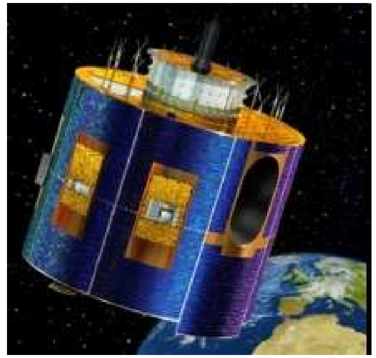
Gravity

# 1.2 Data Assimilation Atmospheric Data II



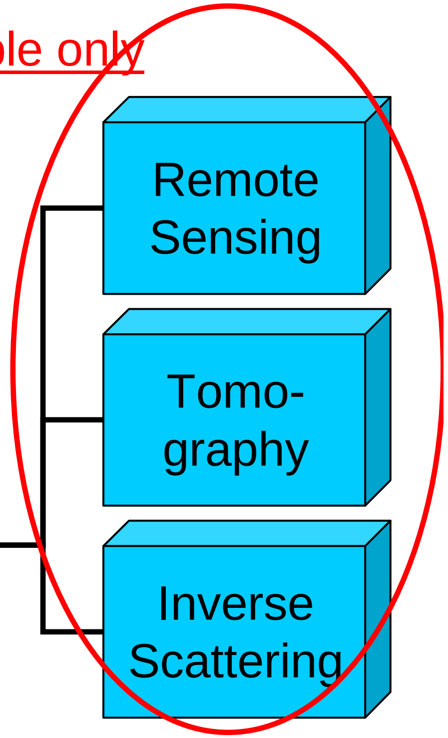

Which Measurements play which role?

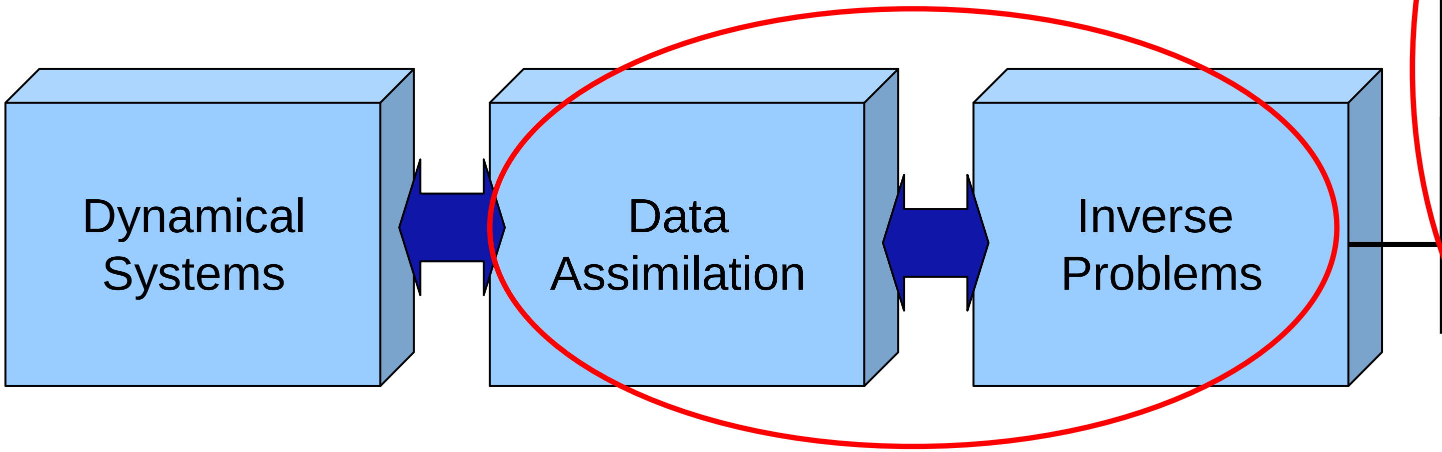
# 1.3 Properties of Observation Operators



