On the application, use and access to ground based aerosol observations through EBAS

Cathrine Lund Myhre

Kjetil Tørseth, Markus Fiebig, Wenche Aas, Ann Mari Fjaeraa, Paul Eckhardt

NILU - Norsk institutt for luftforskning

Norwegian Institute for Air Research

Outline

Motivation

- Some recent examples on the use of data in EBAS
- What is there?

What is EBAS?

- Background and development
- Type of data
- Harmonisation and quality assurance
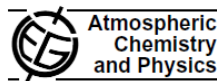- File format

How to find and inspect data

- EBAS web interface
- Formats and information about the files

Access to data

- Data policy
- Single user
- User community

NILU

Some examples of recent papers using data from EBAS

Atmospheric Chemistry and Physics

Introduction to the European Monitoring and Evaluation Programme (EMEP) and observed atmospheric composition change during 1972–2009

K. Tørseth, W. Aas, K. Breivik, A. M. Fjæraa, M. Fiebig, A. G. Hjellbrekke, C. Lund Myhre, S. Solberg, and K. E. Yttri

NILU – Norwegian Institute for Air Research, P.O. Box 100, 2027 Kjeller, Norway
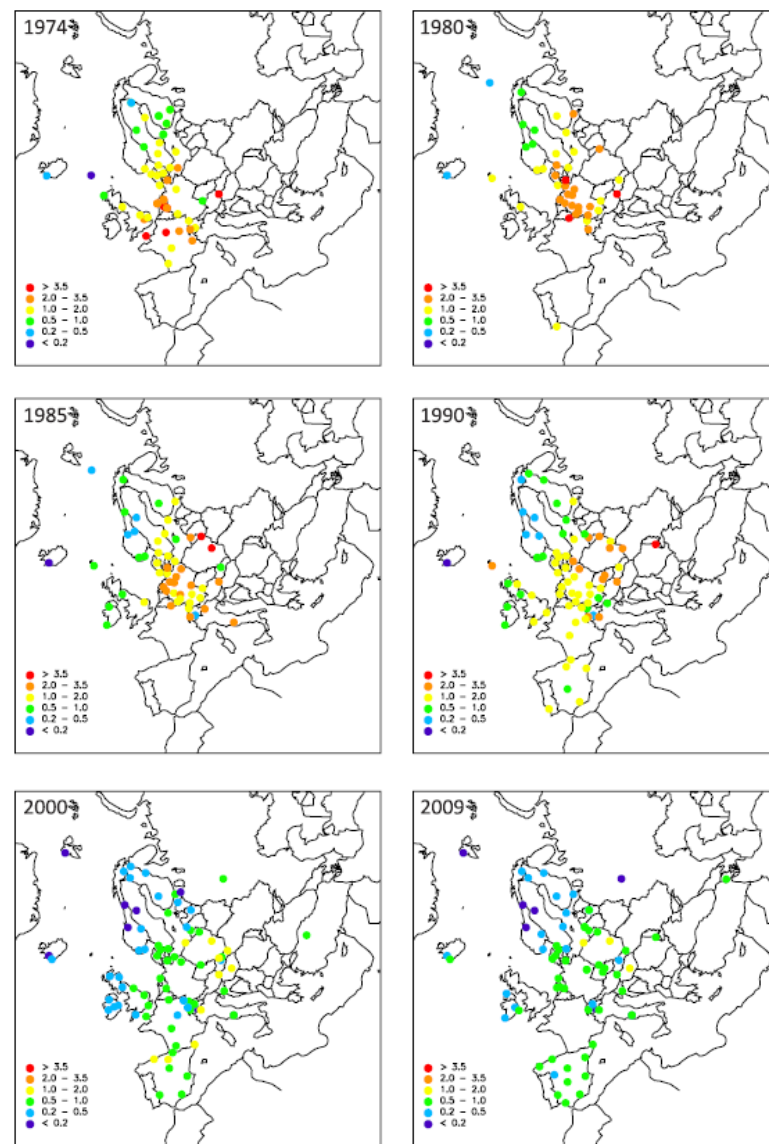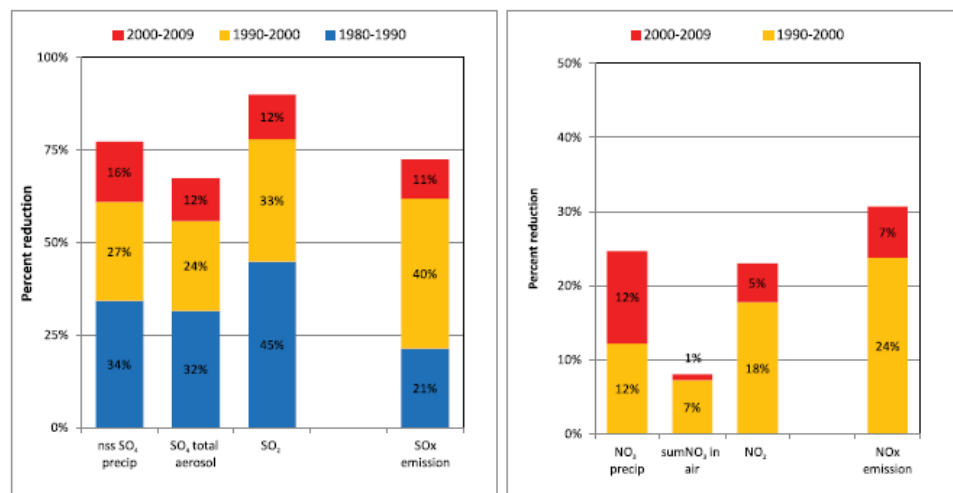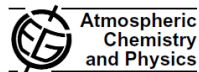
Correspondence to: K. Tørseth (kt@nilu.no)

Fig. 2. Average observed reduction in sulphur and nitrogen components compared to the en...

Fig. 3. Annual mean concentrations of SO_4^{2-} in aerosols from 1974 to 2009. Unit: $\mu g\,S\,m^{-3}$.

