

The EuMetNet Data Rescue portal

Ingeborg Auer (ZAMG), Barbara Chimani (ZAMG), and members of the EUMETNET Expert Team on data rescue^{*})

Introduction

Long time-series are essential for climate research. They allow for analyses of former climate, calculating reanalyses and model evaluations. Time series reaching back to the end of the 18th / beginning of the 19th century are of great importance as a connection between the instrumental period and paleoclimatology. For the best possible resolution in time and space it is essential to use all the available data. For reliable information on climate variability and extreme events, long time series are essential.

But for analyses with high spatial resolution over Europe, the available data are sufficient only since about 1960. Although different "Data Rescue Programs" (e.g. <http://www.climatol.eu/DARE>) have been exercised, there is still a lot of data stored in paper-based archives and therefore not available for climate research but exposed to possible destruction and oblivion. This poster summarizes the results of a EUMETNET Data Rescue Workshop held in Budapest during the 8th Seminar for Homogenization and Quality Control in Climatological Databases in May 2014 (http://www.met.hu/en/omsz/rendezvenyek/homogenization_and_interpolation/programme/).



Austria:
Nearly half of the Austrian data is already digitised and fully quality controlled. But still about 20 % of the data is yet to be digitised. The remainder (about 31%) of the data are at different stages of quality control.



Ireland:
Data before 1961 (temperature) and 1941 (precipitation) has still to be digitised for non-synoptic stations. This means about 150-200 stations for precipitation (some of them very short) and 70-100 stations for temperature. Catalogue for stored daily and monthly precipitation data is finished.



Slovakia:
This initiative includes rescue of metadata and data, a graphical data overview for time-series duration, and quantification of unknown data history on daily and hourly measurements. Quality control is done with a "Graphical interface for controlling and revision of climatological data". Data series are prolonged by combining the data from different stations locations (e.g. caused by relocation). Still the greater part of the historical data is in paper form, but first attempts using character-recognition software have been done.



Great Britain:
Data rescue is still at its beginning, and the different methods of automated versus manual are being considered. Additionally a cataloguing project for getting an overview of the available data is done. The main effort is in retrieving enough digitised data to create gridded monthly datasets.



Finland:
Rescue of metadata is an important issue. Nearly all of the data since 1960 is already digitised, and SYNOP and CLIMATE data since the beginning of the time series. Digitising of data for precipitation stations into a temporary database is underway (if digitising progress goes on with the same speed as until now this will take more than 20 years from now). Quality control is done by a manual control during the digitizing process, and data will undergo automatic quality control afterwards. There is so far no intention of digitising hourly data from aviation and co-operated stations.

Aims:

EUMETNET-Dare complements world-wide running programs and is not intended to duplicate existing work. Therefore the following milestones were defined for EUMETNET-DARE:
1) data inventory of already digitised data and data still only available in paper archives
2) enhancing the data on digitised data sets

For a follow-up project the following tasks are planned:
3) homogenisation of the datasets
4) publishing the data in relevant data-bases

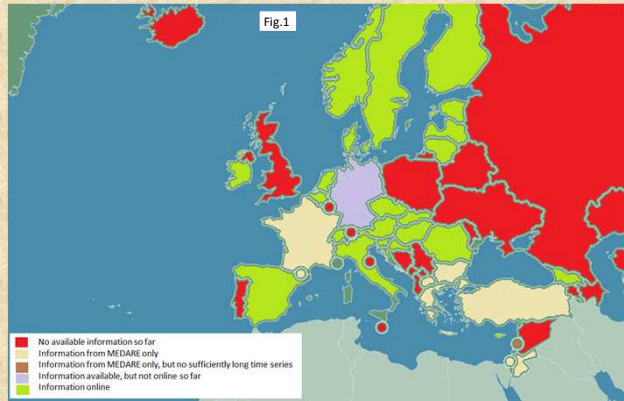
If you know about some long-term time series available in your country or stored in your archive and you can't find this information on the DARE-Homepage, please let us know about it (dare@zamg.ac.at)
If you would like to fill in the questionnaire, please feel free to download it from the EUMETNET-DARE-Homepage.

Activities:

To fulfil task 1 of the current project a questionnaire was sent to all the meteorological services of the WMO region VI (Fig.1). The questionnaire contained questions on the digitised and non-digitised data with a length of ~100 years (~50 years for mountain stations), but also questions concerning the time and amount of money needed for digitising the data and the willingness to provide the data to an international database.

In addition to the information given by the meteorological services, the information displayed on the MEDARE-website (<http://www.omm.urv.cat/MEDARE/>) were included into the available information sheets on the EUMETNET-DARE webpage (<https://www.zamg.ac.at/dare/>). Detailed information on the available data can be found in Fig. 1. As indicated, not all of the countries have returned the questionnaire; and not all of the countries' information can be found in the MEDARE-Database, as this project was focusing on the Mediterranean area. An example of the information given on this website is shown in Fig.2.

The spread in the progress in regard to data rescue in the different countries is quite big. While some of the countries have already digitised all or most of their data and are keeping them in their own database or even make them available via their webpage, other countries are lacking the capacity to start a data rescue progress at the moment. For some of the participating countries their progress in data rescue is displayed on this poster.



