A short introduction to the GEOmon and EBAS databases at NILU, available for the CityZen project

Ann Mari Fjæraa and Sverre Solberg

Norsk institutt for luftforskning
Norwegian Institute for Air Research

NILU

A short introduction to the GEOmon and EBAS databases at NILU, available for the CityZen project

Ann Mari Fjæraa and Sverre Solberg

Contents

A short introduction to the GEOmon and EBAS databases at NILU, available for the CityZen project

1   Databases

1.1   Databases available for CityZen

This document describes two of the databases available for the CityZen project; The EMEP database - EBAS and the GEOmon Distributed Data Centre. The databases are accessible through the web and are located at http://ebas.nilu.no and http://geomon.nilu.no

It should be noted that this description is strictly speaking valid at the moment of writing. A data base and it's content is dynamic and will be continuously extended and developed. The basic concepts should hopefully not change too much, though.

1.2   History of the databases

Both databases are developed and operated by the Norwegian Institute for Air Research (NILU).

EBAS was mainly developed during 1995-2000, and has since 2006 been running on an external server. This allows direct queries through a web interface (first version launched in 2008). The functionality for external users is still under development. We encourage users to make direct contact with the database management group (http://www.nilu.no/projects/ccc/about.html) at NILU to get support for export of larger amounts of data in specific formats, calculation of statistics etc. More functionality can be expected in the period to come as we get more experience with the stability of the system.

EBAS is hosting observation data of atmospheric chemical composition and physical properties submitted by data originators in support of a number of national and international programs ranging from monitoring activities to research projects.

The GEOmon Distributed Data Centre (GDC) was built in 2007-2009 in the GEOmon project and offers access to GEOmon-funded and other data stored at NILU, as well as to external databases through queries of metadata.

2   Available data

2.1   EBAS

All data in EBAS are stored at NILU. In January 2010 close to 18000 public and additionally 16000 restricted datasets were stored in the database. Variables available include greenhouse gases (e.g. CH_4, N_2O, O_3, halocarbons, halon, SF_6),

and aerosols/particles from projects and programmes like EMEP, EUSAAR, EUCAARI and GAW. All EMEP stations are background rural stations. All data are from Europe except for sulphur/nitrogen compounds for which data from North America (US and Canada) are also included.

Species of particular relevance for model evaluation in CityZen:

Gases:

O_3: Hourly timeseries at EMEP and a few GAW sites from around 1990 to present. Of the order of 100 stations, less the first years.

NO: Hourly timeseries at 2-25 EMEP sites from 1991 to present. Few sites before 2000.

NO_2: Daily filter samples (24 h samples from 8 to 8) from of the order of 100 EMEP sites from around 1990 to present, less the first years. Hourly data from 15-30 EMEP sites from 1995 to present.

NOx: Hourly timeseries at approx 10 EMEP sites since 2004. Mostly Spain.

SO_2: Daily filter samples (24 h samples from 8 to 8) from of the order of 100 EMEP sites from around 1980 to present, less the first years.

HCHO + other aldehydes/ketones: Regular samples at 5-10 EMEP sites since 1994 (very few stations with data through all years). These are 8-h (8-16) manual samples with normally two samples per week (working days). Mostly Central Europe.

C_2H_4, C_3H_6, isoprene, toluene and many other NMHC: Regular samples at 5-10 EMEP sites since 1992 (very few stations with data through all years). These are manual grab samples in bottles (approx. 20 min. filling time), mostly sampled two times per week (working days). Continuous data (i.e. every day) with two samples per day (midday/midnight) at Hohenpeissenberg in Germany. Continuous data with 2-h resolution from Rigi, Switzerland.

CO: Hourly timeseries from Jungfraujoch since 1996, else hourly data from a few sites since 2006.

HNO_3, H_2O_2, NH_3, N_2O_5, OH: No data

Aerosols:

SO_4: Daily filter samples (8 to 8) at 100-700 EMEP + North American sites since 1980, increasing number with time. $PM_{2.5}/PM_{10}$ resolution for some of the sites.

NH_4: Daily filter samples of [NH_3(g) + NH_4(s)] at 75-600 EMEP + North American sites since 1980, increasing number with time. $PM_{2.5}/PM_{10}$ resolution for some of the sites.

NO_3: Daily filter samples of [HNO_3(g) + NO_3(s)] at around 100 EMEP sites since 1980, increasing number with time.