Some examples of recent papers using data from EBAS

Atmospheric
Chemistry
and Physics

Aerosol decadal trends – Part 1: In-situ optical measurements at GAW and IMPROVE stations

M. Collaud Coen[1], E. Andrews[2,3], A. Asmi[4], U. Baltensperger[5], N. Bukowiecki[5], D. Day[6], M. Fiebig[7], A. M. Fjaeraa[7], H. Flentje[8], A. Hyvärinen[11], A. Jefferson[2], S. G. Jennings[9], G. Kouvarakis[10], H. Lihavainen[11], C. Lund Myhre[7], W. C. Malm[6], N. Mihapopoulos[10], J. V. Molenar[12], C. O'Dowd[9], J. A. Ogren[3], B. A. Schichtel[13], P. Sheridan[3], A. Virkkula[4], E. Weingartner[5], R. Weller[14], and P. Laj[15]

[1]Federal Office of Meteorology and Climatology, MeteoSwiss, 1530 Payerne, Switzerland
[2]University of Colorado, CIRES, Boulder, Colorado, 80305, USA
[3]National Oceanic and Atmospheric Administration, Earth System Research Laboratory, Boulder, Colorado, 80305, USA
[4]Department of Physics, University of Helsinki, P.O. Box 64, Helsinki, Finland

Atmospheric
Chemistry
and Physics
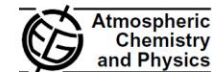
Aerosol decadal trends – Part 2: In-situ aerosol particle number concentrations at GAW and ACTRIS stations

A. Asmi[1,16], M. Collaud Coen[2], J. A. Ogren[3], E. Andrews[3,4], P. Sheridan[3], A. Jefferson[3,4], E. Weingartner[5], U. Baltensperger[5], N. Bukowiecki[5], H. Lihavainen[6], N. Kivekäs[6], E. Asmi[6], P. P. Aalto[1], M. Kulmala[1], A. Wiedensohler[7], W. Birmili[7], A. Hamed[7,8], C. O'Dowd[9], S. G. Jennings[9], R. Weller[10], H. Flentje[11], A. M. Fjaeraa[12], M. Fiebig[12], C. L. Myhre[12], A. G. Hallar[13], E. Swietlicki[14], A. Kristensson[14], and P. Laj[15]

[1]Department of Physics, University of Helsinki, P.O. Box 64, 00014, Finland
[2]Federal Office of Meteorology and Climatology, MeteoSwiss, 1530 Payerne, Switzerland
[3]NOAA Earth System Research Laboratory, Boulder, USA
[4]Cooperative Institute for Research in Environmental Science, University of Colorado, Boulder, USA

Trends in scattering coefficient, 2001-2010
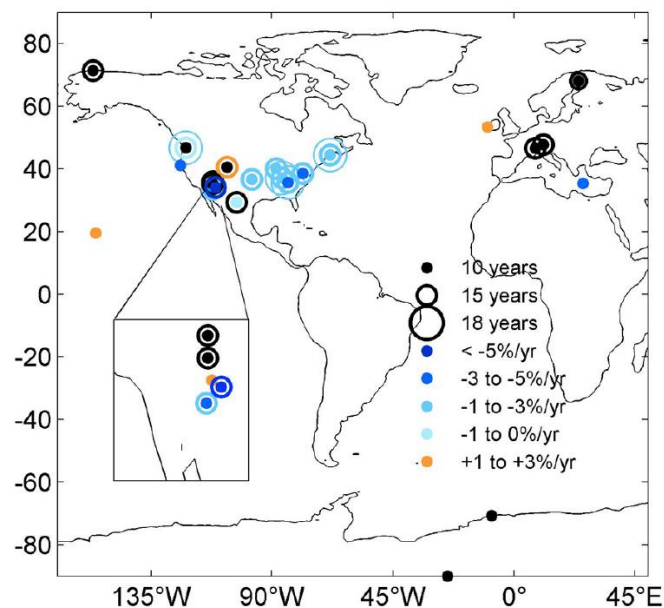
Trends in aerosol particle number concentrations 2001-2010
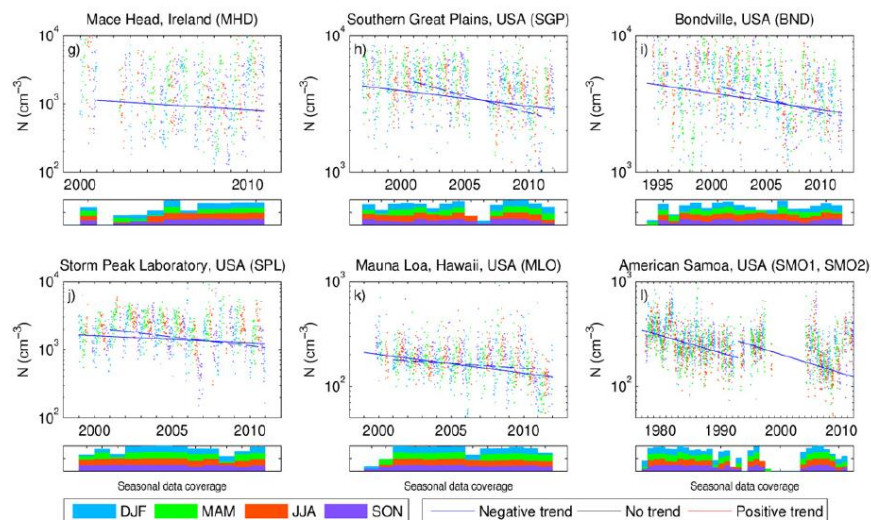
The data base infrastructure EBAS

EBAS is a database infrastructure operated by NILU – Norwegian Institute for Air Research.

It's main objective is to handle, store and disseminate atmospheric composition data generated by international and national frameworks like long-term monitoring programmes and research projects.