For remote sensing data the observation operators $H_j$ :

- ➔ do not measure state variables directly
- ➔ do not measure a quantity which depends on **one** state variable only
- ➔ are not necessarily local, some measure integrated values
- ➔ might be nonlinear or even highly nonlinear



# 2.1 Approaches to Data Assimilation

- ➔ Variational Approaches / Optimization
- ➔ Stochastical Approaches / Ensemble Methods /Particle Filters
- ➔ Hybrid Variational/Ensemble Schemes
- ➔ Combination of Tomographic Methods with one of the approaches above
- ➔ Combination of Source Reconstruction with the above methods

# 2.2 Variational Methods I: 3dVar + 4dVar

Basic Idea: Solve $H x_k = y_k$

based on knowledge about the background $x^{(b)}_k$ at some point in time $t_k$

Option 1) Minimization Problem:

minimize $J(x) = \| x - x^{(b)} \|^2_B + \| Hx - y \|^2_R$ (3dVar)

which is carried out for each time step with x=x(t), y=y(t)

Option 2)

minimize $J(x) = \| x - x^{(b)} \|^2_B + \sum_k \| Hx_k - y_k \|^2_R$ (4dVar)

where $x_k = Mx_{k-1}$ denotes the system state at time $t_k$.

## 2.2 Variational Methods II: Kalman Filter (KF)

- ➔ minimize $J(x) = || x_k - x^{(b)}_k ||^2_{B(k)} + || Hx_k - y_k ||^2_R$
- ➔ Update $B_k=B(k)$ in every time step using the knowledge of H and M (KF)

Theorem:

For linear operators and linear systems the Kalman Filter is equivalent to 4dVar!

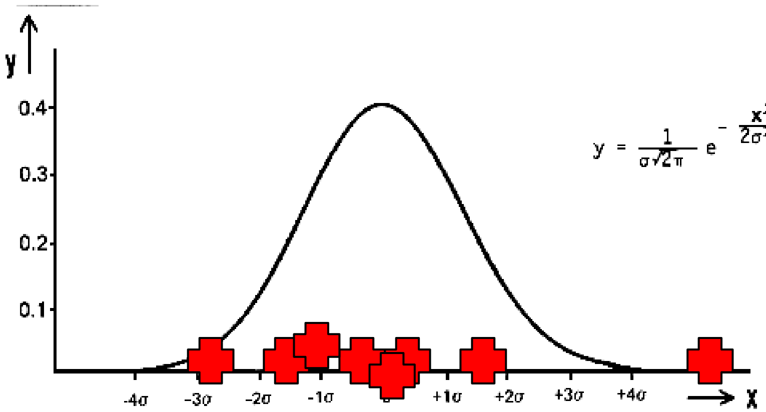
Advantages

- ➔ Iterative Minimization is efficient, can be carried out step by step when measurements arrive

Disadvantages

- ➔ The update of B is difficult to calculate for large-scale systems

# 2.3 Ensemble Methods and Particle Filters




$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$

Advantages:

➔ very flexible, non-Gaussian densities possible

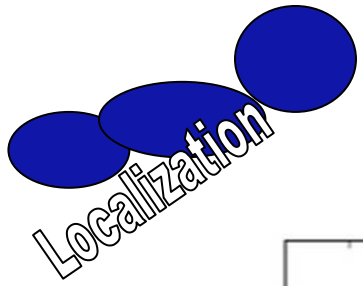
Disadvantages:

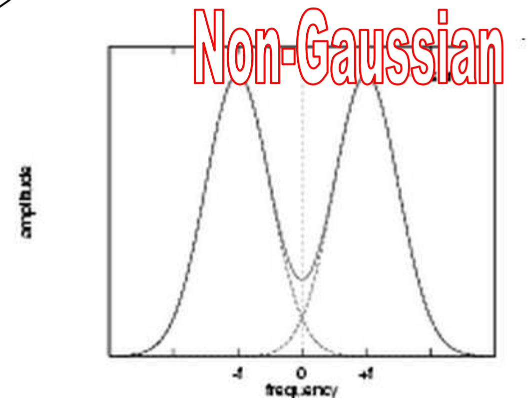
➔ low number of total degree's of freedom! (Needs Localization)

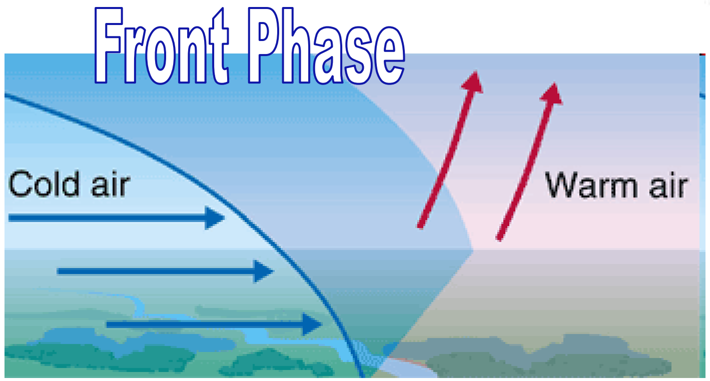
- ➔ Employ an ensemble of states
- ➔ Process the states through the forward model
- ➔ Compare the simulated data from the ensembles with the measurements
- ➔ Use the distance to the measurements to calculate probability weights for the ensemble members
- ➔ Estimate parameters from the posteriori probability distribution using the distribution of the ensemble members/particles

# 2.5 Upcoming Innovation on Filter Algorithms

- Improve Localization, Setup for Localization, Mathematical Analysis
- Treatment of strong non-Gaussian distributions
- Treatment/Reduction of Phase Errors in Assimilation
- Multi-Level approaches
- Adaptive Subspaces



# 2.6 Challenges for further Development

- ➔ Development of ensemble filters/hybrid methods to full maturity
- ➔ Use of a variety of further measurements and quantities for assimilation
- ➔ Time Scale to consider
  (Short-Range, Medium-Range, Long-Range, Real-Time/Very Short-Range)
  - ➔ Integration with Nowcasting Systems and Products
  - ➔ Reduction to Probability Distributions of dependent Quantities for fast further assimilation
- ➔ Local/Global Interaction (COSMO-ICON), High-Resolution
- ➔ System Range: Classical Atmospheric Systems /Atmospheric-Oceanographic Systems / Aerosol Simulation / Other Systems
  - ➔ Extension of purely atmospheric model including assimilation components

# 3.1 Relate Research and Operational Service

**Operation**

Deutscher Wetterdienst

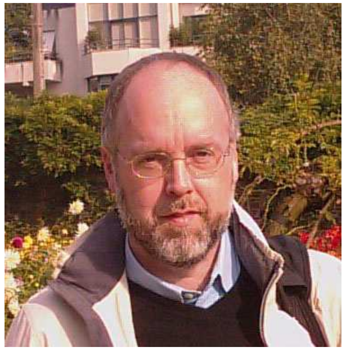
University of Reading

**Research**

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

**Potthast Personal Background**

- DWD (German Meteorological Service) Head Data Assimilation Unit
- Professor Appl. Mathematics, University of Reading, UK
- Apl. Professor Mathematics, University of Göttingen, Germany



# 3.2 Features of University Research

- Driven by curiosity
- Driven by funding needs, pressure to do what is funded
- Often unpredictable
- Often small scale, toy problems
- Small groups (one professor, some students)
- Rather free time disposability
- Open young people, no fixed concepts of life and work, openness for new ideas
- Ability to learn: students become experts in an area within 2-3 years!
- Pressure to earn esteem, publish or perish