SOA and EC/OC: Campaigns 2002/03 and 2006/07 from some EMEP sites. Some data may not yet be public available. Could be used upon request.

Aerosol data (phys/chem properties) from the projects EUSAAR (surface sites) and EUCAARI (both airborne and surface sites) are available and the time series

are presently being complemented. Longer time series for certain sites (e.g. Ispra) may be included if requested by CityZen.

Met.data:

Data from a few countries/years only.

2.2 GEOmon

Datasets made available through GDC are not physically stored in a common database, but provided through links to external databases. Data are global or regional depending on the various programs.

Databases currently available through the GDC

EBAS: EMEP Database (described above), including mainly in-situ measurements from European stations. Focus is on long-range transport of air pollution in Europe, but datasets from other themes and geographical regions are also available.
http://ebas.nilu.no

ESA-CDB: European Space Agency – Campaign Database, currently including mainly atmospheric composition measurements from non-satellite instruments used for calibration and validation (cal/val) of Envisat.
http://nadir.nilu.no/cdb

EARLINET: European Aerosol Research Lidar Network, consisting of 20+ stations throughout Europe providing routine observations of tropospheric aerosol load and properties. Data from EARLINET must be downloaded through the network data portal, currently available at http://svn.earlinet.org/EARLINET
http://www.earlinet.org

AVDC: Aura Validation Data Centre, a NASA database containing non-satellite measurements used for calibration and validation of the NASA Aura satellite.
http://avdc.gsfc.nasa.gov/

NDACC: Network for Detection of Atmospheric Composition Change, containing mainly groundbased remote sensing and balloonborne observations of upper tropospheric and stratospheric atmospheric composition. Focus is traditionally on assessing trends in stratospheric ozone and related substances, but the scope of the network has broadened the last years.
http://www.ndsc.ncep.noaa.gov/data

GAWSIS: Data from WMO-GAW data centre are searchable.
http://gaw.empa.ch/gawsis/

Databases that possibly will be made available through the GDC in the near future

GEOmon Rapid Delivery data (preliminary GEOmon funded data)
ftp://ftp.nilu.no/pub/GEOmon

EEA NRT AQ data (preliminary surface ozone measurements in Europe)
http://www.eea.europa.eu/maps/ozone/map

AIRBASE (an official Air Quality database in Europe, includes urban/industrial areas)
http://air-climate.eionet.europa.eu/databases/airbase/

ICOS/IMECC/RAMCES (CO2 and CH4 data throughout Europe)
http://www.icos-infrastructure.eu/

Of particular relevance for model evaluation in CityZen:

GAWSIS (Surface monitoring data from global remote background sites) – CO, O_3, NO_2 etc.

AIRBASE: EEA's public data base for European Air Quality. Several thousands of surface monitoring sites, including urban/sururban/rural, traffic/industrial/background. O_3, NO_2, PM. Few data before 1995. Most from the latest years.

The issue of representativeness of sites and the selection of a "quality assured, harmonized, and regionally representative data set" for surface monitoring sites is discussed here: http://www.geomon.eu/science/act2/SciAct2_CHE.html

3 How to access the data

The description below shows how to extract and download data from the two data bases EBAS and GEOMON. It should be mentioned, however, that users may presently find ebas rather slow for extracting a selection of many different kinds of data sets. A major upgrade of ebas is foreseen and planned but not yet started. Users wanting a larger set of data may either visit the web page for annual data sets:

http://www.nilu.no/projects/ccc/index.html

and download data sets from there or they are welcome to send a request of data to the data base coordinator (specified on this page) who will help extracting the data sets of interest.

3.1 Data policy

A set of access control procedures has been defined for each of the two databases.

3.1.1 EBAS

The main portal to ebas is through the web interface http://ebas.nilu.no/

EBAS hosts data from a number of international programmes and projects, and the project association to the datasets is used to control the access to data for users. The three different types of access control procedures currently in place is listed below, together with a short description of the various projects and their availability. Any enquiries on how to get access to restricted data can be addressed to ebas@nilu.no

The annual data sets (http://www.nilu.no/projects/ccc/index.html) are available in ASCII text files either in ISO-7168 format or in plain column based files. The

files extracted from the online web service (http://ebas.nilu.no/) may be downloaded in either NASA-Ames 1001 or in "csv" text files.