Status September 2013

Number data sets:

60 000 (1 data set is 1full time series)

Number of sites reported data:

1068

Number of countries:

72

Type of Variables:

713

Different instruments:

95

Regional Coverage by Parameter

Sulphate,

Size Distribution
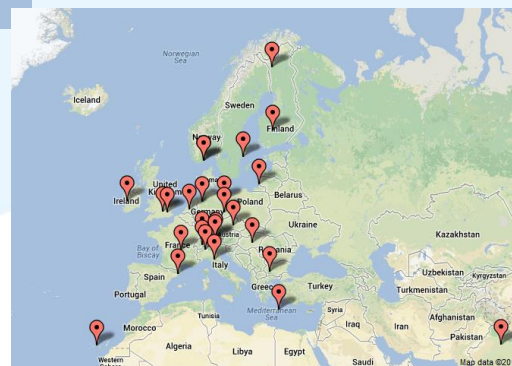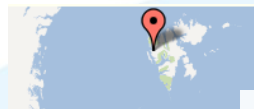
Scattering Coeff.

Absorption Coeff., incl. BC
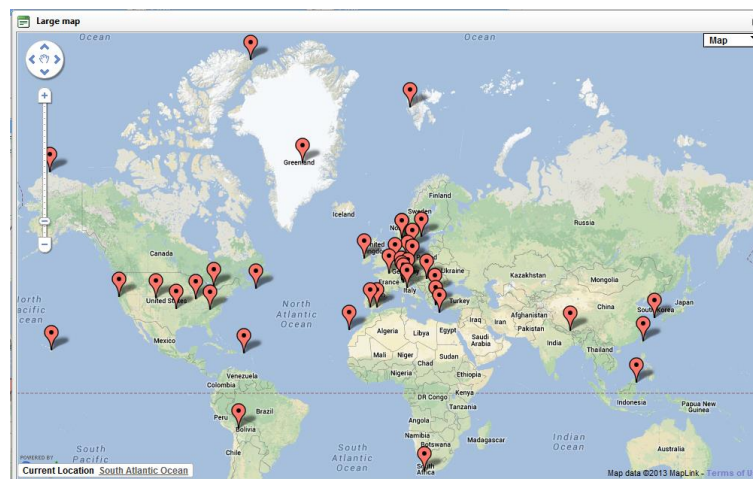
Can be long time series, or shorter periods

What is EBAS?
- **Background and development**
- **Harmonisation**
- **Quality assurance**
- **File format**
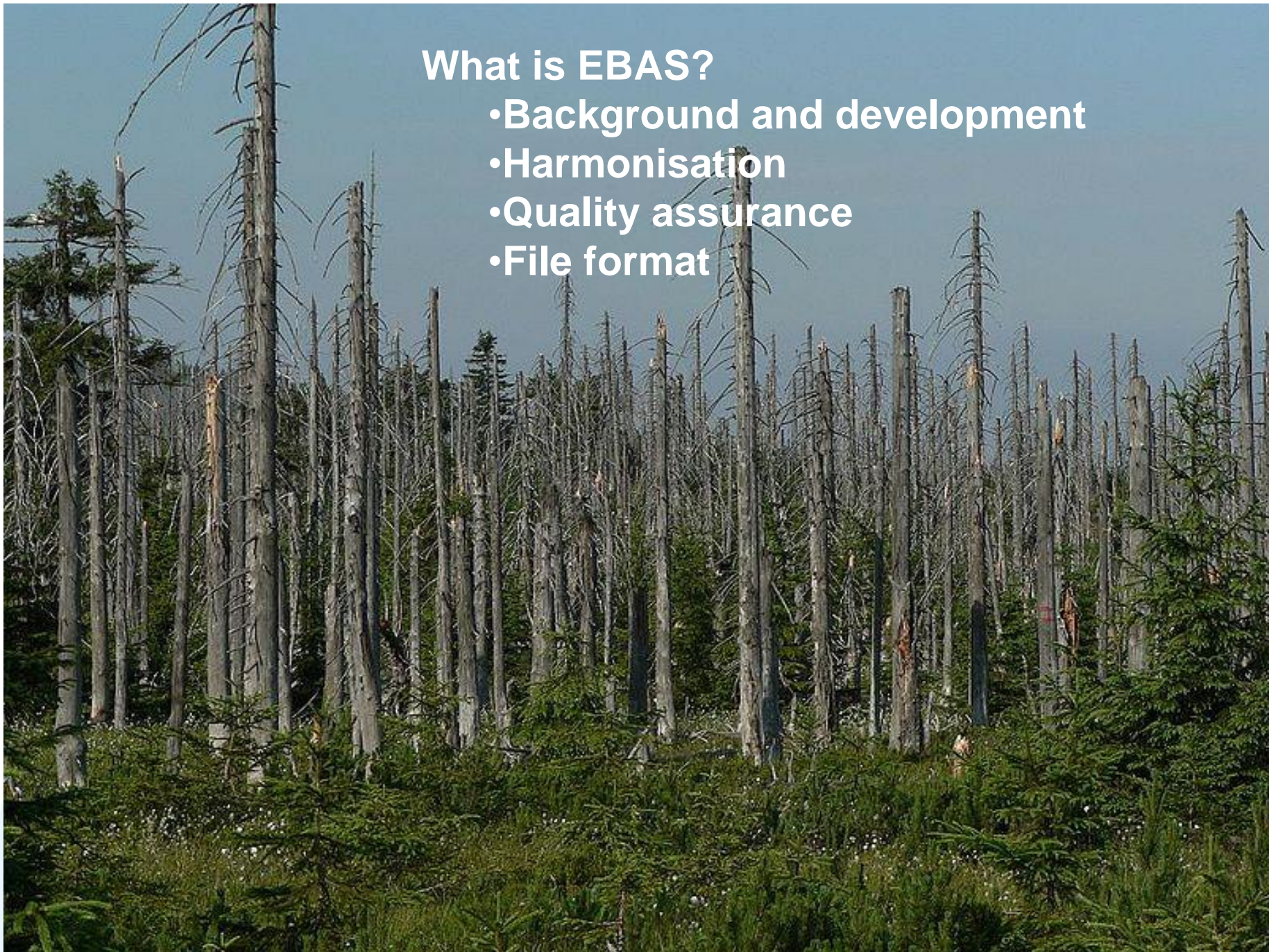
The EBAS History

Convention on Long-range Transboundary Air Pollution (1979, UNECE-CLRTAP, 51 Parties)

- **EMEP: European Monitoring and Evaluation Programme**
- Science based and policy driven programme to solve transboundary air pollution problems.
- National funding and obligations, long term measurements

The main objectives of EMEP are to:

- Provide observational and modelling data on pollutant concentrations, deposition, emissions and transboundary fluxes on the regional scale and identify their trends in time;
- Identify the sources and assess the effects of changes in emissions;
- Improve our understanding of chemical and physical processes relevant to the effects
- Harmonised and QA measurements from regional background

EBAS

- Name is derived from **E**MEP data**BAS**e.
- Today's relational database used since the mid-90s on varying hardware.

