

# Structure and Content of the EARLINET Data Set.

Jens Bösenberg and Holger Linné  
Max-Planck-Institut für Meteorologie, Hamburg, Germany

## Distributed data base:

In the EARLINET project the data were collected by the participating institutions rather than in a central data base. This was done on purpose, because the measurement and evaluation procedures had to be developed and refined while the data set was collected. All data had to be reprocessed repeatedly. These refinements were initiated by requests from users of the preliminary data. The access to preliminary data was a necessary to start the evaluation and interpretation as early as possible and to iterate the process as often as required.

Additionally data were collected, actually "mirrored", at a central location, MPI Hamburg, to ease access and for intermediate documentation. The most recent data set and a "consolidated" data set are always available in suitable form (as archive files or on CD-ROM).

## File content:

Each file contains one of the following:

- a backscatter profile
- an extinction profile plus a backscatter profile with the same height resolution
- a series of backscatter profiles (for an extended period of time)

There are separate files for each station, each observation period (nominally 30 min) and each combination of transmitted and received wavelengths. A common nomenclature is used to designate the file content.

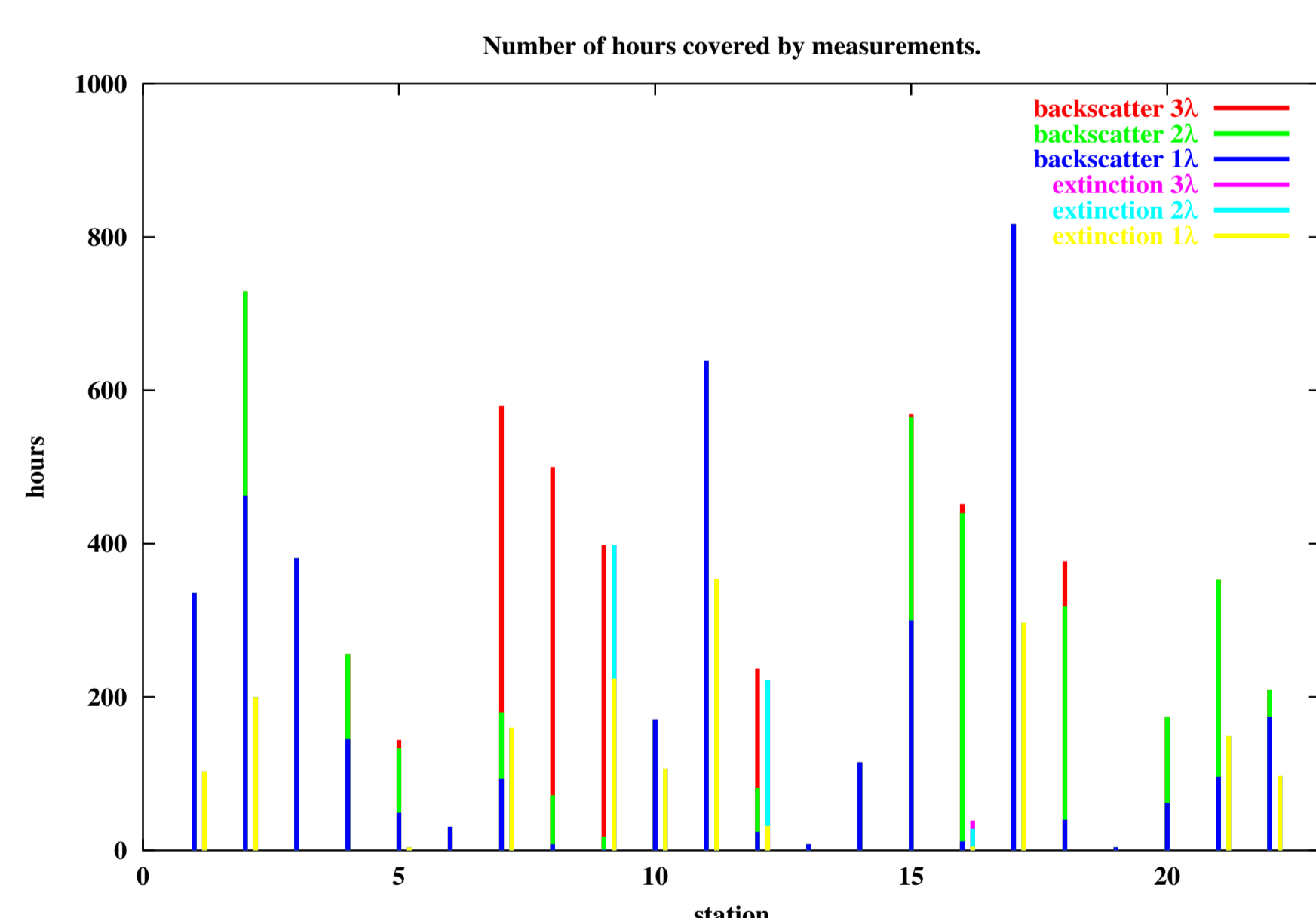
## Data files sorted by categories:

To ease finding of files related to studies specific topics category files have been created containing the file names of profiles addressing these topics. The following table provides an overview over the number of files addressing the various topics:

Category	Number of files
All	13530
Cirrus	382
Climatology	4906
Diurnal cycles	1651
Etna eruption	108
Forest fires	174
Photosmog	99
Rural-urban	568
Saharan dust	1985
Stratosphere	74

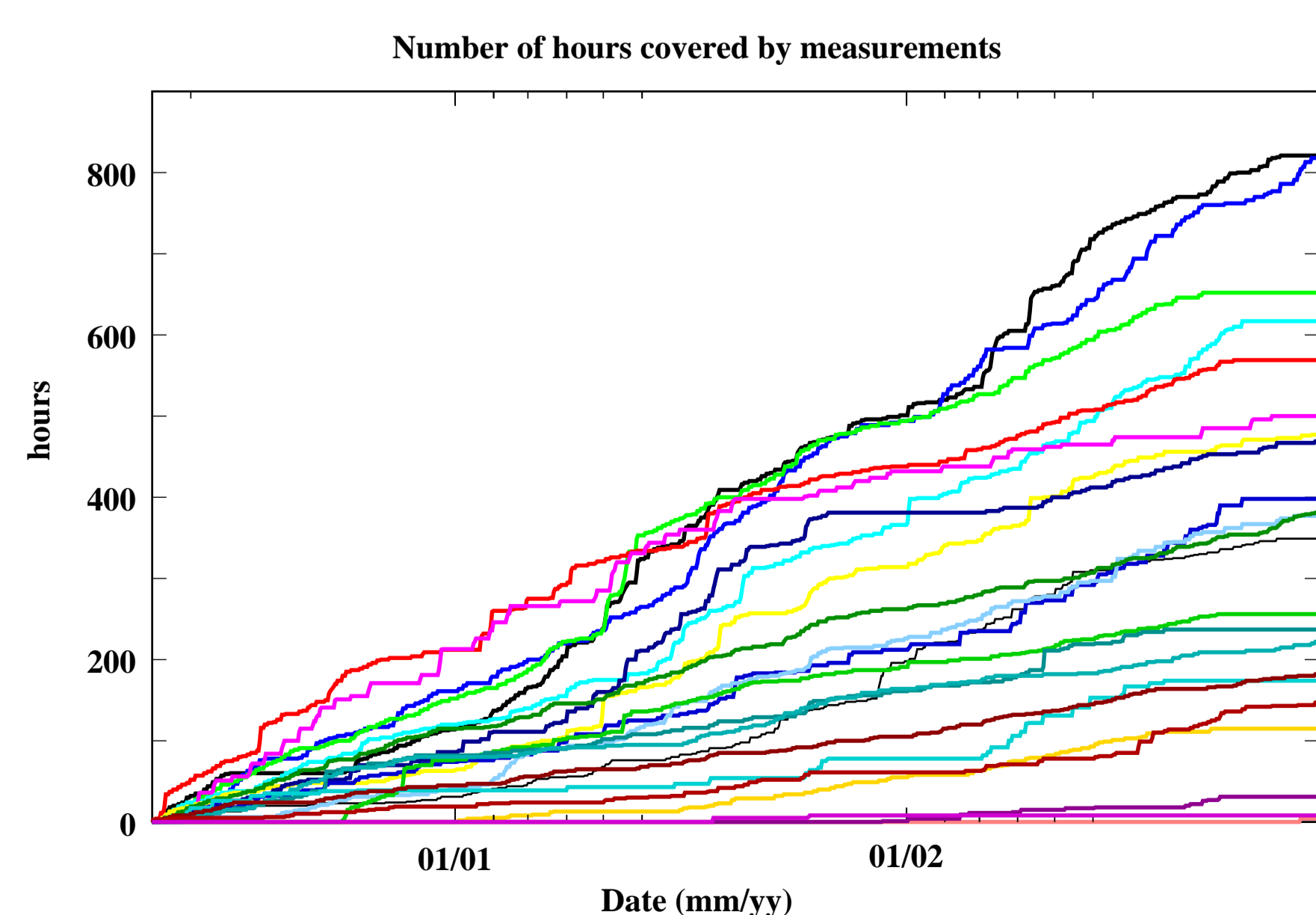
## Data files sorted by station and detection channels:

The following diagram provides an overview over the number of hours covered by measurements at the different stations (arbitrary order, including temporary stations), sorted by number of backscatter and extinction channels. The inhomogeneous coverage in space, time, and channels is obvious.



## Temporal development of the data set:

The number of observations at the different stations is illustrated in the following figure. The steady growth for all stations (although at different pace) is clearly demonstrated.