Public datasets
Datasets associated with projects having a "public" data policy are accessible without login/password. The user will need to approve the general rules for use of data including how to make acknowledgements etc. The public data policy complies with the rules used by international programmes like EMEP, WMO-GAW and others. The data policy disclaimer is shown when a user first enters the web interface and there is a need at the end of the policy document to confirm that the terms are approved. The confirmation needs renewal on a monthly basis. The requirements can at any time be considered through "read disclaimer" link in the login area.

Restricted datasets with simplified procedure
Some projects apply a more restrictive data policy than the "public datasets", but still provide access to data without the need for an approval of users on an individual basis. To access data from such projects the user needs to approve the respective data protocol. The user will then be provided a personal user account for login. Signed protocols are archived at NILU for later reference.

Restricted datasets with approval process
In cases where projects require a restrictive data policy, access is only granted after an individual approval process involving an authorized representative of the respective network/project. Contact persons at NILU are given on the same page.

3.1.2 GEOmon

The main portal to the Geomon data base is through http://geomon.nilu.no/

As the datasets made available through GDC are not physically stored in a common database, the individual databases are to be considered the main archives of the data – GDC simply provides a common portal to access several databases simultaneous.

Access to some or all of the datasets from the contributing databases are restricted, but core metadata that allows search are made freely available (and thus made visible when you search for data) through GDC. This is not always the case for the contributing databases, and you may therefore find more datasets through GDC than you will do through the individual database portals. Information on how to acquire access to restricted data is made available in the data list section (ref. section 3.2.2 below).

Data are available as ASCII text files in NASA-Ames 1001 format or in the format specified by each individual archive.

3.2 How to browse/download data

3.2.1 EBAS

Figure 1 shows the main page of the EBAS web. Lists for selecting nation, station, project, instrument, component and/or matrix are available at the top of the page. A document containing information about its content is available for each list. Below the lists there is a section holding the time span selection, and buttons for plotting the stations on a map and listing data. The login section and a document on data policy and restrictions are found on the right hand side. Additional resources supporting the database are available below the login section. In Figure 2 ozone data from Denmark and Finland are selected as an example. The lists are updated whenever a new variable is selected, and multiple variable can be selected by holding down the Ctrl button. The number of available datasets is shown on the right hand side, below the listing windows.
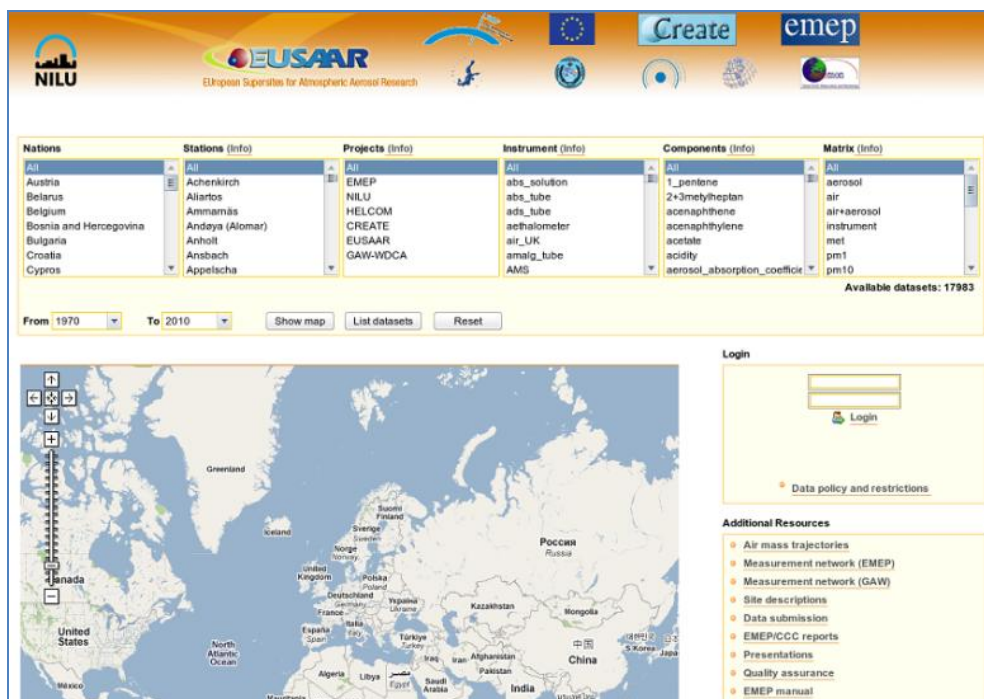
Figure 1: The main page of EBAS.
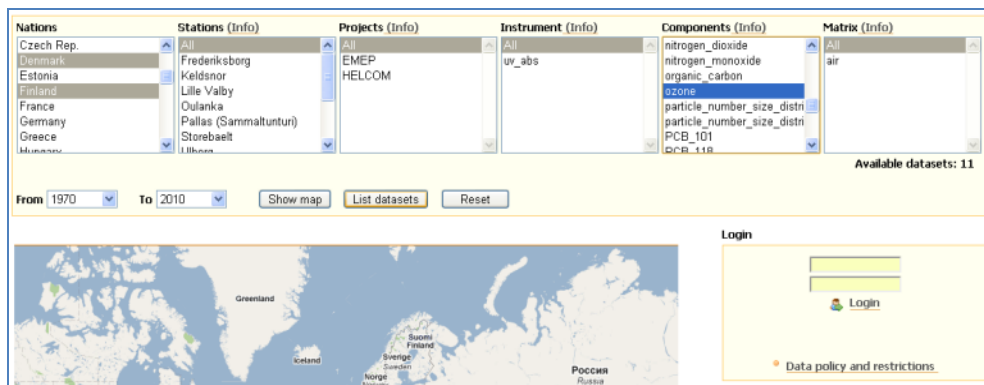
Figure 2: Selecting datasets.