NILU: chemical coordination centre of EMEP

Met.no: Meteorological Synthesizing Centre - West (MSC-W)
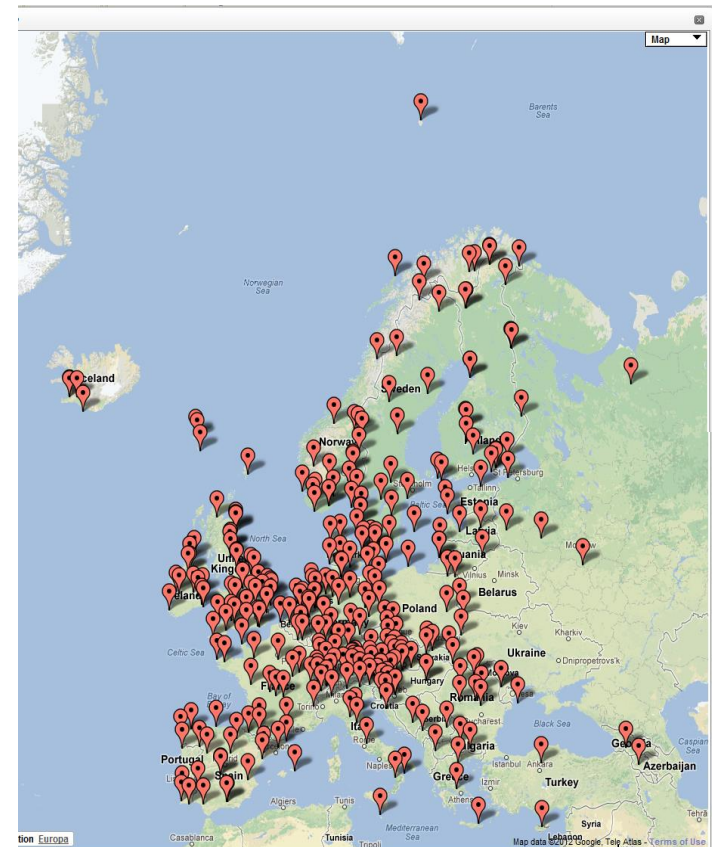
Development of the EMEP measurement programme

EU-projects:
CREATE -> 2003
EUSAAR -> 2011
ACTRIS -> 2014

Global link:
GAW and seek to implement harmonised strategy on observations

Torseth et al, Atm. Chem. Phys. (2012)
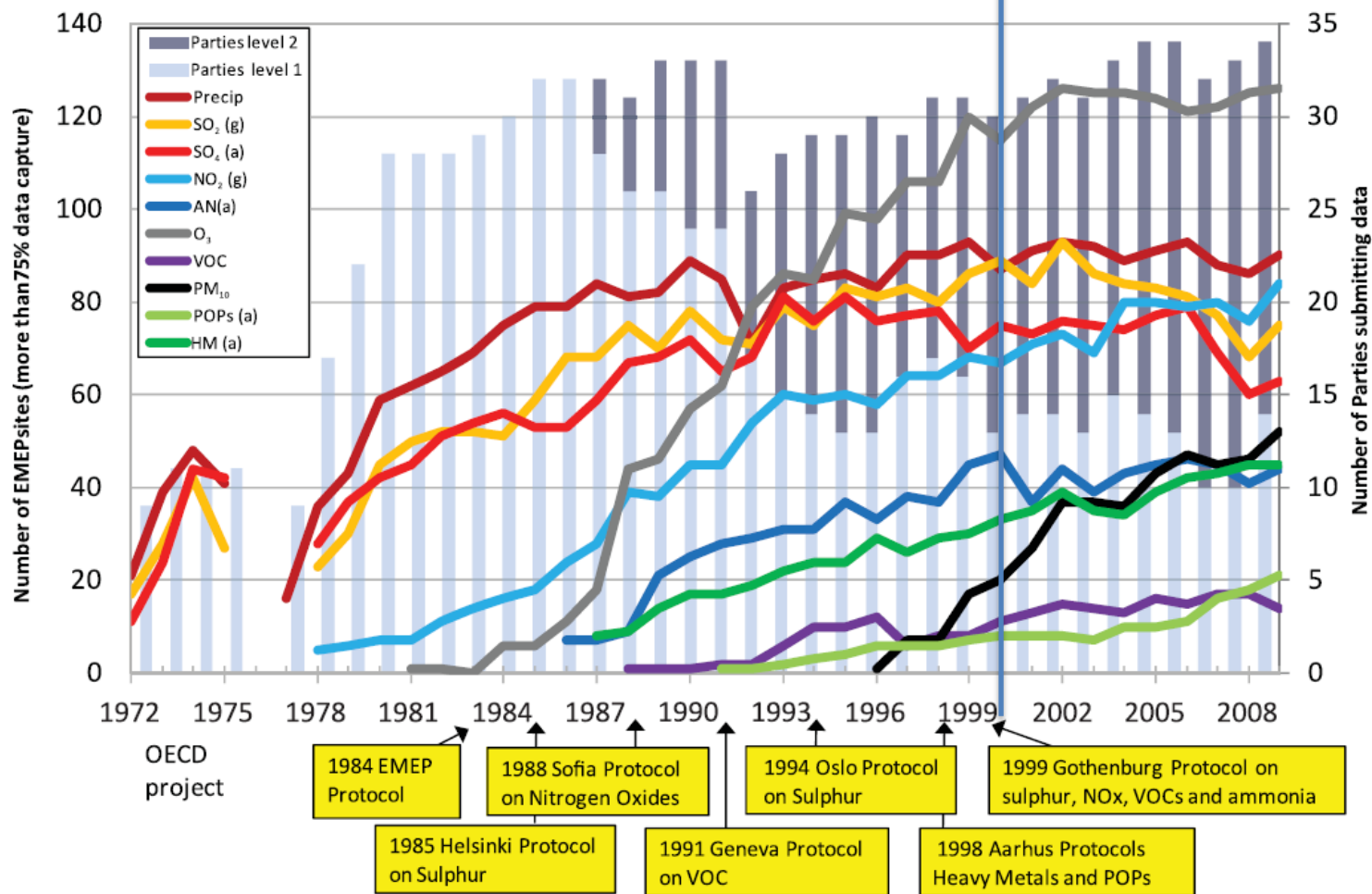
Contribution of other projects and networks to EBAS data infrastructure