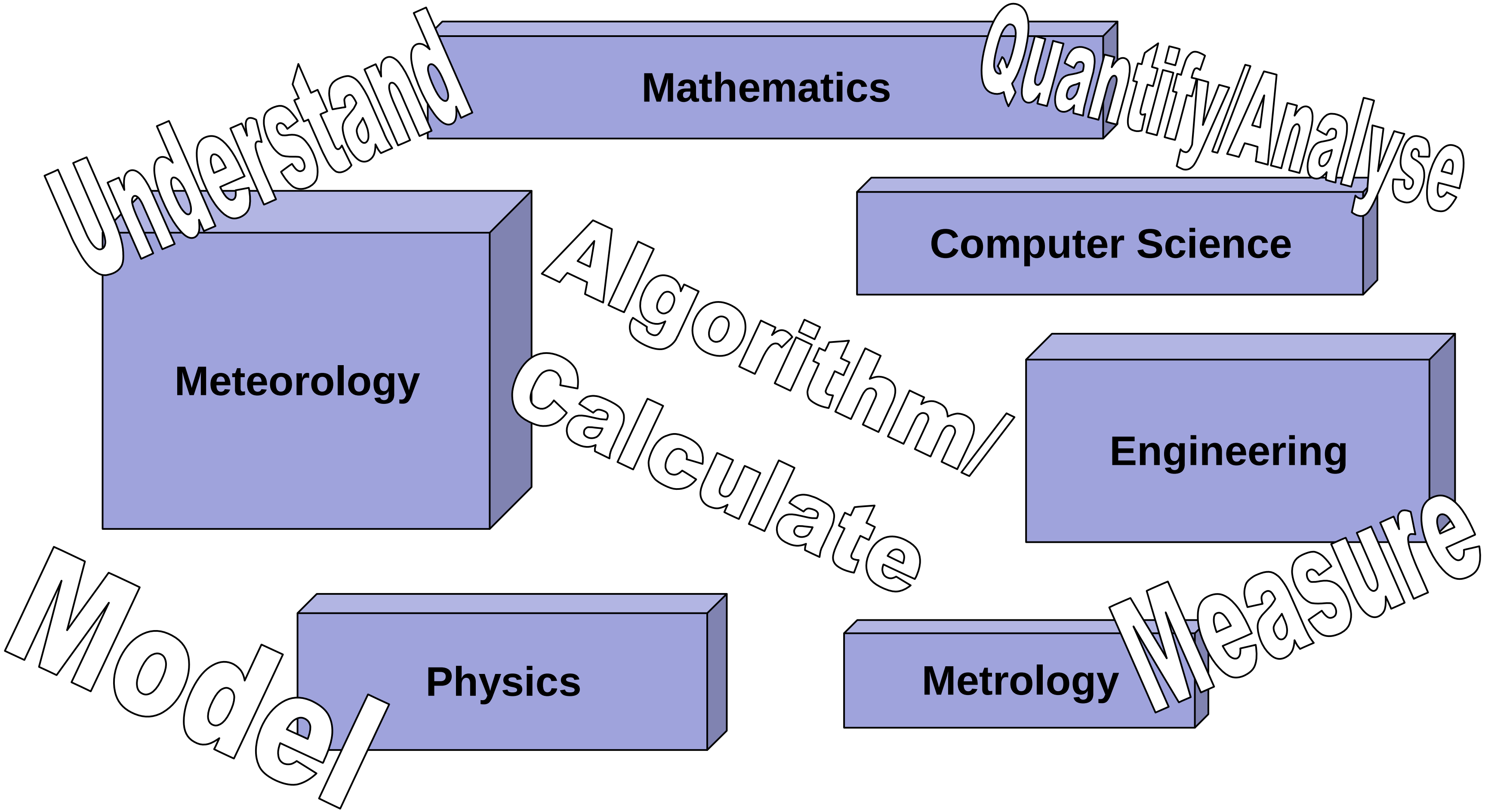
# 3.3 Features of Research in Companies

- Often product driven
- Very applied, less fundamental research
- Short timescales
- Need of return on investment
- Private ownership wants profit and share holder value
- Firm often has limited social responsibility
- Often quite fixed institutional environment
- Mature people, high responsibility
- Not much free time
- Strong pressure to get projects through