After pressing the button for "list datasets", a page as shown in Figure 3 will appear. At the top of the page there is a section for time span, display settings (plotting and simple statistics) and download. The selected variables are listed with metadata (e.g. station and time period). To the left of each dataset there is a checkbox and a small icon. By pressing the small icon the data values will be listed on the screen. By marking the checkbox the selected dataset(s) can be plotted or downloaded. By e.g. selecting the first dataset, timespan one year (e.g. 2006) and pressing 'plot' in the display settings section, the time series plot will appear on the screen. When downloading the data, make sure that the correct time span is selected, the dataset might not have any data for det default value (current year). The downloaded data is stored as a zip file on your local computer. The user may at any time return to the main page by clicking on the "home" icon in the upper left corner or the use reset button below the variable lists.

Figure 3: Listing datasets

3.2.2 GEOmon

The GEOmon Distributed Data Centre is located at http://geomon.nilu.no. The file menu on the left hand side takes the user to the general web pages of the GEOmon project. The GDC web page contains lists for selecting component, location, database, type, platform or matrix. A document containing information about the content is available for each list. Figure 4 gives an example of how to select and list datasets. The lists are updated automaticly when a variable is selected. Multiple variable can be selected by holding down the Ctrl buttom. The number of available datasets is shown on the right hand side, below the listing windows. The 'list datasets' button is located in the lower left corner of the page. Metadata details are available for each dataset when listing the data.
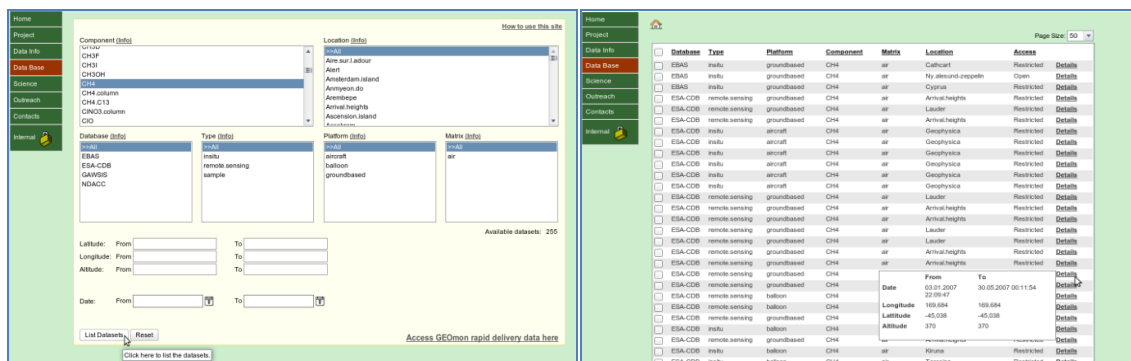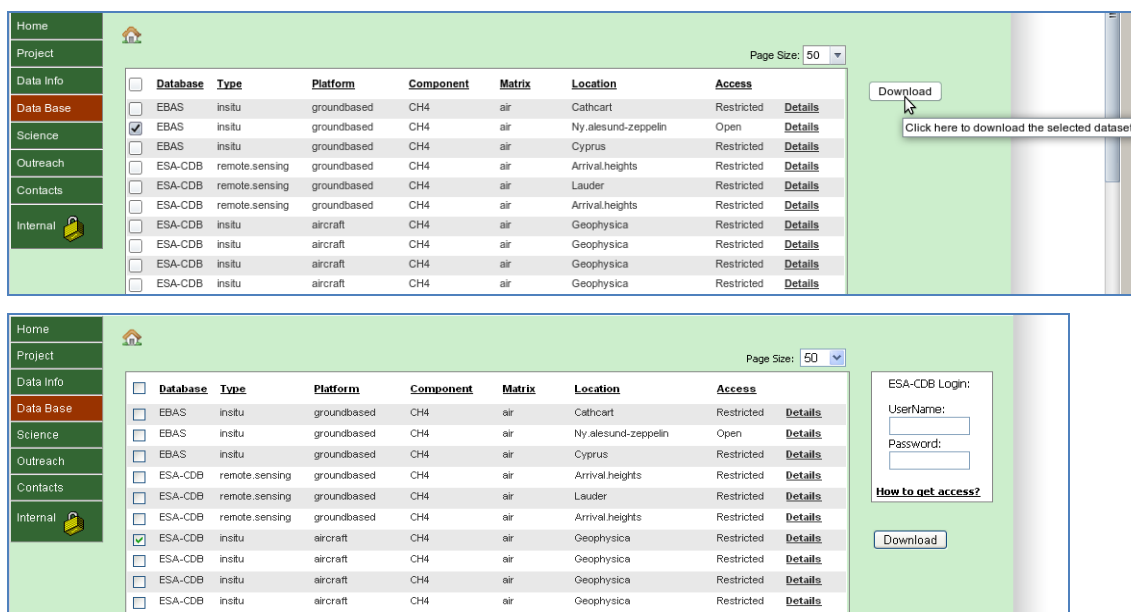
Figure 4: Selecting and listing datasets.