Since the year 2000 EBAS also used widely used by other projects and frameworks.

Dataset distributed on framework, September 2013

Legend (left):
- ACTRIS
- GAW-WDCA_NRT
- IMPROVE
- NADP
- EUCAARI
- NOAA-ESRL_NRT
- EANET
- Others (CREATE, GUAN, InGOS etc)

Legend (right):
- EMEP
- GAW-WDCA
- EUSAAR
- CAMP
- WMO-PCSAG
- NILU
- sel_HTAP
- AMAP

Pie chart values: 32%, 13%, 9%, 5%, 5%, 4%, 4%, 4%, 3%, 3%, 2%, 2%, 2%, 2%, 2%, 8%
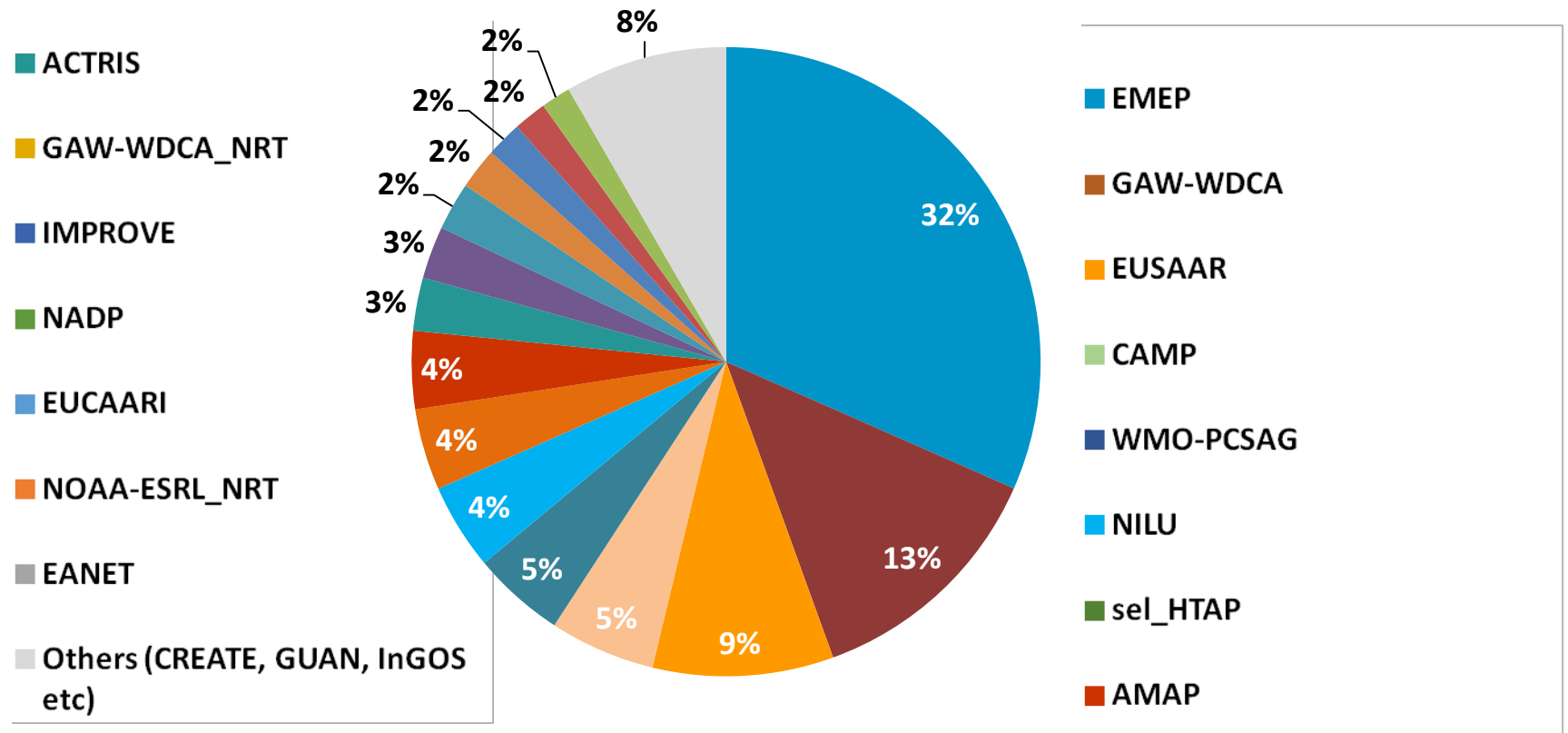
Harmonisation and QA

Important objective:

- Harmonised QA measurements across countries, regions, and even on global scale

Requires:

- refere   loped
- tory in
- arious   variables with
  ent th

-

-

Why Using NASA Ames 1001 Format for Reporting?

1. **Simplicity**

   - Pure ASCII text, human readable, readily opened or edited by simple means (any editor or spreadsheet application).

   - Explanation relatively short, yet contains necessary metadata.

