

## Introduction

NILU has set up an interactive database for observational data through the FP5 EU project RETRO (REanalysis of the TROpospheric chemical composition over the past 40 years). A core activity of the project is a 40-year simulation by four global chemical transport models and the collection of historical observations will be used for extensive validation of model output. The observational dataset is coherently formatted using a common scheme for metadata descriptions and automatic validation routines may therefore easily be applied. The collection is furthermore suitable for assessment of trends on its own.



Figure 1: The Retro web portal for observational data. A login account is required to enter the restricted part of the web pages and a user may have access to data from one or several projects.

## Description of the technical system

The database is set up as an automatic system that handles all upload and download of data through a web-interface. The database architecture is based on the ESA Campaign Database that has been used for archiving and dissemination of ENVISAT Cal/Val data since 2002. Data are converted to and archived in the HDF4.1r3 format, and the metadata definitions from the ENVISAT Cal/Val activity are reused and expanded for Retro. Information is provided on-line and data are stored in a standardized format. Tools for reformatting of data are provided through the web portal. An overview of the database architecture is given in figure 2.

The database provides:

- Documentation
- On-line metadata tables
- Data conversion tools
- Data upload facilities
- Visualisation tools
- Search interfaces
- On-line visualisation of data
- Data retrieval
- Secure data access

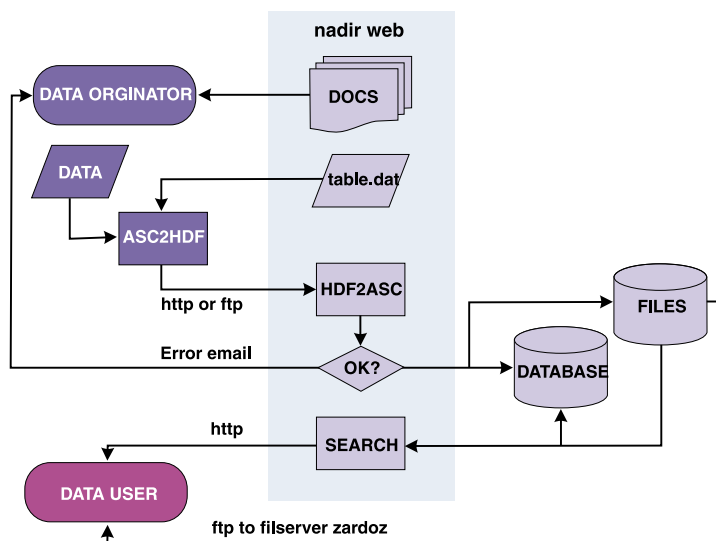


Figure 2: Schematic diagram of data flow from data originator (dark purple modules) to file collection and index database and back to data user (pink module).

## Collection and reformatting of data

A major task for NILU was the reformatting of old data into the new common format. This consisted of converting about 11500 ozone sonde files from the NILU NADIR data centre. These were mainly from European stations. In addition, more than 50000 ozone sonde files from the World Ozone and UV Data Centre (WOUDC / WMO-GAW) were reformatting and uploaded to the database. A simple quality control procedure was applied ozone sonde data to exclude unreliable soundings. All the data from the Climate Monitoring and Diagnostics Laboratory (CMDL / NOAA) data centre were also downloaded and the complete series of CO data and most of the surface O<sub>3</sub> were converted and uploaded.

Finally, a program has been developed to export data from the European Monitoring and Evaluation Programme (EMEP) database on observations. The program is fully generic and features advanced search criteria and options for filtering and sorting data from the monitoring database. All data are automatically exported into HDF files. The program will be used to export O<sub>3</sub>, NO<sub>2</sub>, NH<sub>3</sub>, N<sub>2</sub>O, SO<sub>2</sub>, etc. data from the database. The measurements are quality assured and dates back to the early seventies.

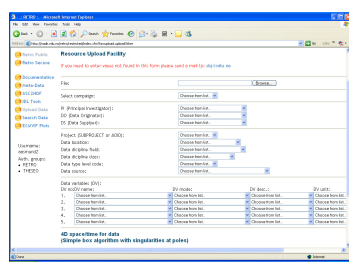


Figure 3: Facility to upload non-HDF data to the Retro database. The user must define a number of required metadata parameters from lists of allowed values through drop-down menus. Files are automatically renamed according to the database file-name conventions and on certain key metadata provided by the user. The file is automatically inserted in a directory structure and registered in the database.

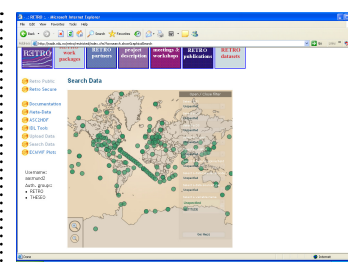


Figure 4: Graphical file search – station mode. A world map shows the location of data from stationary data sources. The user may zoom into the map, filter out data through the built in selection tool and click on the dots to acquire the observations.

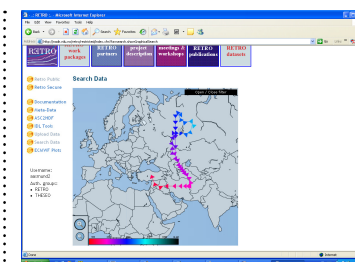


Figure 5: Graphical file search – trajectory mode. The image shows geo-location from moving scalar data sources, in this case only a test-file. The map displays the location, including altitude, of the trajectory. The user may click on the trajectory to acquire the data file.

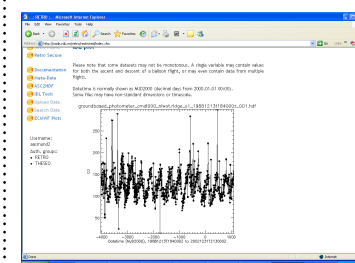


Figure 6: On-line visualisation of data files is supported through the web portal. A data file containing a time series of CO observations from Niwot Ridge (US) (NOAA CMDL) is here plotted by the built-in IDL routines. Contour plots are used for 2D visualisation of data arrays.

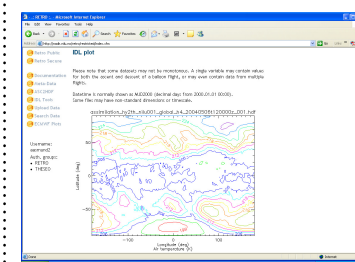


Figure 7: On-line visualisation of data files is supported through the web portal. A data file containing a global stratospheric field of Air Temperature at 475K is displayed by the built-in IDL routines. 2D slices of up to 4D data arrays may be visualised.

## Acknowledgements

RETRO is EU funded project EVK2-CT-2002-00170. The authors would like to acknowledge and thank the NOAA CMDL, WOUDC, EMEP and NADIR data centres for providing access to observational data to the scientific community.

## Conclusions

The database, and its underlying structure, is suitable for archiving and indexing complex data sets from an unlimited number of projects and from a multitude of geophysical sciences in a common context. The database will store model output and observational data, including satellite data, for the Retro project.