Czech Republic:
Data rescue has been done for 10 years, with different progress in different regions. All data since 1961 is already digitised. It is estimated that about 95% of the long-term series have already been completely digitised. But data quality control is still ongoing and will take several years. Problems concerning data distribution might be solved with a new law, perhaps by 2020.



Romania:
Data and metadata are addressed in the effort of data rescue. Sources are the national meteorological archives and also other documents from the library. The data is archived in a web-based (internal) database, and the digitisation of the metadata is also done via a web interface. For some of the stations, for which it is known that long time-series should exist, it was not possible to recover the whole dataset.



Hungary:
The majority of the data from climate stations since 1961 has already been digitized. Earlier measurements are stored on microfilms, microfiches and in paper format. 15 long-term stations for temperature and additionally more than 100 for precipitation (from 1901 and 1 mountainous station from 1952) are digitised, homogenized, completed and quality controlled. The digitization is a permanent process, besides the recording of precipitation on a regular basis. Data rescue took place partly within different projects; recently about 1 522 780 records of sunshine duration and climate data were digitized in the CarpatClim-Project.

Fig. 2

name	temporal resolution	start	end	data (year) (year)	data (T) (year)	start	end	data (month) (year)	data (precip) (year)	start	end	data (year) (year)	data (precip) (year)	start	end	data (year) (year)	data (precip) (year)	start	end	data (year) (year)	data (precip) (year)	start	end	data (year) (year)	data (precip) (year)				
Innsbruck-Ulm	daily, monthly	1877	active	yes	no	1882	active	yes	no	1877	active	yes	no	1877	active	yes	no	1877	active	yes	no	1922	active	yes	no	1916	active	yes	no
München-Melk	daily, monthly	1860	active	yes	no	1860	active	yes	no	1900	active	yes	no	1860	active	yes	no	1860	active	yes	no	1900	active	yes	no	1860	active	yes	no
Salzburg	daily, monthly	1874	active	yes	no	1874	active	yes	no	1880	active	yes	no	1874	active	yes	no	1874	active	yes	no	1860	active	yes	no	1860	active	yes	no
Sonneberg	daily, monthly	1884	active	no	no	1884	active	no	no	1890	active	no	no	1884	active	no	no	1884	active	no	no	1880	active	no	no	1880	active	no	no
Freising	daily, monthly	1890	active	no	no	1890	active	no	no	1900	active	no	no	1890	active	no	no	1890	active	no	no	1880	active	no	no	1880	active	no	no
Kremsmünster	daily, monthly	1874	active	no	no	1874	active	no	no	1874	active	no	no	1874	active	no	no	1874	active	no	no	1887	active	no	no	1887	active	no	no
Wien-Neudorf	daily, monthly	1872	active	no	no	1872	active	no	no	1872	active	no	no	1872	active	no	no	1872	active	no	no	1880	active	no	no	1880	active	no	no
Grac-Ulm	daily, monthly	1894	active	no	no	1894	active	no	no	1894	active	no	no	1894	active	no	no	1894	active	no	no	1911	active	yes	no	1896	active	yes	no
Hochob	daily, monthly	1881	1944	no	no	1880	1944	no	no	1880	1944	no	no	1880	1944	no	no	1880	1944	no	no	1881	1944	yes	no	1924	1944	no	no
Kanzenbach	daily, monthly	1938	active	yes	partly	1938	active	yes	partly	1938	active	yes	partly	1938	active	yes	partly	1938	active	yes	partly	1938	active	yes	partly	1938	active	yes	partly
Villacher Alpen	daily, monthly	1921	active	yes	no	1921	active	yes	no	1921	active	yes	no	1921	active	yes	no	1921	active	yes	no	1931	active	yes	no	1931	active	yes	no
Faßberg	daily, monthly	1936	active	yes	no	1936	active	yes	no	1936	active	yes	no	1936	active	yes	no	1936	active	yes	no	1936	active	yes	no	1936	active	yes	no
Langeoan Airfield	daily, monthly	1952	active	no	no	1952	active	no	no	1952	active	no	no	1952	active	no	no	1952	active	no	no	2007	active	no	no	2007	active	no	no
Schöckben	daily, monthly	1848	active	yes	no	1847	active	yes	no	1847	active	yes	no	1847	active	yes	no	1847	active	yes	no	2007	active	no	no	2007	active	no	no
Garm	daily, monthly	1958	active	no	no	1958	active	no	no	1958	active	no	no	1958	active	no	no	1958	active	no	no	1996	active	no	no	1996	active	no	no
Kufstein	daily, monthly	1846	active	yes	no	1846	active	yes	no	1846	active	yes	no	1846	active	yes	no	1846	active	yes	no	1992	active	yes	no	1992	active	yes	no
Landsack	daily, monthly	1866	active	no	no	1866	active	no	no	1895	active	no	no	1866	active	no	no	1866	active	no	no	1980	active	no	no	1897	active	no	no



Spain:
All data from about 1961 are already digitised. Scanning of the longest manuscript data reports (~1860-1960) and digitisation of monthly precipitation and temperature data for this period is also finished. Current efforts, limited by staff shortage, include imaging of radiosonde data from before 1970 and records on paper strips, as well as the digitisation of additionally found data. The next tasks will be imaging and cataloguing all other data documents and analogue recordings, and their digitisation.