2. **Reduce Format Confusion**

   - Don't increase number of existing formats (NASA Ames, Narsto, NetCDF, HDF, …) unnecessarily as long as metadata can be transported in old format, even though some features are old-fashioned.

   - Existing libraries can be used to handle files.

3. **Keep threshold low**

   - More modern, binary formats exist (NetCDF, HDF), but need special editors and steep learning curve to assemble.

   - NASA Ames can be assembled with simple tools rather quickly.

NILU

How to find and inspect data?

- **EBAS web interface**
- **Plot and download**
- **File structure and formats**
- **Flagging system**

Free tropospheric site Jungfraujoch, 3580 meter

The EBAS Web-Interface 1 / 3

http://ebas.nilu.no

Web-interface since 2009, linking also other tools.

EBAS web-interface functions:
- Search datasets by criteria: Framework, country, station, matrix, instrument type, component.
- Visualise distribution of stations on map.
- Manage access restricted data
- Plot, browse, compare datasets
- Download data

The EBAS Web-Interface 2 / 3

Search result page of EBAS web-interface:

- Lists datasets that meet search criteria set on home page.
- Datasets that are present, but access restricted, are displayed in grey.
- Time period for plotting or download to be selected on top (select appropriate radio button!).

The EBAS Web-Interface 3 / 3
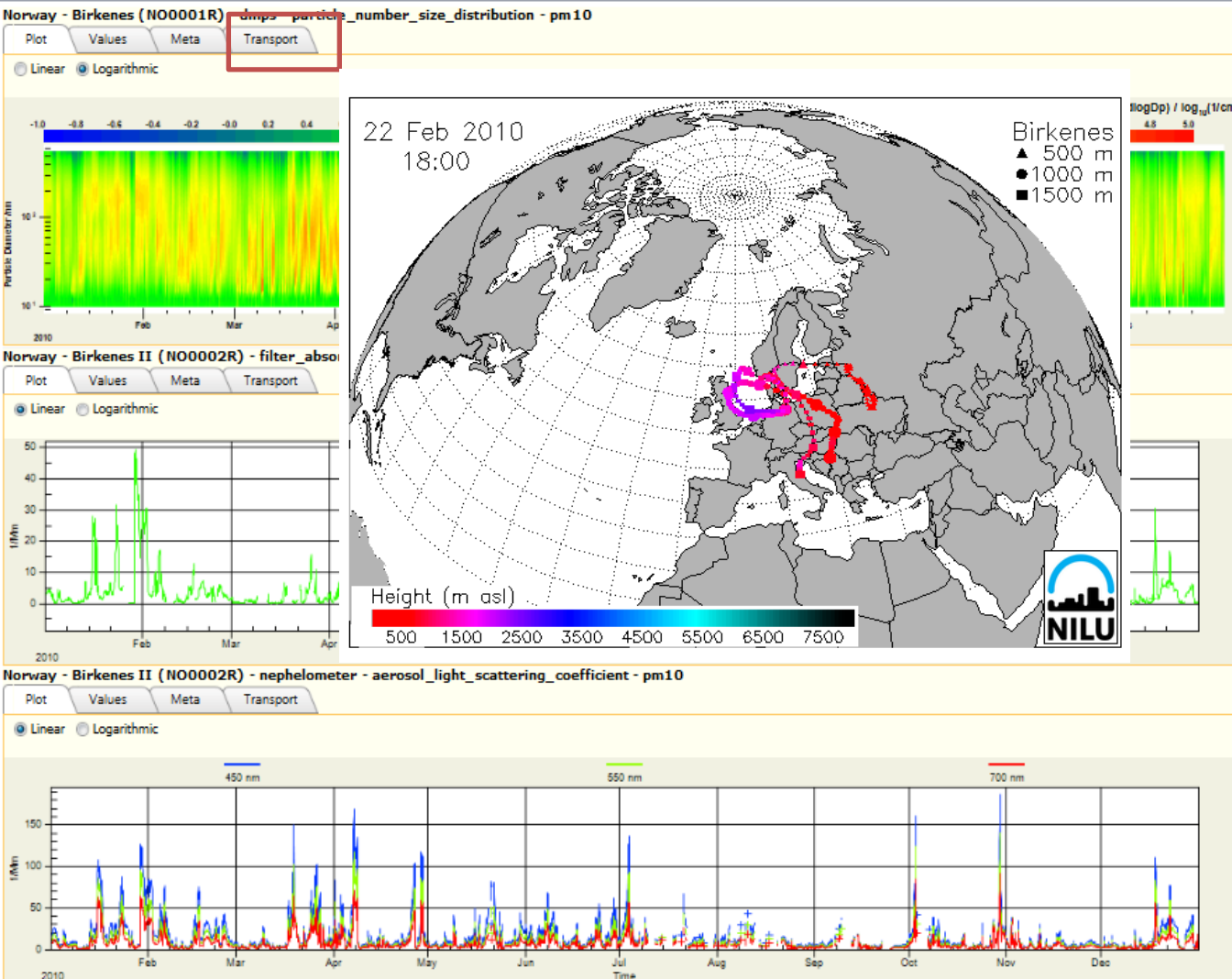
Plot page for selected datasets:

- Screen, evaluate, compare between instruments, compare between stations, …
- Download datasets (data is automatically grouped by instrument).
- FLEXTRA trajectories

What if you want to download and use the data?

You will get a multi column zip file containing:

- 1 file for each instrument selectied in the list per year of data
- E.g. Birkenes organic carbon: 10 years of data -> 10 files
- Web interface not developed for download of larger data volumes

Comprehensive meta data describe the measurements to document the quality
How to interpret this?