# 3.4 Features of Work in Operational Centers

- ➔ Driven by public or state interest/mission
- ➔ Usually rather applied research
- ➔ Large-Scale
- ➔ National responsibility as key point
- ➔ Basic funding on a steady basis
- ➔ Strong focus on operational needs
- ➔ Some time for research, but limited
- ➔ Long-term goals and mission
- ➔ Steady development with mature scientists
- ➔ Institutions provide often quite inflexible support

## 3.5 Sciences needed for Atmospheric Data Assimilation



# 4 Networking

**We need intensive and efficient networking for our work**

- ➔ **Within our consortium (*operational centers*)**
- ➔ **With public research institutions and universities**
- ➔ **With scientists from all related fields**

- **Using the knowledge about the particular environment our partners work in**
- **Knowing the way they operate**

## 4.1 German Data Assimilation Network (being developed)

- German Meteorological Service (Deutscher Wetterdienst DWD)
- Max-Planck Institute for Meteorology (MPI), Hamburg
- University of Bremen,
- Alfred Wegner Institute for Polar and Oceanographic Research (AWI), Bremen
- University of Potsdam
- German Research Center for Geoscience (GFZ), Potsdam
- Berlin Free University, Technical University, Humbold University
- Ludwig Maximilian University Munich
- German Aerospace Center (DLR), Oberpfaffenhofen-Wessling (Munich)
- University of Bonn
- University of Karlsruhe
- University of Hohenheim


**Heinz ERTEL Center**

For Atmospheric Research, Germany

# 4.2 Data Assimilation Development Architecture (DATA) I

1. Routine System at DWD (inhouse, operational assimilation and forecast)
2. Parallel Routine System at DWD (inhouse, preparation of systems for operation)
3. Experimental System (Numex) at DWD (inhouse, testing of system components which are under development)
4. Data Assimilation Testbed Sites (internal + external, Cosmo-DE, Cosmo-EU, ICON DA sites which are used for studies, research and development)
5. Medium Scale and Idealized Systems Repository (internal + external, medium scale and approximate setups which can be used for development of observation operators and methods)
6. Toy Systems Repository (internal + external, toy size systems for teaching and learning, development and study)