## File format:

Files are written in the netcdf format. Advantages of this format are that it is:

- self-describing (through header information)
- flexible (don't care about properties you don't use)
- open for later additions (old programs continue to work with the new files)
- platform independent (for all standard operating systems)
- professionally designed and maintained by NCSA
- including read and write procedures for standard programs
- available at no cost (for research institutes)
- widely used

Mandatory header information:

```
=====  
netcdf headermandatory {  
  dimensions:  
    Length = UNLIMITED; // (5 currently)  
  variables:  
    float Altitude(Length);  
      Altitude:AltitudeUnits = "m";  
    float Backscatter(Length);  
      Backscatter:BackscatterUnits = "1/(m*sr)";  
    float Extinction(Length);  
      Extinction:ExtinctionUnits = "1/m";  
  
  // global attributes:  
    :System = "MPI UV-system";  
    :Location = "Lindenberg, Germany";  
    :Longitude = "12.345 E";  
    :Latitude = "54.321 N";  
    :Altitude = "123 m";  
    :Emissionwavelength = "351 nm";  
    :Detectionwavelength = "351 nm";  
    :Detectionmode = "analog";  
    :Zenithangle = "0";  
    :Shotaverages = "10000";  
    :Rangeresolutionraw = "15 m";  
    :Altituderesolutionevaluated = "60 m up to 1000m, 500 m up to 5000 m";  
    :Startdate = "19970701";  
    :Starttime = "13:00:00 UT";  
    :Stoptime = "13:30:00 UT";  
    :Evaluationmethod = "Fernald";  
    :Input_parameters = "reference value 1.e-7 1/(m*sr) at 6800 m";  
    :Comments = "do not believe these dummy data!";
```

Optional header information (including the mandatory):

```
=====  
netcdf headeroptional {  
  dimensions:  
    Length = UNLIMITED; // (5 currently)  
  variables:  
    float Altitude(Length);  
      Altitude:AltitudeUnits = "m";  
    Altitude:LongName = "Height above sea level";  
    float Dustlayerheight;  
      Dustlayerheight:AltitudeUnits = "m";  
    Dustlayerheight:LongName = "Top of dust layer above sea level";  
    float Mixinglayerheight;  
      Mixinglayerheight:AltitudeUnits = "m";  
    Mixinglayerheight:LongName = "Top of mixing layer above sea level";  
    float Backscatter(Length);  
      Backscatter:BackscatterUnits = "1/(m*sr)";  
    Backscatter:_FillValue = 777.777;  
    float Extinction(Length);  
      Extinction:ExtinctionUnits = "1/m";  
    Extinction:_FillValue = 777.777;  
    float Depolarisation(Length);  
    Depolarisation:_FillValue = 777.777;  
    float ErrorBackscatter(Length);  
      ErrorBackscatter:BackscatterUnits = "1/(m*sr)";  
    ErrorBackscatter:_FillValue = 777.777;  
    float ErrorExtinction(Length);  
      ErrorExtinction:ExtinctionUnits = "1/m";  
    ErrorExtinction:_FillValue = 777.777;  
    float ErrorDepolarisation(Length);  
    ErrorDepolarisation:_FillValue = 777.777;  
    float Rayleigh_Extinction(Length);  
    Rayleigh_Extinction:_FillValue = 777.777;  
    float WaterVaporMixingRatio(Length);  
    WaterVaporMixingRatio:WaterVaporMixingRatioUnits = "g/kg";  
    WaterVaporMixingRatio:_FillValue = 777.777;  
    float ErrorWaterVapor(Length);  
    ErrorWaterVapor:WaterVaporMixingRatioUnits = "g/kg";  
    ErrorWaterVapor:_FillValue = 777.777;  
  
  // global attributes:  
    :System = "MPI UV-system";  
    :Location = "Lindenberg, Germany";  
    :Longitude = "12.345 E";  
    :Latitude = "54.321 N";  
    :Altitude = "123 m";  
    :Emissionwavelength = "351 nm";  
    :Detectionwavelength = "351 nm";  
    :Detectionmode = "analog";  
    :Zenithangle = "0";  
    :Shotaverages = "10000";  
    :Rangeresolutionraw = "15 m";  
    :Altituderesolutionevaluated = "60 m up to 1000m, 500 m up to 5000 m";  
    :Startdate = "19970701";  
    :Starttime = "13:00:00 UT";  
    :Stoptime = "13:30:00 UT";  
    :Evaluationmethod = "Fernald";  
    :Input_parameters = "reference value 1.e-7 1/(m*sr) at 6800 m";  
    :Comments = "do not believe these dummy data!";
```

## Access of data:

Access will be granted according to the EARLINET data protocol agreement. The full text is available at <http://lidarb.dkrz.de/earlinet>.

## Important rules are:

1. All scientists involved in EARLINET have equal and complete access to the database produced during the project.
2. If data from other partners within EARLINET are to be used in a publication, joint authorship should be offered.
3. Each Principal Investigator has the right to refuse that his data be used in a publication if either the data is suspected of having insufficient quality or if the originator plans to use his work in a separate publication.
4. Use of data by third parties is possible only by written consent of all partners that have contributed to the data set to be used.
5. Data provided by third parties are subject to the dissemination rules of the provider.
6. All data will be made freely available 2 years after the end of the contract. This does not affect the intellectual property rights as expressed above in paragraphs 2 to 5.
7. Access to the EARLINET data base is possible only after having signed the protocol.

Requests should be addressed to Dr. Jens Bösenberg, [boesenberg@dkrz.de](mailto:boesenberg@dkrz.de)

*The support of this work by the European Commission under grant EVRI-CT-1999-40003 is gratefully acknowledged.*