EBAS file structure and formats

NASA Ames 1001;

- assume that you read the file header and each column is described here

- Description of all metadata included in the header at: http://ebas-submit.nilu.no/
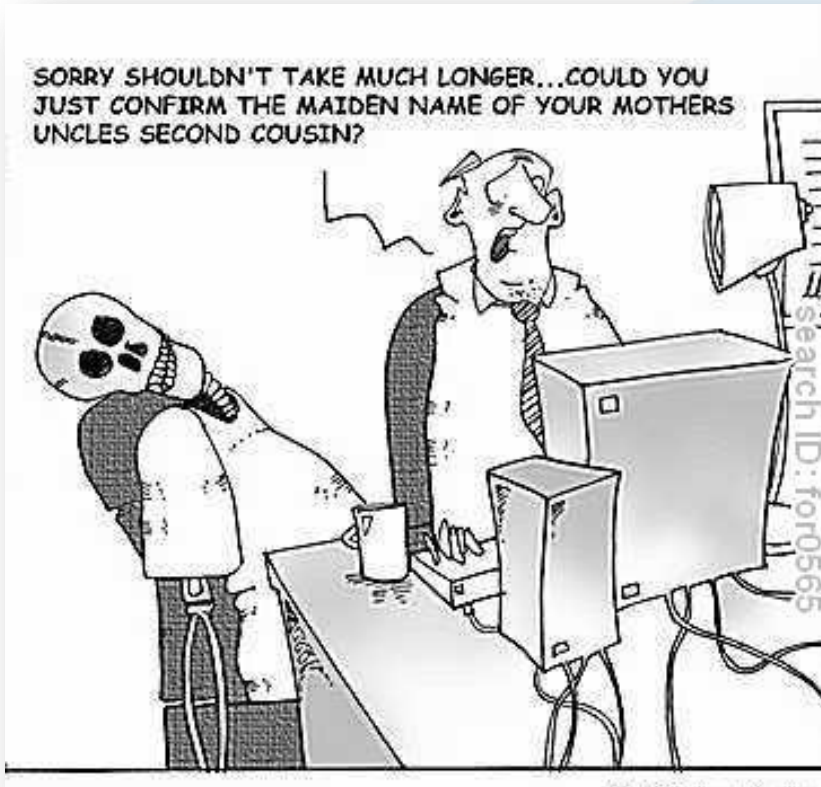
The flagging system

All flags are grouped in three categories:
V (valid measurement), I (invalid measurement) or H (hidden and invalid measurements).

Convention on Long–Range Transboundary Air Pollution

emep

Co–operative programme for monitoring and evaluation of the long–range transmissions of air pollutants in Europe

List of flags used in the EMEP data base

All flags are grouped in three categories: V (valid measurement), I (invalid measurement) or H (hidden and invalid measurements).

Flag	V/I/H	Description
Group 9: Missing flags		
999	I	Missing measurement, unspecified reason
990	I	Precipitation not measured due to snow-fall. Needed for historic data, should not be needed for new data
980	I	Missing due to calibration or zero/span check
900	H	Hidden and invalidated by data originator
Group 8: Flags for undefined data elements		
899	I	Measurement undefined, unspecified reason
890	I	Concentration in precipitation undefined, no precipitation
Group 7: Flags used when the value is unknown		
799	I	Measurement missing (unspecified reason), data element contains estimated value
798	V	Measurement missing (unspecified reason), data element contains estimated value. Considered valid.
797	V	Data element taken from co-located instrument
784	I	Low precipitation, concentration estimated
783	I	Low precipitation, concentration unknown
782	V	Low precipitation, concentration estimated
781	V	Value below detection limit, data element contains detection limit
780	V	Value below detection or quantification limit, data element contains estimated or measured value.
771	V	Value above range, data element contains upper range limit
770	V	Value above range, data element contains estimated value
750	I	H^+ not measured in alkaline sample
741	V	Non refractory AMS concentrations. Don't include compounds that volatalises above 600 deg C
740	V	Probably biased gas/particle ratio
701	I	Less accurate than usual, unspecified reason. (Used only with old data, for new data see groups 6 and 5)
Group 6: Mechanical or instrumental problem		
699	I	Mechanical problem, unspecified reason
681	I	Low data capture

NILU

Access to data

- Data policy
- Single user
- User community

Public, open, easy access is the goal...

Who owns the data in the data base?

Every dataset in EBAS created within ACTRIS, EMEP, GAW (and other programs) is owned by the partner/data providers who created this dataset.

Data policy: The general rule is that data policy is on program level

Regulate the us ▮▮▮▮▮▮▮▮▮▮ **the data providers and programs int**

- Facilitate the ▮▮▮▮▮▮ viders to ensure proper use of data w
- Lot off effort ▮▮▮▮▮ providers
- **Reduce misi** ▮▮▮▮▮ een data use, data analysis, dep
- Make the fur ▮▮▮▮▮ ortant for future funding situation

How is it handled presently?

Most data in EBAS are available free of charge for non-commercial and scientific use

- No login or registration is necessary

Data sets have project/network labels!

- Data policy documents etc
- Project association of a data set is included in the import and export file

Access to data and proper use

Main rule for the open data: "you as data user accept that an offer of co-authorship will be made through personal contact with the data providers or owners whenever substantial use is made of their data. In all cases, an acknowledgement must be made to the data providers or owners and to the project name"

Guidelines for offering co-authorship vs acknowledgement ("rule of thumb")

Offer co-authorship at an early stage;

- When you need site or instrument information to understand, analyze or interpret the data

- Type of data can be relevant: relatively new variables/parameters/methods normally requires offering of co- authorship

Acknowledge the program and EBAS;

- Model validation using large data volumes, aggregated to monthly mean values etc

- Data from regulatory monitoring under e.g. EMEP (e.g. Sulphate, nitrate etc)

NILU can always guide in what to do and provide advice

- Know the intention of the programs and the culture and "politics"

If there is any doubt:
contact the PI in the second line of the file, or NILU

Access to data for AEROCOM user

Few data sets;

- use the web interface!

Alternative 1: for single user of large data volumes

I.   Use the web interface to browse and decide about the data wanted (variable, components, time period, sites). NILU group can assist and provide advice for various purposes (start e.g. with Markus Fiebig or me)

II.   Formulate a request specifying the data sets you want to ebas@nilu.no

III.   NILU personel perform an extract from the data base of the exact specified files, and generate a zip-file (Paul Eckhardt, NILU).

IV.   An e-mail with the data policy corresponding to the project affiliation of files decribing the condition of use will be sent to the user.