Figure 5: Downloading data.

Datasets that have open access can be downloaded by selecting the checkbox to the left of the dataset(s) and then clicking the download button on the right hand side, as shown in the upper panel of Figure 5. When selecting datasets with restricted access the user will be prompted a login window to the underlying external database, as shown in the lower panel of Figure 5.

The user may at any time return to the main page by clicking on the "home" icon in the upper left corner or use the reset button below the variable lists. The downloaded data is stored as a zip file on your local computer.

4   Acknowledgement

REPORT SERIES Technical report	REPORT NO. TR 3/2010	ISBN: 978-82-425-2200-9 (printed) 978-82-425-2201-6 (electronic) ISSN: 0807-7185	
DATE	SIGN.	NO. OF PAGES 9	PRICE NOK 150.-
TITLE A short introduction to the GEOmon and EBAS databases at NILU, available for the CityZen project		PROJECT LEADER Sverre Solberg	
		NILU PROJECT NO. U-108155	
AUTHOR(S) Ann Mari Fjæraa and Sverre Solberg		CLASSIFICATION * A	
		CONTRACT REF.	

REPORT PREPARED FOR The partners of the EU FP 7 project CityZen

ABSTRACT This technical report gives an overview of the data bases EBAS and GEOmon and explains how to extract data of relevance for the CityZen project.

NORWEGIAN TITLE En kort innføring i NILUs databaser EBAS og GEOmon tilgjengelig for prosjektet CityZen.		
KEYWORDS Database	CityZen	EBAS

ABSTRACT (in Norwegian) Denne tekniske rapporten gir en oversikt over databasene EBAS og GEOmon og forklarer hvordan man kan laste ned data som er relevante for prosjektet CityZen.

* Classification   A   *Unclassified (can be ordered from NILU)*
                     B   *Restricted distribution*
                     C   *Classified (not to be distributed)*

NILU is an independent, nonprofit institution established in 1969. Through its research NILU increases the understanding of climate change, of the composition of the atmosphere, of air quality and of hazardous substances. Based on its research, NILU markets integrated services and products within analyzing, monitoring and consulting. NILU is concerned with increasing public awareness about climate change and environmental pollution.

NILU is an independent, nonprofit institution established in 1969. Through its research NILU increases the understanding of climate change, of the composition of the atmosphere, of air quality and of hazardous substances. Based on its research, NILU markets integrated services and products within analyzing, monitoring and consulting. NILU is concerned with increasing public awareness about climate change and environmental pollution.

NILU Norsk institutt for luftforskning
Norwegian Institute for Air Research