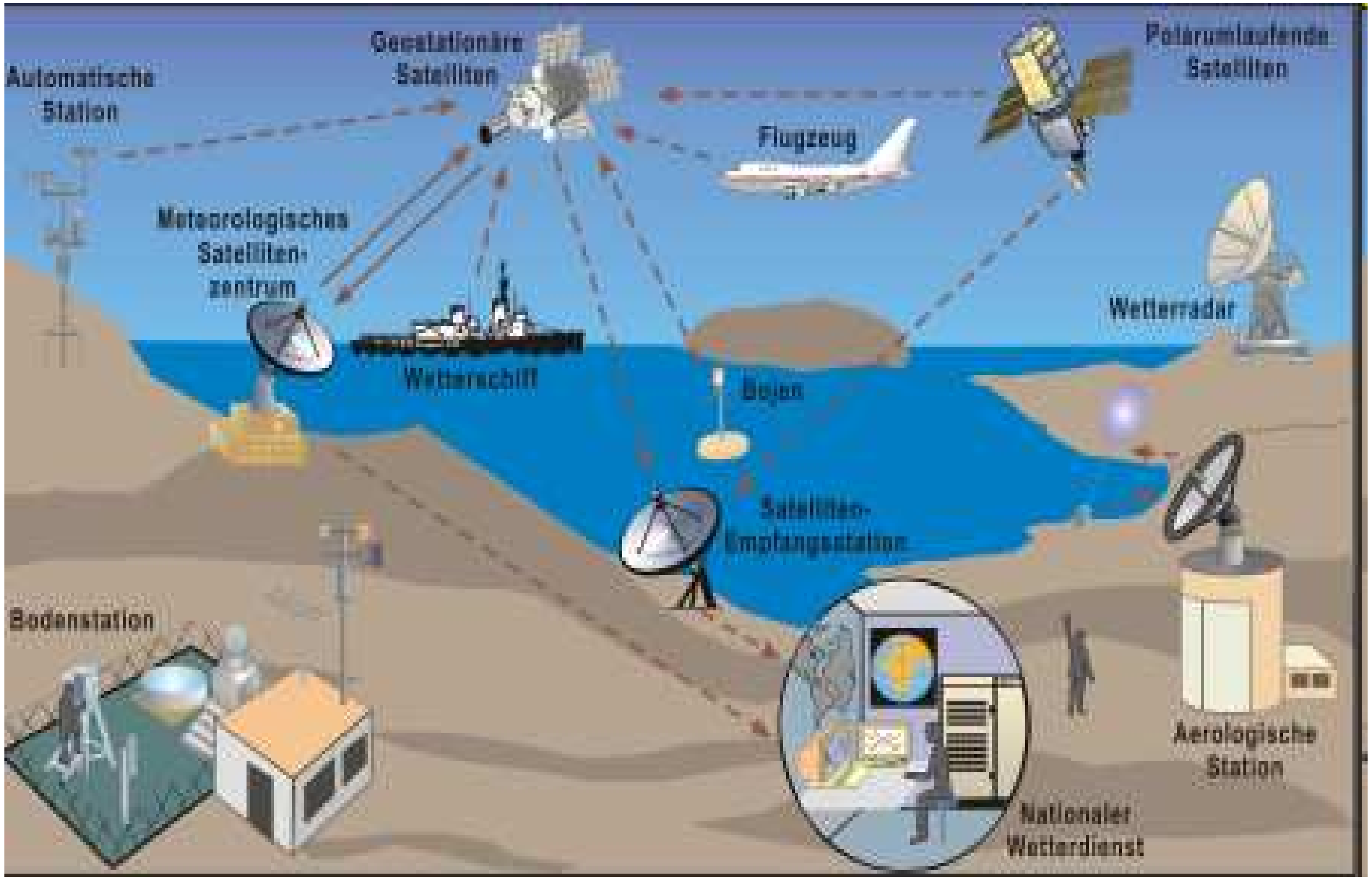
## 4.2 Data Assimilation Development Architecture (DATA) II

- ➔ We need a flow of competences and knowledge in both directions,
  - ➔ *Between Operational Centers (within COSMO, internationally)*
  - ➔ *Operational Centers to the Universities*
  - ➔ *Universities to the Operational Centers*
- ➔ Provide both the simulation components as well as the data and data assimilation components (algorithms, measurement operators etc)
- ➔ Allow external development based on our system
- ➔ Use well defined interfaces to enable the integration of external developments with minimal expense in terms of workload and time
- ➔ Provide huge benefit both for universities using professionally developed systems as for the operational centers extending their range