V.   After acceptance of the condition of use, a zip-file will be made available at a ftp server at NILU containing all data in NASA Ames format, including the metadata for each file.

VI.   Report back to NILU about use of data in publications

NOTE: Re-distribution of data to third users are not in accordance with the data policies, direct contact with NILU is encouraged if this is an issue

Access to data for AEROCOM model community

Alternative 2: for AEROCOM community

E.g. AEROCOM experiment

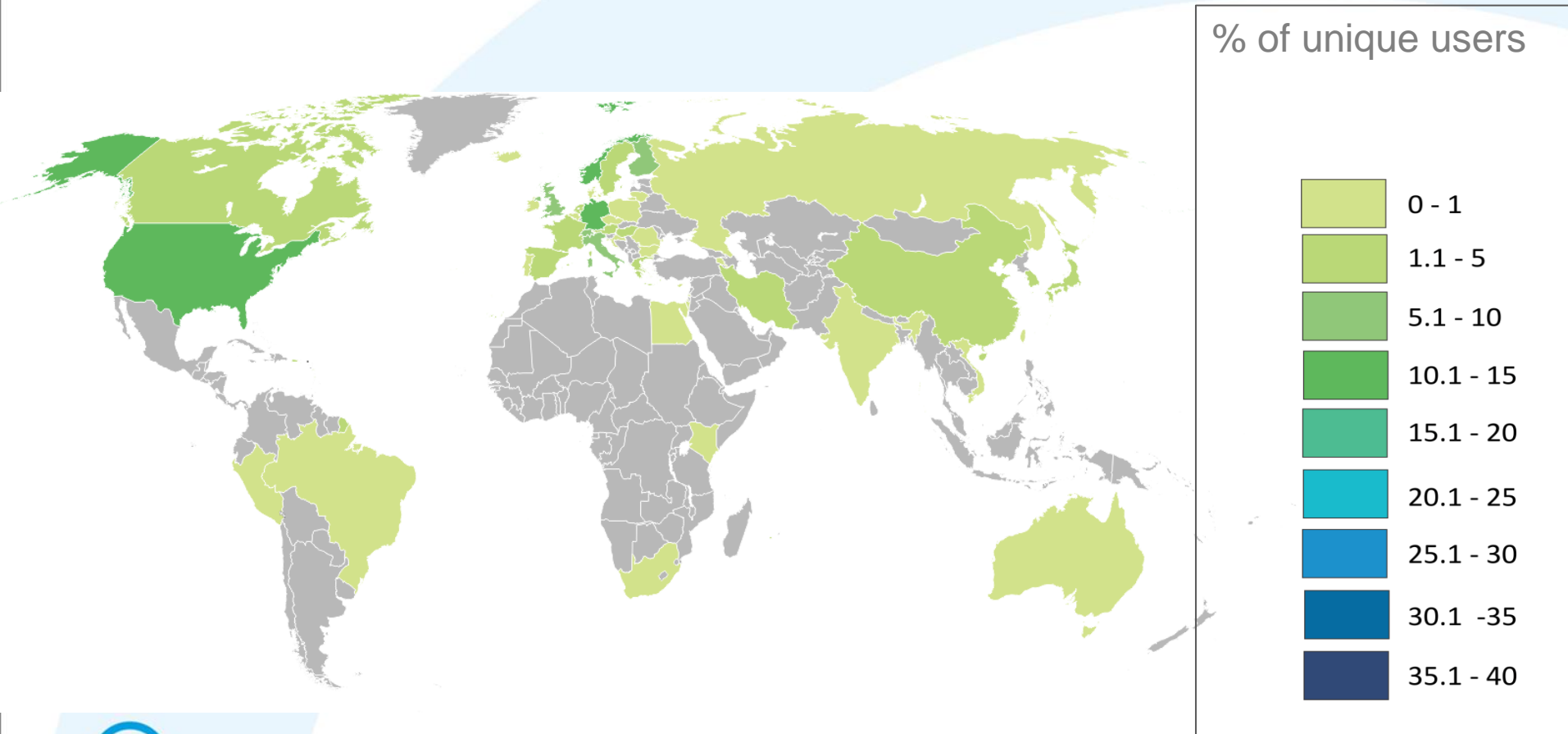MoU with AEROCOM community; for larger groups using exact same data sets

This implies that the data sets can be transferred to AEROCOM INTERFACE for further analysis

MoU containing description of e.g.;

- The experiment

- Specification of the data sets

- Transfer of data to AEROCOM server

- The data processing by AEROCOM

- Data policy and condition of use

- Description of report to NILU upon the use of the data by the AEROCOM community

NILU

On the use of situ data non-regulatory data (aerosol + trace gases) variables since 1. april 2011

1092 users from 49 Countries, have downloaded ca 450 000 data sets data since 1. april 2011 (ACTRIS, GAW-WDCA, EUSAAR)

% of unique users

	0 - 1
	1.1 - 5
	5.1 - 10
	10.1 - 15
	15.1 - 20
	20.1 - 25
	25.1 - 30
	30.1 - 35
	35.1 - 40

NILU

Summary

EBAS Data base:

- ✓ http://ebas.nilu.no

Information about meta data for most instruments and methods:

- ✓ http://ebas-submit.nilu.no

Flagging system:

- ✓ http://www.nilu.no/projects/ccc/flags/index.html

Access to large data volumes:

- ✓ ebas@nilu.no

Advice? Contact

- ✓ Kjetil Tørseth: kt@nilu.no
- ✓ Cathrine Lund Myhre: clm@nilu.no
- ✓ Markus Fiebig: mf@nilu.no

Thank you!

Cathrine Lund Myhre: clm@nilu.no

Why cannot NILU re-distribute data from the database?

No re-distribution of data because:

- Not "our data", associated to programs
- No proper acknowledgment
- Data base is dependent on **trust** to continue to receive data
- Data can be misused and misinterpreted

Solution:

Develop mechanisms/interface for getting data to keep better control of the use. This will also avoid duplications and several versions of same data available and distributed. Not good for anyone..

Data sharing agreements/protocols

NILU