# Many Thanks

DWD Internal Courtyard, Offenbach, Germany

## 1.2 Data Assimilation Atmospheric Data II




Synop, TEMP,

Buys

Radiosondes

Airplanes

Radar

Wind Profiler

Scatterometer

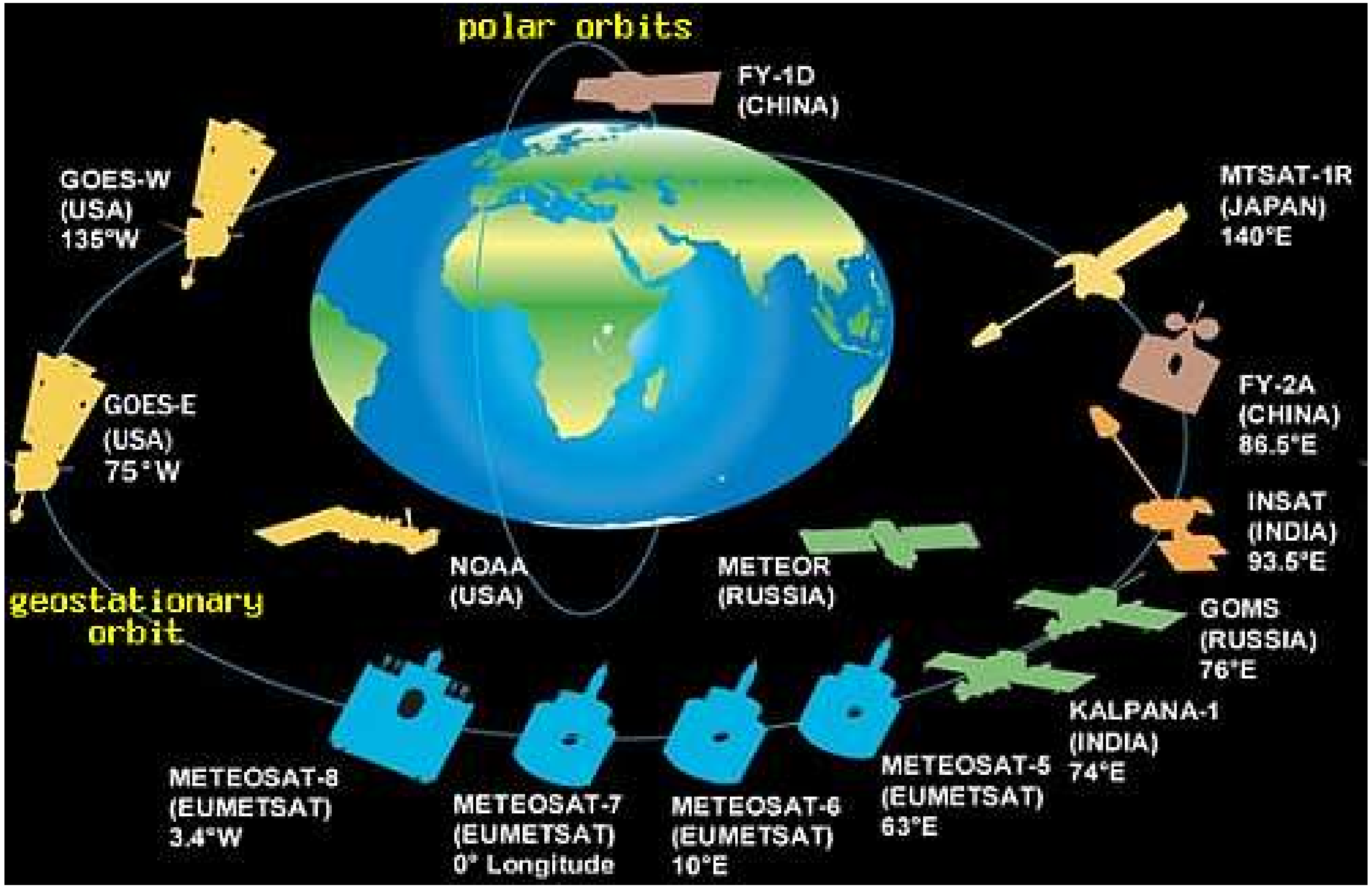
Radiances

GPS/GNSS

Lightning

Gravity

# 1.2 Data Assimilation Atmospheric Data III




Synop, TEMP,

Buys

Radiosondes

Airplanes

Radar

Wind Profiler

Scatterometer

Radiances

GPS/GNSS

Lightning

Gravity